Research Surveys in Linguistics

Lexicalization and Language Change

Laurel J. Brinton and Elizabeth Closs Traugott

CAMBRIDGE

CAMBRIDGE more information - www.cambridge.org/9780521833103

This page intentionally left blank

Lexicalization and Language Change

Lexicalization, a process of language change, has been conceptualized in a variety of ways. Broadly defined as the adoption of words into the lexicon, it has been viewed by some as the reverse process of grammaticalization, by others as a routine process of word formation, and by others as the development of concrete meanings. In this up-to-date survey, Laurel Brinton and Elizabeth Traugott examine the various conceptualizations of lexicalization that have been presented in the literature. In light of contemporary work on grammaticalization, they then propose a new, unified model of lexicalization and grammaticalization. Their approach is illustrated with a variety of case studies from the history of English, including present participles, multi-word verbs, adverbs, and discourse markers, as well as some examples from other Indo-European languages. As a first overview of the various approaches to lexicalization, this book will be invaluable to students and scholars of historical linguistics and language change.

LAUREL J. BRINTON is Professor in the Department of English, University of British Columbia. She is author of *The Structure of Modern English: A Linguistic Introduction* (2000), and *Pragmatic Markers in English: Grammaticalization and Discourse Functions* (1996).

ELIZABETH CLOSS TRAUGOTT is Emerita Professor of Linguistics and English at Stanford University. Her most recent books include *Grammaticalization* (with Paul Hopper, 2003), and *Regularity in Semantic Change* (with Richard Dasher, 2002).

Research Surveys in Linguistics

In large domains of theoretical and empirical linguistics, the needs of scholarly communication are directly comparable to those in analytical and natural sciences. Conspicuously lacking in the inventory of publications for linguists, compared to those in the sciences, are concise, single-authored, non-textbook reviews of rapidly evolving areas of inquiry. The series Research Surveys in Linguistics is intended to fill this gap. It consists of well-indexed volumes that survey topics of significant theoretical interest on which there has been a proliferation of research in the last two decades. The goal is to provide an efficient overview and entry into the primary literature for linguists – both advanced students and researchers – who wish to move into, or stay literate in, the areas covered. Series authors are recognized authorities on the subject matter as well as clear, highly organized writers. Each book offers the reader relatively tight structuring in sections and subsections and a detailed index for ease of orientation.

Previously published in this series

A Thematic Guide to Optimality Theory John J. McCarthy
ISBN 0 52179194 4 hardback
ISBN 0 52179644 X paperback

The Phonology of Tone and Intonation Carlos Gussenhoven
ISBN 0521 81265 8 hardback
ISBN 0521 01200 7 paperback

Argument Realization Beth Levin and Malka Rappaport Hovav
ISBN 0521 66331 8 hardback
ISBN 0521 66376 8 paperback

Lexicalization and Language Change

LAUREL J. BRINTON

University of British Columbia

ELIZABETH CLOSS TRAUGOTT

Stanford University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo

Cambridge University Press
The Edinburgh Building, Cambridge CB2 2RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521833103

First published in print format 2005

ISBN-13 978-0-511-13221-6 eBook (EBL)
ISBN-10 0-511-13221-2 eBook (EBL)

ISBN-13 978-0-521-83310-3 hardback
ISBN-10 0-521-83310-8 hardback

ISBN-13 978-0-521-54063-6 paperback
ISBN-10 0-521-54063-1 paperback

Contents

Figures

Tables

Preface

In the 1990s as historical studies of grammaticalization proliferated and questions arose about the relationship between it and lexicalization, we independently sought to understand better to what extent efforts to maximize the distinctions between the two were justified. At the International Conference on English Historical Linguistics in Santiago de Compostela, September 2000, we discovered that we had somewhat similar concerns and similar ideas, most especially that we were both embracing the realization that what we had polarized (see Hopper and Traugott 1993, 2003; Traugott 1994; Brinton 2002, and, to a lesser extent, Traugott 2005) were in fact very similar in certain respects. Having taken criticisms in Cowie (1995) to heart, Traugott was also concerned about the status of derivation in grammaticalization and lexicalization. Meanwhile, it became clear that many others were making similar efforts to account for the similarities as well as differences between the two processes (e.g., Lehmann 1989, 2002; Ramat 1992, 2001; Wischer 2000; Heine 2003b). The diversity of points of view on the two topics has been a matter of frustration to some, but we view it as an inevitable step in the development of relatively new subfields of linguistics, much as has occurred in the study of syntax or morphology.

Consistent with the aims of this series, *Cambridge Research Surveys in Linguistics*, our purpose in this book is to bring together a variety of scholarly debates concerning the relationship between lexicalization and grammaticalization in language change, with focus on the former. For this reason, the first three chapters present reviews of the literature, which in the case of lexicalization especially contains varied and often conflicting views on how this process is to be conceived. In the last three chapters, we suggest some ways in which these views may be reconciled and present one possible unified approach to lexicalization and grammaticalization. This book is addressed in the first instance to graduate students and established scholars in the field and assumes a general understanding of issues related to diachronic linguistics, and to grammaticalization studies in particular. However, we believe that it could also be used by advanced undergraduates who have a solid grounding in basic linguistics.

In a comparative work on lexicalization and grammaticalization of this nature, it has been necessary to omit a number of aspects of both phenomena

that are of potential interest. For example, we have had little space to discuss the phonological dimension of lexicalization. Moreover, although we have attempted to cover recent research on lexicalization and grammaticalization, we realize that much else may have been done that has not come to our attention. No doubt far more is currently in progress. In particular, we have, for reasons of time and resources, restricted our coverage primarily to work on and in English, with passing reference to other European languages. Therefore, a general understanding of the historical development of English is assumed in the work. Much of relevance has, no doubt, been written on other languages and in other languages. We hope that, despite these limitations of coverage, this volume will provide guidance and inspiration for those who wish to pursue the matter further, especially with reference to non-European languages.

In writing this book we have had to let go of old preconceptions and revise our thinking about lexicalization and grammaticalization; we would like to think we have encouraged others to do so too. We are grateful to Paul J. Hopper, Anette Rosenbach, Scott Schwenter, and Jacqueline Visconti for comments on an earlier draft as well as to three anonymous reviewers of our initial proposal. Isla Reynolds provided careful editorial attention to the manuscript. We would also like to thank Christina Bartels and Kate Brett at Cambridge University Press, who initially conceived of this project with us, and Helen Barton and Alison Powell, who carried the project through, as well as Jacqueline French for copy-editing.

Laurel J. Brinton, Vancouver
Elizabeth Closs Traugott, Berkeley
August 2004

List of abbreviations

ABL	ablative case
ACC	accusative case
Adj	adjective
Adv	adverb
Aux	auxiliary verb
COMP	comparative
Conj	conjunction
Dan.	Danish
DAT	dative case
Det	determiner
Du.	Dutch
EME	Early Middle English
EModE	Early Modern English
Eng.	English
F	feminine
Fr.	French
FUT	future tense
GEN	genitive case
Gk.	Greek
Gm.	German
Gmc.	Germanic
GRAM	grammatical morpheme
HCET	*Helsinki Corpus of English Texts*
Hit.	Hittite
ICAME	*International Computer Archives of Modern English*
IE	Indo-European
INF	infinitive
It.	Italian
Lat.	Latin
LModE	Late Modern English
M	masculine
ME	Middle English
MFr.	Middle French

MHG	Middle High German
ModE	Modern English
ModGm.	Modern German
N	noun
NP	noun phrase
OE	Old English
OED	*Oxford English Dictionary*
OFr.	Old French
OHG	Old High German
ON	Old Norse
PAST	past tense
PDE	Present-day English
PGmc	Proto-Germanic
PIE	Proto-Indo-European
PL	plural
Port.	Portuguese
Prep	preposition
PrepP	prepositional phrase
PRES	present tense
PrP Adj	present participial adjective
PrP Prep	present participial preposition
Prt	particle
PTCP	participle
SG	singular
Sk.	Sanskrit
Sp.	Spanish
Sw.	Swedish
V	verb
1	first person
2	second person
3	third person

Abbreviations of OE texts follow the conventions of the *Dictionary of Old English Corpus*; abbreviations of ME texts follow the conventions of the *Middle English Dictionary*. When citing Old English and Latin we have omitted length marks. In the case of citations from other languages, we have retained them.

1

Theoretical contexts for the study of lexicalization and grammaticalization

1.0 Purpose of the present study

(1) We are celebrating a fascinating holiday today.

is something we might well say to a visitor from abroad, and not think twice about whether *holiday* and *today* or the *-ing* of *celebrating* and of *fascinating* function differently from the point of view of our knowledge of language. However, linguists, grammarians, and others who study and think about language, how it is structured, how we come to know it, and how it changes concern themselves with just such questions. In the introduction to the *Concise Oxford Dictionary of Linguistics*, Matthews says: "Everyone will agree that linguistics is concerned with the lexical and grammatical categories of individual languages" (1997:vi), and this is what our example in (1) is about: *holiday*, *celebrate*, and *fascinating* are usually regarded as "lexical," members of large, "open" classes of forms that are relatively infrequently used and express relatively concrete meaning, while *we*, *are*, and *a* are regarded as "grammatical," members of smaller, relatively "closed" classes of forms that are very frequently used and express relatively abstract meaning. Moreover, *today* is not clearly a lexical or a grammatical form, having partially concrete and partially abstract meaning, and belonging to a rather large set of adverbs. Finally, the *-ing* of *celebrating* and the *-ing* of *fascinating*, although seen as originating in the same grammatical form, are generally understood as having developed differently over time, the former remaining grammatical and the latter becoming lexical. What these differences mean, how this kind of distinction plays out in language change, and what research questions it suggests are among the topics of the present book.

In recent years questions have frequently been raised about the relationship between "lexicalization" and "grammaticalization." The two terms, like many other linguistic terms, have been used to refer ambiguously to phenomena viewed from the perspectives of relative stasis ("synchrony") or of change over time ("diachrony"), to the process and to the results of the process, and also to theoretical constructs modeling these phenomena. According to Lehmann (1995 [1982]:6), the first formulation of an opposition between lexicalization and grammaticalization was Jakobson's (1971 [1959]) characterization of the first as optional, the second as obligatory. Since then, they have been theorized in a number of different ways, sometimes totally independently of each other, sometimes together. One constant in all these uses is pairing of meaning and form, and the extent to which this pairing is systematic or idiosyncratic. The starting point of the present work is to bring together a variety of scholarly debates concerning this relationship in language change, with focus on lexicalization, which has been studied far less systematically than grammaticalization.

The first three chapters are reviews of the literature; the last three propose some solutions. In this chapter, we will briefly introduce the contexts for the study of lexicalization and grammaticalization, most especially on approaches to grammar, lexicon, language change, lexicalization, and grammaticalization. We will not attempt to resolve the differences of opinion. Chapter 2 focuses in more detail on lexicalization, especially the definitions and viewpoints that have emerged during the last fifty years of work in linguistics. Chapter 3 presents recent arguments concerning the similarities and differences between lexicalization and grammaticalization. Chapter 4 suggests one possible integrated approach to lexicalization and grammaticalization that resolves the major debates about their relationship. Chapter 5 addresses some particular problems in the history of English from the perspective of definitions developed in Chapter 4, and Chapter 6 summarizes the book, ending with suggestions for further directions for research.

1.1 Debates concerning grammar and language change

It is impossible to understand how either lexicalization or grammaticalization have been conceptualized without paying attention to underlying assumptions about grammar and its relationship to the lexicon, as well as underlying assumptions concerning the dynamics of language change. A full investigation of these topics would entail a detailed history of linguistics, especially in the twentieth century. Space allows only for some sweeping generalizations here, which, unfortunately, tend to polarize and to be caricatures. However, without some attention to different foundational assumptions, it is often difficult to make sense of the literature or to propose a possible solution to the many issues that have been raised. In Section 1.1.1 we summarize two extreme approaches to grammar, the polar opposites between which much

linguistics has in actual fact been practiced, but which may help frame the varying discourses about lexicalization and grammaticalization.

1.1.1 Approaches to grammar and lexicon: an overview

Toward the end of the twentieth century it appeared that there were essentially two types of linguists – "generative" and "functional" – who, because they were asking fundamentally different questions, often talked past each other (see Croft 1995, 2001; Newmeyer 1998; Darnell, Moravcsik, Newmeyer, Noonan, and Wheatley 1999; Kemenade 1999; Haspelmath 2000a). Although neither group works with a monolithic view of linguistic theory, the functionalist group is more diverse than the generative.[1]

On the one extreme, most formal, generative linguists since the 1960s have sought to answer such questions as "What is the system of knowledge of language?" or "How does this system of knowledge arise in the mind/brain?" (see, e.g., Chomsky 1988:3). The object of study is language as an innate capacity of the individual. The assumption is that the language capacity is computational and syntactic, and by hypothesis optimally structured and ultimately binary in nature. It is a self-contained modular mechanism that does not reflect external factors such as cultural or social systems. Nor does it reflect experiential structures such as vision or production factors such as frequency (this is known as the hypothesis of "autonomous syntax"). The universals of language that are posited are absolute in the sense that one counterexample disproves them (see, e.g., Newmeyer 1998:263). On this view, the grammar of a particular language, whether Swahili, or English, is an "epiphenomenon" of an intrinsic capacity and is of little interest beyond providing empirical evidence for hypotheses about general capacities. And on this view, such traditional questions in historical linguistics as "How did the category auxiliary develop in English?" are uninteresting, or worse, not sensible (see, e.g., Lightfoot 1979, 1999; Hale 1998).

At the opposite extreme, since the 1970s a group of "functional-typological" linguists have sought to answer the question of how speakers can use the "bricolage" or "jerry-built structure[s]" (Bolinger 1976:1) of language to impart information, and to get things done (see, e.g., Hopper 1988). As well as seeing language as a cognitive capacity, this approach privileges language as a device for communication between speakers and addressees. Crucially the assumption is that there is a causal relationship between meaning and linguistic structure, and furthermore that external factors may shape language structure. Language is a human activity, not an epiphenomenon of a static capacity (see Lehmann 1993:320). The prime object of study is language use and how it relates to the grammars of particular languages, and

[1] Croft (1995) provides a useful and detailed discussion of different subtypes of functionalism.

how grammars may vary cross-linguistically. Universals of language are considered to be tendencies, not absolutes, and are usually of a general cognitive nature, not autonomous and not specific to language.

The turn of the present century has seen the emergence of several possibilities for a meeting of minds, as some generative linguists begin to try to account for cognition-based structures (e.g., Jackendoff 1983, 2002), for productivity (e.g., Jackendoff 2002), for the dynamic, emergent properties of the speaker's knowledge of the system (e.g., Culicover and Nowak 2003), and for the variation that undeniably occurs in language (see work on Optimality Theory, e.g., Boersma and Hayes 2001; Lee 2001; Bresnan, Dingare, and Manning 2002). Moreover, some "functional" linguists have sought to formalize their work at least in part (see, e.g., Bybee and Hopper 2001 for frequency studies; Croft 2001 for syntax).

Common to many, but by no means all, theories is the notion of "grammar" (whether at the abstract level of Universal Grammar, or at the more empirical level of the grammar of a particular language) that is distinct from the notion of "lexicon." If such a distinction is made, "grammar" is the set of categories, patterns, and organizing principles evidenced by language, most essentially abstract patterns of semantics, syntax, morphology, and phonology that at least in theory permit infinite combinations. By contrast, the "lexicon" is a finite list (for any individual) of (more-or-less) fixed structural elements that may be combined. The lexicon is typically a theoretical concept, as distinguished from a "dictionary," which is a practical description. Hence, there is discussion of a "mental lexicon" (an abstraction and idealization), not of a "mental dictionary" (Matthews 1997:s.v. "lexicon").[2]

There have been essentially two views of the relationship of the lexicon to the grammar in generative theory of the last fifty years. The first, which Jackendoff (2002) calls the "syntacticocentric approach," assumes that the lexicon is a list of idiosyncratic items which are selected and inserted into syntactic structures (see various versions of generative syntax from "Standard Theory" [Chomsky 1965] through the Minimalist Program [Chomsky 1995]). Phonological and semantic interpretations are derived from the lexicon together with the syntax. The second, proposed by Jackendoff (1997, 2002), provides an alternative architecture: one in which phonological, syntactic, and conceptual structures are parallel components of the faculty of language, and in which lexical items "establish the correspondence of certain syntactic constituents with phonological and conceptual structures" (Jackendoff 2002:131).[3] A key proposal in Jackendoff's

[2] As we will see, while some theories of the lexicon are roughly equivalent to "vocabulary," many are not, since they include grammatical forms such as past tense *-d*.

[3] Other proposals that treat the lexicon as part of the combinatorial architecture of a complex set of parallel structures include Lexical Functional Grammar (e.g., Bresnan 2001), and various types of Head Driven Phrase Structure Grammar

work is that the lexicon is multistructured and includes not only highly idiosyncratic, but also more regular elements. This is more in keeping with many functionalist views of the lexicon, which point to parallels between lexical and grammatical organization, although the regularities may be considered to belong to morphology rather than the lexicon (see, e.g., Bybee 1985, 1988; Langacker 1987; Haspelmath 2002). A more detailed discussion of Jackendoff's views, with focus on the problem of distinguishing types of lexical categories, follows in Section 1.2.

1.1.2 Approaches to language change

Because lexicalization (and grammaticalization) will here be conceptualized primarily as historical processes subject to normal constraints on language change,[4] we will briefly set out some assumptions concerning language change before turning to a more detailed examination of the conception of the lexicon. While a comprehensive examination of theories of language change is far beyond the scope of this introduction,[5] we will mention here a few factors that will help illuminate the debates over lexicalization and grammaticalization.

Historical linguistics was the focal point of attention in the nineteenth century, during which time many foundational ideas of linguistics were developed, most especially the concepts of structure and pattern. Discovery of such sound laws as Grimm's Law, which showed how the Germanic languages differed systematically in consonant articulation from the other Indo-European languages, and the Great Vowel Shift, which showed how later English differed systematically from earlier English with respect to the place of articulation of the long (later tense) vowels, highlighted the ways in which language phenomena are structured.[6] Work on

(e.g., Pollard and Sag 1994) and Construction Grammar (e.g., Goldberg 1995; Fillmore, Kay, Michaelis, and Sag 2003).

[4] There has recently been some confusion about the term "process," depending on whether the term is used restrictively or not. As Newmeyer rightly points out, grammaticalization is not a "distinct process" in the restrictive sense that it is "an encapsulated phenomenon, governed by its own set of laws" (1998:234). He also acknowledges that "process" is most usually used in a non-restrictive sense as a "phenomenon to be explained" (232). We use the term in a slightly different but even more usual non-restrictive way to focus attention on (a) the need for a dynamic perspective, (b) the micro-steps that are obscured by the ">" typical of representations of change.

[5] For the state of historical linguistics at the end of the twentieth century, see Joseph and Janda (2003).

[6] The term "Great Vowel Shift" appears to be attributable to Otto Jespersen, see Jespersen (1961 [1909–1941]:Vol. I, Chap. 8).

comparative reconstruction of proto-languages was made possible by the crucial insight that while change is inevitable, it is not random.

The advent of "structuralism" in the twentieth century shifted the focus of attention from change in patterns over time to pattern and system as manifested in relative homogeneity and stasis, i.e., synchrony (see especially Saussure 1986 [1916]). Insofar as historical work was done in the earlier part of the century, the focus was typically on comparing synchronic stages of a language, or diachrony. Correspondences or "rules" were usually of the form

(2) A > B

This formulation suggested that the structures themselves change, rather than that the representation of these structures differs over time because each generation of speakers has to learn the language anew and uses it in novel ways. More importantly, the structuralist explanation for change was sought in properties of language and languages, in other words, what was thought of as "internal" or "endogenous" change. The formulation in (2) also suggested abrupt change over time, indeed complete replacement of one item by another. However, change always involves variation: older forms and newer forms coexist side by side, in the same speakers as well as in the same community, and a more appropriate formulation is A > A ~ B > B (Hopper and Traugott 2003:49). Even this is misleading, since often, especially in domains that involve meaning, earlier patterns only become restricted or fossilized, not entirely lost. The typical situation is actually (3) where the emergence of B as the only choice may or may not occur:

$$(3)\ A > \left\{ \begin{matrix} A \\ \\ B \end{matrix} \right\} > (B)$$

By the second half of the twentieth century, considerable attention started to be paid to the questions "What is in the arrow?" and "How does change come about?" In a groundbreaking paper calling for integration of synchronic studies of variation with diachronic work, because synchronic variation is the result of and a necessary condition for change, Weinreich, Labov, and Herzog (1968) proposed that the focus of work should be on language variation and change. On this view, study of diachronic correspondences would take a back seat to the solution of several more important problems. These include:

(a) The constraints problem: What is the set of possible changes and possible linguistic conditions for change? Examples include changes in category status, such as the emergence of a new grammatical category (e.g., article, auxiliary verb), loss of an existing grammatical category (e.g., inflectional case), or chain shifts (e.g., Grimm's Law, the Great Vowel Shift).

(b) The transition problem: What are the intervening stages that define the path by which A gives rise to B (and typically coexists with it for at least a while)? Change proceeds by small steps, not large leaps (although accumulations of changes may have cascade-like effects that lead to more substantial change). Change by small steps will be discussed more fully under "gradualness" in Section 1.4.2.
(c) The actuation problem: How does change start, when and where does it start ("actuation") and how does it spread through the system ("actualization")?

One type of change that is particularly important for our discussion of lexicalization and grammaticalization is "reanalysis." In a foundational paper on the topic of syntactic reanalysis, Langacker defined it as covert change: "change in the structure of an expression or class of expressions that does not involve any immediate or intrinsic modification of its surface manifestation" (1977:58). From this perspective reanalysis involves:

(a) change in constituency, or what goes with what (e.g., change in morphological bracketing of [*a*] *napron* > [*an*] *apron*[7]),
(b) a change in category labels (e.g., main verb > auxiliary),
(c) boundary loss (e.g., *be going to* > *gonna*).

Reanalysis is not restricted to morphosyntax: when a lexeme develops new polysemies (e.g., *silly* 'blessed, innocent' > 'foolish'), it has undergone semantic reanalysis.

Another major type of change is analogy: the generalization of a structure (see Kiparsky 1992). By contrast to reanalysis, analogy is overt, and indeed it is often only through analogy that reanalysis can be detected (see Timberlake 1977). Thus, as we will see, when the motion verb construction *be going to* is reanalyzed as a future auxiliary, evidence for this reanalysis comes from use of *be going to* with verbs that do not ordinarily collocate with motion, e.g., verbs of psychological experience such as *like* or *know*.[8]

Weinreich, Labov, and Herzog (1968) made a key distinction between innovation (changes that happen in the individual) and change (changes that spread to others) (see also Milroy 1992; and discussion in Janda and Joseph 2003). In doing so, they proposed a significantly more social view of change than the generative view with which it competed (e.g., Lightfoot 1979). On the generative view, change is equivalent to innovation and is to be found in differences between cognitive states of individuals. Since the focus is on internalized systems, the interest in change lies in how grammars, i.e., internalized sets of patterns and relationships among patterns, change

[7] *Naperon* 'small tablecloth' is a borrowing from OFr.
[8] Detailed discussion of reanalysis and analogy can be found in Harris and Campbell (1995).

(see, e.g., Kiparsky 1968; Kroch 2001). Syntax is privileged as central and autonomous, not at all or only marginally affected by either semantics or phonology, and the child is privileged as the locus of change, since the small child has to learn the language from scratch.

Toward the end of the twentieth century a dialog developed as "functionalist" theorists sought to greater and lesser extents to integrate the advances in formal linguistics with a perspective on language that paid more attention to construal of meaning and to dynamic aspects of language (e.g., Bybee, Perkins, and Pagliuca 1994; Croft 2000). This was often combined with cross-linguistic work on language typology spearheaded by Greenberg (see, e.g., Greenberg, Ferguson, and Moravcsik 1978). Furthermore, there was a shift away from focus on strictly "internal" change to concern for the role of the speaker in the "basic" two-person interactional dyad, and in the community (see especially Milroy 2003). On this view, not only the child as hearer, but also the adult (especially the young adult) as producer, can be the innovator and therefore the catalyst for change (see, e.g., Haspelmath 1999a). We speak of "language change," yet strictly speaking, this is a misnomer – it is not language in general or a language in particular that changes; rather, communities of speakers develop different representations of a system. Much of the work on grammaticalization developed in this theoretical context. Since this book is about lexicalization and grammaticalization as types of linguistic change, our approach is largely functional-typological in orientation.[9]

For historical linguists of a functionalist persuasion, the object of study is how language systems can change over time, as attested by written textual data (also spoken data since audio-recordings have become available). A major contributing factor in the growth of historical linguistics at the end of the twentieth century has been the advent of computerized corpora, which give easy access to information about linguistic contexts for change, frequency, and other factors important in answering the questions posed by Weinrich, Labov and Herzog. For historical work on English the Helsinki Corpus has been a major source of data (see Rissanen, Kytö, and Palander-Collin 1993).[10] Such corpora for the most part reflect not the language of small children, but the rhetorical practices and strategic interactions of speakers with hearers, and have suggested to many that a theory of language change needs to be usage- or "utterance"-based, paying attention to meaning and discourse function (Hopper and Traugott 1993, 2003; Croft 2000; Traugott and Dasher 2002). Furthermore, corpora have confirmed that

[9] For possibilities of combining strictly formal with functional approaches in historical study, see, e.g., Clark (2004).

[10] Kytö (1996) provides a key to the data in the Helsinki Corpus. This corpus and several other corpora of English written and spoken language are available in the *International Computer Archives of Modern English* (*ICAME* 1999).

change is both generational (linked to language acquisition) and communal (linked to networks of speakers [Nevalainen 2004:16; Bergs 2005]).

1.1.3 Summary

To summarize up to this point, we have seen that there are opposing conceptions of grammar: one as a self-contained module guided by a set of language-specific and absolute universals operating independently of contextual factors, and the other as a set of general cognitive tendencies strongly shaped by language-external influences. Minimally, these different approaches admit a distinction between "grammar" and "lexicon." These two conceptions of grammar are paralleled by opposing approaches to language change: one in which change is abrupt, complete, and language internal and may be equated with innovation (from one generation of speakers to another), and the other in which change depends upon variation, proceeds gradually, and is shaped by linguistic and social factors. Studies of grammaticalization and lexicalization have generally been carried out following the latter model of language change.

1.2 Concepts of the lexicon

As a synchronic component of the language faculty, the LEXICON is understood broadly as a finite list of stored forms and the possibilities for combining them. A full conceptualization of this component of language involves distinguishing "lexical category" from "grammatical category" (whether this is a binary opposition or a gradient one). We start this section by introducing some concepts of the units that are stored in the lexicon (1.2.1), then move on to distinctions between categories in the lexicon (1.2.2), and end with some issues concerning gradience and coalescence (1.2.3) and productivity (1.2.4).

1.2.1 The units of the lexicon: holistic vs. componential approaches

A LEXICAL ITEM is the type of unit which belongs to a lexicon, but what kind of unit this is will depend on the theory of the lexicon adopted with respect to whether the units in it are wholes or consist of components of meaning.

Some theories have treated lexical items as unanalyzable wholes (see, e.g., Bloomfield 1933; Chomsky 1965). Lehmann speaks of them as being accessed "holistically": "the holistic approach is to directly grasp the whole without consideration of the parts" (2002:2). Treating an item holistically "means treating it as an entry of the inventory, as a lexical item" (2002:3). However, other theories of the lexicon have been based on the concept of minimal components of meaning. Individual lexical items are the language-particular representations of such components (e.g., *boy* represents + HUMAN, – ADULT, + MALE). The components do not

represent properties of the world (reference), but rather innate properties of the mind that determine the way in which the world is conceived. For example, Fillmore says: "the ultimate terms of a semantic description I take to be such presumably biologically given notions as identity, time, space, body movement, territory, fear, etc." (1970:111).[11] Such semantic components reflect systematic relations that hold among items in the vocabulary of languages and can be used to compare cross-linguistic correlations between meaning and form, known as "lexicalization patterns," and to make generalizations about constraints on these correlations. For example, in the case of a "lexical field," such as verbs of perception, Viberg (1983) proposes that there is a lexicalization hierarchy on two dimensions. One has to do with the number of lexical distinctions available for the verbs of the senses. For example, English distinguishes agentive and experiential *look at*, *see* or *listen*, *hear*, but other languages like Hindi do not, and have only one verb for *look at*/*see* and another for *listen*/*hear*. There are no languages with one verb for *look at*/*see*, and two separate verbs for *listen* and *hear*. On the second dimension, which concerns verbal complexity, "[i]f *look at* (or in a few cases *see*) is expressed by a morphologically complex form, so is *listen to* (or *hear*). But the opposite does not necessarily hold" (Viberg 1983:136). Such theories of lexical components and most especially more recent theories about the way they combine to form lexical representations of complex meanings play an integral role in discussions of "synchronic lexicalization" (see 1.3.1 below). They assume that there are universal semantic components – which being universal are not learned – and various language-specific combinatorial possibilities.

Note that a different term, LEXEME, is also sometimes used. While a lexical item may be understood as any member of the lexicon (whether primarily lexical or primarily grammatical), a LEXEME is typically contrasted with a grammatical morpheme, or GRAM such as PL (= plural) (Bybee, Perkins, and Pagliuca 1994). A lexeme refers to a word considered as an abstraction such as RUN rather than as its various concrete WORD FORMS, or grammatical modifications, such as *run*, *runs*, *ran*, *running*, or FOOT rather than *foot*, *foot's*, *feet* (Matthews 1997; Haspelmath 2002). From this perspective, the lexeme FOOT is a "stem," and plural (the vowel alternation in this case) is a grammatical affix.

[11] See also, with various proposals, Katz and Fodor (1963), McCawley (1968), Bierwisch (1970), Voyles (1973), Leech (1981 [1974]), Gruber (1976), and Lyons (1977). Initial theories about components of meaning have been developed more fully in the light of questions about interfaces between meaning and syntax in, e.g., McCawley (1968), Jackendoff (1983, 2002), Talmy (1985, 2000), Wierzbicka (1985), Dowty (1991), Levin (1993), Levin and Rappaport Hovav (1995), Pustejovsky (1995), among others.

There are different ways of conceptualizing the relationship between units in the lexicon and the combinatorial possibilities they allow. One view is that there are basically two mechanisms in the mind, two ways of knowing: a "memory system" that stores and retrieves individual words, and a "system of symbolic computation" that generates grammatical combinations of words. More specifically, the lexicon, which stores irregularities such as *eat–ate*, *foot–feet*, is distinct from the grammar, which provides a set of regular syntactic and phonological rules of combination such as we find in *walk–walked*, *dog–dogs* (Pinker 1999). Here lexical and grammatical units can be sharply distinguished (according to regularity rather than function, however). Others point out that not all combinatorial possibilities are regular. Some combinations, especially those that have to do with the make-up of the lexicon, are so specific to individual items and constructions that they should be accounted for in the lexicon in terms of "lexical rules." For example, *take (someone) to task* is relatively frozen compared to *take someone to New York* and allows a passive, but not substitution of *to task* by other noun phrases; the past tense of *bring–brought* has only a partial phonological connection with the regular past tense *-d*. Anything stored in long-term memory (including forms like lexical *take to task* and grammatical *-d*) should be considered a "lexical item" (see, e.g., Sag and Pollard 1991; and, for a considerably different view of the lexicon, Jackendoff 2002).

1.2.2 Categories of the lexicon: distinctions between lexical and grammatical categories

As we have seen, the formal units of the lexicon are lexical items. Semantically, the lexicon is said to express "lexical meaning." The adjective LEXICAL has a wider meaning than the nominal terms "lexical item," "lexeme," or "lexicon." Lehmann observes that "lexical" can mean "1) belonging to the inventory, 2) having a specific, concrete meaning" (2002:14). The adjective GRAMMATICAL likewise has wide meaning; it can refer to (1) conforming to the rules of the grammar, (2) having an abstract, structural/functional, or indexical meaning.[12] There are a number of technical concepts in linguistics that highlight these rather different meanings. For

[12] Indexical meanings are those that point to aspects of context. Such contexts may be external to the speaking event (e.g., *that* in *Jill wants that*) or internal to the utterance (*it* in *Jill dropped the pen and picked it up*). They may also involve speaker attitude, such as assessment of where some contentful item should be placed on a value scale (focus particles, e.g., *even* in *Even the NGOs left*), assessment whether a proposition refers to an event prior to, simultaneous with, or expected to be later than time of utterance (tense), or indication of whether a term is intended to be understood as new/non-recoverable or recoverable in the discourse (articles, cf. *a boy*, *the boy*). A useful term is "procedural" (see Blakemore

example, "lexical diffusion" refers to the gradual spread of phonological changes across the vocabulary of a language; here "lexical" implies inventories of lexemes. "Grammaticality" refers to the well-formedness or acceptability of an utterance and relates to the first sense of "grammatical." By contrast, "lexical category" refers to the (relatively) concrete meanings expressed by nouns, verbs, and adjectives and is assumed to contrast with "grammatical category," which refers to the (relatively) abstract, functional meanings expressed by determiners, auxiliaries, complementizers, etc.

The distinction between lexical and grammatical categories is part of the general theory of word classes (or "parts of speech"), which has its roots in the work of Greek and Roman grammarians. Traditionally definitions of the parts of speech have been either notional, relying primarily on meaning (e.g., "Nouns denote persons, places, and things") or a combination of meaning and function (e.g., "Adjectives denote qualities or quantities and modify nouns"). Notional definitions have long been known to be problematic, because semantic classes can be expressed in a number of different ways. In an early attempt to correlate notional and structural, especially morphological, properties of language, Sapir (1920) made a broad distinction between:

(a) Concrete concepts: objects, actions, qualities, which are normally expressed by independent words and "radical elements" (roots), or, if slightly less concrete, by derivational elements.
(b) Relational concepts: more abstract concepts, normally expressed by affixes (e.g., the tense markers *-s*, *-d*), by changes internal to a root (e.g., *run–ran*), or, at the most abstract level, by word order (e.g., *Kim should leave now*, *Should Kim leave now?*).

Lexical categories are largely associated with Sapir's first class, grammatical categories with the second. But the proposed correlations between meaning and form are, as Sapir fully recognized, far from replicable cross-linguistically. While conceptual categories such as PAST, FUT (= future), CAUSE, NEGATION may be universal, one of the basic observations about languages is that inter- and even intra-linguistically, languages often differ with respect to whether a particular meaning is encoded lexically or grammatically, or both. Thus, in English FUT can be expressed by a lexical phrase such as the adverbial phrase *three days from now* (in *I am giving a reading from my novel three days from now*, note the present tense of the verb and the relative referential explicitness of the adverbial phrase) or by a grammatical

1987); however, since this term is specific to Relevance Theory and implies a very particular theory of semantics and pragmatics, it is not used here. Note that some indexicals refer (e.g., pronouns) but do not have contentful meaning (compare *the boy* and *he*).

marker such as the auxiliary *will* (in *I will be giving a reading from my novel*, where futurity is represented explicitly by *will*, but is referentially inexplicit).

In grappling with the problem of parts of speech, much recent generative work has focused on syntactic definitions. Four major universal categories are usually posited: (N)oun, (V)erb, (Adj)ective, and (Prep)osition.[13] These categories are called "lexical" because they dominate lexical items that are semantically "whole," but they are conceptually syntactic features. N, V, Adj, and Prep can enter into complex phrases (NP, VP, AdjP, PrepP). The main motivation for this proposal is to demonstrate the similarities in phrasal structures involving the four categories; for example, according to early versions of the theory, each may be "specified" by a functional category such as determiner, auxiliary, or intensifier (*the* for many nouns, *very* for many adjectives, *right* for many prepositions, e.g., *right in back of the bus*). The four "lexical categories" contrast with "functional categories" like determiners that are not semantically rich, are stressless under most circumstances, and are members of a minor, closed class (see below) (Napoli 1993).

However, few criteria are given to determine how to classify particular items in particular languages as one or other of these parts of speech (Croft 2001). Furthermore, as has often been pointed out, the behavior of these lexical categories may differ from one language family to another. For example, Adjs are treated like Ns in many European, North African, and Australian languages, but like Vs in the languages of North America, East and Southeast Asia (see, e.g., Dixon 1982; Lehmann 1990 discusses the structural consequences for the grammar of absence of the category Adj in a language). In addition, the four major categories are often not specifiable for a particular language. This has led to challenges from a number of functional-typological linguists, some of whom suggest that labels like N, V, Adj, and Prep are categories of particular languages, but are not universal. For example, Bhat (2000) suggests that languages can be distinguished according to the number of open classes of categories that have independent morphological characteristics, or are "lexicalized": three (nouns, verbs, adjectives) or two (nouns and verbs); in the case of the latter, further distinctions can be made according to whether adjectives pattern with nouns or with verbs. Even the universality of N vs. V has been challenged; for example, Sasse (1988) questions whether Iroquoian languages have nouns on the grounds that supposed Ns take a prefix similar to one found on verbs and may have temporal properties such as punctual. This particular question has been answered in detail in Mithun (2000), where it is shown that there are clear formal differences between Ns and Vs in Iroquoian languages. However, because morphological Vs may be

[13] Prep is more appropriately understood as "Adposition," since whether this is a preposition or postposition depends on the word order of the language in question.

used as names for entities, they function as Ns, e.g., *akya:t:ɻ:tǫhk* 'we are related/my cousin.' Although many morphological verbs have been lexicalized, i.e., frozen as Ns, the N–V distinction remains robust (Mithun 2000:419).

The functional-typological solution to the problem of parts of speech focuses not on syntactic but on cognitive, conceptual structures.[14] In this tradition, only N, V, and Adj (not Prep) are considered to be fundamental to linguistic structure. Langacker (1987) takes the position that although some basic cognitive structures are universal, grammars, including semantic structures, are language-specific. Language is a symbolic structure with various interrelated subcomponents; "[t]here is no meaningful distinction between grammar and lexicon. Lexicon, morphology, and syntax form a continuum of symbolic structures" (Langacker 1987:3). A noun is a symbolic structure for what is conceptually a thing, i.e., a concept conceived statically and holistically; a verb is a symbolic structure for a concept that is construed as relational and mentally viewed across time; an adjective (or other modifier) is a symbolic structure for a concept construed relationally, but scanned holistically. The relationship between the construals and their symbolic representations is established conventionally in particular languages. Likewise, grammatical morphemes are symbolic structures (Langacker 1991:3), albeit often more abstract and "skeletal" semantically, and usually non-referential. Using an approach compatible with Langacker's, but different in allowing for universal semantic parameters, Croft (1991, 2001) suggests that parts of speech should be regarded not as categories of particular languages, but rather as typological prototypes on which speakers draw in constructing a grammar.

Some linguists prefer to speak not of lexical (major) classes and grammatical (minor, functional) classes, but of open and closed classes. According to Talmy (2000, I:22), open classes are those that are "quite large and readily augmentable relative to other classes" and closed classes are those that are "relatively small and fixed in membership." For him open classes are Ns, Vs, and in those languages that have them, Adjs. All other linguistic forms are closed class; these include determiners, prepositions, adverbs, particles, and some intonation patterns, such as question intonation. Such a division naturally raises the question of how large an open class is, and how to think about more fine-grained distinctions such as count vs. mass nouns (e.g., *knife*/*two knives* vs. *warmth*/**two warmths*), or transitive vs. intransitive verbs (e.g., *hit* vs. *come*). Talmy's solution is to treat such categories, and also "subject," "indirect object," etc., as covert closed class subtypes; these are "design features" of language and cannot easily be added to. One difficulty with treating Adjs as open class members is that, according to Dixon (1982), in some languages they actually belong to a closed class insofar as

[14] For various proposals, see, e.g., Hopper and Thompson (1985), Langacker (1987), Croft (1991), or Pustet (2003).

morphosyntactic criteria are concerned. Also, there is a continuum between open and closed classes, and membership in any one class may be more or less prototypical. For example, in those languages in which adverbs are of many types and in which there is a large inventory of Adv forms, e.g., English, there is a range from relatively open class, derived adverbs such as *fortunately* to relatively closed class monomorphemic adverbs such as *now*, *just* (as in *just three people came*) (Ramat and Ricca 1994, 1998).

Despite general consensus that some kind of distinction needs to be made between lexical and grammatical classes, where to draw a boundary between them, or whether there is a boundary at all, continues to be a matter of some heated debate. The claim in generative theories such as Government and Binding theory that Prep is "lexical" but in theories of grammaticalization that Prep is "grammatical" serves as a key example of some of the ways in which studies of lexicalization and grammaticalization have run into incompatabilities, especially when basic assumptions have not been laid out. Furthermore, it has long been observed that just as some case relations are more syntactic than others (e.g., subject and object are more syntactic than dative and instrumental), so prepositions are in fact not homogeneous with respect to their syntax and semantics and are better analyzed as comprising both lexical and grammatical subsets (e.g., Lehmann 2002). For example, we need to distinguish between grammatical prepositions like *of* and more lexical ones like *with*. Note that *of* is a "default" preposition in English: in *the painting of the artist*, *of* expresses subject or object or possessive relationships (the artist painted the picture, someone painted a picture of the artist, the painting belongs to the artist), whereas in *the city of Rome* it expresses none of these. By contrast, *with* is more restricted in meaning, as accompaniment, or use, but does not express the fundamentally syntactic relationships of subject and object.[15]

1.2.3 Discrete vs. gradient categories

In an attempt to develop theories of Universal Grammar and of individual grammars as optimal systems, generative linguists have tended to treat categories of grammar as binary, discrete entities, and lexical expressions of those categories as unique. A lexical item with the syntactic feature N is precisely an N; an item with the feature Auxiliary (Aux) is precisely an Aux. It will not be more or less of an N, more or less of an Aux, more or less prototypical of its category. As an idealization, each individual N or Aux (or member of a category) will have all the morphosyntactic behavioral properties of that category; exceptions are accounted for in terms of special semantic or phonological cooccurrence restrictions.

[15] For a detailed study of contemporary uses of *with*, see McKercher (2001).

This idealization was challenged early on in Ross (1972) who argued that N, Adj, V should not be regarded as discrete categories; instead, they should be thought of like the cardinal vowels, what we would now call "prototypes" or idealized clusters of behavioral properties (see, e.g., Rosch 1978; Taylor 1997 [1989]), categories that are distinguished by degree rather than kind (Ross 1972:326). Individual items will show more or fewer of the criterial behavioral characteristics. Like Ross, most functional theorists have been concerned with the gradience not only within but between categories.[16] Some Ns are "more nouny" than others; e.g., *house* has more nominal properties than *home* (e.g., *go to the house*, **go house*, *go home* [*go to the home* has a special meaning, like *to the nursing home*]). Some members of Aux are more auxiliary-like than others, e.g., *must* has more Aux properties than *ought to* (*You mustn't smoke, ?You oughtn't to smoke*).[17] Sometimes degree of categoriality depends synchronically on variety; speakers of British English treat *have to* more like an Aux than US speakers (Brit. *Have you to leave*?, US *Do you have to leave?*). Denison (2001) discusses several examples involving *fun*, *key*, and *designer* showing that they exhibit gradience. For example, in *That was great fun*, *fun* functions as N. Since Ns can modify other Ns (as does *stone* in *The stone wall*), *fun* in *Those were fun times* can be regarded as an N. But since this syntactic slot between Determiner (Det) and N is also available for Adjs, there is potential ambiguity between *fun* functioning as N or as Adj in this position. *Fun* here is gradient between N and Adj, according to Denison. However, it clearly functions as Adj in *Now let's think of someone fun*, and *That was very fun*, *The funnest evening*. Gradience can be considered a factor motivating change, and also as the outcome of changes in usage (see above, 1.1.2).

1.2.4 A continuum of productivity

The notion of a continuum is also related to another phenomenon distinguishing lexical and grammatical categories, namely, the feature of productivity. Broadly construed, PRODUCTIVITY is a design feature of language that allows speakers to produce novel combinations, or it is the speaker's statistical readiness to produce them (see, e.g., Aronoff 1976; Kastovsky 1982; Plag 1999; Jackendoff 2002). Productivity is "fundamentally concerned with the ability of a speaker to produce new forms" (Bauer 1994:3357) and is equated with the frequency, or relative frequency, of a form. In a language like English, the most productive (or "default") items are grammatical (plural

[16] For a critique arguing that gradience within categories is more significant than between categories, see Aarts (2004), and, for historical research questions arising out of this claim, Section 6.2.2.

[17] These properties are often called the NICE properties (behavior in negation, inversion, "code," and emphatic affirmation) (see, e.g., Palmer 1988:14ff.).

inflection for the noun, third singular present tense and past tense inflections for the verb), whereas the least productive ("idiosyncratic") items are certain lexical formatives (such as the derivational prefix *be-* in *befriend*) and most lexical items (roots or stems). Although word formation processes such as clipping, conversion, or blending allow for the production of new lexical items from existing lexical items (see Chapter 2), these processes are relatively unproductive compared to "rule-governed" grammatical processes (Bauer 1994:3356). It is important to recognize that productivity is a "gradient concept" (Bussmann 1996:s.v. "productivity"), ranging from relatively idiosyncratic patterns (e.g., the voice alternation of the fricative in the pair *north–northern*) to relatively regular ones (e.g., the derivation of an adjective from a noun by *-y* as in *earth–earthy*) to highly regular ones (e.g., the derivation of an agent noun from a verb by *-er* as in *sing–singer*). Furthermore, it is highly dependent on the age of the speaker, and the discourse context (a standard example of a nonproductive derivational form is nominal *-th* as in *warmth*, *depth*; however, *-th* is being used innovatively in on-line discourse, cf. *coolth*, *greenth*, *gloomth* [Baayen 2003]).[18] Typically, a continuum is established from "non- or unproductive", to "semi-productive," to "productive"; this continuum corresponds roughly to the continuum from lexical to grammatical. Bauer suggests that "[t]he converse of productivity ... is 'lexicalization'" (1994:3355). However, as will be discussed in subsequent chapters, some derivational morphology is almost as productive as inflectional morphology, e.g., manner adverbial *-ly*, as in *fortunately* (also the nominal prefix *pre-* as in *pre-season)* (Jackendoff 2002:155).

Of particular interest for studies of language use is the interaction of productivity with frequency (see, e.g., Baayen and Renouf 1996). In frequency studies an issue of central importance is the distinction between type and token frequency (see, e.g., Bybee 1985, 2003). TYPE FREQUENCY concerns the number of categories or constructions with which an item cooccurs, e.g., how many verbs occur in "middle alternation" constructions such as *The butcher cuts the meat easily* ~ *The meat cuts easily*, or how many verb types (process, stative, etc.) *be going to* cooccurs with. By contrast, TOKEN FREQUENCY concerns the number of instances of a form, e.g., how often *google* (verb) or *the* is used.[19] Type frequency of word formation may correlate with productivity. However, the correlation is often weak owing to residues of earlier increases or decreases in type frequency; e.g., *-ment* has a high type frequency in contemporary English (it occurs with a very large number of lexical bases, e.g., *government*, *derailment*), but it is currently not

[18] Thanks to Anette Rosenbach for this reference.

[19] Kastovsky (1986) uses a slightly different distinction. He says "productivity" encompasses on the one hand the number and type of constraints on the application of a word formation rule (he refers to this as "rule scope"), and on the other the frequency of the application of the rule in performance.

highly productive in the sense of being likely to be used in the creation of new forms (Haspelmath 2002). Because of the disparities between frequency (most especially type frequency) and likelihood of use in innovation, Haspelmath (2002) prefers to use the term "regularity" rather than "productivity," while Baayen (2003) argues that productivity is best understood in terms of probability rather than of frequency.

1.2.5 Summary

We have seen that problems in establishing clear binary distinctions between lexical and grammatical categories and between major and minor (functional) or open and closed word classes have led scholars to adopt a gradience view of lexical and grammatical and of word class membership. Similarly, scholars have adopted the view that the lexicon does not consist solely of a list of discrete and fully fixed items but represents a continuum from more to less fixed, from more to less fully conventionalized, and from more to less productive items. As we will see below, the continuum models of the lexical/grammatical split and of the lexicon fit better with the historical facts of change, which is often (though not always) gradual in the sense that change occurs by very small steps. Those items that combine fully productively with other forms are "grammatical," those that are more arbitrary are "lexical," some with several historically attested or predictable intermediary stages.

Having outlined some of the conceptual background, we now briefly introduce lexicalization and grammaticalization.

1.3 Lexicalization

The term LEXICALIZATION has been used for two very different phenomena. Synchronically it has been used for the coding of conceptual categories. Diachronically it has been used variously for "adoption into the lexicon" or "falling outside the productive rules of grammar." Much recent work has been on lexicalization in the first sense, with some attention to typology. We will mention some major trends in this work in Section 1.3.1, but will not elaborate on it further in this book. In Section 1.3.2 we turn to lexicalization in the second sense, as a diachronic process. This approach has often been refined in the context of the discussion of grammaticalization, which is the topic of Section 1.4.

1.3.1 Synchronic perspectives on lexicalization

The term "lexicalization" in the synchronic sense refers to the extent to which there are links between conceptual representation and syntax, and how the nature of such links may be formalized. Interest in the topic largely derives from discussion in various frameworks of lexical insertion in a

grammar, such as the proposal that the lexeme *kill* is the representation of a more abstract structure such as CAUSE BECOME NOT ALIVE (McCawley 1968). Much current research is on predicates and cross-linguistic evidence for the possible conflation of complex conceptual structures into a single lexical form (e.g., Talmy 1985, 2000; Jackendoff 1990, 2002), and on the ways in which a small set of event types (e.g., action, motion, change of state) link up to a large set of "roots" (or particular lexemes) with idiosyncratic meaning (e.g., Rappaport Hovav and Levin 1998a, 1998b).

One line of research that has been particularly influential is Talmy's hypothesis that languages fall into two general typological groups with respect to the lexicalization of certain motion event types (Talmy 1985, 2000):

(a) Languages like English (and most European languages other than the Romance languages), Chinese, Finno-Ugric, and Ojibwa that characteristically code (or "lexicalize") a motion event together with Manner in a single verb root (or "lexeme"), and express the Path of motion in a "satellite" to the verb such as a particle, or affix. In English, *The bottle floated into the cave* is the default expression: manner (*floated*) is conflated into the abstract verb MOVE, while the Path (*into the cave*) is profiled separately. *The bottle entered the cave floating* is possible, but is less colloquial, and is perspectivized differently (the bottle appears to be treated as agentive). Similar examples in the domain of stative (i.e., non-motion) location are provided by *The rope hung across the canyon* ('the rope was at the canyon in a hanging manner'), *Paint streaked the rug* ('paint was all over the rug in a streaked manner').
(b) Languages like Romance, Semitic, Japanese, and Polynesian that characteristically lexicalize a motion event together with Path in a single lexeme, and express Manner in an associated form. In Spanish, *La botella entró flotando en la cueva* 'The bottle entered floating to the cave' is the default: motion and Path are conflated (*entró*) while manner (*flotando*) is profiled.

The first group of languages is said to be "satellite-framed," the second kind "verb-framed." These differences have been shown to have a large number of consequences for style in narrative and other kinds of discourse. For example, it is hypothesized that speakers of satellite-framed languages pay more attention to dynamic activities along a path in narrative, whereas speakers of verb-framed languages pay more attention to static descriptions (Berman and Slobin 1994). Another hypothesis is that satellite-framed languages will have a larger inventory of manner of motion verbs (e.g., *creep*, *glide*, *clamber*, *spring*) than verb-satellite languages, at least in informal registers where borrowings are least likely to be used (*enter*, *exit*, *descend*, borrowed into English from Romance, conflate motion and Path).

The groups are clearly not discrete, not only because of the many borrowings that may occur, but also because of the sheer richness of the ways

in which words can be formed in any language. Furthermore, typological changes may occur: Latin, for example, had many characteristics of satellite-framing but its descendants, the Romance languages, are predominantly verb-framing; however, modern Italian is moving toward satellite-framing (Slobin 2004). Nevertheless, the broad sweep of the claims has fairly wide acceptance, and the semantic distinctions identified have helped advance lexical semantics from a marginal to a central role in linguistic theory.

Important synchronic and cross-linguistic claims about lexicalization have also been made in the narrower domains of quantifiers like *all*, and modals like *must*. It was noticed since Jespersen (1917) that certain negative meanings have exclusively phrasal expression cross-linguistically and are not expressible by a single word-sized unit. In particular, Horn (2001 [1989]) and Levinson (2000) have argued that while cross-linguistically the meanings associated with *all*, *some*, and *no(ne)* (= 'all not') are lexicalized in the sense that they have distinct single word forms, **nall* (= 'not all,' which implies 'some') is not. Likewise, forms expressing *necessary*, *possible*, *impossible* (= 'not possible that') appear, but not **nossible* (= 'possible that not,' which implies 'possible'). The generalization is that strong negative scalar items (*all*, *necessary*) incorporate negation morphologically, whereas weak ones (*some*, *possible*) do not. Horn and Levinson argue that because *some* and *possible* implicate *not all* and *not necessary* respectively, there is no need for a separate word meaning *nall* or *nossible*. Building on Horn (2001 [1989]), van der Auwera (2001) has shown that similar, but more flexible, constraints exist for modal verbs. Word order specialization and even single-word lexicalization of NOT NECESSARY do occur; e.g., we have *must*, *may*, *mustn't* (= 'necessary that not'), and also *needn't* (= 'not necessary that,' which implies *may*). However, historical drifts from 'not necessary that' or 'not possible that' to 'necessary that not' and 'possible that not,' i.e., to the stronger negative modal, are not infrequent; hence, modal systems may resemble negative ones with respect to asymmetry of lexicalization possibilities. Note that in these discussions, "lexicalized" expressions crucially do not include complex fixed multi-word expressions, but strictly single word-like entities.

1.3.2 Historical perspectives on lexicalization

Historical perspectives on lexicalization will be discussed in more detail in Chapter 2. For overviews see Brinton (2002), Lindström (2004), Himmelmann (2004), Traugott (2005). By way of introduction, an intuitive and broad sense of the term "lexicalization" is that it refers to adoption into the lexicon, e.g.:

- "[T]he adoption of a word into the lexicon of a language as a usual formation that is stored in the lexicon and can be recalled from there for use" (Bussmann 1996:s.v. "lexicalization").

- "[A] process by which new linguistic entities, be it simple or complex words or just new senses, become conventionalized on the level of the lexicon" (Blank 2001:1603).

Lehmann (1989, 2002) restricts lexicalization to "a process in which something becomes lexical" (2002:14) in the sense of entering the inventory and becoming holistic.

According to another general meaning, "lexicalization" occurs when a form can no longer be accounted for by regular grammatical rules:

- "Whenever a linguistic form falls outside the productive rules of grammar it becomes lexicalized" (Anttila 1989 [1972]:151).
- The stage when "a lexeme has, or takes on, a form which it could not have if it had arisen by the application of productive rules" (Bauer 1983:48).

These general conceptions of lexicalization have been given more precise, and often incompatible, definitions in the literature on word formation (see Chapter 2) and grammaticalization (see below and Chapter 3).

A third interpretation of lexicalization refers to shifts from implied to coded (or conventional) meaning, i.e., from pragmatic to semantic polysemy. For example, *see* in the sense of 'visual experience' can be understood to imply 'understanding' in certain circumstances, e.g., *I see that*. This pragmatic meaning may become a semantic polysemy and eventually the only meaning of the word; cf. the eventual development of PIE **weid-* 'see' not only into terms for seeing (*vision*, *video*, etc., via Latin), but also *idea* (via Greek) (Sweetser 1990). It is certain "bridging contexts" that allow either pragmatic or semantic interpretations of the same form, as observed by Evans and Wilkins (2000:550): "In exploring bridging contexts, the primary question is: What recurrent contexts, and what cultural scripts, allow particular pragmatic extensions to occur with sufficient frequency that they get lexicalized as distinct, but related meanings?" A better term for lexicalization in this sense would appear to be "semanticization" (Hopper and Traugott 2003:235, n. 2 to Chap. 4).

A fourth general meaning pertains to the development of concrete meanings, especially those associated with major classes. This meaning is central to much discussion of the relationship between lexicalization and grammaticalization.

Finally, lexicalization is sometimes equated with semantic change in general (e.g., Blank 2001:1603–1604).

Some of the issues that arise in an attempt to integrate the differing definitions of diachronic lexicalization are:

(1) To what extent does lexicalization differ from processes of word formation such as compounding, derivation, and conversion?

(2) How does lexicalization intersect with or relate to other processes such as idiomaticization, demotivation, routinization, institutionalization?

(3) What is entailed by conceptualizing lexicalization as either a process of fusion leading to decreased compositionality or a process of separation leading to increased autonomy?
(4) Is lexicalization gradual or abrupt?

1.4 Grammaticalization

Like lexicalization, GRAMMATICALIZATION (also called "grammaticization") can be conceived synchronically and diachronically. However, in the case of grammaticalization, the majority of the work has been either historical or typological (with focus on implications for change). What is common to all studies of grammaticalization is the attempt to understand how syntactic and morphological patterns are structured, how and why grammatical categories arise, and how free combinatorial and fixed patterns interact (for a comprehensive overview, see Heine 2003a). The research has been so engaging that some skeptics have assumed that researchers have been seeking to identify a unique kind of change that has an "independent status of its own" (Campbell 2001a:117; see also Newmeyer 1998:237ff.). However, this is not the case. Rather, researchers have sought to show that grammaticalization is interesting "because it is largely irreversible and because we observe strong correlations between phonological, syntactic and semantic-pragmatic changes. It is a macro-level phenomenon which cannot be reduced to the properties of the corresponding micro-level phenomena" (Haspelmath 2004:26).

1.4.1 Synchronic perspectives on grammaticalization

From a synchronic perspective, grammaticalization is primarily a morpho-syntactic, discourse pragmatic phenomenon, to be studied from the point of view of fluid, dynamic patterns of language use at a moment in time. It is a way of organizing patterns of language that differ in minimal, local ways (e.g., Romaine and Lange 1991; Thompson and Mulac 1991; Mair 1994; Ziegeler 1996; Beths 1999; Aijmer 2004). For example, in discussing the synchronic relationship between sentences involving *I think that [Clause]* as in (4a) and those involving "epistemic parentheticals" *I think Ø* 'probably' (4b, c), Thompson and Mulac argue that the epistemic parentheticals are grammaticalized (reduced, more frequent) versions of the main clause:

(4) a. I think that exercise is really beneficial.
b. I think exercise is really beneficial.
c. Exercise is really beneficial, I think.
(based on Thompson and Mulac 1991:313)

They see a direct correlation between frequency/ first person subjects/ *that*-less complements and the grammaticalization of the forms. In (4b, c) *I think* has undergone decategorialization from a complement-taking noun + verb

sequence into a kind of unitary particle with different distributional properties. There is a reversal of the matrix clause/complement clause structure, the original complement clause being reanalyzed as the matrix clause and the original matrix clause now serving as a parenthetical disjunct. Romaine and Lange (1991) see *like* undergoing "recategorization" from a preposition taking a nominal complement to a conjunction taking a sentential complement and then reanalysis as a discourse marker with syntactic detachability and mobility. Mair (1994) argues that with the loss of concrete, visual meaning, *to see* is coming to serve as a grammatical device connecting clauses (as in *Why should people be surprised to see this happen*); it is in a stage of incipient grammaticalization. Aijmer (2004) points to the cross-linguistic grammaticalization of markers of 'ability' (Eng. *manage*, *be able to*; Sw. *lyckas* 'manage to', *få* 'get') as aktionsart markers of 'successful completion,'[20] focusing on the metonymic link between potentiality ('participant-internal possibility') and actuality ('participant-internal actuality') effected by means of a generalized conversational implicature. As Aijmer notes, "[*m*]*anage to* has been grammaticalized directly from a lexical element to a grammatical marker. The development from ability to successful completion may be an instance of a more general pathway whereby process verbs give rise to markers of tense, aspect, and modality in certain contexts" (2004:73).

Some of the factors that go into synchronic studies of grammaticalization are mentioned at the end of the next section, since they are based on empirical evidence from language change.

1.4.2 Historical perspectives on grammaticalization

From the historical perspective, grammaticalization is at the very least "the subset of cross-linguistically recurring correlations across time among semantic-pragmatic, morphosyntactic, and (sometimes) phonological changes" (Hopper and Traugott 2003:1–2). Usually it is thought of as "that part of the study of language change that is concerned with such questions as how lexical items and constructions come in certain linguistic contexts to serve grammatical functions or how grammatical items develop new grammatical functions" (Hopper and Traugott 2003:1).[21] The latter definition owes its origins to work by Meillet, who famously characterized

[20] See Brinton (1988:23ff.) on distinctions among types of aktionsart ("lexical aspect," "situation type"), such as state (*live*), activity (*run*), achievement (*break*), accomplishment (*grow up*).

[21] Functional reanalysis has often been thought of in terms of a shift from lexical to functional head status. For example, Roberts (1993a) has suggested that in Romance former lexical heads (main verbs) become functional heads (auxiliaries) of the construction in which they occur.

grammaticalization (a term that he apparently coined) as "the attribution of grammatical character to a previously autonomous word" (1958 [1912]:131),[22] and to Kuryłowicz, who defined it as "the increase of the range of a morpheme advancing from a lexical to a grammatical or from a less grammatical to more grammatical status, e.g., from a derivative formant to an inflectional one" (1975 [1965]:52). Since its inception, then, grammaticalization has been conceptualized as in some way related to the lexicon. But since its inception it has also been conceptualized as related to discourse phenomena: Meillet (1958 [1912]) included word order change in grammaticalization (see also Claudi 1994; Helftoft 1996). And Givón (1979:209), drawing attention to topic, focus, and other discourse factors, proposed the grammaticalization schema in (5):

(5) discourse > syntax > morphology > morphophonemics > zero

While the suggestion that discourse precedes syntax has been questioned (Traugott 1982), the importance of discourse and especially of constructions has nevertheless become central to most research on developments known as grammaticalization: "grammaticalization does not merely seize a word or a morpheme ... but instead the whole construction formed by the syntagmatic relations of the element in question" (Lehmann 1992:406; see also Lehmann 1993; Traugott 2003).

Standard examples of grammaticalization are: Lat. *cantare habeo* 'sing:INF + have:PRES1SG' > Fr. *chanterai* 'sing:FUT1SG', and Eng. *be going to* (V) 'motion verb + purposive clause' > *be gonna* (auxiliary).[23] We illustrate briefly with an account of *cantare habeo*. A phrasal structure consisting of a verb in the infinitive followed by 'have' in the present tense indicative came to replace the old Latin inflectional future *cantabo* 'I will sing' during the third to sixth centuries AC. The process involved reanalysis of a biclausal structure (infinitive complement dependent on *hab-*) as a single constituent; this allowed fusion across morpheme boundaries, phonological attrition, and semantic analysis to a future marker, as illustrated by:

(6) Iustinianus dicebat: 'Daras.'
Justinian said: 'Give:FUT2SG' (< dare habes)
(Seventh-century Fredegario [Roberts 1993a:234])

[22] There were many precursors to Meillet, including Georg von der Gabelenz, Karl Brugmann, and Henry Sweet, all of whom discussed aspects of fusion (see Lehmann 1995 [1982]; Lindström 2004).

[23] For fuller accounts of the development of the Romance future, see, e.g., Fleischman (1982), Hopper and Traugott (2003:Chap. 3), Roberts (1993a, 1993b). For accounts of the development of the English *go*-future, see, e.g., Bybee and Pagliuca (1987), Pérez (1990), Hopper and Traugott (2003:Chap. 4).

The textual evidence suggests that the construction arose in contexts of saying: if one has something to say, one may be considered obliged to say it; this obligation is future oriented. Over time the implied future came to be salient, and eventually the fused infinitive marker and the 'have' verb came to be part of the regular tense paradigm. That fusion is not inevitable is illustrated by the fact that Sardinian has a phrasal future (in Sardinian the verb 'to have' is *aere*, and the first person singular present is *appo*):

(7) L'appo a fakere
lt-have:PRES1SG to do:INF
'I will do it'
(Modern Sardinian [Roberts 1993a:235, citing Jones 1993])

Unidirectionality from lexicon or construction > grammar and not vice versa is a strong, empirical, and therefore interesting hypothesis; for discussion see, e.g., Haspelmath (1999b), Hopper and Traugott (2003:Chap. 5). Although it is widely acknowledged that grammaticalization is typically unidirectional, the unidirectionality hypothesis has met many challenges, some of which involve lexicalization. These will be discussed in more detail in Chapter 3. Challenges to unidirectionality are the topic of articles in Campbell (2001b; see most especially Janda 2001 and the bibliography in it) and of several contributions to Fischer, Rosenbach, and Stein (2000).

Among factors that are widely discussed in the context of grammaticalization, all contributing to unidirectionality are the following (for a partially different set of "parameters" see Lehmann 1995 [1982]):[24]

(a) DECATEGORIALIZATION: This refers to the shift from one category status to another, correlated with a shift from prototypical membership of a category to less prototypical membership, and maybe eventually to prototypical membership of a new category. Typically, the shift is from a "major" to a "minor" category. Ramat argues that the term "decategorialization" is not ideal, since it "evokes the idea of forms losing any categorial status" (Ramat 2001:398) and suggests the term "transcategorization (or recategorization)" instead.[25] The notion that decategorialization could imply loss of "any categorial status" presumably originates in the generative assumption that categories are discrete and non-gradient. Since proponents of decategorialization privilege

[24] They are weight, cohesion, and variability; viewed paradigmatically these are "integrity," "paradigmaticity," and "paradigmatic variability"; on the syntagmatic dimension they are "structural scope," "bondedness," and "syntagmatic variability" (Lehmann 1995 [1982]:123).

[25] Since Ramat intends transcategorization to cover cases of conversion (e.g., *showcase* [N] > *showcase* [V]), which are not cases of grammaticalization, it is not adopted here.

prototypicality and gradience, not discreteness and uniqueness, this criticism is moot because it stems from a different theoretical orientation.

(b) GRADUALNESS: This notion refers to the fact that most changes occur in very small structural steps, typically with innovative uses coexisting alongside older ones (the A > A ~ B > (B) phenomenon mentioned in 1.1.2) (see, e.g., Plank 1989; Lichtenberk 1991). Not all properties of an earlier lexical source are likely to change at the same rate. In an early characterization of gradualness, Givón, speaking of shift from verb to preposition such as Yoruba *fi* 'take' > 'with' (instrumental) says: "It is highly unlikely that a verb would change suddenly into a preposition by all semantic, morphological and syntactic criteria at once ... In particular, morphological and syntactic behavior is likely to lag behind the more progressive semantic re-analysis, and thus quite often represent vacuous relics of the older semantic situation" (1975:86). Furthermore, textual evidence suggests that many changes involve periods of relative indeterminacy in which it is not clear whether the older or the newer usage is in evidence; in other words, the steps may be tentative at first. Indeed, the first steps may never result in change, in the sense of acceptance by a community of speakers.

As an example of gradualness, most of the instances of *be going to* [V] that could be interpreted as auxiliary uses in the Helsinki Corpus during the sixteenth and seventeenth centuries are ambiguous between motion and future, e.g.:

(8) *Quak*. But ... where is thy Dame?
Ma. Even now departed to hold forth amongst the Congregation of the Righteous, in the full Assembly of the Righteous.
Quak. What to the Hill of Sion, that the wicked do prophanely call the Bull-and-Mouth?
Ma. Yea, verily; for having on the sudden a strong Impulse by the operation of the Spirit, she said unto me, Mary, and I answered I am here; whereupon she answered and said, she **was going to instruct** our Friends.
(1685 Samuel Pepys, *Penny Merriments*, p. 147 [*HCET*])

Ignoring the larger context of the conversation, one might assume that *was going to* is an example of the future use, but in the context of questions about where the subject is, and responses concerning where she has departed to, this is equally or even more likely to be an example of the motion construction. Two interpretations are possible in context, and it does not matter in any essential way whether speakers and hearers share the same structural analysis; in other words, this is a "bridging context" (see Evans and Wilkins 2000:550; Heine 2002:84;

Enfield 2003:30 and 1.3.2 above). Change has clearly occurred only when *be going to* is used in a new context which is not ambiguous (Heine 2002:85 calls this a "switch context"). The small, local changes, when identifiable as steps in change over time, occur abruptly with respect to grammars – *be going to* cannot be both motion verb and auxiliary at the same time in the grammar (they have to be two variants), but there are usually indeterminate uses for a time prior to the change, or particular instances of use may be indeterminate after the change has occurred. In this sense gradualness is the historical correlate of gradience (see 1.2.2). "Gradualness" is thus a primarily diachronic term characterizing changes from one state to another over time. It should be contrasted with "gradience," which is a synchronic term characterizing the continuum between one linguistic category and another (see 1.2.3).[26]

A second sense of gradualness is frequentistic: changes spread through the system slowly, often gaining considerable momentum over time (see (h) below). A third meaning occasionally found in the literature is that grammaticalization changes take a considerable length of time to develop. For example, Bruyn (1996) argues that grammaticalization in creole situations is unique in that the time-depth for grammaticalization is often very short and therefore gradualness is often not evidenced. Fairly rapid changes can, however, also be found in non-creole situations, e.g., the development of modals in English one after another in rapid succession.

(c) FUSION and COALESCENCE: Prototypical examples of grammaticalization involve boundary loss and morphological/phonological fusion, or "bonding" (Bybee 1985; Bybee, Perkins, and Pagliuca 1994). Grammaticalization may also involve the loss of phonological segments, or coalescence.[27] Both fusion and coalescence are clear in the development of Lat. *cantare habeo* into Fr. *chanterai*. Originally free verbs (both

[26] Yet another term to be distinguished from both is "gradability." This is a semantic term referring to 'more or less,' and has to do with the nature of lexical structure: most adjectives are gradable, e.g., *sad* (*very sad*, *sadder*), though some are not (e.g., *utter*, **very utter*) (see, e.g., Paradis 1997; Aarts 2004).

[27] Fusion and coalescence are sometimes treated as synonyms (see, e.g., Matthews 1997). However, they are clearly separate processes. Even within this two-way distinction a number of further fine-grained distinctions need to be made, depending on the language under discussion. For example, Andersen (1987), in discussing the development in Polish of forms of the copula 'be' into inflections ("desinences") of the preterite indicative and subjunctive of all verbs and the present tense of 'be,' distinguishes three types of "univerbation" (combining into one word, see 2.3.1): morphological univerbation (fusing of syntactic and morphological boundaries, which reduces autonomy), prosodic univerbation (stress shifts resulting in new syllabification), and segmental univerbation (morphophonological reduction).

habeo cantare and *cantare habeo* are attested) come to be fixed in the latter order (fusion), and syllables are eventually lost (coalescence). Coalescence is also illustrated by the development, textually attested from the early 1900s on, of the auxiliary *be going to* as *be gonna*. But not all grammaticalization involves overt fusion and coalescence. For example, although the core auxiliaries of English, *may*, *can*, *must*, etc., are highly constrained in their syntax, they can occur without a following verb, as in *Yes*, *you may*. Examples like auxiliaries do, however, show greater dependency with adjacent constituents than the main verbs from which they derive. This kind of situation has led Haspelmath to define grammaticalization as "a diachronic change by which parts of a constructional schema come to have stronger internal dependencies" (2004:26) where by "internal dependency" he means dependency internal to constituent structure.

(d) TYPOLOGICAL GENERALITY: Cross-linguistic evidence shows that although change never has to happen, grammatical categories are regularly recruited cross-linguistically from similar sources, see, e.g., Bybee, Perkins, and Pagliuca (1994) on cross-linguistic sources for tense, aspect, and modality. Heine and Kuteva (2002) provide a major source of information about common grammaticalization patterns in the languages of the world. Since earlier and later forms coexist, this means that similar kinds of polysemies are repeatedly found cross-linguistically.[28]

(e) METAPHORIZATION and METONYMIZATION: In early work on grammaticalization when semantic change was considered, it was conceptualized in terms of metaphor. Thus, the semantic changes of *have* from 'possession' > 'obligation' > 'future' and of *go* from 'motion' > 'future' were construed as metaphors. While the result of grammaticalization is often synchronically metaphorical, textual evidence for the development of many grammatical formatives out of lexical and constructional material is metonymic in the sense that it is highly context-bound and arises out of the implicatures in the speaker–hearer communicative situation. Metonymy is a cognitive process in which "one conceptual entity ... provides access to another conceptual entity" (Kövekses and Radden 1998:39) and points to or "indexes" relations. This is illustrated by (8), where contexts of motion (departure) with intention (to speak to the congregation) index futurity. Heine, Claudi, and

[28] Some interesting testable hypotheses have been made regarding patterns that do *not* occur; e.g., Talmy claims that grammatical categories do not express color or Euclidean geometric concepts such as "fixed distance, size, contour, and angle" (2000, I:28) but rather relativistic, topological and topological-like concepts like adjacency, region, partition. Therefore, *corner in time* is not a likely source of case, although it is an attested lexical metaphor (Traugott 1986).

Hünnemeyer (1991) model a theory of grammaticalization in which metaphor plays a major role, especially such conceptual shifts as PERSON > OBJECT > SPACE > TIME > QUALITY. Traugott and König (1991) outline an alternative and complementary conceptual metonymic approach according to which implicatures from lexical items in specific, repeatedly used contexts come to be semanticized. Brinton (1988) was one of the first works in which the role of metonymy (in the development of aspectual forms) was explored.

(f) SUBJECTIFICATION: The development of grammatical forms conceptually involves the recruitment of material to express the grammatical relations that the speaker envisions. In the case of the main verb use of *be going to*, the motion meaning is predicated of the subject of the sentence, and the direction is anchored in both the subject's and the speaker's perspective (it alternates with *be coming to*); but in the case of the auxiliary use, future meaning is based entirely in the speaker's (not the subject's) perspective. Langacker (1990:2) gives the example:

(9) An earthquake is going to destroy that town.

Here the earthquake is neither going anywhere nor intending anything, and the perspective is entirely that of the speaker. The development of raising constructions generally involves subjectification of this kind (Langacker 1999). This view of subjectification is the historical version of the synchronic difference pointed out by Benveniste (1971b [1958]) between the subject of the sentence ("sujet énoncé") and subject of the utterance ("sujet d'énonciation"). A more comprehensive view of subjectification is proposed in Traugott (1982, 1995a) which includes changes of this type but also extends to the development of connectives like *besides*, which serve not only to combine clauses, but also to evaluate the connection, cf. also *but* (< OE *on butan* 'on the outside,' later 'except') or *only* (earlier 'exclusively, solely,' later 'but, except').

(g) BLEACHING: Loss of lexical content meaning has long been considered criterial in the development of grammatical forms, and is known as "bleaching." It is important to note, however, that in cases of grammaticalization, a new, more abstract, grammatical meaning develops that replaces the content meaning. Again, *be going to* is a paradigm example: motion from one place to another distant from the speaker is lost, but futurity shifts from an implied to a central meaning, i.e., is assigned semantic rather than pragmatic function in the course of grammaticalization (Sweetser 1988).

(h) FREQUENCY: Over time grammatical items come to be more frequent than the lexical constructions from which they derive (see, e.g., Bybee 2003). A much-noted characteristic of grammaticalization – growth in type frequency – involves the cooccurrence of the unit undergoing grammaticalization with an increasingly large number of other unit

types. For example, the number of verb types that the *be going to* construction cooccurs with increases as the auxiliary becomes established. Indeed, grammaticalization of the *be going to* construction may be identified as having occurred precisely when it cooccurs with a verb or a subject that is inappropriate to the source meaning. The motion with intention meaning excludes inanimate subjects in the main clause and statives in the purpose clause. When *be going to* begins to be used with inanimate subjects like *earthquake*, and stative verbs like *know* in *She's going to know the answer soon*, the type frequency of the construction increases. It leads to increase in token frequency (the number of instances of a form) and contributes to "bleaching" (see (g) above).

Correlations among factors (a)–(h) have been used in work on synchronic grammaticalization to develop hypotheses about the time-depth of particular grammatical forms and about their lexical origins. For example, Bybee, Perkins, and Pagliuca (1994:300) point out that as grammatical morphemes are constantly being renewed in similar ways, languages will provide several alternatives for expressing more or less the same concept (cf. the English futures *will*, *be going to*), a variation phenomenon that Hopper (1991) has called "layering." Since each instance of grammaticalization develops independently, there will often be overlap as well as discreteness of meaning between the older and newer forms; this discreteness of meaning reflects the historical origin of the form. Thus, while both *will* and *be going to* express future, the latter is associated with planned future, which is not surprising considering its origins in a purposive construction. Since token frequency "corresponds to the generality of the schema, which in turn corresponds to a higher degree of grammaticization" (Bybee 2003:612), token frequency can be used to hypothesize the historical time-depth of a particular grammatical morpheme – the more frequent an item, the more grammaticalized it is. While this principle is a useful heuristic when more and less lexical uses of the same form are compared, it needs to be considered with caution when competing forms are compared; competition between older and newer layers of grammatical morphemes often results in the decline or even loss of the older one. *Shall* is a case in point; the auxiliary verb use of *shall* and *will* precede *be going to* by over five hundred years, but the older form *shall* is now highly restricted (structurally to questions like *Shall we go*?, stylistically to legal language) and nearly obsolete in American English. A further criterion is that the older a grammatical morpheme the more likely it is to be fused and short (Bybee, Perkins, and Pagliuca 1994:47); earlier *will* (> *'ll*) certainly illustrates this point in comparison to *be going to*, except in some of the latter's reduced spoken forms such as [gɔnə]).

Some of the issues raised by diachronic grammaticalization studies that are relevant to diachronic lexicalization are:

(1) Are there only certain factors that clearly distinguish grammaticalization from lexicalization?
(2) Are there other factors that can be seen as defining properties of both grammaticalization and lexicalization?
(3) Is some lexicalization the same as "degrammaticalization"? If so, would one expect to find the reverse of factors (a) to (h) in the case of lexicalization?

These issues will be discussed primarily in Chapter 4.

1.5 Conclusion

This chapter has set out some preliminaries necessary for the study of lexicalization and grammaticalization from a historical, functionalist perspective. By examining conceptions of the lexicon, we have seen that the distinction between lexical and grammatical, between open and closed or major and minor class, and between unproductive and productive must be understood as gradient. Following an examination of the diverse ways in which lexicalization has been understood in the literature (Chapter 2) and of the often conflicting ways in which lexicalization and grammaticalization have been related (Chapter 3), we will find the notion of gradience and gradualness central to the integrated model we present in Chapter 4. We will examine some case studies from the history of English which highlight problematic areas at the interface between lexicalization and grammaticalization (Chapter 5). The book will end with a brief set of concluding remarks and questions for further research (Chapter 6).

2

Lexicalization: definitions and viewpoints

2.0 Introduction

Lexicalization is the process by which new items that are considered "lexical" (in terms of the theory in question) come into being. Several broad definitions can be found in the literature (Brinton 2002):

(a) ordinary processes of word formation (see 2.1),
(b) processes of fusion resulting in a decrease in compositionality (see 2.3),
(c) processes of separation resulting in an increase in autonomy (see 2.4)

Definition (a) occurs primarily in discussion of dynamic processes available to speakers synchronically that give rise to nonce forms, which may or may not come to be conventionalized and accepted later by other speakers (a process called "institutionalization," see 2.2). Definitions (b)–(c) occur primarily in discussion of change. In almost all the work considered in this chapter, a conceptual opposition is assumed between relatively free forms organized by rules and principles (the grammar) and fixed forms occurring in inventories (the lexicon). In particular, there is recognition of some version of a continuum or scale from relatively free, syntactic types of expression to highly compressed types of expression. This is a matter of COMPOSITIONALITY, or the extent to which "the meaning of ... expressions (in particular sentences) can be reconstructed from the meanings of their individual elements and their syntactic relationship to one another" (Bussmann 1996:s.v. "principle of compositionality"). From a historical perspective, continua of this kind invite the researcher to focus on either the beginning (moving away from syntax and productive rules) or on the endpoint "a process that leads from something that is not a (or, one) lexeme to a lexeme, i.e., to something that 'belongs in the lexicon'" (Lessau 1994:s.v.

"lexicalization"). The former perspective can be thought of as "falling outside the productive rules of grammar," the latter perspective as "adoption into the lexicon," as discussed in Chapter 1. In Moreno Cabrera's terminology, lexicalization is thus "syntactogenetic" (affecting syntactically determined words, phrases, or sentences) and "lexicotelic" (going from syntax to the lexicon) (1998:218).

Despite the apparent neatness of rubrics such as (a)–(c) above, there is actually considerable difference of opinion as to how to analyze particular changes within each category. In this chapter we report on a variety of definitions and viewpoints, some of which contradict each other. In Chapter 4 we attempt to resolve the disparities from the perspective of one consistent theory of grammar and lexicon.

2.1 Ordinary processes of word formation

In the context of very broad characterizations of lexicalization, little or no distinction has been made between word formation and lexicalization. Traditionally, lexicalization refers to routine processes of word formation, such as compounding, derivation, and conversion.[1] These are usually construed as the productive principles enabling a speaker's dynamic, creative language use to produce new words, i.e., form–meaning pairs with some degree of referential or content meaning. They increase the vocabulary of a language, and at the same time enhance the resources available for any one particular lexical field. For example, van der Auwera (2002:20) cites productive processes of word formation, such as compounding and derivation, that produce *songwriter*, as cases of lexicalization. This is the broadest definition and from a historical perspective, probably the least satisfactory because it tells us little or nothing about what kinds of changes the products of different kinds of word formation can undergo over time.

Many researchers recognize this problem and define lexicalization more restrictively, regarding lexicalization as concerned with the semantic particularities and idiosyncrasies identifiable within the broader framework

[1] "Word formation" clearly presupposes some notion of "word." As Anderson says, "there really is no satisfactory resolution" to the problem of defining the term "word," "since it involves several mutually independent (and sometimes conflicting) criteria" (1985a:4). At least two concepts of word are used: word as grammatical unit (roughly "part of speech" – this allows "word" to include complex phrases such as *mother-in-law*, PL *mothers-in-law*), and word as intonational unit (roughly a unit that can occur between pauses, e.g., Eng. *isn't*, Eskimo *iqalussuarniariartuqqusaagaluaqaagunnuuq* 'it is said that we have admittedly got a strict order to go out fishing sharks') (Anderson 1985b:151). Our focus will be on the word as a grammatical unit.

of word formation. For example, both Quirk, Greenbaum, Leech, and Svartvik (1985) and Huddleston and Pullum (2002) include sections on lexicalization within their chapters on word formation[2] and differentiate it from regular word formation. While Quirk et al. say that there is a process called "conversion," by which one can use a noun as a verb (e.g., *to carpet*) (1985:1520), they discuss under lexicalization the fact that although in principle any noun can be used as a verb, there are specific meaning orientations associated with such uses (1527–1530). More strikingly, lexicalization is defined in Huddleston and Pullum as "words that are or were earlier morphologically analysable but which could not be formed with their present meaning by the current rules of word-formation" (2002:1629).

In Chapter 4 we will argue that word formation and lexicalization should be treated as separate phenomena. Nevertheless, since so much work on lexicalization draws on information about word formation, in this section we outline the main types that have been said to be synchronically relevant to lexicalization and that are referred to in the literature on historical lexicalization and grammaticalization.[3]

2.1.1 Compounding

As a general process, compounding involves the unifying of two or more autonomous words to form a third (Bauer 1983:11), e.g., *blackboard*, *wallpaper*. After unification, the parts are no longer fully independent; for example, in *black* + *board* (Adj + N) > *blackboard* (N), there is stress shift to the first syllable and loss of semantic motivation (blackboards may be black, blue, green or brown, not, however, white). For inflected languages, it is preferable to think of the combination of roots or stems rather than words (where "root" is a form that carries primary lexical meaning and cannot be analyzed further, and "stem" is a lexical form minus inflectional morphology). Most languages have productive synchronic rules for compounding. They may involve phonological changes ranging from slight (e.g., placement of stress and reduction of unstressed vowels as in *sómebòdy* < *sóme bódy*) to more extensive ones (as in *cupboard* [kʌbərd], *forecastle* [fovksəl]) which lead "straight to the effacement of the complex character of the construction" (Blank 2001:1602). Over time phonological and other changes may obscure the original productivity.

[2] Chapter 19 on "Lexical word-formation" in Huddleston and Pullum (2002) is primarily by Laurie Bauer and Rodney Huddleston.

[3] The list here is by no means comprehensive; for example, ideophones are not discussed (see Childs 1994 for an outline of syntactic, semantic, and pragmatic as well as phonological aspects of ideophones, and research questions about them).

2.1.2 Derivation

Like compounding, derivation results in the formation of new lexemes. While in a compound lexeme two elements are roots or stems, in a derived lexeme only one of two elements is a root or stem, the other a derivational affix. The distinction between compounds and derivational morphology is, however, sometimes difficult to determine because many derivational affixes evolve from roots in compounds (see 2.3.2), or because some derivational affixes have been given independent status, e.g., *ism*, *ology* (Bauer 1983:35) (see 2.4). Derivational morphemes are usually said to be of two main types: "purely" semantic/lexical (e.g., *un-* 'not' in *unhappy*, 'reversal' in *undo*), and those related to grammatical meanings (e.g., agentive *-er* in *swimmer*; in It. the collective noun markers *-i*, *-a*). The pure semantic type clearly belongs to word formation, such as the suffix *-ship* forming abstract nouns from other nouns (e.g., *championship*, *membership*) or the prefix *mis-* expressing the meaning 'wrong, wrongly' (e.g., *mistrust*, *misdeed*). Our concern here will be primarily with the derivational morphemes that have grammatical consequences. Many such morphemes have dual functions, as what Kastovsky calls semantic "labels" and as markers of "syntactic recategorization," e.g., *-age* 'charge, fee' (label) as in *anchorage*, *poundage*, and *-age* 'result of V' (syntactic recategorization) as in *breakage*, *wastage* (1986:596).[4]

While derivation is usually considered crucial to lexicalization, the other principal type of bound element, inflection, is not. Since derivation is associated with lexicalization and sometimes with grammaticalization, and inflection with grammaticalization,[5] it is useful to consider what factors are adduced in distinguishing derivational from inflectional morphology. As in other areas, there is considerable disagreement about what the exact distinctions are,[6] but there is general agreement that absolute distinctions cannot be made, cross-linguistically at least (see Corbett 1987). For the present purposes, the following characteristics may be mentioned (see especially Bauer 1983; Luraghi 1998).

Grammatical derivative morphemes typically:

(a) are not obligatory (agentives in English may be marked with *-er* [*swimmer*, *jogger*] or not [*advocate*, *chairperson*]); they are "semiproductive" in

[4] See Fleischman (1976) for a detailed analysis of how the derivational morpheme *-age* (< Graeco-Roman *-aticu*) was borrowed into Spanish, Portuguese, Italian, and English from French as a part of French words and then became available in derivation based on native stems.

[5] Himmelmann (2004:28), however, suggests the emergence of derivational morphology may be a process separate from both lexicalization and grammaticalization.

[6] See, e.g., Aronoff (1976); Bauer (1983); Anderson (1985a, 1985b); Bybee (1985); Dressler (1989); Plank (1994); Dalton-Puffer (1996); Haspelmath (2004).

that they can be relatively freely used to make new lexemes (*-ness*, *-er*), but many gaps in the system remain,
(b) compete with other derivatives having the same function, e.g., nominalizers like *-ness* (*pettiness*) or *-ity* (*rigidity*), but not **pettity* or *rigidness*,
(c) may involve variables in a relatively open system, e.g., Bauer (1983:23, citing Guilbert 1975:182–183) points out that in French suffixes may be added to a verbal base to create an action nominalization (*-tion*, *-isation*, *-age*, *-issage*, *-ement*, *-ture*), an agent nominalization (*-eur*, *-ateur*, *-euse*, *-atrice*, *-oir*), an adjectivalization (*-ée*, *-ant*, *-atif*, *-able*), or a new verb (*-ailler*, *-asser*, *-onner*),
(d) change class membership, e.g., *petty*, *divine* (Adj), *pettiness*, *divinity* (N), and produce new lexemes or "stems" (syntactic recategorization).

By contrast, inflections typically:

(a) are obligatory, e.g., in English, marking verb agreement with the subject (*she walks*, *they walk*),[7]
(b) compete only marginally with other inflections having the same function, e.g., the past participle *-en* of strong verbs (e.g., *see*/*saw*/*have seen*, *ride*/*rode*/*have ridden*), competes only recessively with the past particle *-t*/*-d*/*-əd* of weak verbs (e.g., *sleep*/*slept*/*has slept*, *dream*/*dreamed*/*has dreamed*). This is especially clear in the case of verbs that were originally weak and have been analogized in some dialects to strong verbs with vowel alternation in the past tense, but still retain *t*/*d* past participles (e.g., *dive*/*dove*/*have dived*/**have doven*),
(c) involve variables in a relatively tight, closed system (e.g., singular/plural),
(d) do not change class membership or produce new lexemes.

While (d) typically holds for inflectional morphology, there are a few types of inflections that are problematic and exemplify clearly the gradient nature of derivation vs. inflection (see Haspelmath 2002:230–235). One of these types is the participle, which is a verbal inflection as in *we are running*, but also appears to be a deverbal adjectival inflection in many languages. For example, in (1) (from Haspelmath 2002:230), the participle *pfeifende* 'whistling' behaves like an adjective with respect to word order and syntactic modification of the noun, but also like a verb in that it combines with a locative modifier *im Wald* 'in the wood,' and a manner adverb *laut* 'loudly.'

[7] Consonant cluster deletion and other phonological rules notoriously affect such obligatory inflections, sometimes giving rise to zero inflection, cf. in African American Vernacular English, the frequent reduction of past *-t*/*-d*/*-əd* after consonants in weak verbs (e.g., *she walk-Ø yesterday*), but not in suppletive and other forms (e.g., *did, wen(t)*).

(1) der im Wald laut **pfeif-end-e** Wanderer
the in.the wood loudly whistle-PTCP-M.SG hiker
'the hiker who is whistling loudly in the forest'

This kind of dual-functioning inflection could be considered to be (i) an anomalous word-class changing inflection, (ii) an inflection showing dual behavior, i.e., having the internal syntax of verbs with respect to their dependents, but the external syntax of adjectives with respect to the other elements in the clause (Haspelmath 2002:232), (iii) "conversion" of verb to adjective (see 2.1.3 immediately below) (Huddleston and Pullum 2002:1644). Yet again, in a study of historical developments in Italian, Luraghi (1998) considers the relationship to be that of inflection vs. derivation (see 3.3.1). We will come back to the question of participial adjectives in greater detail in Section 5.1.

2.1.3 Conversion

"Conversion" in the sense relevant to this book is "functional shift" from one category to another (Bauer 1983:226–230).[8] In English it is usually equated with zero derivation. Conversion typically involves derivation from one major class item to another, as in the case of:

(2) *(a) run* (N) < *(to) run* (V)
(a) runaway (N) < *(to) run away* (V + Prt)
(a) private (N) < *private* (Adj)[9]
paper (shredder) (Adj) < *paper* (N)
tow-away (zone) (Adj) < *(to) tow away* (V + Prt)
(to) calendar (V) < *(a) calendar* (N)
(to) lower (V) < *lower* (Adj)[10]

Querying whether such conversions constitute instances of lexicalization, not merely semantic change in context, Blank (2001:1604) decides in the affirmative. He sees such shifts as particularly common in creoles, e.g., Mauritian Creole *volor* '(to) steal' (V) < Fr. *voleur* '(a) thief' (N).

[8] Matthews (1997:s.v. "conversion") cites a second meaning of conversion: use of a lexical unit which primarily belongs to one syntactic subclass with the character of another, and exemplifies with the use of transitive *cook* as an intransitive in *Dinner is cooking*. Bauer (1983:227) cites use of countable nouns as non-countables, and vice versa, cf. *some tea*, *two teas*, and suggests these are not cases of word formation but of syntactic usage (see also Huddleston and Pullum 2002). This second meaning does not concern us here.

[9] Denison (2001) points out, however, that Adj > N and N > Adj have been considered unusual. Marchand (1969 [1960]) and Quirk et al. (1985) treat Adj > N as in *a bitter (ale), a daily (paper)* as ellipses, not word formation.

[10] Note *lower* (Adj) has a comparative inflection.

Some phrasal compounds may be included among examples of conversion, e.g., *forget-me-not*, *love-in-a-mist* (Bauer 1983:207); these clearly have specific, idiosyncratic meanings, in these cases, flower names. Moreno Cabrera (1998) notes that such changes are typically "concretions" and metonymic in nature. Also included among (partial) conversions are pairs involving phonological variants that reflect earlier sound changes in the language (*belief–believe*), or stress shifts (*fréquent* (Adj)–*frequént* (V), *ímport* (N)–*impórt* (V)) (Bauer 1983:229).

In a language with little inflection, conversion may involve not only zero derivation from one major class to another but also from a more minor to a more major class:

(3) *(to) off* (V) < *off* (Adv)[11]
(to) down (V) < *down* (Adv) (cf. *to down a drink*)
up (Adj) < *up* (Adv) (cf. *the up train*)
if (N)< *if* (Conj) (cf. *ifs, ands, and buts*)

In languages with inflection, conversion may be derivationally more complex, as in the case of the German verb *duzen* from the pronoun *du*, the French verb *tutoyer* from the pronoun *tu* (also a metalinguistic use). An alternative analysis has been given for the subset of prepositions. Ziegeler suggests that they are "better seen as a type of metonymy, motivated by a situation in which the particle or prepositional form, *up* or *down*, can in time serve to stand for the entire verb phrase" and cites examples like Singaporean English *to off the microwave* (Ziegeler 2003:229).[12]

When conversion involves a shift from minor (closed, nonlexical, grammatical or functional) class to major (open, lexical, fully referential) class, e.g., V *(to) off* < Adv *off*, it has been widely recognized as lexicalization (and degrammaticalization) (see, e.g., Ramat 1992, 2001; Newmeyer 1998, who refers to "upgrading" from functional to lexical category; Traugott 2005).[13] Likewise Hopper and Traugott (1993:49) define lexicalization as the development of a fully referential lexical item from a nonlexical, or grammatical item, such as the development of the verbs *up*, *down* or nouns *upper*, *downer* from the homophonous particles *up*, *down* (note that *upper* and *downer* are not instances of simple conversion, but rather of conversion and derivation).

[11] Himmelmann (2004:39n. 5) points out that it is not clear whether verbal and adjectival *up* and *down* are derived from the adverb or the preposition. He also points to the use of *up* in compounds, cf. *up-and-coming*, *up-beat*.

[12] Ziegeler also considers the possibility that such uses arise out of ellipsis (2003:229n. 4).

[13] By contrast, Lehmann (2002) excludes conversion from lexicalization on the grounds that it does not involve the unification of two or more morphemes.

Other examples of conversion that have been adduced as instances of lexicalization include the "deeper syntactic reinterpretation" of the pronoun *that* as a conjunction (Anttila 1989 [1972]:151). However, Giacalone Ramat regards this as a case of "increased grammaticalization" (1998:120), and Norde points out that it may be regarded as a non-directional change, a lateral conversion on the same level (i.e., from functional to functional category) (2001:234). Newmeyer cites work by Gelderen (1996, 1997) that suggests the shift of *for* from preposition to complementizer and of *there* from demonstrative to expletive may also involve "upgrading" from less to more lexical status (1998:272–273).[14] Huddleston and Pullum (2002) point to the conversion (largely in the Early Modern English period, see Görlach 1991:109) from participle and adjective to preposition, e.g., *during*, *concerning*, *opposite*. Ramat calls forms such as *during* "transcategorization" or "recategorization" because they have "entered a restricted category, namely adpositions" (2001:397) but then goes on to say that "[s]trictly speaking lexicalization may also be considered to be an instance of transcategorization" (398). Here lexicalization is apparently equated with conversion.

There are substantial semantic constraints on conversion. For example, conversion from N > V is constrained by salient semantic and use-based characteristics of the parent noun, such as location, agent, instrument: e.g., to *bottle* is to put something in a bottle, to *mine* is to remove something from a mine, to *water* is to put water somewhere, to *milk* is to take milk from somewhere, to *dust* is either to remove dust from or place dust on a location, to *cook* is to do cooking, and to *hammer* is to use a hammer as an instrument (see, e.g., Clark and Clark 1979; and, with a different perspective on how to account for the semantic differences, Aronoff 1980). Quirk et al. (1985:1528) likewise call attention to pragmatic factors that have to do with world knowledge and familiar use, e.g., *to paper a wall* is usually understood as 'putting wall-paper on a wall,' although it is conceivable that it would mean putting newspaper or linguistic articles on a wall. They further note that conversion may be specific to a particular item, e.g., *cup one's hands*, not **mug one's hands*, *carpet the bedroom*, not **rug the bedroom*: "in both these examples, it is the lexical item of more general meaning that has lent itself to conversion" (1529). There may be pragmatic constraints on conversion as well. For example, with respect to conversion from major to minor class, Haspelmath has suggested that such examples as *ifs, ands, and buts*, are words "taken out of their construction and employed metalingistically" (1999b:1064n.1); that is, they are quotations, and therefore function rather differently from regular lexical items.

Despite the semantic and pragmatic idiosyncrasies noted above, conversion is a very general type of word formation. Even borrowed words can also

[14] The examples are somewhat dependent on the theory of syntax involved, and deserve further investigation.

undergo conversion, cf. *Fenshui your commute* (Bay Area Rapid Transit advertisement). How general it is is emphasized by Bauer who points out that many minor class forms may undergo conversion to several different major classes, as in the case of the use of *down* as a verb (*she downed her drink*), a noun (*she has a down on him*, *a first down*), and an adjective (*a down day*). He concludes:

> all form classes seem to be able to undergo conversion, and conversion seems to be able to produce words of almost any form class, particularly open form classes (noun, verb, adjective, adverb) ... The commonness of conversion can possibly be seen as breaking down the distinction between form classes in English and leading to a system where there are closed sets such as pronouns and a single open set of lexical items that can be used as required.
>
> (Bauer 1983:226–227)

As he admits, this suggestion is highly speculative. It is, however, of considerable potential interest to studies of changes in the organization of grammar and lexicon.

2.1.4 Clipping and ellipsis

Sometimes an existing form is abbreviated, or "clipped", e.g., *flu* < *influenza*, Gm. *Uni* < *Universität*, Fr. *fac* < *faculté*. Typically, one (relatively) accented syllable is selected (4a); however, occasionally nonstressed syllables can also be selected (4b) (Hock and Joseph 1996:302). If abbreviation of a phrase occurs (4c), this is called ellipsis:

(4) a. *fridge* < *refrigerator*
mike < *microphone*
phone < *telephone*
b. *bus* < *omnibus* 'for all' (Lat. *omn-* 'all' + *-ibus* DAT.PL)
fan < *fanatic*
c. *drive-in* < *drive-in theatre*
narc < *narcotics agent*
pub < *public house*
roll-away < *roll-away bed*
sci fi < *science fiction*

Both clipping and ellipsis have been adduced as cases of lexicalization (see Anttila 1989 [1972]:151; Blank 2001:1604, 1605). According to Blank, clipping concerns the deletion of one or more syllables from multisyllabic words, whereas ellipsis leads to the formal reduction of a complex word or phrase. In ellipsis the semantics of the omitted element is absorbed into the remainder by metonymy; thus in earlier work ellipsis was often regarded as a primarily semantic change (e.g., Ullmann 1962). However, ellipsis clearly involves both form and meaning (Hock and Joseph 1996:175) as in:

(5) *canary* < bird from the Canary Islands
damask < *Damask cloth* (cloth from Damascus)
denim < Fr. *serge de Nîmes*
jeans < *Gene* (= Genoa) *cloth*

Hock and Joseph point out that words that arise by ellipsis of this kind may themselves be subject to compounding, e.g., *blue jeans*, or metonymic extension, e.g., *canary* for a yellow color (1996:175).

A more elaborate type of ellipsis is cited by Morena Cabrera (1998:216): Sp. *el hombre de los bigotes* 'the man of the moustache' > *el de los bigotes* 'the of the moustache' > **el de bigotes* 'the of moustache' > *el bigotes* 'the moustache, i.e., the one wearing the moustache.'

2.1.5 Blending

Blending, which involves the fusing of words into a single lexeme, or "portmanteau word," by a process of compounding and clipping, is also treated as lexicalization (see, e.g., Blank 2001:1605–1606). Blending may destroy the integrity of both elements (6a), or it may clip off the beginning or end of the one element leaving either the first (6b) or the second element (6c) intact:

(6) a. *bit* < *b(inary)* + *(dig)it*
infomercial < *info(rmation)* + *(com)mercial*
heliport < *heli(copter)* + *(air)port*
b. *filmography* < *film* + *(bi)ography*
skyjacker < *sky* + *(hi)jacker*
c. *docudrama* < *docu(mentary)* + *drama*
permapress < *perma(nent)* + *press*
blog < *(we)b* + *log*

When the two words end or begin with the same sound(s), as in the case of *guestimate* < *guess* + *estimate* or *netiquette* < *net* + *etiquette*, the formation can be seen as proceeding in either direction.

2.1.6 Back formation

Lexicalization may also encompass back formation, or the creation of a morphologically simple form from a word which is analyzed (frequently incorrectly), as a morphologically complex word on the basis of analogy with derivational (7a) and inflectional (7b) patterns existing in the language (Hock and Joseph 1996:163–164):

(7) a. *emote* < *emotion*
enthuse < *enthusiasm*
laze < *lazy*

orate < *orator*
orientate < *orientation*

b. *pea* < Fr. *pease* (SG)
sherry < Sp. *Xeres* (SG placename, nativized as *sherries*, referring to the fortified wine imported from Xeres (de la Frontera))
(ice/roller) skate < Du. *schaats* (SG)

In (7b) the root-final *-s* of the borrowed term has been analyzed as an Eng. *-s* plural.

2.1.7 Initialism/acronym

Another process sometimes included in lexicalization is the formation of acronyms, which conflates the initial sounds (or letters) of words within a complex word or phrase into a unified lexeme. Acronyms may be represented as fully formed, unified lexemes in lower case (8a), but they are more often represented in upper case and hence are less obviously lexemes (8b):

(8) a. *laser* < *l(ight) a(mplification) (by) s(timulated) e(mission) (of) r(adiation)*
radar < *ra(dio) d(etecting) a(nd) r(anging)*
scuba < *s(elf)-c(ontained) u(nderwater) b(reathing) a(pparatus)*
Erasmus < *E(u)r(opean community) a(ction) s(cheme for the) m(obility of) u(niversity) s(tudents)* (Blank 2001:1605)
b. *AIDS* < *a(uto) i(mmune) d(eficiency) s(yndrome)*
NATO < *N(orth) A(tlantic) T(reaty) O(rganization)*

A related process is the production of initialisms, in which the initial letters of words within a complex word or phrase are articulated as letters forming a word. Writing conventions (the use of upper case, or lower case with periods) suggest that initialisms are not viewed as fully formed lexemes, e.g.:

(9) *IBM* < *I(nternational) B(usiness) M(achines)*
ID < *id(entification)*
OED < *O(xford) E(nglish) D(ictionary)*
PBS < *P(ublic) B(roadcasting) S(ystem)*
VIP < *v(ery) i(mportant) p(erson)*
TV < *t(ele)v(ision)*

The choice of writing convention may make a difference in meaning, e.g., *AM* < *amplitude modulation* vs. *a.m.* < Lat. *ante meridiem*. Occasionally, however, initialisms seem to be treated as unified lexemes, e.g.:

(10) *emcee (MC)* < *m(aster) (of) c(eremonies)*
veep (VP) < *v(ice) p(resident)*

2.1.8 Loan translation

Loan translation or "calquing" may be mentioned as a case of lexicalization (see Blank 2001:1606). Calques, which are often compounds or idioms, are literal (morpheme-by-morpheme) translations of terms from another language. In the source language, the terms may themselves be derived forms (see 11a) or partial borrowings (see 11b), and once calqued may undergo normal processes of change such as phonological reduction. Calques may also be incomplete ("blended") translations from the source language; for example, Blank cites Fr. *lecteur compact disque* < Eng. *compact disk player*, where *lecteur* is Fr. for 'reader' (not a translation of *player*). Old English was a period in which loan translation was used extensively for increasing the vocabulary, a characteristic shared by Modern German. For example, in his Grammar, Ælfric used the loan translations in (11a); some Modern German examples are given in (11b):

(11) a. OE *fore-set-nys* 'in front-set-ness' (< Lat. *praepositio* 'preposition')
OE *ge-tacn-ung* 'en-sign-ing' (< Lat. *significatio* 'meaning')
OE *betwux-aleged-nys* 'between-put down-ness' (< Lat. *interiectio* 'interjection')
b. Gm. *Fernseher* 'far-seer, television'
(cf. Gk *tēle* 'far' + Fr. *vision*)
Gm. einmütig 'one-minded, unanimous'
(cf. Lat. *un-* 'one' + *animus* 'mind') (Anttila 1989 [1972]:140)

While the technical terms in (11a) have been replaced by the Lat. terms from which they were translated, other OE calques have survived (and undergone phonological and semantic change), like *gospel* < OE *gospell* < *god* + *spell* (< Gk. *euangelion* 'good tidings') or *Holy Ghost* < OE *Halig Gast* (< Lat. *Spiritus Sanctus*).

2.1.9 Coinage or root creation

A rare or exceptional process – often contrasted with routine processes of word formation – is "coinage" or "root creation." Rather than making use of existing linguistic elements, coinage involves the invention of a new root morpheme. Some consider coinage to be totally arbitrary, others allow for some motivation. For example, Bussmann writes of "the first-time creation of an unmotivated ... i.e. non-complex and completely arbitrary ... connection between expression and content" (1996:s.v. "coining"), whereas McArthur says that a coinage may be either "motivated" or created "ex nihilo" (1992:s.v. "root creation"). Motivated coinages include onomatopœic or echoic words such as *gulp*, *hiss*, *blab*, *vroom*, or *zap* and those involving some resemblance to a pre-existing form, such as *splosh* or Kerouac's *sploopsing* (see Quirk et al. 1985:1524). In the case of ex nihilo

coinages, as McArthur notes, there is "no lexicological way of accounting for the formation of a word" (1992:s.v. "root-creation"). While everyday examples of root creation in Present-day English (PDE) are difficult to adduce (*quark*, *granola*, *Kodak*, and *googol* are the standard examples), "literary coinages" (such as Shakespeare's *multitudinous*, *dwindle*, Milton's *sensuous*, *oblivious*, Spenser's *blatant*, *askance*, Lewis Carrol's *chortle*) are generally less restricted in context of use. McArthur observes that coinage is common in fantasy literature (e.g., *hobbit*). He concludes that "[t]he limits of actual or apparent root-creation are hard to establish, because it shades into ... conventional processes of word-formation" as coinages become the base for more complex forms (1997:s.v. "root-creation").

2.1.10 Metalinguistic citation

Finally, word formation may include what can be termed "metalinguistic citation," or a speaker's ability to pick up any piece of linguistic material and make it into a word, as in *There are two e's in my name*, where the letter *e* is pluralized and behaves as a noun. Heine, who prefers to call this process "revalorization," recognizes this as a type of lexicalization: a process which "serves to assign a segment of text, irrespective of whether it is a meaningful entity or not, the status of a referential, lexical form" (2003b:166). He includes in "segment of text" sounds, letters, grammatical forms, as well as full sentences.

2.1.11 Summary

In conclusion, we have seen that, because ordinary processes of word formation create new lexemes in language, they are treated as instances of "lexicalization." As we will discuss more fully below, historically, compounds and derivations are most obviously cases of lexicalization because they involve processes of fusion which serve to erase or efface boundaries between independent morphemes and give rise to unified lexemes over time. Ellipsis and clipping, though they do not involve processes of fusion, do serve to create monomorphemic forms by eliminating morphemes, and hence morphemic boundaries. More controversy surrounds conversion and back formation as exemplars of lexicalization; although they give rise to new lexemes, these processes do not involve the loss of morphemic boundaries. Nonetheless, conversion, especially conversion from minor to major (functional to lexical) class, is often treated as lexicalization since it involves the shift from less > more lexical.

Although the processes of word formation discussed in the previous sections are typically treated as part of the synchronic component of the language, Quirk et al. observe that "a description of word-formation must

of its nature be diachronic" because "the utterance of a new word raises the possibility that an addition has been made to the total means of expression available" (1985:1525). Thus, in the following section, we explore the processes by which new words are spread and become added to the inventory, and in Section 2.3, we turn to processes of word formation which are more universally regarded as diachronic in nature.

2.2 Institutionalization

How does a new form innovated on the basis of word formation rules come to be conventionalized as part of the accepted vocabulary of a community? This question is usually answered in terms of INSTITUTIONALIZATION, which is sometimes regarded as a precursor of lexicalization, sometimes identified with it. The focus here is not strictly on structural processes; rather, "institutionalization" refers to the spread of a usage to a community and its establishment as the norm. It is "the integration of a lexical item, with a particular form and meaning, into the existing stock of words as a generally acceptable and current lexeme" (Lipka 2002 [1990]:112, who refers to similar definitions by Quirk et al. 1985:1522 and Bauer 1983:48; see also Lipka 1994).

When first used for a particular occasion, an innovation is known as a NONCE WORD or NONCE FORMATION.[15] A nonce formation is a new complex word coined on the spur of the moment or made up casually on the fly out of phrases, e.g., *slow-growther*, *skilled man-hours*, *eating-habit-wise*, or *bandwagon-jumper-on* (referring to a type of person). In contrast to a coinage, a nonce word is formed by applying regular word formation rules. It serves an immediate communicative need of, or solves a problem for, the speaker, whether it be economizing, filling in a conceptual/lexical gap, or creating a stylistic effect (Bussmann 1996:s.v. "nonce word"). A nonce word must be understood from its context.

As a nonce formation comes to be accepted by part or all of the speech community (it is "institutionalized"), it becomes a new word of the language, or a NEOLOGISM (Bussmann 1996:s.v. "neologism"; Matthews 1997: s.v. "neologism"). In this process, it often comes to be limited, specialized, or fixed in meaning. It may narrow to a subset (perhaps one) of its possible meanings, become relatively independent of context, and be included in the dictionary along with more generic meanings (Ryder 1999:305–306). Bauer gives the example of (British English) *telephone box* being institutionalized in the meaning 'telephone kiosk' as opposed to 'box shaped like a telephone,' 'box which functions like a telephone,' or other possible meanings (1983:48).

[15] A single instance in the historical record is known as a "hapax legomenon."

A more recent example is the use of clipped *hood* (< *neighborhood*) to refer to inner city, largely African American, neighborhoods.[16]

For some scholars, institutionalization is the step which precedes lexicalization. Bauer suggests a path of lexicalization as follows (1983:45ff.; cf. Ryder 1999:305–306):[17]

nonce formation > institutionalization > lexicalization

Institutionalization occurs "when the nonce formation starts to be accepted by other speakers as a known lexical item," and lexicalization occurs when the formation "has, or takes on, a form which it could not have if it had arisen by the application of productive rules" (Bauer 1983:48). At the institutionalized stage, a word is created by productive processes and listed with its full derivational history, but over time various features are lost, the word diverges from the expected pattern in unpredictable ways and begins to behave as if it were a monomorphemic entity (Bauer 1992:566).[18] Ryder (1999) points out further that the meaning of the lexicalized form is not completely predictable, it is independent of context, and it is given in the dictionary.

The focus of Bauer's and Ryder's analysis is the form and the meanings associated with it. In discussing institutionalization, Quirk et al. (1985:1525ff.) comment on a different aspect of institutionalization, one associated with coding of meaning. Like Bauer, they propose three stages, the first at which the entity/activity/quality is unfamiliar, the second at which there is some discussion of and familiarization with the notion, and the third at which it is widely accepted as viable. These stages correspond to three formal structures: sentence or periphrasis, nominalization, word. The authors suggest that lexicalization occurs when "a notion that could previously be discussed only in sentences and periphrases which varied from person to person" is expressed in a single lexical item (1526), for example, *engines powered by electricity* may become familiar enough for speakers to choose an expression involving productive word formation such as compounding, e.g., *electrically-powered engines*; later, when the idea is institutionalized, a word formation process may be selected that allows for expression by a single word, e.g., the nominalization *electrification*.

While Bauer and Ryder see institutionalization as a precursor of lexicalization, others see the two as coextensive. Blank (2001:1597–1599) explicitly

[16] The shift from nonce to institutionalized status was observable in writing during the 1990s with the shift from '*hood* (with apostrophe) to *hood* (without).

[17] Clark and Clark (1979), however, suggest a path from innovation to varying degrees of "idiomatization," see below.

[18] Because institutionalization is the result of productive processes and lexicalization of unproductive processes, the former is considered to be "opposed to lexicalisation, although both create established words" (Bauer 1988:246).

rejects the distinction between institutionalization (conventionalization of a complex word which remains semantically completely predictable) and lexicalization (subsequent demotivation or idiomatization of the complex word). He argues that "the sense of a word formation is never completely predictable from the meaning of its components" (1598) and that idiomatization is present from the start and is not due to the lexicalization process (1600). He notes that even an apparently nonlexicalized form such as *musical theory* is idiomaticized to some degree, as are more clearly lexicalized forms such as *rattlesnake*, *crybaby*, and *wheelchair*. Likewise, Pawley and Syder argue that "What makes an expression a lexical item, what makes it part of the speech community's common dictionary, is, firstly, that the meaning of the expression is not (totally) predictable from its form, secondly, that it behaves as a minimal unit for certain syntactic purposes, and third, that it is a *social institution*" (1983:209; italics original). The lexical item serves as a conventional label for a conventional concept; thus, we have *backache* and *headache*, but not *footache* and *thighache*. According to Pawley an expression gains institutional status when it becomes a member in a set (e.g., *go to school*, *take a day off*), it acquires customary status (e.g., *back door*, not *rear door*), or legal status (e.g., *breaking and entering*), it becomes a speech act formula (e.g., *excuse me, may I come in*?), or is used as an acronym (e.g., *ASAP*) (1986:105–106).

Institutionalization may go by the terms "petrification" (Leech 1981 [1974]:225–227), "routinization" (Haiman 1994), "canonization" (Morita 1995:471), or "usualization" (Blank 2001:1596). Institutionalized forms may also be said to be "received" or "frozen" (see Bauer 1983:48n.). While "canonization" and "usualization" are clearly related to adoption by a group of speakers in a community as a norm, the other terms are especially frequently found in connection with the concept of lexicalization as fusion, to which we now turn.

2.3 Lexicalization as fusion

The usual synchronic perspective on lexicalization as fusion (usually referred to as "conflation" or "coding") concerns alternative ways of coding the same concept (see 1.3.1). Little attention has been paid to this perspective from a historical point of view. Here meaning is kept constant and the expression changes. An example would be discussion of whether causative comes to be expressed phrasally (*make red*, *make brown*), derivationally (*redden*), or monomorphemically (*brown*).

In contrast, the usual historical perspective on lexicalization as fusion involves tracing the development of a form from a more complex to a simpler sequence (sometimes, but not always, the meaning may change considerably). A typical example is fusion resulting from compounding and coalescence, as in the case of *garlic* from OE *gar* 'spear' + *leac* 'leek.'

The fossilization or petrification of constructions into inseparable units is key to the notion of the lexicon as an inventory (Lyons 1977:547; Bauer 1983:48). We start with relatively uncontroversial examples of fusion of phrasal, syntagmatic constructions (2.3.1), then move on to fusion that involves word formation (2.3.2) and phonological developments (2.3.3). Finally, we introduce some primarily semantic and pragmatic viewpoints on fusion (2.3.4).

2.3.1 Syntagm > lexeme

One of the commonest conceptions of lexicalization is the unification, or UNIVERBATION, of a syntactic phrase or construction into a single word. Lehmann (1995 [1982]:7–8) identifies Žirmunskij (1966) as one of the first to highlight unification of syntactic phrases as the source of many lexical items. Some representative descriptions of lexicalization as univerbation follow:

- "the phenomenon that a complex lexeme once coined tends to become a single complete lexical unit, a simple lexeme. Through this process it loses the character of a syntagma to a greater or lesser degree" (Lipka 2002 [1990]:111),
- "[t]he process of *unification* (whereby a word combination develops into a single (compound) word)" (Lessau 1994:s.v. "lexicalization," Lessau's emphasis),
- "the process whereby independent, usually monomorphemic, words are formed from more complex constructions" (Traugott 1994:1485),
- "when a phrase or a syntactically-determined lexical item ... becomes a full-fledged lexical item in itself" (Moreno Cabrera 1998:214),
- "related to desyntacticization, in the sense of a syntagmatic structure losing its syntactic transparency and merging into a single lexical item" (Wischer 2000:364),
- "a process by which complex words formally and semantically [lose] their motivation" or "a process by which complex words become simple words" (Blank 2001:1603).

Hagège (1993:171ff.) defines lexicalization as the process "by which two elements that, when used in the sentence framework, were in the past or are today free units, or two elements of which one is already frozen into an affix, are associated through compounding and derivation respectively, and yield a new complex unit." Also, Biber, Johansson, Leech, Conrad, and Finegan (1999:58–59) argue that lexicalization occurs when a multi-word lexical unit comes to function as a single grammatical unit, or part of speech. Unlike free combinations of words, these show limited possibilities for substitution, tend to be written as single words, and may acquire meanings which are not predictable from their individual parts.

Univerbations of older provenance often involve some degree of phonological reduction and are morphologically and semantically opaque (12a), while those of more recent provenance may be relatively transparent both in form and in meaning (12b) and be considered to be "idioms" (see 2.3.4):

(12) a. *each* < OE *a-gelic* 'ever-like' (*OED*)
either < OE *a-hwæþer* 'each of two' (*OED*)
handicap < *hand in the cap*
hobnob < *hab nab* < OE *hab ne-hab* 'have not-have' (*OED*)
topsy turvy < *top so turvy* (< ME *terve* 'turn'? [*OED*])
wherewithal 'necessary means' < *where with all*[19]
willy-nilly 'whether desired or not, haphazardly' (< OE *will* 'want' *ye/he nill* 'not want' *ye/he* (*OED*)
It. *forse* 'perhaps' < Lat. *fors sit an* 'case be that' (= 'be the case that') (Giacalone Ramat 1998:122)
It. *adesso* 'now' < Lat. *ad ipsum* 'to itself-ACC' (Giacolone Ramat 1998:122)
Sp. *correveidele* 'gossipy' < *corre, ve y dile* 'run, see and say (it) to him/her' (Ramat 2001:396)

b. *dyed-in-the-wool* 'inveterate'
more or less 'somewhat'
mother-in-law
ne'er-do-well[20]
nuts-and-bolts 'practical details'
out-of-pocket 'paid out in cash'
stick-to-it-iveness 'persistence'[21]
Fr. *aujourd'hui* < *au jour d'hui* 'on the day of today' (Blank 2001:1602)
Gm. *demzufolge* 'consequently' < *dem zu folge* 'that-DAT to follow' (Ramat 2001:394)
Gm. *Langeweile* 'boredom' < *lange Weile* 'long time' (Blank 2001:1601)

A rather specialized type of fusion is involved in "delocutivity" (see Benveniste 1971a [1958]), in which "a whole utterance is transformed into a more or less complex word expressing a contiguous concept" (Blank

[19] Note the conversion to a nominal form. See van der Wurff (2002) on the development of Adv *withal* 'in addition, notwithstanding, therewith' > Prep *withal* 'with'; he would seem to see the development of the adverbial form as a case of lexicalization and the development of the prepositional form as grammaticalization occurring in the context of infinitives.

[20] Note the frozen archaic form *ne'er*.

[21] The syntagma "stick-to-it" has become the root for derivation by adjectivalizing *-ive* and nominalizing *-ness*.

2001:1602, 1604), as in It. *non so che* 'I don't know' > *nonsoche* 'something which is difficult to explain' or Sp. *vuestra merced* 'your honour' > *Usted* 'you (formal)'. Eng. *goodbye* < *God be with you* represents another such example. Blank suggests that this transformation of a speech act into a word, though a "marginal" phenomenon, is "maybe" the "most typical" case of lexicalization (1604), perhaps because he views it as both a process of conversion (see 2.1.3) and of fusion (see further 4.2.1).

2.3.2 Complex > simple lexeme

Many who associate lexicalization with word formation restrict their attention to precisely those processes of word formation that have become synchronically unproductive, and thus result in frozen, irregular, unpredictable, or idiosyncratic forms. For Bauer, lexicalization occurs when complex forms must – for semantic, morphological, or phonological reasons – be listed separately rather than generated freely by a productive set of rules. A lexicalized form "cannot be derived by the synchronically applicable rules of derivational morphology, for whatever reason" (Bauer 1978:6). Blank (2001:1601) points out that the loss of word formation rules may serve to make complex words difficult to analyze, hence demotivated and lexicalized, as in It. *terremoto* < Lat. *terrae motum* 'earth-GEN motion,' i.e., 'earthquake' (the Lat. genitive is opaque in Italian; see also Gm. *Sonnenschein* 'sun-GEN shine' with obsolete genitive). Blank asserts that such fusion in compounds is particularly observable in adverbs, conjunctions, prepositions, and pronouns, though he gives no examples (1602). Lexicalization is thus a diachronic process that leaves idiosyncratic traces (Bauer 1983:50). It "interferes with regular word-formation processes and overlaps their result" (Lipka 2002 [1990]:111).

Perhaps the clearest example of lexicalization in the sense of unproductive word formation processes is the case of fused or amalgamated compounds in which significant phonological and morphological changes or losses from the vocabulary have made the parts of the compound opaque, as in the frequently cited *lammas* < OE *hlaf* 'loaf' + *mæsse* 'mass.' Fused compounds include those in which one root is recognizable, but the other is obscure (13a), and those in which neither root in the compound is recognizable (13b):

(13) a. *cobweb* < OE (*atter*) *coppe* 'spider' + *web* 'web'
earwig < OE *eare* 'ear' + *wicga* 'one that moves'
mermaid < OE *mere* 'sea' + *mægd(en)* 'maiden'
mildew < OE *mele* 'honey' + *deaw* 'dew'
b. *gospel* < OE *god* 'good' + *spell* 'tidings'
gossip < OE *god* 'god' + *sib(b)* 'relation'
halibut < OE *halig* 'holy' + *butte* 'flat fish'

lord < OE *hlaf* 'loaf' + *weard* 'guardian'
orchard < OE *wyrt* 'herb' (or *ort* < Lat. *hortus*) + *geard* 'yard'
(*American Heritage Dictionary*, *OED*)[22]

Blank (2001:1600) cites the "berry" forms of Eng. (*cran*-, *boysen*-) or Gm. (*Him*-, *Brom*-), in which parts of a compound have been lost from the lexicon as in (13a). Lehmann (1989:13) cites the MHG compound *adal-aar* 'edel-Adler' ('noble eagle') > monomorphemic ModGm. *Adler* 'eagle' as an instance of lexicalization, and *husband* (< *hus* 'house' + ON *bōndi* 'dweller' [present participle of *būa* 'dwell']) are cited in Huddleston and Pullum as an "extreme case of lexicalisation" (2002:1629) (these are examples of the type in (13b)). Likewise, Bauer (1983:51–52) sees as exempla of "phonological" lexicalization the separation of *hus*- in *husband* from *house*, which results from phonetic change, or of *-day* [dɪ] from *day* [daɪ] (in the names for the days of the week, in some dialects), which results from change in the phonological features of a morpheme due to stress in a compound.

Another aspect of the behavior of compounds which has been subsumed under lexicalization is what Bauer (1983:59) calls "syntactic lexicalization," when it is not possible to predict syntactic behavior from general principles such as word order patterns, as in the case of *pickpocket*, *scarecrow*, or *spoilsport*.

Many derivational morphemes originate in compounded roots, including some that are quite productive (14a) in English, and others that have become unproductive (14b):

(14) a. *-ly* < OE *lice* 'body, likeness'
-ment 'action' < Fr. *–ment* '-ly' (adverbial) < late Lat. *–mente* 'in that manner-ABL' < Lat. *–mente* 'mind-ABL' in such phrases as *obstinate mente* 'with obstinate mind'[23]
b. *-ric* < OE *rice* 'realm'
-dom < OE *dom* 'setting, jurisdiction'

The development of derivational morphology from compounds has typically been considered a prototype instance of morphologization of syntactic structure (see, e.g., Anderson 1992) and of grammaticalization (see, e.g., Ramat 2001; but see 2.4 below). However, derivational affixes which have become unproductive, such as *-ric* or *-dom*, are typically also included under the rubric of lexicalization (see, e.g., Blank 2001:1602).[24] As instances of the subtype of lexicalization he calls "morphological," Bauer cites derivatives such as *eat*/*edible* or *right*/*rectitude*, which are morphologically related but have different roots, one being productive and the other nonproductive

[22] For further examples see Skeat (1887:420–430).
[23] We thank Harm Pinkster for this etymology.
[24] The role of derivational morphology in grammaticalization is problematic (see Cowie 1995), as discussed in Section 3.4.

(1983:53–54). The nonproductive root plus affix is said to be lexicalized and must be listed separately in the lexicon, e.g., *eat* + *-able* = *edible*.

2.3.3 Demorphologization and phonogenesis

Among changes discussed in the literature on lexicalization (and on grammaticalization) is "demorphologization" or "demorphemization," the movement from morphology either into phonology or into syntax (see Joseph and Janda 1988:198–202).[25] The shift into syntax is discussed in Section 2.4. Here we consider some of the changes known as DEMORPHOLOGIZATION whereby a morpheme loses (most of) its grammatical-semantic contribution to the word and becomes an indistinguishable part of the construction of the word, while retaining part of its original phonological substance (see Hopper 1990:154). If the perspective taken is primarily that of phonology and of "the process whereby new syntagmatic phonological segments are created out of old morphemes," it can be called PHONOGENESIS (see Hopper 1994:31). Examples (derived from Hopper 1990, 1994) of "morphological residues" resulting from such changes may be observed in:

(15) *alone* < *all* + *one*
about < OE *on* + *be* + *utan* 'on/at + by + outside'
besides < OE *be* + *sidan* + *-es* 'by + side-GEN'[26]
eleven (cf. Gothic *āinlif* 'one-left')
filth (containing the original abstract noun-forming suffix **-þi-*)[27]
handiwork < OE *hand* + *ge* + *weorc* 'hand + prefix + work'
seldom (where *-om* is said to reflect the original dative plural *–um*)[28]
sooth 'true' (deriving from a present participle **?sonts*, cf. Dan. *sand* 'true')
stand (where *-n-* represents an older nasal infix)
swine/swill/swig (where *sw-* derives from PIE **sū* 'pig')

Hopper observes that often phonogenesis creates either basic vowels or apical consonants (s, t, d, θ, l, r, n). The latter can be found in old comparatives in *-r*

[25] Joseph and Janda (1988) regard most examples adduced prior to their article as somewhat problematic, however.
[26] Rissanen (2004) treats the development of *besides* as a clear case of grammaticalization.
[27] Bauer (1983), however, cites *-th* as an example of "morphological lexicalization."
[28] According to the *OED*, *seldom* derives from OE *seldan*, which was altered to *seldum* by analogy with adverbial datives such as *whilum*; thus the ending of *seldom* is not an original dative ending. (Thanks to Minoji Akimoto for pointing out this correction.)

(*near*, *rather*), old strong past participle in *-n* (*forlorn*, *rotten*), old genitives in *-s* (*once*, *against*, *huntsman*), or Gmc. present participles in *-ende* (*friend*, *fiend*) (Hopper 1994:40–41). Van der Auwera gives the example of *twit* < OE *æt-witan* 'at blame, reproach,' where *t* has become phonological (2002:21).

Lehmann (1989:13) describes bound morphemes becoming submorphemic and being integrated into the root as lexicalization, as in the case of the unproductive, formerly morphemic *-t* in Gm. words such as *Ankunft*, *Flucht*, and *Sicht*. Although not focusing on the lexicalization aspect, Skeat (1887:430) speaks of "petrifications" in which grammatical inflections are preserved long after the notion of the inflection has passed from them, as in *whilst* < *while-s-t* with genitival *-s* and excrescent *t*, or *since* < *si-n-ce* where the *n* represents an original dative *-m* (OE *siððam* > *siððan*) and the *-ce* represents an original genitive *-s*.

Greenberg (1991) discusses a possible development of definite articles in which they contract until they are found in fossilized form in only a few lexical forms. The article is lexicalized, Greenberg explains, "in the sense that synchronically it is a part of the host morpheme" (1991:301). An example in English is *nother* < *an other* (as in *a whole nother*). The incorporation of definite articles into borrowed forms, such as *alcohol* < Arabic *al koh'l* or Sp. *alcalde* 'mayor' < Arabic *al qadi* 'the judge,' which Blank (2001:1605) treats as a subtype of clippings, likewise bear a resemblance to Greenberg's (1991) examples. According to Greenberg (1991), if the reduced forms of articles later acquire new grammatical functions, they are "regrammaticalized." In the grammaticalization literature, REGRAMMATICALIZATION potentially refers to three possible circumstances: (a) a form without any function acquires a new grammatical function; (b) a form is reinterpreted in a new grammatical function; and (c) a form which has lost its function regains it (see Anttila 1989 [1972]:150; Heine, Claudi, and Hünnemeyer 1991:4, 262n.11; Lessau 1994:s.v. "regrammaticalization"; A. Allen 1995).

Ramat defines "a process whereby linguistic signs formed by rules of grammar are no longer perceived (parsed) in this way" but simply as "lexical entries" and calls this "lexicalization" (1992:550–551), citing examples such as comparatives that have lost their grammatical status (*elder*, *mayor*) or participles that are no longer part of the verbal paradigm (*shorn*, *cloven*) (many of his examples are the same as Hopper's [1990, 1994]). Ramat distinguishes this from "relexicalization," giving the example of the Lahu verb *lá* 'to come,' used first in serial verb constructions and then as a bound deictic particle (i.e., as a grammatical element); it is finally incorporated into the verb and becomes synchronically unanalyzable and opaque as part of the lexical root (554). The difference from examples like *elder* is that in the Lahu case the incorporated element started out as a lexical item in a syntagmatic construction, became grammatical, and then became lexical once more. Thus, Ramat observes, in addition to Givón's well-known observation that "[t]oday's morphology is yesterday's syntax" (Givón

1971:413), we need to recognize that "today's grammar may become tomorrow's lexicon" (Ramat:1992:557).

Closely related to phonogenesis are cases of PHONOLOGIZATION in which phonologically conditioned alternations may eventually split into new phonemes as a result of erosion of segments. Hopper argues that whereas phonogenesis changes morphemic elements into segmental phonemes on the syntagmatic level, phonologization is a paradigmatic process affecting the inventory of phonemes in a language. A well-known example is the emergence of front rounded vowels in German as a result of "umlaut" (also known as "*i*-mutation"): back vowels followed by *-i*, *-j* had phonologically conditioned fronted variants; when *-i* and *-j* were lost by erosion, the fronted vowel ceased to be conditioned, cf. ModGm. *schön* 'beautiful' < OHG *sconi*. Such changes often lead to new paradigmatic contrasts; thus, in the class of PGmc. nominals taking a plural marker **-iz*, umlaut led to marking of plural by vowel alternation as in *foot–feet* (PGmc. **fōt* 'foot' + **-iz* PL). Note that when new paradigmatic contrasts arise as a result of phonological changes that have become opaque, the morphological structure has not changed, but the condition for the phonological variation is no longer phonological; it is morphological and this process is often known as MORPHOLOGIZATION of phonological rules (see Joseph and Janda 1988:196–197; Anderson 1992:339ff.). Such changes may also produce morphologically opaque lexical pairs; for example, in PGmc. the causative was formed derivationally by past singular of the verb + *-j-*; umlaut led eventually to the formation of non-causative/causative pairs like *sit*/*set* (the latter is < PGmc. *sat* 'sit-PAST' + *j-* CAUSATIVE); see also *lie*/*lay*, *sit*/*set*, *fall*/*fell*, or *drink*/*drench* or the N/V pair *stink*/*stench* (Hopper and Traugott 1993:223n.2; Newmeyer 1998:263–264).

Thus, we have seen that there are two different types of fusion that have been treated as "lexicalization," as they both produce unified lexemes. In one, the individual parts which are fused remain distinguishable but form a single lexeme, as in the case of syntagmas fused together as words (e.g., *run-of-the-mill* 'ordinary') or of dependent roots reduced to bound affixes (e.g., OE *had* 'state, rank of person' > *-hood*). In the other, the individual parts are fused together in a single lexeme and have coalesced in such a way that they are indistinguishable, as in the case of fused compounds (e.g., *hussy* < OE *hus* 'house' + *wif* 'wife') or of phonogenesis (e.g., *awake* < OE *on* + *wacan*). Such forms are not subject to synchronic morphological analysis.

2.3.4 Idiomaticization and demotivation

Here we consider some of the semantic and pragmatic aspects of fusion that are typically discussed in the literature on lexicalization, specifically idiomatization and demotivation.

IDIOMATICIZATION is associated with routinization, which leads to univerbation, compacting, obliteration of boundaries, and simplification.

Difficulties have been encountered in attempts to provide a definition of "idiom" (see, e.g., Clark and Clark 1979; Nunberg, Sag, and Wasow 1994), though typically three features have been seen to be predominant:

(a) semantic opacity or noncompositionality: it is impossible to deduce the meaning of an idiom such as *shoot the breeze* from 'shoot' + 'the' + 'breeze,'
(b) grammatical deficiency: an idiom does not permit the syntactic variability characteristic of free combinations, such as passive (**the breeze was shot*), negation (?*didn't shoot the breeze*), internal modification (**shoot a strong breeze*, **shoot breezes*, **shoot some breeze*), or topicalization (**the breeze he shot*),
(c) lack of substitutability: synonymous lexical items cannot be substituted (**shoot the wind*, **fire at the breeze*) nor can items be reversed or deleted.

Pawley (1986) cites two major criteria associated with idioms: (a) syntactic restrictions, including transformational defectiveness (e.g., *theatre-goers*, but not *school-goers*), the lack of underlying structure (e.g., *it's easier said than done*), and syntactic ill-formedness (e.g., *long time no see*, *once upon a time*) (109–110), and (b) arbitrariness, including semantic idiomaticity (*drop a brick* 'commit a social gaffe'), arbitrary selection of one meaning (e.g., *bullet hole* 'the entry place for a bullet,' not 'a hole shaped like a bullet'), and arbitrary selection of one form (111–112). Nunberg, Sag, and Wasow argue further that the defining characteristic of idioms is not noncompositionality but rather conventionality, or the fact that the meaning or use of idioms "can't be predicted, or at least entirely predicted, on the basis of a knowledge of the independent conventions that determine the use of their constituents when they appear in isolation from one another" (1994:492).

It has come to be recognized that many idioms are not entirely opaque semantically nor invariant syntactically and morphologically (see, e.g., Fillmore, Kay, and O'Connor 1988). Rather, idiomaticity must be recognized as a graded concept, with combinations ranging from more to less idiomatic in nature. Nunberg, Sag, and Wasow (1994) make a distinction between what they call "idiomatic phrases," such as *saw logs* = 'snore,' where idiomatic meaning is not distributed over the parts, and "idiomatic combinations," such as *spill the beans* = 'divulge the information,' where the parts carry identifiable portions of the idiomatic meaning. The latter allow many grammatical modifications, such as adjectival modification (*leave no legal stone unturned*), quantification (*touch a couple of nerves*), negation (*spill no beans*), pluralization (*drop a hint*/*hints*), passivization (*the decks were cleared*), and so on not allowed by the former. Bauer sees some idioms as "partially fossilized" or semantically specialized, their interpretation being dependent on textual and/or extralinguistic factors. He calls such forms "received forms" and believes that they can be "thawed out" or explained in terms of their constituent elements (Bauer 1978). And Blank

(2001:1599–1600) distinguishes three degrees of idiomaticity in compounds, ranging from Gm. *Hundekuchen* 'dog biscuit,' in which the literal meanings of the combined words are preserved but the meaning of the complex word is not fully predictable, to *wheelchair*, which is still based on the usual meaning of the components but refers to untypical or specialized representatives of the designated category, to forms involving metonymic or metaphoric transformations, such as It. *bocca di leone* 'snap-dragon.'

Despite a certain lack of consensus about the meaning of "idiom," the identification of lexicalization with idiomaticization is widespread (Bauer 1983; Lehmann 1989; Lipka 1992; Bussmann 1996; Wischer 2000; Huddleston and Pullum 2002; Traugott 2005). Indeed, according to Lehmann (2002:14) idiomaticization IS lexicalization in the sense of coming to belong to an inventory, and Moreno Cabrera (1998:214) points to idioms as the best example of lexicalization. Lipka (1992:97) cites examples such as *wheelchair*, *pushchair*, and *trousersuit*, which have specific and unpredictable meanings. Bussmann considers idiomaticization to be the diachronic element of lexicalization, which occurs when "the original meaning can no longer be deduced from its individual elements" or "the original motivation of [a] unit can only be reconstructed through historical knowledge," as in the case of *neighbor*, *cupboard*, or *mincemeat* (1996:s.v. "lexicalization"; also see "idiomaticization," "motivation").

In an early study of word formation and semantics Kastovsky defined lexicalization as "die Eingliederung eines Wortbildungs- oder syntaktischen Syntagmas in das Lexikon mit semantischen und/oder formalen Eigenschaften, die nicht vollständig aus den Konstituenten oder dem Bildungsmuster ableitbar sind. Idiomatisierung und Demotivierung beziehen sich dabei auf entsprechende semantische bzw. formale Eigenschaften" ('the integration of a word formation or syntactic construction into the lexicon with semantic and/or formal properties which are not completely derivable or predictable from the constituents or the pattern of formation. Idiomatization and demotivation thus refer to corresponding semantic and formal attributes respectively') (1982:164–165). Here and elsewhere lexicalization is generally associated not only with institutionalization and idiomatization but also specifically with semantic DEMOTIVATION (Lipka 1992:107–109; Wischer 2000:358). Bauer identifies a subtype of lexicalization which he calls "semantic lexicalization" (1983:55–59), instancing compounds such as *blackmail*, *mincemeat*, *townhouse*, and *butterfly* or derivatives such as *unquiet*, *gospel*, and *inspector* which lack semantic compositionality (because semantic information has been either added or subtracted). Anttila (1989 [1972]:151) adduces examples such as *sweetmeat*, *nutmeat*, *Holy Ghost* 'spirit,' *widow's weeds* 'clothes,' and *fishwife*, which are morphologically transparent but semantically opaque, as instances of lexicalization. Viewing "demotivation" as integral to lexicalization, Blank (2001:1603) includes traditional examples of semantic change under

lexicalization, which he defines very broadly as "a process by which new linguistic entities, be it simple or complex words or just new senses, become conventionalized on the level of the lexicon." Thus, generalizations or specializations such as *holy day* 'religious feast' > *holiday* 'day off from work, vacation' or OE *steorfan* 'to die' > *starve* 'to die of hunger' constitute lexicalization in Blank's view (2001:1600, 1604).[29]

A related view of semantic lexicalization is discussed by Norrick (1979) under the rubric of "lexicalization of a pragmatic function." For example, he suggests that *skunk* "lexicalizes" a function as a term of abuse and that *sorry* and *pardon* are lexicalized as formulas for apologizing (cf. *thanks*, *please* that lexicalize other speech act functions). As pointed out in Section 1.3.2, this phenomenon is best thought of as a case of semanticization rather than lexicalization.

2.4 Lexicalization as increase in autonomy

While many of the examples cited in the lexicalization literature involve increase in fusion and loss of autonomy, not all do. Here we review examples that involve movement out of morphology into the lexicon in ways that involve promotion, not demotion, of autonomy and compositional structure.

As we saw in Section 2.3.3, the movement from morphology into phonology is sometimes subsumed under the name "demorphologization" (Joseph and Janda 1988; Hopper 1990) and sometimes "phonogenesis" (Hopper 1994), as in the forms *never*, *nor*, *naught*, etc., where the original OE *ne* affix is reduced into a phonological element. The shift from a morphological element into a clitic or word in a syntagma has also been termed "demorphologization" (Ramat 2001) or "syntacticization" (Klausenberger 2002). If an inflectional affix becomes a clitic this is known as "cliticization" (e.g., perhaps the genitive *-s* in English); if a clitic comes to be an independent word, this is known as "decliticization" (e.g., the emphatic *ep* in Estonian); and if a derivative morpheme becomes a word, this is known as "syntacticization" (e.g., use of *ism* as a noun). Note that the structural source or starting point in each case is an affix.

Cases of CLITICIZATION which are not also demorphologization occur when an independent item comes to form a phonological unit with the constituent that follows or precedes it. Typically it occurs during the grammaticalization of a lexical item; e.g., auxiliary *will*, along with other auxiliaries, came to have shorter, dependent forms such as *'ll, 'ld* in the sixteenth century (McElhinney 1992) as in *She'll do it*, *I'd have done it if*

[29] Blank's definition of lexicalization is sufficiently broad that it also encompasses loanwords (2001:1606). These increase the lexical resources available to speakers, and enrich lexical fields, and therefore have to do with lexicalization as coding (see 6.2.3).

I could, but it may also involve already grammatical items like pronouns. These are examples of increased fusion. However, the example of a clitic arising out of an inflection and therefore gaining autonomy is an instance of cliticization which is also an instance of demorphologization. A frequently cited example is the case of the development of the so-called group genitive clitic in English (e.g., *the mayor of the city's decision*) from the inflection -(*e*)*s*. The history of this development is, however, much in dispute. According to Janda (1980; see also 1981), the group genitive clitic develops via the *his* genitive (e.g., *the mayor his city*), which he considers to be a reanalyzed form of the inflection interpretable as an enclitic (i.e., not directly from the inflection). Janda argues that the spread of the group genitive exactly parallels that of the *his* genitive. Thus, he sees a development from inflection to word and back to clitic. He suggests that the demise of the *his* genitive is due to the increasing non-syllabicity of the form as well as to stylists' attacks on it. C. Allen (1997) contests Janda's view, citing an impressive array of data. She argues that the *his* ("separated") genitive is merely a variant of the inflected ("attached") genitive, sharing the same distribution; the group genitive is a direct development of the inflected genitive in its spread to all noun classes. Group *his* genitives do not appear until after group inflected genitives are well established. She notes that only later are there attempts to treat the form of the *his* genitive as a pronoun (hence, the appearance of *her* and *their* in this structure). C. Allen (2003) further argues that there is no evidence that the inflected genitive is lost in early Middle English nor that the separated genitive serves as a model for the group genitive. In later Middle English inflected and separated genitives share an inflection-like quality while group genitives have a clitic-like quality. She concludes that it may not be possible to make a sharp distinction between inflection and clitic in this case (see also Norde 2001 on a comparable development in Swedish). Rosenbach (2002) adds to C. Allen's analysis that a major factor in the development of -*s* was probably the development of the determiner system. Whatever one's position on this complex change and its relation to the so-called *his* genitive, this example does illustrate the increased autonomy of an erstwhile inflectional form.

According to an early definition, DECLITICIZATION occurs when clitics "emerge, or re-emerge, as independent words" (Jeffers and Zwicky 1980:223). While questioning the validity of their examples, Lehmann likewise uses the term with reference to situations when a form which was a clitic at a former stage acquires an autonomous use (1995 [1982]:18). This movement from morphology to syntax is typically understood as degrammaticalization (see, e.g., Lessau 1994; Klausenburger 2002:31), or, less often, as lexicalization. For example, Campbell (1991) in discussing the example of decliticization in Estonian treated below equivocates in classifying the starting forms as "clitic[s]/suffix[es]" and saying that they have been either

"decliticized" or "lexicalized." Unproblematic examples of decliticization are difficult to adduce. Despite Campbell's claim that instances of decliticization "while not frequent, are reasonably well known" (1991:295) and Newmeyer's belief that the process "seems to have happened repeatedly" (1998:270), it is rather the case that the same few examples are cited again and again. The least problematical cases of decliticization in the literature include the following: the deaffixation of the abessive affix *-taga* in Lapp and its reanalysis as a clitic, as well as its further evolution into a postposition and free adverb in Enontekiö Lapp (Nevis 1986a); the attainment of independent status by the question clitic *-es* and the emphatic clitic *-ep* in Estonian (Nevis 1986b; see also Campbell 1991:290–292); the acquisition of independent, sentence-initial connective status by a number of clause-final "enclitic particles" in recent Japanese (Matsumoto 1988); and the unbounding of the future enclitic *-to*/*-nto* in Ilokano (an Austronesian language) and of the enclitic *-is* 'also' in Hungarian, an independent form meaning 'indeed' that can be reduplicated (Rubino 1994),[30] and the occurrence of the 1P.PL suffix *-mid*/*muid* as an independent form *muide* (replacing an earlier independent pronoun *sinn*) (see Bybee, Perkins, and Pagliuca 1994:13–14). Doyle (2002) suggests the Irish development may have been partly motivated by a typological change from synthetic (inflectional) to analytic (periphrastic) word order in Irish, a suggestion also made by Norde 2001 for the rise of the clitic *'s* genitive in English.

Generally, however, examples of decliticization are quite controversial. Jeffers and Zwicky's (1980:223–224) argument that Indo-European relative/interrogative/indefinite forms such as Lat. *quis*/*quid*, Gk. *hos*, Sk. *yas*, and Hit. *ku-is*/ *ku-it* develop from the PIE clitics **-kwe*/*-kwo*, **-ye*/*-yo* is contested by Lehmann (1995 [1982]:18–19). He argues that the clitic status of these forms in PIE is far from certain, that the reconstruction may be wrong, and that the syntax is inconsistent. A second example cited by Jeffers and Zwicky (1980:224) of unaccented finite verbs in PIE is likewise summarily dismissed by Lehmann (1995 [1982]), who points out that such verbs were never clitics. Kroch, Myhill, and Pintzuk's (1982) suggestion that the "demonstrable decliticization" of the subject pronoun in forms such as *hastow* or *wiltow* after about 1550 in English is similarly questionable. Brinton (2004) argues that the change does not involve decliticization since *-tow*/*-tou* are reduced forms, not clitics. Their loss in the sixteenth century involves replacement of the reduced forms by full forms, a replacement motivated phonologically by loss of the sandhi rule assimilating the interdental fricative *-th* to a preceding dental. Strictly speaking, the subject pronoun is not really "replaced" since it always existed side by side with the reduced form; what we really have is loss of a later form, a phenomenon that Haspelmath (2004:34) calls "retraction" (see 3.2.2).

[30] Rubino (1994) was unavailable to us, but see Campbell (2001b:128) and Newmeyer (1998:271–272) for discussions of this article.

Likewise, to Janda's argument that the Sp. first person ending *-mos* has become the clitic *-nos* in regional dialects of Spanish (1995; 2001:287–288, 301), Klausenburger (2002:40) responds that this is a case of replacement of *-mos* by *-nos*, i.e., a supplanting of one form by another made overt by stress shifts, not a true case of syntacticization in the sense of the development of the one form into the other.

Klausenberger regards the development of clitics out of more grammatical material as "(partial) syntacticization" (2002:39). Full syntacticization that is also considered an instance of lexicalization is associated with changes in which an inflectional or a derivative morpheme attains independent status as a word with concrete meaning, e.g., *ology* or *bus* (as noted in 2.1.4, the latter was originally part of the Lat. DAT-PL *-ibus* inflection). The change from an affix to an autonomous word is seen by Anttila (1989 [1972]:151) as a "particularly clear case" of lexicalization. A limited repertoire of forms is cited in the literature (see examples in Anttila 1989 [1972]:151; Ramat 1992:549–550; Ramat 2001:393; Newmeyer 1998:269; Heine 2003b) including the following:

(16) a. *ade* 'fruit juice' (< *lemonade*, *orangeade*)
ism 'doctrine, theory, practice' (< *fascism*, *socialism*, *communism*)
ology 'subject of study' (< *zoology*, *sociology*)
onomy 'area of knowledge' (< *economy*)
ocracy 'form of rule or influence' (< *aristocracy*, *bureaucracy*)
itis 'sickness' (< *appendicitis*, *bronchitis*)
burger (< *hamburger*)
hood (< *neighborhood*)
b. *bi* (< *bi-sexual*)
ex (< *ex-husband*, *ex-wife*)
teen (< *teenager*)

Alternatively, these forms may be seen as cases where the lexical base has been clipped off, what in Huddleston and Pullum are termed "foreclippings," in which the beginning is clipped (16a), or "post-clippings," in which the end is clipped (16b) (2002:1635). They acquire semantic concreteness by metonymy with the clipped segment (see Moreno Cabrera 1998) but are typically treated as names for abstract umbrella categories (fruit juice in the case of *ade*).

Forms such as *ology* or *bus* may be understood as instances of lexicalization not only because they contribute new forms to the lexicon but also because they involve a bound morpheme acquiring status as a lexeme.

2.5 Conclusion

We have seen that the conception of lexicalization as the production of neologisms (new entries in the inventory) encompasses a wide variety of linguistic processes, both synchronic and diachronic. These may create,

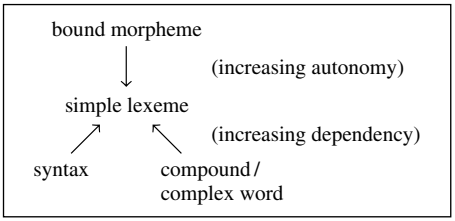


Figure 2.1 Changes typically encompassed in work on lexicalization

modify, combine, or separate existing units, and thus lexicalization would seem to include opposing directions of change leading to greater or lesser dependency and greater or lesser compositionality, as represented in Figure 2.1.

Whether "lexicalization" should be as all encompassing as this, or whether it should be more narrowly defined, remains to be addressed. First, however, it will be necessary to explore the relationship between lexicalization and the phenomenon that it is most often seen to contrast, or correlate, with, namely, grammaticalization. This will be the topic of the following chapter.

3

Views on the relation of lexicalization to grammaticalization

3.0 Introduction

Lexicalization is often discussed in isolation from grammaticalization, especially in studies of word formation. However, it has increasingly been the case that it has entered into grammaticalization studies. One area in which the linking of lexicalization and grammaticalization is especially apparent is in work on fusion, including what has been called freezing, univerbation, or bonding, depending on the type of item that undergoes boundary loss. Fusion of syntagmatically free items into fixed phrases and sometimes further reduction by coalescence is typical both of certain types of lexicalization (cf. *gate-crasher*, *blackbird* [pronounced with syllabic *r*]) and of grammaticalization (cf. Eng. *within*, Fr. *chanterai*). For this reason, the same data are sometimes argued to be instances of lexicalization or grammaticalization or both. A few examples are discussed in Section 3.1. When fusion is conceptualized in the larger context of discussions of unidirectionality – a theoretical construct that has emerged most prominently in work on grammaticalization – parallels between lexicalization and grammaticalization readily emerge (3.2). These parallels concern loss of compositionality, both fusion of originally separable morphemes on the dimension of form and idiomaticization on the dimension of meaning; many examples have been cited in Chapter 2. Unidirectionality encompasses other changes in the grammaticalization literature as well, most especially "less to more grammatical" status of form–meaning pairs. Where changes have been identified that do not conform to unidirectionality, they are often considered to be lexicalizations that are the opposite, reverse, or "mirror images" of grammaticalization (see 3.3), and hence a kind of degrammaticalization (3.3.1). However, increasingly, lexicalization has been seen as different from

degrammaticalization in certain respects (3.3.2). An area which has proved particularly problematic is derivational morphology. We reserve discussion of this issue till Section 3.4.

3.1 Some examples of fusion and coalescence treated as either lexicalization or grammaticalization

Fusion typically involves the loss of original morphological boundaries, and coalescence involves phonological change (assimilation, reduction, loss, etc.). Perhaps the most frequently cited problematic examples of phrases that have become fixed and fused are the following:

(1) *today* < OE *to* + *dæge* 'at day-DAT'
Gm. *heuer* 'this year' < OHG *hiu jaru* 'this year-DAT'

Meillet (1958 [1912]:138–139) cites *heute* 'today' < *hiu tagu* 'this day-DAT' as an example of grammaticalization.[1] This analysis has been rejected by Hopper and Traugott (1993:23) and Joseph (2003:477), although for very different reasons. Hopper and Traugott see the unification of *heute* as demonstrating the emergence of a new lexical item (cf. also Lehmann 1995:1263), while Joseph finds *heute* equally as grammatical before it was reduced as after, and hence neither lexicalization nor grammaticalization. Giacalone Ramat (1998:121–122) cites *heute* as an example of the process of lexicalization (i.e., the emergence of new lexical items via univerbation) occurring at the end of the grammaticalization process. In any case, we may note that *today* vs. *tomorrow* provide good examples of the locality and gradualness of change, since the morphological status of *day* is more transparent than that of *morrow*, which is now archaic or obsolete (Hopper 1994:36).

The fusion exhibited by derivational affixes derived from roots in compounds (see 2.3.2) is also the subject of controversy and inconsistency (see Lindström 2004; Himmelmann 2004). For example, Lehmann (1989:12) cites the change of "ein ehemaliges Lexem zum Derivationsaffix," as in the case of OHG *haidus* 'Gestalt' > MHG *-heit* 'abstract' as an instance of lexicalization. Likewise, Blank (2001:1602) includes the reanalysis of compounds as affixes, e.g., OE *had* 'state (rank, order) of N (= person)' > ME *-hood* 'collectivity of individuals in state of N' (Dalton-Puffer 1996:80) or late Lat. *mente* 'in that manner-ABL' > Fr. *–ment* '-ly (adverbial),' among types of lexicalization as it gives rise to new affixes that may become productive in derivation. In contrast, Ramat finds OHG *haidus* > MHG *-heit* a "fine example of grammaticalization" (1992:558n.). He also sees the development of *-dom* < OE *dom* or of *-ly* < OE *lice* as the "exact converse" of the

[1] However, unlike *hiu jaru*, *hiu tagu* appears not to be attested in OHG.

lexicalization of *-ade* > *ade*, which involves a productive affix becoming a lexical item (1992:551). Similarly, Hopper and Traugott (1993) cite Eng. *-hood*, *-dom*, and *-ly* as examples of reanalysis leading to the creation of productive affixes (40) and describe Fr. *-ment* as a "straightforward instance of grammaticalization" involving the rise of a new grammatical formative out of a formerly autonomous word (130–131). However, neither Lehmann nor Hopper and Traugott are entirely consistent in their treatment of this phenomenon: Lehmann points in passing to the development of *-ly*, *-mente* as involving "grammaticalizations of nouns" (1995 [1982]:87), while Hopper and Traugott propose a "cline of lexicality" showing the development *a basket full* > *cupful* > *hopeful* that would seem to subsume these changes under lexicalization (1993:7; see Cowie 1995 for questions about the status of this cline).

There also exists little unanimity concerning the analysis of the development of some fixed phrases. Since they result in frozen, if complex forms, they would seem to constitute prima facie examples of lexicalization. However, since they also involve items that are "grammatical" (closed class, functional), they may at the same time seem to be examples of grammaticalization. Among the important types of freezing subject to differing interpretations are phrasal and prepositional verbs (e.g., *point out*, *think about*), complex prepositions (*except for*, *depending on*), correlative coordinators (*both ... and*, *either ... or*), complex subordinators (*as soon as*, *in order that*), and "inserts," or discourse markers (*you know*, *I mean*) (see Biber, Johansson, Leech, Conrad, and Finegan 1999:58–59).

Many consider complex prepositions, such as *ahead of*, *in case of*, or *on top of* cases of lexicalization (see Quirk, Greenbaum, Leech, and Svartvik 1985:669–673). For example, Lehmann argues that Modern Castilian (ModCast) complex prepositions such as in (2) undergo reanalysis characteristic of lexicalization in that they are subtracted from the rules of syntax (unified) and their internal structure is obscured (2002:9–11):

(2) ModCast *bajo* 'below' < OldCast *baxo de* 'under from'
ModCast *desde* 'since' < OldCast *des de* < Vulgar Lat. *de ex de* 'from out of'
ModCast *ante* 'before' < OldCast *ante de* '(in) front of'
ModCast *hacia* 'towards' < OldCast *face a* 'fac(ing) toward'

For Lehmann, lexicalization must begin with a complex form since only such a form can undergo "renunciation of its internal structure" (2002:13). Lexicalization necessarily involves unification: "[T]he coalescence of two grammatical morphemes must be called lexicalization" (13). Therefore, in Lehmann's view, when the internal structure comes to be no longer relevant, as in the case of ModCast *desde*, or even only partially relevant, as in the case of Eng. *himself* ('reflexive' < *him self* 'him himself'), lexicalization has occurred, even though the forms that make up the unit are usually

considered members of grammatical categories (preposition, pronoun). Central to Lehmann's (2002) argument is the claim that there are grammatical and lexical members of every word class, so that the primary prepositions such as *de* are "grammatical" and the secondary prepositions such as *desde* are "lexical." Such unified lexicalizations may then undergo grammaticalization and become regularized in the sense that their autonomy is reduced, e.g., prepositions are highly constrained paradigmatically.

While Lehmann (2002) envisions lexicalization preceding grammaticalization in the case of complex prepositions (and other similar phenomena) on a one-way linear "path," Ramat (1992:553–554) suggests that the English prepositions *among* (< OE *on gemang* 'in crowd'), *beside* (< OE *be sidan* 'by side'), and *instead of* (< ME *in stede of* 'in place of') are part of a spiral from the lexicon (via syntax) to the grammar and back to the lexicon. He notes that the change from the analytic source (the phrase) to the synthetic target (complex preposition) can be seen as grammaticalization, with the targets clearly belonging to the grammatical words of English. Being autonomous words "they are, however, part of the English lexicon and cannot be considered on a par with other grammatical means like affixes, ablaut, vowel harmony, etc." (554). Ramat thus sees both grammaticalization and lexicalization at work in these examples, conceding that "the boundary between lexical and grammatical units is not neat" (555). Quirk et al. (1985) likewise equivocate in respect to complex prepositions (as well as phrasal and prepositional verbs, and complex subordinators). Because "word" is, according to them, the unit of lexicology, they classify complex prepositions as cases of "phrasal lexicalization" that lie outside word formation. These structures, despite some amount of grammatical and semantic integrity and even phonological coalescence, retain a degree of separateness or internal autonomy (e.g., *in place of*/*in the place of*, but *instead of*/**in the stead of*). They are "treated as a grammatical phenomenon ... though without ignoring the lexicalization aspect" (Quirk et al. 1985: 1530n.). Furthermore, they establish a "scale of cohesiveness" for complex prepositions ranging from those which behave "in every way like a simple preposition," e.g., *in spite of*, to those which behave "in every way like a set of grammatically separate units," e.g., *on the shelf by* (*the door*) (671).

Yet another view is that the development of complex prepositions provides a "fairly uncontroversial example" of grammaticalization (Traugott 2003:636; see also Tabor and Traugott 1998:244–253) as they involve decategorialization of the nominal, generalization to a larger class of complements, and syntactic reanalysis as functional items, all of which are typical of grammaticalization.

Opinion is equally divided concerning the lexical or grammatical status of what Quirk et al. term "multi-word verbs," a category encompassing phrasal verbs such as *turn up*, prepositional verbs such as *defer to*, and phrasal-prepositional verbs such as *face up to* (1985:1150). In defining a

"multi-word verb" as a "unit which behaves to some extent either lexically or syntactically as a single verb," Quirk et al. clearly equivocate about the status of these forms (1150). They include both the phrasal verb *give up* and the prepositional verb *see to* among the phenomena which share grammatical and lexical characteristics. In support of lexical status, they note that the particle or the preposition may form an "idiomatic combination" (1154, 1159) with the verb, even when separated from it, and that such idiomatic combinations must be distinguished from free combinations in which the parts retain distinct meaning (1152). In a discussion of "phrasal lexicalization," Nevalainen likewise notes the intermediate status of verb plus particle (adverb, preposition) combinations, which "are often treated in grammar rather than lexis because they do not always have the grammatical and semantic integrity of a single word" (1999:423). Huddleston and Pullum (2002:272ff.), though they discuss multi-word verbs in their chapter on clausal complementation, explicitly describe them as "lexicalized" (283–284) and see all multi-word verbs as "verbal idioms" (273).[2] Claridge notes that "a large proportion" of the multi-word verbs are lexicalized (2000:43), whereas Biber et al. are unequivocal that both phrasal and prepositional verbs are lexicalized (1999:58–59). Ramat argues that the forms such as Gm. *sich interessieren an*/*für* 'be interested in' "have to a large extent lost their grammatical or syntactic function and form part, so to say, of the verbal lexical entry" (1992:553). Finally, in the case of the combination of verb plus preposition in Sp. *abusar de* 'to abuse,' *creer en* 'to believe in,' *parecerse a* 'to resemble', Lehmann does not consider them to be either lexicalized or grammaticalized because the constructions "remain discontinuous" and there is no evidence of reanalysis (2002:12).

"Composite predicates" or "complex verbs" of the type *lose sight of*, *take action*, *make use of*, *give a sigh*, *have a cry*, *do the laundry* also pose a problem of analysis. Those of prime interest here involve V (article) N (Prep), in which a meaningless, or so-called "light," verb, typically *give*, *make*, *have*, *take*, or *do*, is paired with a deverbal N, as in *give an answer*, *make a change*, *have a look*, *take a seat*, or *do harm* (*to*) (see Brinton 1996a for a discussion of composite predicates in Modern English). Traugott, for example, concludes that historically the development of composite predicates is "best construed as lexicalization" (1999:259). In contrast, Huddleston and Pullum treat these forms as grammatically productive, including them in their chapter on clausal complements (2002:290–296). Quirk et al. likewise discuss composite predicates in their chapter on "Complementation of verbs and adjectives," seeing them as one type of prepositional verb (1985:1158–1159, 1211–1212). However, they later point out in their chapter on word formation that such phrasal lexical items, though exhibiting some

[2] Chapter 4 of Huddleston and Pullum (2002) on clause complements, in which multi-word verbs are discussed, is primarily by Rodney Huddleston.

degree of separateness, are characterized by fusion and are thus equivalent to single words; for example, they remark that "on the one hand, *lose sight of* can be viewed as a single lexical item; on the other hand, it comprises three grammatical words, one of them (*lose*) with some morphological variation" (1530n.). They conclude that composite predicates and other phrasal lexical items illustrate the "gradience between grammar and lexicon, including a gradience in lexicalization" (1530). Likewise, Algeo sees composite predicates as "somewhere near the middle of the magnetic field of language ... where grammar and lexis meet" (1995:203).

It has also been argued that phrasal discourse markers, or what Quirk et al. call "comment clauses,"[3] such as *I think*, *y'know*, *it seems*, are fixed and partially fused by a process either of lexicalization or of grammaticalization (1985:1112–1118). For example, Schiffrin speaks of *I mean* and *y'know* as "lexicalized clauses" (1987:319). Aijmer sees "conversational routines" such as *thank you* or *I'm sorry* as lexicalized, at the same time recognizing "degrees of lexicalization on a scale of frozenness" resulting from the degree of fixedness, analyzability, institutionalization, and non-literal meaning (1996a:10). Wischer (2000:364) argues that ME *methinks* was both lexicalized and grammaticalized, in contrast to *I think*, which is not lexicalized, only pragmaticalized (see 5.5). Krug (1998) argues that the development of the discourse marker *innit?* from *is it not?* (through an intermediate stage *in't it*) is an example of lexicalization as it involves the form becoming invariant (always *is* and *it*, regardless of the subject, operator/verb, and tense in the main clause), inseparable, and morphologically opaque as well as losing phonological substance and syntactic transparency.[4] However, it is more common to see phrasal discourse markers as resulting from grammaticalization as they typically involve decategorialization and reanalysis as functional elements within the domain of discourse (e.g., Brinton 1996b; also Traugott 1995b).

Our final example of coalescence, in which there is considerable disagreement about whether lexicalization or grammaticalization (or neither) is involved, is phonologization (see 2.3.3). For Bybee the development of monomorphemic words, resulting from phonological change and morphological loss (e.g., causative *lay*) is a case of lexicalization (1985:11–12; see also Traugott 1994:1485). However, Hopper considers phonologization to be an even later stage of grammaticalization than phonogenesis (1994:31).

[3] A variety of names is used to refer to items serving pragmatic functions, and there is little agreement concerning the inventory of items included in this category (see Brinton 1996b). One possible typology of "pragmatic markers" is provided by Fraser (1996).

[4] Krug (1998) also sees the construction as being desemanticized, ie., losing contentful meaning; this view might argue against lexicalization.

3.2 Similarities between lexicalization and grammaticalization

The possibility of describing the same phenomena as either lexicalization or grammaticalization suggests certain similarities between these two processes. Beyond the fact that both lexicalization and grammaticalization constitute an important part of the language user's faculty for language building (see Hagège 1993), what is particularly interesting is that fusion and demotivation are significant factors both in vocabulary building and in grammaticalization.

Given these similarities, Wischer sees lexicalization and grammaticalization as parallel processes but as operating on different levels of the language (2000:364). According to Wischer (2000), while lexicalization involves a syntagm becoming more lexical (*mother-in-law*), grammaticalization involves a syntagm becoming more grammatical (*be going to*, *in front of*). Similarly, while lexicalization involves a lexeme becoming more lexical (OE *hlaf-weard* 'loaf guardian' > *lord*), grammaticalization involves a grammatical item becoming more grammatical (e.g., function word > clitic) (365). Both processes involve "gradual phonetic reduction, syntactic reanalysis, demotivation, fossilization, conventionalization" (364).

However, according to Wischer (2000), the changes in lexicalization and grammaticalization are semantically very different. When a collocation or word is lexicalized, "a specific semantic component is added," but when a term is grammaticalized, "specific semantic components get lost and an implied categorial or operational meaning is foregrounded" (364–365). For example, conversion of *up* to a verb or of *if* to a noun illustrates instantaneous adding of the various types of referential or event-related meaning that are appropriate to N and V categories. Similarly Ramat (2001:397) cites Lazzeroni's (1998) observation that derivational morphemes such as *-ade*, when converted into a noun, build a hyperonym; i.e., a term indicating a syntactic class, in this case of fruit juice, i.e., a content meaning is added. By contrast, in typical cases of grammaticalization, content is not enriched, but is "bleached" (it gradually becomes backgrounded as grammatical meanings are enriched). Nevertheless, the content meaning typically continues to constrain the new grammatical item for some time; e.g., the differences between *will*, *shall*, and *be going to* "can be understood as continuations of their original lexical meanings" (Bybee and Pagliuca 1987:117). This is a characteristic that Hopper (1991) has termed "persistence."

We have seen that the word formation processes known as compounding (see 2.1.1) and blending (see 2.1.5) involve the subtype of fusion known as "univerbation," the unification of two or more autonomous words to form a third; univerbation is also involved in lexicalizations of phrases into lexemes (see 2.3.1) or of complex into simple lexemes (see 2.3.2). Much work on

lexicalization also refers to "demotivation," the loss of identifiable compositional forms, and to idiomatization, the loss of identifiable compositional meaning (see 2.3.1, 2.3.4). While these kinds of phenomena are frequently found in the literature on lexicalization, and at least univerbation is considered by some to be criterial for lexicalization (cf. Kastovsky 1982; Lehmann 2002; Lipka 2002 [1990]), they have received relatively little attention and have roused little controversy. However, in the literature on grammaticalization various types of unification, demotivation, and idiomatization often take center stage within the context of a substantially more all-encompassing hypothesis: that of unidirectionality (see 1.4.2). Some have seen unidirectionality as criterial for grammaticalization (e.g., Lehmann 1995 [1982]; Haspelmath 1999b), a position that has proved to be highly controversial. Given the centrality of the hypothesis of unidirectionality in grammaticalization studies, and its possible relevance to lexicalization studies, it is prudent to investigate more fully what underlies it.

3.2.1 Unidirectionality

Haspelmath (1999b) outlines several types of arguments for unidirectionality in grammaticalization:

(a) iconicity,
(b) the competing motivations of economy and clarity,
(c) learners' tendency to select optimal structures,
(d) speakers' tendency to use novel, "extravagant" expressions.

To these four may be added other proposals:

(a) fulfilling a need to express abstract domains of cognition in terms of concrete domains,
(b) speaker–hearer negotiation of meaning,
(c) frequency effects.

Givón (1975) provided an essentially iconic argument for the irreversibility of unidirectionality in grammaticalization,[5] that is, an argument based on the principle that semantic relations are matched by the formal patterns used to express them.[6] The hypothesis is that contentful, specific (lexical) information will be represented by fuller form than more abstract, skeletal, and bleached (grammatical) material. The latter may become eroded over time, and furthermore, it is "unlikely that a more crucial portion of the information contents of the utterance ... will be entrusted to such a reduced morpheme" (Givón 1975:96). This hypothesis accounts for synchronic matches between form and function but does not account for change, for

[5] Givón refers to grammaticalization as "lexical re-analysis" (1975:49).
[6] For detailed studies of iconicity, see Haiman (1980, 1983).

example, for why phonologically reduced forms are not bulked up (see Haspelmath 1999b).

While the competing motivations of economy and clarity are mentioned already in von der Gabelentz (1901 [1891]), they were more fully articulated by Langacker (1977), who distinguished various types of economy, with focus on signal simplicity (reduction) and perceptual optimality (periphrastic expression). Speakers seek various types of optimality (1977:104–116):

(a) signal simplicity, which leads to boundary loss and morphological bonding,
(b) constructional simplicity, which leads to decrease in markedness,
(c) code simplicity, which may lead to lexical periphrasis (rather than inflection), and
(d) reuse of old items with new meaning.

Haspelmath says there are two problems with this account: why do the two motivations of economy and clarity not cancel one another out and lead to stasis or bidirectionality, and why is it that "only some of the words are reduced dramatically" (1999b:1052)? The first criticism is valid only if speakers and hearers are regarded as equal partners in language change. However, Langacker's model assumes that the speaker is the privileged partner in language change: it is speakers who use language creatively (1977:107). We may also note that it is only in speaking that hearers can reveal innovations they have made (see Traugott and Dasher 2002:279).

In the generative tradition, in which explanation for change focuses on learnability and linguistic competence, unidirectionality is seen to be highly problematic since acquisition involves "a random 'walk' through the space defined by the set of possible parameter values" (Roberts 1993a:252). While the concept of "random walk" leads Lightfoot (1999), for example, to conclude that unidirectionality is a romantic concept with no substance, Roberts (1993a) himself seeks to explain the wide attestation of unidirectionality as the result of reanalysis that leads to structural simplification in the sense of requiring fewer movement operations (see also, in a weaker version, Roberts and Roussou 2003). Haspelmath (1999b) points out that the child has no means to compare less and more complex structures, and that the explanation from parameter setting does not account for reduction and bleaching.

Haspelmath puts forward an alternative proposal, based in a theory of performance and language use, rather than of competence, arguing that unidirectionality, and more specifically the irreversibility of unidirectionality, falls out from speakers' attempts to be "extravagant" (1999b, 2000b). Desiring to talk in ways that will make them noticed and socially successful, speakers use novel (usually periphrastic) ways to express what has become routinized in speech because it is frequently used. This novel expression, once adopted by others, in turn becomes routinized. Haspelmath draws on a variety of work, including Keller (1994) on "invisible hand" effects,

Lehmann (1985, 1993, 1995 [1982]) on continua from lexical to grammatical poles of expression, as well as Bybee (2003) on frequency effects. The "invisible hand" effects are the outcomes of collective actions of individuals following similar maxims or heuristics,[7] i.e., unintended outcomes of intentional acts of speech. Speakers are goal-oriented in the sense that they want to communicate, but not in the sense that they have a particular structure in mind.

"Successful communication" is undoubtedly a goal. However, the textual data suggest that the initial steps in grammaticalization are often indeterminate. The "bridging contexts" for change may be imperceptible at first (see, e.g., the example of *be going to* in 1.4.2), though eventually when the new expression has come to be accepted and is used as the basis for various kinds of analogical extension it may become "noticeable." It may then even be subject to prescriptive attitudes. Grammaticalization can be achieved in part because of relatively unconscious acts of creativity. Therefore, the hypothesis is problematic that grammaticalization is motivated by speakers' tendency to use novel, "extravagant" structures, or even what Hopper and Traugott (1993) and Lehmann (1995 [1982]) have more weakly called "expressive" ones.

Others again have conceptualized goal-orientation in a more intentional way. For example, Heine, Claudi, and Hünnemeyer (1991) have argued that grammaticalization is unidirectional because there is an egocentric "need" (28–29) for expressing certain grammatical functions; grammaticalization is "the result of a problem-solving strategy according to which concepts that are more immediately accessible to human experience are employed for the expression of less accessible, more abstract concepts" (51).[8] Such appeals to "communicative need" have, however, been rejected by Bybee on the grounds that what is expressed grammatically in one language may not be expressed grammatically in another, and that sometimes there are several grammatical expressions of very similar notions, e.g., conditional (see Bybee 1985:202–205; Bybee, Perkins, and Pagliuca 1994:297–300). Bybee, Perkins, and Pagliuca suggest that the reason for unidirectionality is that "the grammatical framework, consisting of both word order regularities and grammatical morphemes, provides a means of facilitating production through automation" (1994:298–299).

[7] The maxims include such practical guidelines as:

(a) "Talk in such a way that you are not misunderstood"
(b) "Talk in such a way that you are understood" (i.e., either "[t]alk in such a way that the other can recognize your intentions," or "[i]ndicate your intentions in such a way that the other can recognize them") (Keller 1994: 94, 98).

[8] See also Traugott (1980) and Givón (1982) on speakers' "need" to specify new relations or strengthen those that have eroded.

For Haspelmath, as well as Heine, Claudi, and Hünnemeyer, and Bybee, Perkins, and Pagliuca, unidirectionality is a strong, almost inviolable tendency in grammaticalization, motivated by language use. Counterexamples are sporadic and do not pattern in significant ways. On the one hand, unidirectionality is motivated at least in part by speaker production, which accounts for the repeated renewal of already extant expressions and for the tendency to imply as much as possible in context. On the other hand, it is motivated by hearer perception and hearer's tendency to infer as much as possible from context. What has become clear in recent years is that these motivations are not at all as unique to grammaticalization as was at first assumed in the grammaticalization literature. Vocabulary is renewed, content items undergo fusion and coalescence, and semantic specificity becomes more opaque (less motivated) over time. Lexicalization even allows for some instances of "bleaching" in the sense of loss of contentful meaning; e.g., loss of the meaning 'rank, state, condition' in the collective noun forming suffix *-hood* (< OE *had*) or bleaching of the meaning 'honey' in the first half of the compound *mildew* (< OE *mele*). Furthermore, like grammaticalization, lexicalization is governed by egocentric tendencies. This is particularly obvious in the case of the recruitment of spatial terms such as *put (it) to (someone)* to refer to acts of speaking and eventually to performative function (Traugott and Dasher 2002:Chap. 5).

In other work that addresses issues regarding similarities and differences between lexicalization and grammaticalization with respect to univerbation, Himmelmann focuses on the fact that both share "a common point of origin, i.e., the spontaneous and productive combination of lexical items in discourse" (2004:36). He conceptualizes the combination of two lexical elements A and B in a syntactic context X and associated with a semantic-pragmatic reading K (formalized as (X) A B | K). "Context" is understood to be of three types: host-class, syntactic, and semantic-pragmatic. Examples of expansion follow (based on Himmelmann 2004:33):

(a) "Host-class expansion" involves elements that collocate in terms of parts of speech, e.g., Adj – N, V – Adv. A grammaticalizing form will increase its range of collocations with members of the relevant category. For example, the range of *be going to* expanded from activity to stative verb subclasses, and the range of the passive marker *be* + past participle was expanded from non-progressive (*the house was built*) to progressive contexts (*the house was being built*) in the eighteenth century.[9]

[9] Prior to that time the passive non-completive construction was of the type *the house was building*.

Himmelmann formalizes host-class expansion as $A_n > A_{n+x}$ (e.g., stative Vs > stative + activity Vs, common Ns > common + proper Ns).[10]

(b) "Syntactic expansion" pertains to larger contexts such as argument positions (subject, object), or clause final/medial/initial position. For example, emerging articles develop first in core subject and object argument positions, only later, if ever, in adpositional ones. As grammaticalization proceeds, syntactic contexts may change; e.g., as deverbal prepositions like *during*, *notwithstanding* developed, they came to precede nouns rather than follow them (Kortmann and König 1992:675). Syntactic context expansion is formalized as $X_n > X_{n+x}$.

(c) "Semantic-pragmatic expansion" pertains to semantic polysemies and increase in pragmatic contexts. For example, when articles arise out of demonstratives they may take on uniqueness functions dependent on the larger situation, including encyclopedic knowledge (e.g., *the president*), and may be used in associative anaphoric contexts (cf. *a house – the front door*, where the definite article is licensed by association with *house*), both of which contexts are unavailable for demonstratives. Semantic-pragmatic context expansion is formalized as $K_n > K_{n+x}$.

Combined, the formula in (3) is said to present "a definition of grammaticization as context expansion" (b represents a grammaticalized element):

(3) $(X_n)\ A_n\ B \mid K_n > (X_{n+x})\ A_{n+x}\ b \mid K_{n+x}$

(Himmelmann 2004:33)

Himmelmann (2004) proposes that in grammaticalization all three contexts expand. By contrast, in lexicalization the first does not, while the second and third may expand, remain the same, or even narrow. While both grammaticalization and lexicalization undergo univerbation according to Himmelmann, only grammaticalization involves unidirectionality. This is because all contexts expand in grammaticalization,[11] but not in lexicalization. Specifically, lexicalization often involves reduction of host-class contexts. For example, when Gm. *grosser Wurf* 'big throw' became a fixed phrase, idiomaticized and univerbated in the meaning 'great opportunity,' *gross* became restricted to the host *Wurf* (other uses of *gross* continued as before). The syntactic contexts in which this phrase was used remained roughly the same. However, since the lexicalized phrase is metaphorical, the semantic-pragmatic contexts were extended. In other instances

[10] Himmelmann (2004) uses the arrow →. We have replaced this by > for consistency with our text.

[11] Himmelmann is focusing on processes involved in the earlier development of a newly grammaticalized form. When the item comes to be in competition with yet newer forms it may, of course, become recessive and its contexts may decrease till phonogenesis or even total loss occurs.

semantic-pragmatic contexts may be narrowed, e.g., Gm. *Hochzeit* 'wedding' < OHG *dio hoha gezit* 'the high time,' which could be used for all sorts of festivities. Himmelmann (2004) formalizes such narrowing as $K_n > K_{n-x}$ (36).[12]

Lehmann (2002) and Himmelmann (2004) restrict the term lexicalization to forms that undergo univerbation. In this respect both consider lexicalization to be only partially unidirectional: it is morphologically unidirectional, but by no means as systemically subject to unidirectionality as grammaticalization. By contrast, researchers like Ramat (1992), who do not identify lexicalization with univerbation, see it as undergoing changes very unlike grammaticalization, including clipping (*ism*) leading to greater autonomy, and semantic changes such as abstract > concrete, which we can see in V > N shifts that convert an action into an entity (*a kick*), an action into a person (*a spy*), an action into its result (*a crease*), an action into an instrument (*a lock*), an action into a place (*a sink*); here the assumption is that persons and things are more concrete than actions (cf. Heine, Claudi, and Hünnemeyer 1991:50).

3.2.2 "Renewal" and "revival" in the context of unidirectionality

The repeated adoption into the lexicon of constructions (e.g., *blackboard*), phrases (e.g., *forget-me-not*), and even whole sentences (e.g., *Who done it?*) has been regarded as non-controversial. Replacements in the content domain may involve layered renewals and replacements of vocabulary types as language typologies change, as in the case of the replacement of inseparable prefixes by a postpositional prepositional verb pattern (e.g., of OE *wiþcweþan* by *speak against*) (Kastovsky 1992:375–376).

A related, but not identical, type of adoption into the grammatical pole of the inventory has, however, raised a number of issues concerning unidirectionality. This is the adoption of periphrastic expressions. As conceived by Meillet (1958 [1915–16]), Lehmann (1985), and Haspelmath (1999b) among others, grammaticalization often originates in periphrastic expressions; this conception can be characterized as a "periphrasis first, reduction second"

[12] Lehmann (2002) uses a partially similar formalism to emphasize different aspects of lexicalization and grammaticalization. Specifically, starting with a construction [XY]Z, he says that in lexicalization the whole construction Z is affected, and the internal structure becomes irregular and gets lost; in grammaticalization, on the other hand, a constituent X of the construction Z becomes the focus of change; the internal relations between X and Y become more "strict and constrained" (Lehmann 2002:13). Here the focus is on the nature of the univerbation and on internal structure, whereas for Himmelmann the focus is on context. In this respect, lexicalization and grammaticalization are more similar to each other in Himmelmann's than in Lehmann's construal.

model of grammaticalization (Haspelmath 2000b:790). In this view, periphrases represent the youngest layer in the process of renewal and reinforcement (Hopper 1991:22–24). By contrast, others adopt a "reduction first, periphrasis second" model of grammaticalization in which periphrasis may appear to be a counterexample to unidirectionality. Van der Auwera, for example, claims that reinforcement is a type of degrammaticalization; e.g., *keep V-ing* > *keep on V-ing* shows an increase in both syntax and lexical meaning (2002:24–25).

This "problem" arises if grammaticalization is thought of in terms of cycles from full expression to reduced expression to zero, which is then "rectified" by new full expression (see Givón 1979; Heine and Reh 1984). But in actual fact what we find in the data is that novel expressions compete with existing ones prior to the elimination of older forms (Lehmann 1985). The novel expressions are typically drawn from at least partially different lexical sources. Usually they meet the requirements of word order at the time that the new expression is developed.

Another issue in work on unidirectionality has been how to think about the "revival" or reinvigoration of an extant but rarely used older form rather than renewal by a different, competing form. Haspelmath treats examples like this as "retraction" (2004:33–34) (the newer form recedes; for a detailed example from the development of concessive subordinating conjunctions with *ob-* in Gm., see de Groodt 2003). Many of the examples cited in the literature arguing against unidirectionality in grammaticalization apparently involve such revivals. For example, Newmeyer (1998:272) cites Gelderen's (1997) demonstration that *man* was primarily an indefinite pronoun in OE but after a period of alternation with the lexical noun came to be used primarily as the latter; seemingly becoming less grammatical over time. However, the issue is that the pronoun did not give rise to the noun, and therefore this example does not appear to be a problem for unidirectionality. Another putative example is *dare*. According to Beths (1999) this verb is attested in OE as a main verb with a directional PrepP, and very occasionally with *to* (a phenomenon that Beths considers evidence for main verb status in OE). Although most of its uses from OE on show auxiliary properties (as in *I daren't go*), clear main V-like uses began to appear in ModE (as in *I don't dare (to) go*); according to Beths, the auxiliary verb use has died out in PDE. However, the British National Corpus (see Aston and Burnard 1998) suggests otherwise: in the vast majority of cases *dare* takes the bare, not the *to*-infinitive, and the predominant negative strategy is typically *n't* (*do*-negation is used for present contexts only four times), so it clearly does not behave exclusively like a main verb (Krug 2000:200–201). Krug regards *dare* as the most marginal of the "emerging auxiliaries," *going to*, *got to*, *want to*, *have to*, *need to*, *ought to*, and *dare* (note no *to*) (236). Neither *man* nor *dare* were originally grammatical forms that later come to have lexical status. Rather, the older major class form has always coexisted with the

minor class form. Another example involves the grammaticalized perfect (as in *have written a book*) and the 'conclusive perfect' with stative/resultative meaning (as in *have a book written*). In OE and ME, both orders coexist, and the meanings are distinguishable only by context. The grammaticalization of the perfect with postposition of the object by the end of the sixteenth century allows the order with pre-position of the object to become reserved for stative/resultative meaning and leads, following a period of dormancy, to the increased frequency of this construction in the eighteenth century. Thus, it is not the case that the grammaticalized perfect gives rise to the lexicalized stative/resultative; rather, after a period of coexistence, these forms become syntactically and semantically differentiated (see Brinton and Stein 1995).

3.3 Differences between lexicalization and grammaticalization

In an attempt to maximize the distinction between grammaticalization and lexicalization, some scholars have associated lexicalization almost exclusively with clippings, conversions, acronyms, and other instantaneous changes, many of which were discussed in the previous chapter (e.g., *-ism* > *ism*, *up* [Adv] > *up* [V]). Since these changes may be seen as involving movement from less > more lexical and hence from more > less grammatical, they are putative counterexamples to the hypothesis that grammaticalization involves unidirectional changes from lexical, referential, open class items (and constructions including them) to grammatical, indexical, closed class items (see, e.g., Ramat 1992, 2001).[13]

Most importantly for our purposes, it has been queried whether some instances of lexicalization may indeed be "reversals," "inversions," "converses," or, more vividly, "mirror images" of grammaticalization. This question has been answered in different ways. Some argue that lexicalization (or some subtypes of it, such as clipping and conversion) is the reversal of grammaticalization and hence a type of "degrammaticalization"; the corollary is that "delexicalization" is a type of grammaticalization (e.g., Ramat 1992, 2001; van der Auwera 2002). Others, however, point to considerable differences between degrammaticalization and lexicalization (e.g., Norde 2001; Lehmann 2002). Unsurprisingly, such differences in position reflect competing views on what lexicalization and degrammaticalization are. They also reflect emphases on different aspects of grammaticalization itself. In this respect, it is useful to recall the bipartite kind of definition of grammaticalization that originated with Kuryłowicz: "grammaticalization consists in the increase of the range of a morpheme advancing from a lexical to a

[13] These types of changes therefore also figure crucially in discussion of cautions regarding historical reconstruction (see, e.g., Koch 1996).

grammatical or from a less grammatical to a more grammatical status" (1975 [1965]:52). Givón refers to the second part of the definition ("from a less grammatical to a more grammatical status") as "secondary grammaticalization" (1991:305). Traugott (2002) extends the terminology and uses "primary grammaticalization" to refer to the first part of the definition ("from a lexical to a grammatical ... status"). Heine refers to forms at the stage of primary grammaticalization as G1, those at a more grammatical, i.e., secondary, stage as G2, and lexical forms as L (2003b:165).

One thing that seems to be well agreed on is that the question of whether lexicalization is the reversal of grammaticalization or not can only be answered at the level of abstract schemas, such as Givón's discourse > syntax > morphology > morphophonemics > zero (1979:209; cited in 1.4.2). As Andersen (2001) has pointed out, the level of schemas is the level of the analyst who can have a perspective on decades or even millennia of change (where they are attested) in a way that no speaker can analyze usage in his or her own lifetime. There are virtually no examples of a specific token of grammaticalization being reversed along a path identical to its initial development (Wischer 2000; Norde 2001; Heine 2003b). In other words, one does not expect to find, and one does not find, Fr. (*je*) *chanterai* reversing its historical trajectory, e.g., **chanterai* > *cantare habeo* (with all the in-between steps). As Norde points out, this would be logically impossible, given the kinds of phonological changes and boundary erasures involved (2001:260); it is also semantically and pragmatically impossible to recover bleached or lost meanings (Wischer 2000). Instead, one expects, and finds, that an old change type may be renewed, as in the recruitment of a verb with generalized meaning and an infinitive for the periphrastic future, e.g., *je vai chanter* (in this case the word order is different, reflecting changes in word order from Latin to Romance, but the change type is similar).

One of the very few examples of an apparent true reversal appears to be the development in Pennsylvania German of auxiliary *wotte* into a main verb 'wish, desire' (this, however, does not involve undoing phonological changes relevant to secondary grammaticalization, only initial changes relevant to primary grammaticalization). The form *wotte* originates in the preterite subjunctive of *welle* 'would,' which developed along with many other modal auxiliaries in the language from a main verb 'want' into an auxiliary in ways familiar in grammaticalization at the stage of primary grammaticalization (Burridge 1998, 2002). The development of *wotte* into a main verb is very unusual, however: "all the morphosyntactic changes which it underwent during its transition to auxiliaryhood" have been undone (Burridge 2002:219). This instance of degrammaticalization appears to be motivated by preemption of a form for the ideological purposes of the community. Burridge proposes that it is a kind of euphemism, i.e., hypercorrection (see Janda 2001), adopted by the community of Mennonites as a way of avoiding expressing a wish too bluntly. Such preemption is not

unique to grammaticalization. One can take many linguistic (and non-linguistic) items and use them as group identity markers. Examples in the phonological domain include Labov's (1972) study of fishermen in Martha's Vineyard, and Eckert's (2000) study of Detroit school children; in both cases, specific groups used the raised phonetic variants of /ai/ and /au/ as markers of identity. An example in the syntactic domain is the use of invariant *be* among adolescents who speak Vernacular African American English (Rickford 1999). It would appear, then, that the choice of *wotte* is part of a general social process, and the fact that it reverses grammaticalization is as idiosyncratic as the fact that the identity choices of some speakers in Martha's Vineyard and Detroit reverse a typical phonological shift.

Although exact reversals of grammaticalization would thus appear to be impossible, it has nonetheless been argued that lexicalization can be viewed as a type of degrammaticalization. In the next two sections we present arguments that have been made for this view (3.3.1) as well as arguments that have been made against this view (3.3.2).

3.3.1 Lexicalization as a type of degrammaticalization

Like lexicalization, DEGRAMMATICALIZATION has been widely used in different, often contradictory, ways.[14] These include the polar opposites of "upgrading" of grammatical elements into the lexicon and "downgrading" from more to less grammatical, which may lead to loss of grammatical meaning, a resulting "empty morph," and eventually complete disappearance. Very useful overviews and critiques of different definitions are provided in Heine (2003b) and Ziegeler (2003); see also Norde (2002). We will be concerned only with those views of degrammaticalization that pertain to lexicalization (though not always under the name "degrammaticalization").

Kuryłowicz himself refers to lexicalization as a "reverse process" to grammaticalization (1975 [1965]:52). One example that he gives is of the PIE derivational nominal affix *-a* signaling collective meaning, which, he says, was grammaticalized in Latin as the neuter plural, and then once more became a derivational affix with collective meaning in Italian.[15] Another example he gives is of the restriction of the *be*-perfect in English to a few archaic constructions with specialized, largely adjectival meaning such as *He is gone*, *She is finished* (note the metaphorical adjectival meaning of *He is far-gone*).

[14] Haspelmath (2004:22) points out that Lehmann (1995 [1982]:16–19) invented the term degrammaticalization to name something that in Lehmann's view did not exist. It has served as a highly useful catalyst for research.

[15] Lehmann (1995 [1982]:17), however, rejects Kuryłowicz's example, pointing out that the Italian *-a* forms are frozen and idiomaticized without a modicum of productivity (see 3.4. below).

Other possible examples of the shift from inflectional > derivational cited in the literature are the developments of the Hungarian present participle *–ó* into derivational morphology (Moreno Cabrera 1998) and of the Latin present participle *-nt* into derivational morphology in many Romance languages. According to Luraghi (1998), the Latin present participle, though obligatory in relevant aspectual constructions, was from the beginning not a prototypical inflection since it required adjectival inflection (4a), and like an adjective could be used as a verbal noun (by ellipsis) (4b):

(4) a. centuriones ... fortissime **pugnantes** conciderunt.[16]
centurions ... bravestly fight-PTCP-PL fell
'The centurions were killed while they were fighting (Adj) with the greatest bravery'
(c. 50 BC, Caesar *De Bello Gallico* 6.40 [Luraghi 1998:356])

b. eorum ut quisque primus venerat, sub muro
they-GEN as each first arrived, under wall
consistebat suorumque **pugnantium**
took:stand their:own-GEN fight-PTCP-PL-GEN
numerum augebat.
number-ACC increased
'As soon as each one arrived, he went to the walls and added himself to the number of the fighting (men) (N)'
(c. 50 BC, Caesar *De Bello Gallico* 7.48 [Luraghi 1998:356])

This non-prototypicality was the source for instability, and speakers in different areas took a number of different approaches to this instability. In French, the reflexes of Latin participles are participles (inflections) or gerunds (regular nominal derivations). In Spanish, they are fully lexicalized as idiosyncratic parts of Ns (e.g., *calmante* 'sedative' < *calmar* 'to soothe', *intermitente* 'flashing light' < Lat. *intermittens* 'letting go'; see Moreno Cabrera 1998:218–219). In Italian, some have become lexicalized as invariant adverbial prepositions, which arose in ablative absolute constructions, cf. *durante* 'during.' However, original (but recessive) inflectional uses, as well as derivational and highly idiosyncratic lexical uses, all coexist in Italian (Luraghi 1998:356–360). Despite such layering, Luraghi concludes that the history of these present participles is a case of degrammaticalization because the general direction of change is toward derivation and lexicalization.

In the past decade, the view of (a subset of) lexicalization as a case of degrammaticalization has come to include much beyond the shift from inflectional > derivational. In one of the first papers to undertake a survey

[16] This is cited, mistakenly, as 'considerunt' in Luraghi's text.

of degrammaticalization, Ramat (1992) discusses the development of suffixes into substantives in English. Most of these suffixes are clippings of borrowed derivative morphemes like *-ism*, *-ade*, *-ology*. The lone example of an erstwhile inflection is the much-cited *bus* (from the Lat. DAT-PL *-ibus* in *omnibus* 'for all'). Citing Lehmann's claim that "[a] sign is grammaticalized to the extent that it is devoid of concrete lexical meaning and takes part in obligatory grammatical rules" (Lehmann 1995 [1982]:viii), Ramat says that "*bus*, *ism*, *ade* are grammatical formatives that [have become] devoid of any grammatical function, that is, separated from their grammatical rules ... and [have] acquired concrete lexeme status with their own autonomous lexical meaning. LEXICALIZATION IS THUS AN ASPECT OF DEGRAMMATICALIZATION – or more exactly: degrammaticalization processes may lead to new lexemes" (1992:550; capitals original).

Ramat (1992:551–556, 2001) includes among instances of degrammaticalization the loss of grammatical categories such as the dual, a type of change that is, at the same time, frequently cited as the last stage of grammaticalization (Givón's [1979] final stage zero). There are also instances where loss results in unproductive freezing as part of a stem, or an old morpheme may be reused in a restricted, new way. Ramat's examples of lexicalization/demorphologization/degrammaticalization in this article range from restriction of originally independent morphemes (what Hopper 1994 calls phonogenesis) to reanalysis of morphophonemic alternations as new phonemes (Hopper 1994 calls this phonologization) (see 2.3.3). Examples of the former that Ramat includes are comparative markers that have lost their grammatical status (*elder*, *mayor*), and past and present participles that no longer strictly belong to verbal paradigm (*cloven*, *shorn*, *during*, Gm. *während*, Fr. *pendant*, It. *durante*). Examples of the latter are loss of morphology due to umlaut (*drink*/*drench*) and palatalization (the originally Indo-European*-*sk*- inchoative in *wash*). Of the Latin inchoative *-sk*-, Ramat says it was "downgraded to a particular subsector of the paradigm (only to singular and third plural), and, at the same time, 'promoted' to allomorphic variation of some particular lexical entries", e.g., It. *finisco* 'I finish'" (1992: 552).[17]

A fuller argument is developed in Ramat (2001), where Figure 3.1. is presented.[18] Although the figure does not cite the term "degrammaticalization," the examples are discussed with reference to it. For Ramat (2001), degrammaticalization encompasses demorphologization (in the sense of upgrading) and

[17] This example has been used to argue for a number of differences among "upgrading." Greenberg (1991) calls this "regrammaticalization" (see Chap. 2, n. 27); Lass (1990) and Giacalone Ramat (1998) cite it as "exaptation."

[18] An earlier model given in Ramat presents a spiral movement from lexicon$_1$ > (syntax >) grammar > (syntax >) lexicon$_2$ (1992:555). He says that movement to lexicon$_2$ constitutes lexicalization. The model is not fully explicated, and exactly what developments are envisaged at each stage is not clear. The model

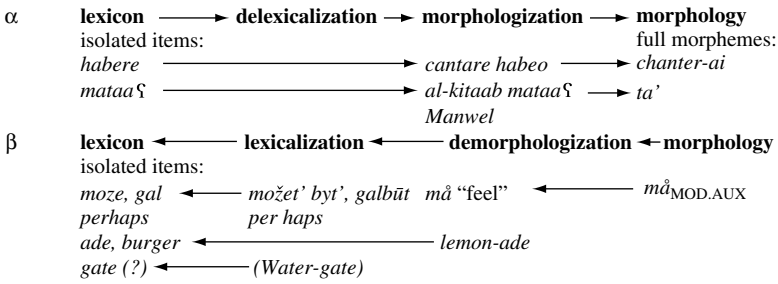


Figure 3.1 Continuum between grammar and lexicon

(Reprinted, by permission, from Paolo Ramat, "Degrammaticalization or transcategorization?" In Chris Schaner-Wolles, John Rennison, and Friedrich Neubarth (eds.), *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler. Presented on the Occasion of his 60th Birthday*, © 2001 by Rosenbach and Sellier, Torino, p. 394.)

lexicalization. Note that the leftmost and rightmost columns refer to components of grammar (lexicon and morphology), whereas the intermediate ones refer to types of change. The examples cited are as follows:

(a) grammaticalization of pre-Maltese *mataaʕ* 'possession,' which developed into *ta'* GEN via constructions like *al-kitab mataaʕ Manwel* 'the-book property-of Manuel' > Mod. Maltese *il-ktieb ta' Manwel* 'the book of Manwel' (Haspelmath (1999b:1057).

(b) fusion, coalescence, and lexicalization of Polish *może* and Lithuanian *gal* 'perhaps, may be' (adverb),[19] via the constructions *možet' byt*' 'it may be' and *galbūt* 'it may be,' respectively; also of *perhaps* from Lat. *per* 'by' + late ME *hap* 'chance' + PL.

(c) lexicalization of Sw. *må* 'feel' from 'may' (auxiliary) (by conversion); and "liberation" of *ade* from *lemonade*, *burger* from *hamburger*, and, by hypothesis, *gate* 'political scandal' from *Watergate* (by clipping) (cf. also *Irangate*, *petrolgate* cited in Ramat 1992:550).

seems to suggest that an example such as *orangeade* begins in lexicon$_1$ and then "moves" directly to grammar (where *-ade* belongs to grammatical well-formedness rules?); *-ade* then moves directly to lexicon$_2$ where it is substantivized (lexicalized) as an independent word *ade*. On this view, *orangeade* is the only source for *-ade*; however, it seems more likely that *orangeade* itself did not give rise to *-ade*; rather, this derivational morpheme, which was borrowed with French nouns, came to be relatively productive (as in *orangeade*, *lemonade*, etc.) and then came to be used independently as a noun, for example in crossword puzzles. Likewise, according to Ramat's (1992) model, *stead* begins in lexicon$_1$, then moves to syntax as *instead of*, then to grammar as an adverb *instead*, then back to lexicon$_2$ (as a lexical item listed in the dictionary?). However, *stead* came to have its new functions only in certain very specific contexts (not in *homestead*, for example).

[19] Ramat and Ricca say *gal* originated as 'it can' (3 SG) (1994:297).

lexical item	grammatical formative	
no grammatical function	lower degree of grammatical function	higher degree of grammatical function
→→→ *grammaticalization / delexicalization* →→→		
←←← *lexicalization / degrammaticalization* ←←←		

Figure 3.2 Lexicalization and degrammaticalization as overlap categories.
(From Johan van der Auwera, "More thoughts on degrammaticalization." In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 2002, p. 21, Table 1. With kind permission of John Benjamins Publishing Company, Amsterdam/Philadelphia. www.benjamins.com)

Note that β allows lexicalization to be conceived both as "adoption into the lexicon" in the sense of adopted into the set of independent major class items (the arrow to lexicon) and as "falling outside the productive rules of grammar" (the arrow from morphology).

For van der Auwera (2002:20), as well as Ramat, degrammaticalization and lexicalization are one and the same, as seen in Figure 3.2. But, as van der Auwera says, degrammaticalization "looks at it from one end" (the starting point), and lexicalization "from the other" (the result) (2002:20). Moreover, the processes differ in their scope, or extent. Lexicalization corresponds with only one part of degrammaticalization, what van der Auwera (2002) calls "wide degrammaticalization." An example like *if* changing from Conj > N (an instance of conversion) is, in his view, both lexicalization ("the making of a lexical item out of something other than a lexical item") and degrammaticalization ("the undoing of a grammatical formative into something other than a grammatical formative") (21).[20] By contrast, "narrow degrammaticalization" is "the undoing of a grammatical formative into a grammatical formative with a weaker degree of grammatical function" (21), as in the case of the English genitive *-s*, which has shifted from an inflection to a clitic at least on some analyses. Furthermore, van der Auwera usefully makes explicit what is often only implicit in discussion of reversals and mirror images: that what is traditionally termed lexicalization (i.e., ordinary processes of word formation such as compounding and derivation as in *songwriter*) does not constitute degrammaticalization at all since a grammatical formative is neither undone nor made less grammatical. Similarly, not all grammaticalization is delexicalization, cf. Gm. *zum* < *zu dem* (i.e., some cases of coalescence involve grammatical items) (van der Auwera 2002:21).

[20] Van der Auwera (1999) argues that the Dutch and German inseparable prefixes are both more grammaticalized and more lexicalized (restricted to a fixed list of lexemes, non-productive) than *if* when used as a noun.

As was mentioned above, Ramat (2001) includes among his examples of degrammaticalization instances of loss of morphological transparency such as occurred in *elder* (loss of comparative adjectival morphology), a phenomenon that Hopper (1994) has called phonogenesis (see 2.3.3). Joseph asserts that phonogenesis, if it is related to grammaticalization at all, must be seen as degrammaticalization since demorphologization is a process distinct from grammaticalization (2003:477). Likewise, van der Auwera cites phonogenesis as an instance of degrammaticalization; however, he does not call it lexicalization (2002:21). Finally, Giacolone Ramat describes phonogenesis, exemplified by the loss of the Lat. comparative affix in It. *signore* or Fr. *seigneur*, as the final stage of both grammaticalization and lexicalization (1998:121).

Clearly the conceptual relationship between lexicalization and degrammaticalization is ill-defined. We turn now to arguments that the two are distinct.

3.3.2 Lexicalization as distinct from degrammaticalization

In this section we discuss arguments that have been made against those sketched in the preceding one. Such arguments pertain to the nature of the lexicon, instantaneous vs. gradual change, and semantic-pragmatic issues.

For Ramat (1992, 2001) and van der Auwera (2002) degrammaticalization appears to be the same thing or almost the same thing as a subtype of lexicalization, and therefore the opposite of grammaticalization: if grammaticalization involves a shift from lexical > less grammatical > more grammatical status, then a shift from more grammatical > less grammatical > lexical status is degrammaticalization. This is at least in part because they think of the lexicon as the repository of stored items that are contentful as opposed to abstract and grammatical.

However, if the lexicon is seen as the repository of stored information, not as the repository of content information, then the equation of all types of degrammaticalization with lexicalization is cast into doubt. For Lehmann (2002), lexicalization is the fusion and coalescence of two or more morphemes; it destroys the regular syntactic construction, eliminates its internal structure, and leads to irregular internal relations. If both these morphemes are grammatical, lexicalization still occurs, because the key to lexicalization is fusion and coalescence ("reduction"). Lexicalization involves not only morphological bonding but also irregularity, opacity, and idiomatization; speakers have access to lexical items "holistically" (see Figure 3.3.). Hence, in Lehmann's view the mirror image of lexicalization is not degrammaticalization but FOLK ETYMOLOGY, the "bestowing [of] structure onto a hitherto opaque expression" (2002:13). In folk etymology, a lexical item (which may be historically complex, but which has become opaque to speakers) is reanalyzed and given a morphological structure that it did not have before and that appears to be at least partially more transparent (see Anttila 1989 [1972]:92–93, 142). For example, Fr. *andier* 'metal support for a fireplace

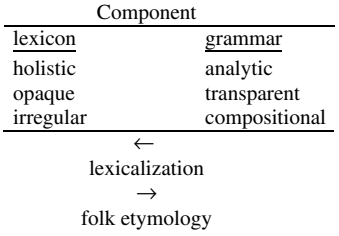


Figure 3.3 Lexicalization
(Based on Christian Lehmann, "New reflections on grammaticalization and lexicalization." In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 2002; p. 14, Table 5. With kind permission of John Benjamins Publishing Company, Amsterdam/Philadelphia. www.benjamins.com)

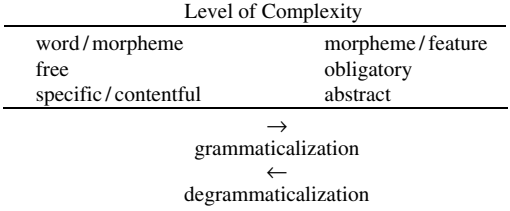


Figure 3.4 (De)grammaticalization
(Based on Christian Lehmann, "New reflections on grammaticalization and lexicalization." In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 2002, p. 15. Table 6. With kind permission of John Benjamins Publishing Company, Amsterdam/Philadelphia. www.benjamins.com)

grid' was borrowed as *andire* and given the folk etymologies *andiron endiron*, and *handiron*; *asparagus*, which was borrowed from Greek, was reanalyzed as *sparrow grass* (Hock and Joseph 1996:18, 174); OE *samblind* 'half blind' was reanalyzed as *sandblind*, OE *weddlac* 'pledge-gift' was reanalyzed as *wedlock*.

By contrast, grammaticalization is in Lehmann's (2002) view the change into a grammatical formative that involves "analytic access to a unit" (i.e., morphological analysis). Degrammaticalization is a process that bestows more specific, concrete meaning on a form with an erstwhile abstract, general meaning, gives "autonomy to a hitherto dependent expression," and requires "an enhanced measure of creativity"; in other words, a form undergoing degrammaticalization, which may have been obligatory, becomes freer, more subject to syntactic manipulation, as well as more specific in meaning (see Figure 3.4.). Degrammaticalization is therefore "magnitudes rarer than grammaticalization" (Lehmann 2002:14–15). Here Lehmann follows Ramat (1992) in defining degrammaticalization as the acquisition of

grammaticalization	degrammaticalization	lexicalization (of grammatical items)
gradual	*gradual*	*abrupt*

Figure 3.5 Grammaticalization, degrammaticalization, and lexicalization
(Based on Muriel Norde, "The final stages of grammaticalization." In Ilse Wischer and Gabriele Diewald (eds.), *New Reflections on Grammaticalization*, 2002, p. 48, Figure 2. With kind permission of John Benjamins Publishing Company, Amsterdam/Philadelphia. www.benjamins.com)

contentfulness (a change-type that he excludes from lexicalization). His examples appear to be conversions like *up the ante* or clippings such as *ade*. They do not, however, involve the kind of degrammaticalization associated with phonogenesis, e.g., the loss of morphological distinctions in *elder*, or those associated with phonologization, e.g., causative *lay*.

In rejecting the mirror image view of lexicalization and grammaticalization, Norde (2001, 2002) focuses not on fusion so much as on gradualness. She notes that both lexicalization and degrammaticalization, as conceived by Ramat (1992), involve counter-directional leftward movement along the cline:

content word > grammatical word > clitic > inflectional affix

but are fundamentally different processes (2002:47). For degrammaticalization to be the converse of grammaticalization, it should occur in stepwise fashion, i.e., gradually. However, there is no evidence of gradualness in Ramat's examples of *bus*, *ade*, *ism*, which he sees as "the exact converse" of typical examples of grammaticalization such as OE noun *dom* > derivational morpheme *-dom* (1992:551). On the contrary, they are abrupt and may involve a straight jump from affix to content word. The same can be said of Lehmann's (2002) examples of degrammaticalization. Norde (2002) suggests the schema in Figure 3.5. Note that not all lexicalization is regarded as abrupt; in the Figure abruptness is regarded as a defining characteristic only of lexicalizations of grammatical forms (Norde's examples are affixes > nouns, e.g., *-ism*).

Likewise Haspelmath (2004) argues that degrammaticalization must be limited to gradual changes. Showing that many of the examples of degrammaticalization cited in the literature are not actual reversals of grammaticalization at all, he coins the term "antigrammaticalization"[21] to describe true reversals of unidirectionality (at the level of schematic types). He defines ANTIGRAMMATICALIZATION as "a change that leads from the endpoint to the starting point of a potential grammaticalization and also shows the same intermediate stages" (2004:27–28). He argues that an antigrammaticalization must occur in the same context as the grammaticalization and must involve intermediate stage(s). All of his examples involve frequently cited changes

[21] The prefix *anti-* is Gk. meaning 'against'; by contrast, the prefix *de-* is Lat. meaning 'off, from.'

from bound to free status (e.g., Eng. and Mainland Scandinavian GEN *-s*, Irish 1 PL subject suffix *–muid* > independent pronoun); none involves a shift to fully concrete lexical status.

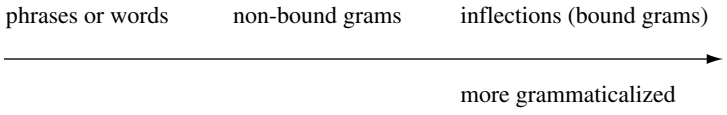
Only after we have reconciled the many different definitions of lexicalization will it be possible to resolve the question of the relationship between lexicalization and (de)grammaticalization. These topics will be taken up in Chapter 4.

3.4 Status of derivation

One area that illustrates some of the complexities in the debates about the relationship of lexicalization and grammaticalization as historical processes is the position of derivation vis-à-vis inflection. Meillet (1958 [1912]) and Kuryłowiz (1975 [1965]) include derivational morphology as a step toward inflectional morphology. This might suggest a possible cline of the type:

phrase > compound > derivation > inflection

However, others have been less sure about hypothesizing a generalizable diachronic cline from derivation to inflection. For example, Bybee sees the "derivation-inflection cline" as a primarily synchronic and typological phenomenon. Derivation is less general, more particular to the root than inflection; hence, derivational forms may split off from their basic form, e.g., something can be *dirty* without involving real *dirt*, and *soiling* a cloth does not require the agent to be *soil* (1985:88). In contrast, inflection changes the base in only minimal ways. Bybee, Perkins, and Pagliuca point out that perfective grams that derive from derivational sources (e.g., spatial terms meaning *up*, *out*, *into*, etc.) are completive; they do not become inflections (these derive from anteriors) (1994:88–90). Furthermore, some aspectual types such as iterative, continuative, and frequentative are more likely to be expressed derivationally than inflectionally (174). For them, as for most researchers on grammaticalization, the diachronic grammaticalization scale has three main positions (based on Bybee, Perkins and Pagliuca 1994:40):



This is a diachronic cline in which form–meaning pairs are tracked. Bybee (1985) and Bybee, Perkins, and Pagliuca (1994) also posit a synchronic "scale of fusion" which has five main "expression types" (adapted from Bybee, Perkins, and Pagliuc 1994:41):[22]

[22] Enger (2002:93), arguing that the history of ON clitic pronoun **sik* 'self' > *-s(t)* involves not only the development of inflectional morphology from a clitic, as is usually supposed, but also derivation, posits a cline with four positions:

syntactic – non-bound grams – inflection – derivation – lexical

This is a scale of fusion in coding (derivation is close to the lexical end because boundaries are often fused) and the lexical end is most fused because for Bybee, Perkins, and Pagliuca it ideally involves monomorphemic forms like *set*, *lay*, where phonologization has led to total loss of the derivative causative (see 2.3.3). It is the intersection of the two scales that is the subject of their book, but no fully clear resolution is offered concerning the extent to which derivation gives rise to inflection, or vice versa. Considerable ambivalence is also found in Hopper and Traugott (1993) as has been discussed in detail in Cowie (1995).

Building on Williams (1981), Haspelmath (1992, 2002, 2004) suggests a radically different way of thinking about the problem. He says that "[i]nflectional patterns do not show stronger internal dependencies than derivational patterns"; shifts from inflectional to derivation status, insofar as they are attested, are changes "internal to the morphology" and therefore not cases of antigrammaticalization (2004:32). For the same reason, shifts from derivational to inflectional status are not cases of grammaticalization.

In general, however, the consensus seems to be that derivation and inflection, which prototypically do have different functions (see 2.1.2), form a continuum not only synchronically but also diachronically. The degree of fusion does not seem to be the determining factor since fusion occurs in both grammaticalization and lexicalization. Rather, the issue is increase in productivity and context-expansion. In general, derivation is less productive than inflection in the sense that it is more idiosyncratic and less likely to be obligatory.

3.5 Conclusion

In this chapter we have seen how disparately the relationship between lexicalization and grammaticalization has been construed. As both lexicalization and grammaticalization may involve fusion and in that sense, at least, be unidirectional (less > more bound), the same examples of fusion are conceptualized variously as lexicalization and grammaticalization by different researchers. On the one hand, those who view the lexicon as consisting of "contentful" forms generally see lexicalization (more grammatical > less grammatical > lexical) as the polar opposite of grammaticalization (lexical > less grammatical > more grammatical) and therefore an instance of degrammaticalization. On the other hand, those who view the lexicon as consisting of learned/stored items – both grammatical and lexical – have, for different reasons, rejected the "mirror-image" view of lexicalization and

grammatical word – clitic – derivational affix – inflectional affix. As Enger says, the generality of such a cline needs to be investigated.

grammaticalization and have uncoupled lexicalization from degrammaticalization, for example, by seeing the former as abrupt and the latter as gradual.

A troublesome area relating to fusion involves the relation of inflection to derivation. In Section 5.4 the example of the problematic relation between derivation and inflection, both synchronically and historically, is discussed in respect to adverbial *-ly*.

4

Toward an integrated approach to lexicalization and grammaticalization

4.0 Introduction

The previous chapters have set out a number of definitions of lexicalization and a number of ways in which lexicalization has been construed in relation to grammaticalization. Many of these definitions and views have been contradictory. This chapter will attempt to reconcile some of the differences and present a coherent view of lexicalization viewed as change in form, internal structure, and meaning.

We have seen that in its broadest sense "lexicalization" as a historical phenomenon has been conceptualized in the literature as "institutionalized adoption into the lexicon," or the creation of neologisms. In this view "lexicon" is understood as an inventory of both "lexical" and "grammatical" units (see 1.2.2). Adoption may be from the lexicon, from morphology, or from syntax (yielding partially new forms); it may even be adoption from another language (borrowing) or invention that yields completely new forms (coinages). Therefore, all changes in the inventory of forms in a language – both those that are prototypically seen as lexicalization and those that are prototypically seen as grammaticalization – fall under the rubric of "adoption into the lexicon", i.e., lexicalization. Forms resulting from routine processes of word formation, including blends such as *telemarketing* (< *telephone* + *marketing*), derivations such as *biodegradable*, compounds such as *cable television*, conversions such as *text* (V) < *text* (N), fixed phrases such as *get it together*, clippings such as *disco* (< *discotheque*), acronyms such as *yuppie* (< *young urban professional*), deliberate coinages such as *xerox*, or back formations such as *baby-sit* (< *baby-sitter*), all represent additions to the inventory of forms in the language. Similarly, forms resulting from

grammaticalization, such as the auxiliary *must* (< OE main verb *mot-* 'be allowed to'), the definite article *the* (< demonstrative *that*), the infinitival marker *to* (< Prep *to*), the complex preposition *instead of* (< Prep + NP + Prep construction), and the development of the dental *-d* past in Gmc., instead of vowel alternation (cf. *ride–rode*), all represent additions to the inventory of forms in the language. This is because grammatical items are form–meaning pairs that are presumably stored in memory in much the same way that lexical items are and must similarly be learned. As these examples show, from the perspective of "adoption into the lexicon," fusion and coalescence, although frequently involved, need not be. Fusion concerns internal tightness of collocation and fixing of sequences (e.g., *pins-and-needles*, not **needles-and-pins*, or *hous-es*, not **-es-house*), and coalescence involves reduction of phonological sequences via regular phonological changes subsequent to fusion (e.g., ME *bot* 'boat' + *swein* 'mate' > *boatswain* 'petty officer on a ship' > *bosun*, *have to* > *hafta*). Reverse processes of separation such as clipping (*ism*, *bus*) may also yield new items in the lexicon. Moreover, new stored items can be the result of instantaneous change, as in the case of conversion (V *window* < N *window*), or of more gradual changes, as in the case of coalescence (*hussy* < OE *hus* 'house' + *wif* 'woman,' *gotta* < *have got to*).

The notion of "adoption into the lexicon" is useful for thinking synchronically about the end result of change. However, the lexicon is not homogeneous at any level of abstraction from individual to community. It contains some forms which are productive, some which are not, some forms which are referential, some which are not (see 1.2). Conceiving of one all-encompassing process by which forms enter the lexicon obscures the differences among types and functions of forms and moreover obscures the processes by which they come to be structured within the inventory. To account for such differences it is necessary to distinguish between lexicalization in a more narrow sense and grammaticalization. And it is useful to refer to the "inventory" rather than the "lexicon." This eliminates confusion between lexicalization in the broadest sense and lexicalization in the narrow sense. Types of "adoption into the inventory" can be differentiated as lexicalization or grammaticalization, depending on the function of the adopted item, and shifts in function within the inventory can be accounted for. After examining some crucial assumptions concerning a model of grammar in terms of which this can be accomplished 4.1, we will present definitions of lexicalization and grammaticalization within the larger concept of "adoption into the inventory." Our focus is on both the processes that lead to this adoption and its outcomes. Alleged "reversals" of lexicalization and grammaticalization are discussed in Section 4.3, and degrees of parallelism between lexicalization and grammaticalization are pointed out in Section 4.4.

4.1 Basic assumptions

The integrated approach to lexicalization and grammaticalization that we propose here makes certain assumptions about the synchronic model of grammar and about the nature of language change.

The model of grammar assumed is not heavily theory-dependent, but is dynamic, allows for constructions, gradience, and degrees of productivity, a model that is in general conceptually compatible with Jackendoff (2002). This means that phonology, syntax, and conceptual structures are linked. Gradience among lexical items is reflected in lexical entries that subcategorize syntactically more or less strictly for constituents and for constructions (see Jackendoff 2002:177). Members of major, open lexical classes (N, V, Adj) are free to combine in ways permitted by the syntax, while members of minor, closed grammatical classes are minimally free and sometimes obligatory. Semantically, members of major classes are typically contentful and referential, members of minor classes are functional, usually indexical, and sometimes non-referential.

A major problem for theories of grammar has been where to locate word formation. Although it is generally understood as encompassing "the basic elements, combinatory principles, and semantic function of new word formations" (Bussmann 1996:s.v. "morphology") and seen as a subbranch of morphology, views on word formation depend crucially on conceptions of both the lexicon and morphology, and the ways in which the latter interfaces with syntax, phonology, semantics, and the lexicon.[1] Further complicating the issue is the fact that, as a result of processes of language change, there is in any language a synchronic continuum between highly productive "principles of free combination" that "build things of word size or smaller" (Jackendoff 2002:158), and far more restricted, semi-productive or even idiosyncratic combinations (see 1.2.4). The free combinations have been called "lexical rules" or "lexical formations" (see, e.g., Matthews 1974:41), designations which suggest that they operate within the lexicon. However, as Jackendoff has pointed out (2000:158), it is highly misleading to use the term "lexical" for such free combinations, since they do not involve storage in long-term memory (what we call the "inventory"). By contrast, restricted combinations must be learned and can therefore be considered to be stored in the inventory. The model of grammar assumed here makes a similar distinction. Combinations of the type known as productive word formation operate outside of the lexicon. As a productive synchronic phenomenon, word formation is seen as preceding, and being independent of, lexicalization; lexicalization may (but does not necessarily) result in semiproductive

[1] For a summary of some of the different positions regarding the form of word formation rules and their grammatical principles, see Motsch (2003).

Table 4.1 *Schema of correlations of categories along continua*

Level	Continuum		
Lexicon	Lexical		Grammatical
Category	Open / Major		Closed / Minor
Syntax	Free		Obligatory
Semantics	Contentful		Functional
Morphology	Nonproductive	Semiproductive	Productive

forms, such as restricted derivational morphemes, while grammaticalization may (but does not necessarily) result in forms that serve as default affixes such as inflections.

Since productivity is so central to our understanding of lexicalization and grammaticalization, it is worth pausing to review our assumptions regarding this issue. We understand productivity as the "[a]bility of word-forming elements to be used to form new linguistic expressions" and is a gradient concept (Bussmann 1996:s.v. "productivity"). In a synchronic perspective, on one end of the continuum are independent lexical items (e.g., *dog*, *paper*) or those made up of morphemes that combine infrequently or with rigid constraints (such as *-t* in *theft*); these are "nonproductive" forms. On the other end of the continuum are items that combine with others essentially by default, typically in a morphosyntactic construction (e.g., obligatory inflections); these are "productive." Between these two ends of the continuum are items that combine less or more regularly to recategorize the lexical items with which they combine (e.g., grammatical derivational forms like *-ness*, *-ly*) and those that combine with relatively few constraints to change the semantics of the lexical items with which they combine (e.g., semantic derivational *un-* in *unhappy* or *-ful* in *hopeful*); these are "semiproductive". The nonproductive to productive continuum correlates in general with a continuum from lexical to grammatical items, from open to closed classes, from major to minor classes, from items that are maximally free (available for creative use in syntax) to maximally obligatory, and from contentful (conceptual, relatively concrete, usually referential [in some possible world, cf. *unicorn*]) to functional (relatively abstract, often indexical).

The set of correlations of categories along continua is summarized schematically in Table 4.1. This schema obscures the fact that there is a significant area of overlap synchronically in the semiproductive area, where we find both content items like *feel for* and grammatical ones like *apart from*.[2] It also obscures the fact that there is a cline of grammatical elements

[2] Talmy (2000, I:22–23) argues that only open classes have lexical complexes such as *spill the beans*, *have it in for*, on the grounds that grammatical ones like *apart from* are made up of combinations of closed class forms. However, the distinction seems arbitrary (see *have it in for*, which is composed of at least three grammatical items),

from phrases to function words and clitics to affixes that does not align so much with the right side of the table as with the middle. We can identify three levels of grammaticality with respect to degrees of fusion with external elements:[3]

G1 = periphrases, e.g., *be going to*, *as far as*, *in fact* (in their early stages),
G2 = semi-bound forms like function words and clitics, e.g., *must*, *of*, *'ll*, genitive *-s* (many function words are cliticized in some positions, but free in others, e.g., *of*, which can be stranded as in *That's all I can think of*),
G3 = affixes such as derivational morphology that changes the grammatical class of the stem, e.g., adverbial *-wise* (fairly productive); most especially inflectional morphology (very productive; sometimes default), including zero inflection (Bybee 1994).

. Being functional, a grammatical item has a morphosyntactic relation with another item (the external host). Therefore, the cline G1–G3 is a cline of grammaticality with respect to degree of fusion with an external host: periphrases are, at least in their early stages, relatively free with respect to position, but, as grammatical items, are characterized by some internal fusion and functional meaning; semi-bound forms by contrast are variably bound to their hosts, and affixes are invariably bound.

The three levels of grammaticality are, of course, highly schematic, and obscure more fine-grained syntactic and semantic gradience as has been identified, for example, for deverbal prepositions by König and Kortmann (1992:684). While forms in a series like *preceding – failing – notwithstanding – except – bar* (as in *bar none*) are all morphologically G2 in PDE (though *notwithstanding* was originally a phrase) and all function as Preps, criteria of collocation and substitution show that there is a gradience from most > least verbal/most prepositional. *Except* and *bar* (along with other deverbal prepositions like *past*, *ago*) are no longer associated with their source verbs and, like prototypical adverbs (Ramat and Ricca 1994) are monomorphemic. The members of the three types of grammaticality arise via rules of morphology, syntax, and phonology.

and a continuum between open and closed class is closer to the empirical facts (see also Langacker 1987; Jackendoff 2002).

[3] Note that G1, G2, and G3 are not isomorphic with Heine's (2003b:163) G1 ~ G2 ~ G_o. The latter are stages of grammaticalization such as are undergone by a form–meaning pair like *do* or *take*, whereas G1, G2, and G3 are conceived as schematic synchronic form-fusion types; any one form–meaning pair representing this form–fusion type may or may not develop into the next type. Furthermore, Heine's G_o is a "grammatical form which has no more grammatical meaning" (e.g., *do* in *Did she go?*), not necessarily an affix.

Table 4.2 *Synchronic clines of lexicality and grammaticality*

Nonproductive				
L3	L2	L1		
	←	Semiproductive	→	
		G1	G2	G3 Productive

Lexical elements likewise do not align entirely with the left side of the table but extend from the middle to the left. Since lexical items are representations of major categories, and relatively free, they form a cline defined primarily with respect to degree of fusion in internal structure. Lexical items range from fully transparent to less transparent to idiosyncratic. We postulate a cline of lexicality, as follows:

L1 = partially fixed phrases, e.g., *lose sight of*, *agree with*,
L2 = complex semi-idiosyncratic forms, e.g., *unhappy*, *desktop*,
L3 = simplexes and maximally unanalyzable idiosyncratic forms, e.g., *desk*, *over-the-hill*.

Like the grammaticality cline G1–G3, this lexicality cline L1–L3 is schematic and obscures fine-grained distinctions: although they are all L2, an item such as *greenhouse* is semantically and morphologically more transparent than *sweetheart* or *strawberry*.

The two clines thus model the heterogeneous texture of the inventory. We may represent the two clines as in Table 4.2.

With respect to language change, a fundamental assumption is that it occurs only in language use, and only in context (Croft 2000) (see 1.1.3). We assume the child as language learner is not only a hearer and passive processor of input but also an active producer of output; both input and output are utterances used in context of situations including speech situations. Likewise, the adult as creator of novel expressions, or interpreter of them, is a producer of or listener to utterances. Linguistic items, whether on the open or the closed end of the continuum, are used in linguistic contexts (utterances), and in social contexts (speaker–hearer interactions). As subtypes of linguistic change, both lexicalization and grammaticalization occur in context and are both therefore subject to the pragmatics of the interactive speech situation. Of course, speakers' ability to use word formation principles can lead to innovation out of syntactic context (as in metalinguistic game-playing or inventing tradenames, etc.), and the ability to cite morphemes out of context can result in metalinguistic citation of grammatical items like *not* or even genitive *-s*. However, language does not change in isolation. Therefore, the changes that are the topic of this book must be conceptualized in the context of utterances (clauses) or writing (sentences), not isolated items.

Despite the assumption that both lexicalization and grammaticalization occur in context, it is the practice of many scholars to cite examples out of context. We have for the most part been no exception so far; however, examples in context will be provided in Chapter 5. Citations out of context tend to obscure the saliency of gradience and indeterminacy in change. In particular, they obscure the difference between what we have been calling "instantaneous, abrupt" changes and "gradual" changes. It is true that for an individual, if a reanalysis occurs, that reanalysis is instantaneous. However, because change occurs in the context of a construction, the continued use of the old as well as the new construction by an individual typically gives rise to intraspeaker variation, and use or non-use of the new construction by different speakers within the community inevitably gives rise to interspeaker variation. This is what leads to indeterminacy. It is characteristic of both grammaticalization and lexicalization.

Finally, we assume that innovation is only a prerequisite to change; change involves institutionalization, however this is construed in a particular community at a particular time (see 2.2). Motivations for acceptance are social. However, caution must always be exercised not to project contemporary social factors onto the past, when populations were smaller, perhaps largely not literate, and social mores as well as available genres may have been different (see Bergs 2005).

4.2 Definitions revisited

In order to provide a coherent view of lexicalization and grammaticalization as the processes used in speaker–hearer interaction and their products, we find it necessary to select aspects of extant definitions and to revise them in certain ways.

4.2.1 Lexicalization

One definition of lexicalization is Kastovsky's conception of lexicalization as "die Eingliederung eines Wortbildungs – oder syntaktischen Syntagmas in das Lexikon mit semantischen und/oder formalen Eigenschaften, die nicht vollständig aus den Konstituenten oder dem Bildungsmuster ableitbar sind" ('the integration of a word formation or syntactic construction into the lexicon with semantic and/or formal properties which are not completely derivable or predictable from the constituents or the pattern of formation') (1982:164–165; quoted in 2.3.4). Another is Lipka's definition of lexicalization as "the phenomenon that a complex lexeme once coined tends to become a single complete lexical unit, a simple lexeme. Through this process it loses the character of a syntagma to a greater or lesser degree" (2002 [1990]:111]; quoted in 2.3.1).

Combining these two definitions we propose the following:

> Lexicalization is the change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the item may become more lexical.

This definition is consistent with a broad conception of lexicalization as "adoption into the inventory," but also concerns modifications within the inventory. Here we pursue the implications of conceptualizing lexicalization in this way:

(1) Lexicalization is understood as a historical change that results in the production of new lexical/contentful forms. It is not simply a process of adoption or incorporation of unchanged elements into the inventory. Hence, for example, simple borrowing without formal or semantic change is excluded from lexicalization. So too are word formation processes which are largely transparent. But treating the literal phrase *Forget me not*! as a name for a plant is a case of adoption into the inventory since it involves a non-literal meaning that must be learned.

(2) The input to lexicalization may be anything stored in the inventory, including word formations such as compounds (e.g., OE *furh* 'furrow' + *lang* 'long' > *furlong* 'unit of measure'), syntactic constructions (e.g., *run-of-the-mill* 'ordinary'), and even grammatical items (e.g., Fr. *tutoyer* 'to refer to addressing someone with the informal second person pronoun'). Items in the input tend to be highly specified semantically.

(3) Lexical items once formed often undergo further change toward the lexical pole of the lexical–grammatical continuum; that is, additional morphological, phonological, and semantic change may modify existing lexical forms. This is a change from "less" to "more" lexical on a scale of lexicality (L1 > L2 > L3) (e.g., from an analyzable compound to an unanalyzable simplex, as in the case of *bosun*). This type of change encompasses cases of phonogenesis in which the original morpheme that becomes an indistinguishable part of the word is lexical (e.g., *neighbor* < OE *neah-ge-bur* 'near + dweller') as well as those in which it is grammatical (e.g., *forlorn* [original past participle *-en*], *near* [comparative *-r*]). Phonologization, such as *drench* or *stench*, where an original causative has been incorporated into the root, is also included (see 2.3.3).

(4) The output of lexicalization is a "lexical," i.e., contentful item that is stored in the inventory and must be learned by speakers. Therefore, when lexicalization involves changes from grammatical to lexical the item comes to be semantically contentful (thus, *up* + COMP > *upper* after conversion to N is lexicalized with the meaning 'part of a shoe above the sole').

(5) The output of lexicalization can be a form of any complexity.[4] Formally, contentful items range from fixed or idiomatic phrases (L1), to compounds and derived forms (L2), to lexical simplexes and idiosyncratic, fossilized forms (L3).

(6) Lexicalization is a gradual change in the sense that it is non-instantaneous, and proceeds by very small and typically overlapping, intermediate, and sometimes indeterminate, steps.[5] Compounds often exhibit gradual phonological change as the roots freeze and coalesce (e.g., *nostril* < OE *nosþyrel* 'nose + hole'), and roots are sometimes downgraded to derivations (e.g., *-dom* 'condition' < OE *dom*[6] 'jurisdiction, setting,' as in *star-dom*); the formal status of forms may also be ambiguous (e.g., *overturn* may be analyzed as a compound or a derivation).

(7) Lexicalization typically involves fusion (erasure of phrasal or morphological boundaries), as exemplified by shifts from syntagm > lexeme (e.g., *out-of-hand*) and complex form > lexeme (e.g., OE *bere* 'barley' + *ærn* 'house' > *barn*). Often, as in the latter example, there is also reduction of phonological sequences (coalescence).

(8) Lexicalization often involves semantic and pragmatic idiomaticization, i.e., the semantic components lose their compositionality (e.g., *black market* refers to neither a market nor anything literally black). The new meanings are often highly idiosyncratic, sometimes more abstract (e.g., *nuts-and-bolts* 'working components [often of an agreement]),' sometimes more specific (e.g., ModE *bailiwick* 'a person's specific area of interest or skill' < ME *bailiffwic* 'jurisdiction of a bailiff,' and see *upper* cited in 4. above).

(9) Lexicalization typically involves decrease in pattern productivity and may involve decrease in token productivity (this is host-class reduction, see Himmelmann 2004 and 3.2.1).

A problematic tendency in the grammaticalization literature has been to classify whatever does not obviously conform to the unidirectional model of grammaticalization as lexicalization (or degrammaticalization, see 3.3.1) when in fact the examples fall variously under the rubrics of word formation, lexicalization, and grammaticalization. The shift of a minor to a major word class (*to up*), though it results in greater "lexicality" in the sense of more contentful meaning and membership of a major class, is a shift better understood as the word formation process of conversion rather than as lexicalization, because it is instantaneous and the meaning derived by

[4] It need not be monomorphemic, contra Lehmann (2002).

[5] This is contra Hopper and Traugott (1993), in which an attempt was made to maximize differentiation between lexicalization and grammaticalization.

[6] See *Domesday Book* ('book of statutes').

conversion is predictable.[7] If significant semantic change in the new category use occurs, a form developed by conversion may, of course, like any other product of word formation, undergo lexicalization (i.e., it will have to be learned separately from its source). For example, the phrase *God be with you* was converted from a blessing into a farewell greeting (a kind of delocutive, see 2.3.1); over time it lost its internal motivation, was reduced phonologically, and became *goodbye* (see Arnovick 1999:97ff.). Likewise, the acquisition of independent lexical status by originally bound forms (e.g., *pro*, *ism*) may be understood as the word formation process of clipping, not lexicalization. Note that word formation can be a very deliberate, self-conscious process, as any new product name (e.g., *uncola*) demonstrates, while lexicalization is not (cf. *barn*, *bailiwick*). In sum, word formation is not lexicalization. Most particularly, products of the often-cited types of word formation, conversion and clipping, do not qualify as instances of lexicalization.

What, then, qualify as cases of lexicalization?

(a) fused syntactic phrases, accompanied by idiomatization (*bread-and-butter* 'necessities of life'), and sometimes undergoing morphophonogical change (*handicap* < *hand in cap*),
(b) fused compounds, such as *mildew* < OE *mele* 'honey' + *deaw* 'dew,'
(c) phonogenesis, such as *handiwork* < OE *handgeweorc*, *mayor* < Lat. *major* 'great' + 'or' COMP,
(d) phonologization, such as *drink*/*drench*, and
(e) creation of semantic, non-category-changing affixes, such as *-hood* < OE *had* 'rank.'

What makes these diverse processes instances of lexicalization is crucially that the output is new or modified forms which are semantically contentful/"lexical," not functional/indexical/"grammatical." Semantically, the items differ from their immediate sources by being more idiomatic and less compositional; morphonologically they are more fused; with respect to productivity, their ability to collocate with host-classes becomes reduced.

As noted in point 1. above, simple borrowing is not lexicalization. However, once borrowed, items may then be subject to either word formation processes or lexicalization within the borrowing language. Examples of lexicalization operating on borrowed words or phrases include cases of phonogenesis (e.g., Sp. *alcoba* 'alcove' < Arabic *al* 'the' + *qubba* 'vault') or of fusion and loss of compositionality (e.g., Eng. *vis-à-vis* 'compared with' < Fr. *vis-à-vis* 'face-to-face'; Eng. *window* < ON *vind* 'wind' + *auga*

[7] However, it should be noted that some conversions (which are typically understood as instantaneous) may actually develop gradually, as the category status of a form in a particular context may be indeterminate (cf. *fun* in *a fun party*, and discussion in 1.2.3).

'eye'; Eng. *je ne sais quoi* 'attribute that is difficult to describe' < Fr. *je* 'I' *ne* 'not' *sais* 'know-PRES 1.PL' *quoi* 'what').

4.2.2 Grammaticalization

One definition of grammaticalization is "[T]he change whereby lexical items and constructions come in certain linguistic contexts to serve grammatical functions, and once grammaticalized, continue to develop new grammatical functions" (Hopper and Traugott 2003:18; see 1.4.2). A somewhat different understanding is encapsulated in Himmelmann's definition of grammaticalization as context-expansion of all three types: host-class, syntactic, and semantic-pragmatic; of these, the first is definitional of grammaticalization as opposed to lexicalization, in his view (2004; see 3.2.1)

In order to highlight the relation of grammaticalization to lexicalization, we propose the following definition:[8]

> Grammaticalization is the change whereby in certain linguistic contexts speakers use parts of a construction with a grammatical function. Over time the resulting grammatical item may become more grammatical by acquiring more grammatical functions and expanding its host-classes.

If we pursue the implications of this definition, we briefly observe the following:

(1) Grammaticalization is conceived of as a historical change that results in the production of new functional forms. It is not simply a process of adoption or incorporation of unchanged elements into the inventory.

(2) The input to grammaticalization may be anything stored in the inventory, ranging from strings (*be going to*), to constructions (*let's* 'hortative' < *let us* 'allow us'; *naught* < OE *na* 'no' + *wiht* 'anything'), to lexical items (*may* < OE *magan* 'have the strength to,' *once* < OE *an* 'one' + *-es* 'adverbial genitive') and grammatical items (infinitive *to* < Prep *to*). However, items in the input must be semantically general.

(3) Grammaticalized items once formed often undergo further change toward the grammatical pole of the lexical–grammatical continuum (e.g., additional ordinary morphophonological processes of fusion, etc.). This is a change from "less to more grammatical," on a scale of grammaticality (G1 > G2 > G3).

(4) The output of grammaticalization is a "grammatical," i.e., functional form. In advanced cases it may become semantically contentless, i.e., "bleached," and even non-referential (e.g., *do* in *Did she leave?*), or phonologically contentless but meaningful (e.g., zero inflection).

[8] Note that this definition does not require all grammatical items to have originated in lexical items, as unidirectionality hypotheses have implied.

(5) The output of grammaticalization can be a form of any complexity. Formally, such items range from grammatical constructions or periphrases (G1), to function words and clitics (G2), to inflections (G3).

(6) Grammaticalization is gradual in the sense that it is non-instantaneous and proceeds by very small and typically overlapping, intermediate, and sometimes indeterminate, steps.

(7) Grammaticalization typically involves fusion with a host, sometimes followed by coalescence/reduction of phonological sequences: e.g., Old Hungarian *vila* 'world' + *bele* 'guts/core + directional' > *vilagbele* 'into the world' > Mod. Hungarian *világba* 'world' + 'directional case marker' (Anttila 1989 [1972]:149); also the fusion of the periphrastic Lat. *(cant)-are habeo* 'INF + have + 1SG' and subsequent phonological reduction as a tense affix as in Fr. *(chant)-erai* 'FUT1SG'). Frequently the fused and reduced forms come to serve as part of a more general pattern of grammatical marking, such as a case or tense paradigm, a phenomenon that Lehmann (1995[1982]:135) calls "paradigmaticization."

(8) Grammaticalization also typically involves loss of concrete and literal meanings (idiomaticization, bleaching) counterbalanced by the strengthening and eventual semanticization of more abstract and general meanings contextually derived in bridging contexts (see 1.4.2) (e.g., the contextually derived implicatures of futurity derivable from certain uses of the literal motion expression *be going to* come to be salient).

(9) Since grammaticalization always involves "host-expansion," it also involves increase in pattern and token productivity.[9]

To reiterate a central point, not all instances of fusion or phonological reduction are instances of grammaticalization. Only those which yield functional, closed class items (e.g., *be going to*, *perhaps*, *instead of*, Gm. *heuer*) may be considered grammaticalization; the more reduced they are, the more these items tend to become members of grammatical systems such as paradigms. Those instances of fusion and reduction that result in an open class item must be seen as lexicalization, even if a grammatical item is lost in the process (e.g., *alderman* 'a member of a municipal legislative body' < OE *eald* 'old' + *-*ira* COMP + *mann* 'man'; It. *lista nozze* [N + N] 'list wedding = wedding list' < *lista di nozze* 'list of wedding' [N + Prep + N]);[10] such instances of fusion tend to be more idiosyncratic.

[9] There is considerable debate about whether or not at least token frequency precedes or follows grammaticalization (for different points of view see, e.g., Bybee 2003, and papers in Lindquist and Mair 2004, especially Mair 2004).

[10] Rossella Tereni, p.c. These forms are not fully compounds since they allow plural on the first noun (*liste nozze*). Because of the plural formation, they are close to the phrasal end of L1 (cf. also *mothers-in-law*).

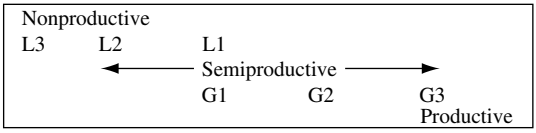
4.2.3 Lexicalization and grammaticalization in language change

The interest in unidirectionality and pragmatic factors in grammaticalization, as well as the rhetorical stances taken by some practitioners who in the 1980s and 1990s emphasized the difference of their approach from that of many generative grammarians, has suggested to some that grammaticalization is conceived as in some way unique and separate from "normal" language change (see especially Newmeyer 1998). However, it is doubtful that anyone working on grammaticalization seriously thought of it as unique. Everything we have discussed suggests that grammaticalization (or indeed lexicalization) are subtypes of language change subject to general constraints on language use and acquisition. Lexicalization involves processes that combine or modify existing forms to serve as members of a major class, while grammaticalization involves decategorialization of forms from major to minor word class and/or from independent to bound element to serve as functional forms. Both changes may involve a decrease in formal or semantic compositionality and an increase in fusion.

The perspective adopted in this work has focused on both lexicalization and grammaticalization as they concern adoption of items into the inventory, and modifications of items within the structure of the inventory, with respect to such factors as function, productivity, fusion, coalescence, compositionality, and degree of contentfulness. This focus is compatible with, but backgrounds, well-known aspects of grammaticalization as a process of morphosyntactic change, such as the development of complementation structures and of morphological paradigms. Lexicalization is likewise compatible with larger perspectives on the lexicon, such as changes in the relation between conceptual structure and expression, but to date these have received very little attention.

In conceptualizing adoption into the inventory, we assume that speakers make use of general rules of grammar, phonology, and word formation. Using these general rules, they routinize strings or idiomaticize extant form–meaning pairs, a process which may lead to the innovation of new units in the individual's inventory. These "units" are not necessarily single morphemes. A unit consists of a syntactic class membership, a phonological string, and a semantic component, each or all of which can change. A unit may be either "lexical" (referential/contentful) or "grammatical" (functional/non-referential) and may correspond to any level of lexicality (L1–L3) or any level of grammaticality (G1–G3). As a language user makes use of this unit, it may change status either toward the lexical/contentful/nonproductive pole ("lexicalization") or toward the grammatical/functional/productive pole ("grammaticalization"). Note that as speakers adapt units, they will not necessarily move them all the way down the cline; that is, not all units will "come to completion" at the stage of L3 or G3. Lexicalization and grammaticalization occur only if the form is accepted by speakers (institutionalized); nonce formation is excluded.

Table 4.3 *Diachronic change along clines of lexicality and grammaticality*



Since speaker–hearer interaction typically drives forms in the direction of greater lexicality, or greater grammaticality, from the perspective of change, the synchronically bi-directional arrows of Table 4.2 are diachronically unidirectional, as in Table 4.3.

4.3 "Reversals" of lexicalization and grammaticalization

Much of the literature on lexicalization has been concerned with articulating lexicalization as some kind of reversal or "mirror image" of grammaticalization (see 3.3). Once a form is lexicalized, speakers may use it in a less lexical way; once a form is grammaticalized, speakers may use it in a less grammatical way; neither strategy is widely attested, however. Given the approach adopted in Section 4.2, we suggest a more restrictive view of such reversals than has been customary in much of the literature.

4.3.1 Delexicalization/antilexicalization

In Section 3.3.1, we saw that grammaticalization, when understood as the shift from lexical to grammatical status, has sometimes been equated with "delexicalization" (see, e.g., van der Auwera 2002).[11] Alternatively, if it is seen in conjunction with morphologization, delexicalization can be understood as part of the shift from lexicon to morphology (see, e.g., Ramat 2001).

We have argued that the shift from lexical to grammatical is a shift in function, not a direct reversal of lexicalization. Therefore, if "delexicalization" is to have any useful meaning as the reversal of lexicalization, it must refer to a change that leads from more to less lexical on the continuum of lexicality, thus from L3 (simplex) > L2 (compound, derivative) > L1 (fixed phrase). It involves an increase in compositionality. Analogizing on Haspelmath's terminology of "antigrammaticalization" (see 3.3.2 above and 4.3.2 below), we will use the term ANTILEXICALIZATION for such a reversal.

[11] Van der Auwera (2002) does, however, observe that grammaticalization can occur without delexicalization; this would typically be secondary grammaticalization from G2 to G3.

While such changes are rarely attested, "folk etymology" provides one example (see Lehmann 2002). As discussed in Section 3.2.2, in folk etymology a form is made more transparent whose morphological and semantic structure has become opaque, either because of losses from the language or borrowings into the language. Folk etymology creates an analyzable, complex form from what has become an unanalyzable simplex; the meaning thus becomes more compositional, and changes in phonetic form lead to increased morphological structure, hence increased morphological compositionality. For example, the first half of OE *angnægl* (< OE *ang* 'painful', cf. Gm. *Angst*), which had become opaque, was altered to resemble the known word *hang* and thus made transparent, with a new form and a new meaning; the original compound became compositional again as *hangnail*. Some instances of back formation may qualify as antilexicalization as they involve the attribution of morphological structure to otherwise unanalyzable forms and subsequent gain of semantic compositionality. For example, the verb *jell* results from analysis of *jelly* (< ME *gelee*) as consisting of *jell* + *-y* (incorrectly assumed to be from OE *-ig* > ModE *-y* as in *sleepy*). Once back formed the verb *jell* acquires the meaning 'to take definite shape'. Similarly, the form *burger* results from giving *hamburger* (originally denoting a ground beef steak, ultimately from Gm. *Hamburger* 'one/thing from Hamburg') the morphological structure *ham* + *burger*. The form *burger* then acquires the meaning of 'sandwich with bun and cooked meat patty' and becomes a productive element in compounding, giving *cheeseburger*, *fishburger*, *mushroomburger*, etc. Finally, the form *alcoholic* has been falsely analyzed as *alc* + *-oholic* and yields the new derivational form *-aholic*/*-oholic* with the meaning 'addicted to or compulsively in need of' found in *workaholic*, *chocoholic*, etc. These appear to be shifts from L3 > L2. Several innovations in contexts of metalinguistic usage that have become widely adopted, such as *history* > *herstory*, *revision* > *re-vision*, count among examples of L3 > L2. Most examples of L2 > L1 would appear to be on the fly innovative jokes (such as intending or interpreting an idiom like *off the wall* literally as well as figuratively in context), but not institutionalized changes.

4.3.2 Degrammaticalization/antigrammaticalization

As we saw in Section 3.3.1, lexicalization has sometimes been equated with a shift from grammatical to lexical status and has been associated, in part or whole, with degrammaticalization. For example, van der Auwera (2002) understands lexicalization as part of "wide degrammaticalization" (though he observes that lexicalization can occur without grammaticalization), and Ramat (2001) sees the shift from morphology to lexicon as encompassing demorphologization and lexicalization (which he equates with degrammaticalization). However, many of the examples cited cannot be considered "reversals" in our schema because they do not involve

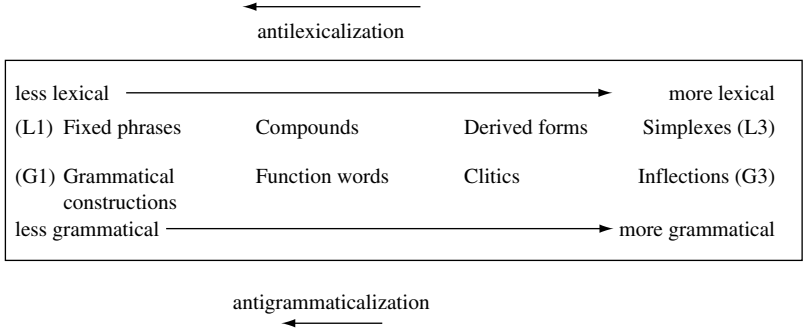


Figure 4.1 Flowchart of antigrammaticalization and antilexicalization

gradual change (e.g., conversions like Pennsylvania Gm. *wotte*, or clippings like *-ism* > *ism*, *ex-* > *ex*, *pro-* > *pro(s)*).

Adopting Haspelmath's (2004) term, ANTIGRAMMATICALIZATION, we argue that the reversal of grammaticalization refers exclusively to changes that lead gradually from more to less grammatical on the continuum of grammaticality, i.e., changes from G3 (inflection) > G2 (clitic, function word) > G1 (periphrasis). This type of change results in increased autonomy of the form and often decreased productivity (from fully to semiproductive). This is "narrow degrammaticalization" in van der Auwera's (2002) sense, or movement from a higher to a lower degree of grammatical function. Examples of antigrammaticalization are relatively rare; they include shifts from inflection to clitic, "cliticization" such as English and Scandinavian genitive *-s* (G3 > G2), and infrequent examples of "decliticization," or shifts from clitic to independent word, such as Estonian *-es* > *es* and perhaps ME *-tow*/*-tou* > *thou* (see 2.4).

Antigrammaticalization does not encompass loss of a grammatical category (as sometimes claimed, see Ramat 1992). Likewise, it does not refer, as Páez Urdaneta (1982) argues, to the loss of function at the propositional or textual level and acquisition of new functions at the conversational level, that is, "performative" or sociostylistic functions; as we will show, this is, in fact grammaticalization (of an adverbial phrase to function as an index of discourse function, see 5.5).

Antigrammaticalization and antilexicalization are presented schematically in Figure 4.1.

4.4 Degrees of parallelism between lexicalization and grammaticalization

To what extent, then, are lexicalization and grammaticalization, which we have portrayed above as complementary processes, similar, and to what

extent are they different? What – apart from the output (lexical or grammatical form and meaning) – distinguishes lexicalization and grammaticalization?

4.4.1 Stronger parallels

As shown in Section 4.2, at a global level of abstraction, lexicalization and grammaticalization have relatively strong parallels. However, from a more fine-grained perspective, these parallels may obscure significant differences in outcome. Both lexicalization and grammaticalization undergo:

(a) **Gradualness** (see 1.4.2).
(b) **Unidirectionality** (see 3.2.1). Both fusion (point c below) and coalescence (point d below) are factors in unidirectionality. However, lexicalization leads to more contentful, grammaticalization to more abstract, functional meaning.
(c) **Fusion**. In lexicalization, fusion involves the freezing and fixing of collocations, e.g., *strike a balance*, *hit the mark*. Grammaticalization involves a similar freezing and fixing of collocations, e.g., *take a* {*walk*, *bath*, *bite*, *fall*, *look*, *nap*}.
(d) **Coalescence**. Coalescence is the reduction of phonological segments subsequent to fusion. In the case of lexicalization, coalescence tends to lead to phonological reductions and to irregular formations such as are the result of phonogenesis (*alderman*) and phonologization (*sit–set*), whereas in the case of grammaticalization it may lead to reduced forms (e.g., *want to* > *wanna*) and to regular patterns such as paradigms. The conclusion that we must reach, given the discussion in preceding sections, is that while the literature often associates coalescence with grammaticalization (e.g., Bybee, Perkins and Pagliuca 1994), it is at least as important in lexicalization as defined here (see also Lehmann 2002; Lipka 2002 [1990]; Himmelmann 2004, among others) but has different effects. Grammaticalization may result in zero forms because it may participate in or trigger systemic paradigms, lexicalization never does.
(e) **Demotivation**. idiomatization, or loss of semantic compositionality. These are also factors in unidirectionality. In the case of lexicalization loss of compositionality tends to lead to increase in semantic specificity, contentfulness, and idiosyncracy, whereas in the case of grammaticalization it leads to more general and abstract grammatical meaning.
(f) **Metaphorization and metonymization**. Both these general kinds of semantic change have been identified with lexicalization and grammaticalization, sometimes in rather different senses. Grammaticalization has been variously interpreted as primarily metaphorical (see especially Claudi and Heine 1986; Heine, Claudi, and Hünnemeyer 1991) or as primarily metonymic (see, e.g., Brinton 1988; Traugott and König 1991). According to Heine, Claudi, and Hünnemeyer

(1991), the metaphors involved are very general "experiential metaphors" (50), arranged along a scale of abstraction which is PERSON > OBJECT > ACTIVITY > SPACE > TIME > QUALITY (48). The arrangement of the metaphorical categories "is unidirectional ... and can be defined in terms of 'metaphorical abstraction' ... concepts that are more immediately accessible to human experience are employed for the expression of less accessible, more abstract concepts" (51); e.g., expressions of future (TIME) are conceptualized in terms of motion (SPACE) as in *be going to*. The relationship between the categories is bridged in the process of change by metonymic extensions that arise in context (73–74). A more specific account of the types of implicatures and "conceptual metonymies" that are crucial to the development of grammatical markers in context has been developed in the Invited Inferencing Theory of Semantic Change (Traugott and Dasher 2002; see also Enfield 2003 on bridging). The assumption is that metaphor is in general the outcome of pragmatic extensions of meaning that become widely accepted and eventually semanticized in the context of speaker–hearer interaction (see many papers in Barcelona 2000 on the metonymic basis of most metaphors).

Seeking to maximally distinguish lexicalization and grammaticalization, and building on Heine, Claudi, and Hünnemeyer's (1991) primarily metaphorical account of grammaticalization, Moreno Cabrera (1998) identifies grammaticalization with metaphorical change, lexicalization with metonymic change. He postulates a "Metonymical Concretion Hierarchy" for lexicalization that is the mirror image of Heine, Claudi, and Hünnemeyer's "Metaphorical Abstraction Hierarchy" for grammaticalization (216). However, this polarization is problematic, even if we set aside the insight that most metaphor has its basis in metonymy. For one, Moreno Cabrera calls only a very limited set of change types lexicalization, namely, "the process creating lexical items out of syntactic units" (214). These are (i) derivation by ellipsis (215–216) (e.g., Sp. *el bigotes* 'the (one wearing a) moustache,' *el corneta* (M) 'the bugler' < *la corneta* (F) 'the bugle' perhaps via *el de la corneta* 'the (one) of the bugle'; see 2.1.4), and (ii) the lexicalization in several European languages of phrases for 'not forgetting' to designate the flower that is in Eng. *forget-me-not* and in Sp. *nomeolvides* 'not-me-forget' (217). The instances of ellipsis are not lexicalizations but word formations in our view; and the one class of *forget-me-not* terms does not warrant the claim that lexicalization is based on metonymy. Moreno Cabrera sees the types of lexicalizations he cites as "highly context-dependent" and then concludes, "therefore concretion and ambiguity, being the two main characteristics of context-dependent elements, will play the starring role in lexicalization processes" (225–226).

As we have shown, both lexicalization and grammaticalization are highly context-dependent. Insofar as lexicalization involves metonymy

(and eventually metaphor), the concretion does not result from context-dependency, but from the fact that lexicalization involves a shift to the contentful, referential pole of language. The terms metonymization and metaphorization may be appropriate for both lexicalization and grammaticalization, but they do not necessarily refer to the same types of metonymy or metaphor. Metonymy in lexicalization is more likely to be driven by social custom and encyclopedic knowledge (see *forget-me-not*). In grammaticalization it is driven by more strictly linguistic meaning (see *be going to*).

4.4.2 Minimal parallels

While lexicalization and grammaticalization correlate to a greater or lesser extent in the ways outlined above, they hardly correlate at all in the following respects:

(g) **Decategorialization**. Hopper defines decategorialization as the process by which forms "lose or neutralize the morphological markers and syntactic privileges characteristic of the full categories Noun and Verb, and ... assume attributes characteristic of secondary categories such as Adjective, Participle, Preposition, etc." (1991:22).[12] Decategorialization is a defining characteristic of grammaticalization since it the mechanism by which lexical items become functional. Although decategorialization may occur in certain kinds of word formation such as conversion (cf. *fun* N > Adj) and compounding, and in processes that are the input to lexicalization, it is not a characteristic of lexicalization in general.

We may note that fusion (point c) and decategorialization (point f) are particular subtypes of the much larger mechanism of change known as "reanalysis." Fusion is the type identified as "cohesion" in Harris and Campbell's (1995:61) characterization of reanalysis. Decategorialization encompasses three other subtypes: change in constituency, change in hierarchical structure, and change in category labels.[13] Some linguists have attempted to identify reanalysis with grammaticalization (e.g., Roberts 1993a). However, although there can be no grammaticalization without some reanalysis, the two are not equatable (Haspelmath 1998; Detges and Waltereit 2002; Hopper and Traugott 2003:58–63).[14] For example, while

[12] Hopper's inclusion of Adjective among the "secondary" categories is, however, problematic, at least for most European languages.

[13] Harris and Campbell cite a fifth type of reanalysis: changes in grammatical relations. This is not discussed here.

[14] Haspelmath (1998) argues that grammaticalization does not require reanalysis, but using a restrictive definition of reanalysis, he discounts category change as reanalysis and also claims that grammaticalization does not involve an

clipping is a case of reanalysis, it is not a case of grammaticalization. Others, while not attempting to equate grammaticalization with reanalysis as a whole or with any specific subtype of reanalysis, nevertheless regard the erasure of phrasal and morphological boundaries leading to fusion and phonetic reduction as prototypical in grammaticalization, at least in its later stages (e.g., *hafta* < *have to*) (see Bybee, Perkins, and Pagliuca 1994). However, fusion in the sense of phonological bonding to the host is not a strictly necessary characteristic of grammaticalization (cf. stranded forms of the Prep *at* as in *Where are you at*?), nor is phonological reduction – there is very little phonological reduction in the stem of Vs that become auxiliaries in English (e.g., *must*, *may*). We must conclude that grammaticalization necessarily involves reanalysis of category status (decategorialization) and changes in constituency; it often involves fusion in the sense specified here of morphological and phonological coalescence, but it does not require it.

(h) **Bleaching**. Grammaticalization is often characterized by bleaching (weakening of meaning through generalization, most especially loss of contentful meaning), at least in late stages. By contrast, lexicalization is most often characterized by concretion (addition of concrete meaning), though some lexical items may lose contentful meaning and acquire pragmatic meaning, as in the use of animal terms as insults, cf. *You dog*. Items that can undergo grammaticalization tend to have quite general meanings (e.g., terms for 'thing,' 'go,' 'come,' 'behind'), while items that lexicalize often have highly specialized meaning (e.g., *black market*). It should be noted, however, that bleaching alone is a misleading concept: while content meaning may be reduced or even lost in grammaticalization, pragmatic and indexical meaning is added (Hopper and Traugott 2003:94–98). As Sweetser says with respect to the development of *go*-futures in various languages: "[W]e lose the sense of physical motion (together with all its likely background inferences). We gain, however, a new meaning of future prediction or intention – together with *its* likely background inferences" (1988:392, italics original).

(i) **Subjectification**. Many instances of grammaticalization have been shown to involve subjectification, the anchoring of meaning in the speaker's assessment of the situation (see 1.4.2). Since grammaticalization involves shifts toward more abstract, less referential, markers, the prime function of which is to represent the speaker's perspective on the situation or to get others to do things, it is necessarily the case that subjectification is characteristic of grammaticalization. Among the best-

ambiguous stage; we, however, have shown that such an ambiguous (or "bridging") stage is typical of grammaticalization.

known examples are the shift from the modality of obligation ("deontic") to that of conclusion ("epistemic"), as in the case of *must* ('it is required that' > 'I conclude that'). Subjectification is, however, not unique to grammaticalization (cf. the development of speech act verb meanings of lexical items like *promise*) (Traugott and Dasher 2002). It is not typical of lexicalization, however, since the latter characteristically involves concrete, referential meaning.

(j) **Productivity**. Items that grammaticalize become more productive in the sense that the grammaticalizing element occurs with increasingly large numbers of categories, i.e., with increasing type frequency. The shift is from a less to more productive pattern (see Lehmann's "paradigmaticization" and Himmelmann's "host-class expansion"). By contrast, increasing lexicality, which obscures boundaries and makes lexical items unanalyzable, reduces pattern productivity (see 1.2.4).

(k) **Frequency**. Items that grammaticalize are used "in more contexts and for a larger set of lexical items"; therefore grammaticalizing items always become more token frequent than their source (Himmelmann 2004:37). However, since lexicalization does not involve host-expansion, nor generalization across context, increased token frequency is not expected.

(l) **Typological generality**. Grammaticalization patterns tend to be cross-linguistically replicated (Bybee, Perkins, and Pagliuca 1994; Heine and Kuteva 2002) and may affect whole semantic classes; e.g., body parts often become case markers, verbs of intentionality become modals, verbs of motion become future markers. Lexicalizations tend to be irregular in that they are more likely to be language- or at best area-specific, and not to have systemic effects.

4.5 Conclusion

We may summarize this section with Table 4.4., showing degrees of parallelism between lexicalization and grammaticalization.

Clearly, lexicalization is far less constrained by various types of linguistic processes than grammaticalization is. The answer to the question why exactly grammaticalization is so constrained and also so systematic cross-linguistically needs to be sought in the function of grammatical items. As mentioned in the preceding section, grammatical items are highly abstract, schematic markers, the prime function of which is to represent the speaker's perspective on the situation or to get others to do things. Therefore, a lexical item must be (or have become) relatively non-specific in content before it can be recruited to serve such an abstract and schematic grammatical function. An ambiguous bridging context, in which both the old and the new meaning can occur, is a prerequisite for grammaticalization. Typically

Table 4.4 *Parallels between lexicalization and grammaticalization*

		Lexicalization	*Grammaticalization*
a	Gradualness	+	+
b	Unidirectionality	+	+
c	Fusion	+	+
d	Coalescence	+	+
e	Demotivation	+	+
f	Metaphorization/metonymization	+	+
g	Decategorialization	−	+
h	Bleaching	−	+
i	Subjectification	−	+
j	Productivity	−	+
k	Frequency	−	+
l	Typological generality	−	+

'+' characteristic of '−' not characteristic of

an item that has the potential to grammaticalize in such a context is relatively unspecific in meaning, and thus can be enriched by the pragmatics of the context. Once the contextual pragmatics in a bridging context become salient in a community, different meanings and structures may be intended by speakers or understood by listeners (reanalysis). If this innovation is spread to minimally different new contexts (analogical host expansion) and adopted by other speakers (institutionalization), change has occurred. Expansion to new hosts leads to increased type frequency/productivity and also increased token frequency. Increasing token frequency is not only a result of but also a contributor to further grammaticalization.

5

Case studies

5.0 Introduction

In this chapter we discuss a number of examples from the history of English which illustrate some particularly gray areas between lexicalization and grammaticalization. We start with the development of present participles (5.1), then move on to multi-word verbs (5.2), composite predicates (5.3), adverbs formed with *-ly* (5.4), and discourse markers (5.5). Insofar as possible we suggest solutions, though several particular questions still remain.

5.1 Present participles

Participles figure in several discussions of lexicalization and grammaticalization because the forms often have a variety of functions. We have mentioned two such instances: the possible development of the Latin inflectional present participle into an Italian derivational morpheme (see Luraghi 1998, cited in 3.3.1), and the cline in English of deverbal prepositions ending in *-ing* (4.1). Here we briefly consider issues related to present participial adjectives (PrP Adjs) in English such as *fascinating*, present participial prepositions such as *during*, conjunctions (Conjs) such as *concerning*, and degree adverbs like *piping* as in *piping hot*.[1]

[1] Similar issues are involved with past participles, cf. the use of *broken* as V in the passive *The door was broken by the vandals* and as Adj in *The door looked broken to me*, while *granted* or *given* can be used as prepositions and conjunctions, as in *Given the disagreement . . . , Given that there is disagreement . . .*

In PDE it is usual (see Huddleston and Pullum 2002:80–81[2]) to contrast verb-forms and present participial Adjs in uses such as:

(1) a. They were **entertaining** the kids with their puppets. (V-form)
b. They were very **entertaining** puppets. (PrP Adj)

By contrast,

(2) Anne was **entertaining**.

is ambiguous (she could have been entertaining guests, or she could have seemed amusing). The V-form is aspectual while the present participial Adj is not. In PDE V-forms can have objects and complements (1a, 3a) while present participial Adjs cannot; complements of present participial Adjs require prepositions (3b).

(3) a. You're **fascinating** us. (V-form)
b. The puppets are **fascinating** to us. (PrP Adj)

Furthermore, unlike V-forms, most present participial Adjs can be introduced by *very* as in (1b), and by *seem* (*the puppets seem fascinating*); they can also undergo derivation (e.g., *unentertaining*, *entertainingly*). Present participial Adjs are said in Huddleston and Pullum to be cases of conversion, but to represent an unusual type in that the form that is converted is inflected, and the process is "much more productive" than is common in conversion (2002:1644). Likewise present participial Preps like *during* are said to have arisen by conversion (610).

We may note that the participial Adjs are for the most part paired with "psychological verbs" (known as "Psych-Verbs"; see Levin 1993:189–190 for a list of these Psych-Vs). Some Psych-Vs, such as *amuse*, *entertain*, *fascinate*, are transitive, with two arguments, a causer of/stimulus for the change in psychological state as subject and an experiencer as object (Levin 1993:191). Other Psych-Vs have non-psychological polysemies (which are often more concrete earlier meanings); these may have a V-form that can serve as a modifier of NPs. However, the present participial Adj form is available only for the Psych-V meaning, as can be seen by comparing polysemous V-*ing* forms of *revolt* 'rebel/disgust.'[3] In the sense 'rebellious' it is a V-form which does not allow *very* (4a) or *seem* (4b), but in the sense 'repulsive' it is a present participial Adj, which does:

(4) a. The **revolting** students picketed the President's office. (V-form or PrP Adj)
b. The students/sandwiches seemed **revolting**. (PrP Adj)

[2] Chapter 3 on "The verb" in Huddleston and Pullum (2002), in which present participial Adjs are discussed, is primarily by Rodney Huddleston.

[3] Some might prefer to regard this pair as synchronically homonymous; the sense 'disgust' is a metonymy from *revolt* to the effect on those against whom the revolt was conducted. Present-day homonymy is independent of the point being made here.

The modifying V-form is construed contingently (revolting/being rebellious around the time of the picketing); the present participial Adj form is construed as representing a cognizing person's general and atemporal opinion. Not every member of the set of Psych-Vs in Levin's (1993) list appears to have a present participial Adj form, especially more recent and colloquial ones like *bug*, *miff*, but also some older ones like *irk*, *rankle*. Furthermore, participial Adjs are for the most part restricted to the subclass of Psych-Vs in which the stimulus is the subject and the experiencer the object, e.g., *amuse*. In another subclass in which the experiencer is the subject, e.g., *admire*, *love*, *support*, many of the semantically positive members have a present participial Adj form, but most of the negative ones do not (e.g., *deplore*, *hate*). This suggests that the present participial Adjs are not conversions of the inflected V as claimed in Huddleston and Pullum (2002:1644), but rather *-ing* derivations of a subset of Psych-Vs.

5.1.1 Present participles in Old English and the rise of the progressive

In OE, the present participial morpheme, *-ende*, was inflectional and agreed in number, case, and gender with the N modified. It could be used to express "limited or perpetual meaning" and sometimes possibly habitual meaning (Mitchell 1985, I:274, 275).[4] Like Lat. *-nd*, it was a hybrid with both verbal properties (a V-stem and traces of aspectual meaning), and adjectival properties (agreement, and modifier status). Most V-stems that occur with *-ende* in OE are intransitive and third person. They typically belong to such semantic classes as verbs of rest (*sittan* 'sit'), movement (*faran* 'go'), state (*libban* 'live'), or change of state (*growan* 'grow'); transitives are largely verbs of communication (*secgan* 'say,' *singan* 'sing') (Nickel 1966; Mitchell 1985, I:276). Most importantly for contrast with the later development of the progressive, the present participial morpheme could readily be used with statives such as *wunian* 'last, dwell' and *habban* 'have.' Some examples are:

(5) a. seo þridde yld wæs ða **wuniende** oð David
the third age was then lasting until David
'the third age lasted up to (the time of) David'
(c. 1000 *ÆlfSig* 270f.1 [Miller 2002:268])

b. . . . if heo hwæt digeles on hyre
. . . if she anything obscure:GEN in her
hæbbende sy
having be-SUBJUNCTIVE

[4] When the participle is coordinated with a simple present or past, however, it is sometimes difficult, according to Mitchell, to ascertain whether *-ende* has a specific meaning (1985, I:274).

'. . . if it (the text) has anything obscure in it'
(c. 1000 *ÆCHom* I.388.29 [Warner 1993:96])

V-*ende* forms appear regularly along with Ns and Adjs as predicate adjuncts of the object of Vs of perception, knowing, judging, naming, etc. (Visser 1963–1973, I:552–569), but from a modern standpoint many seem to be only marginally adjectival because they are typically not gradable. When they are coordinated with or appositional to Adjs, these Adjs too are often only marginally gradable or non-gradable in OE:

(6) syðþan he aldorþegn **unlyfigendne,/** þone deorstan
then he nobleman unliving, the dearest
deadne wisse
dead knew
'then he knew that the nobleman was not living, the dearest (lord) was dead'
(c. 750 *Beo* 1308–1309 [Visser 1963–1973, I:556])

Even though *unlyfigendne* is not (or only marginally) gradable, the derivational *un-* prefix suggests adjectival status. Denison gives the following as criteria for Adj status in OE (1993:373, citing Nickel 1966): modification by Advs such as *hu* 'how,' *swa* 'so,' *to* 'too' (but not *swiðe* 'very' which could also modify Vs), prefixation by *un-*, and inflection by degrees of comparison (7):

(7) þær he **hattra** and **beornendra** wæs
where he hott-er and burning-er was
'where he [it, my arm] was more ardent'
(c. 890 *Bede* 5 3.394.5 [Mitchell 1985, I:649])

Present participial Adjs may serve as the base for manner Advs such as OE *unawendlice* 'unchangeably.'

Much has been written about the replacement of the V-*ende* form by V-*ing* and the development of *be* + *-ing* into the progressive (see, e.g., Mossé 1938; Nickel 1966; Visser 1963–1973; Denison 1993; Warner 1995) and we will not attempt to repeat it here. While the details of the development are still in question, there is general agreement that it is an instance of grammaticalization. Contributing factors included, in OE, the use of adjectival V-*ende* with a copula in predicative and appositive constructions, and of an agentive nominal marked by the derivative *-end* (e.g., *lærend* 'teacher', *hælend* 'savior') in predicate constructions. In ME *-ing* replaced *-ende* partly under the influence of another nominal derivative, *-ung*/*-ing*, which had a "default" meaning of 'act/process of doing X' (see Dalton-Puffer 1996:93). Other contributing factors included the development of the perfect (*have* + *-en*), and possibly the influence of Celtic and Latin. The copula *be* in construction with *-ing* came to be reanalyzed as part of the verbal Tense/Aspect/Mood system; that is, it became an auxiliary in an increasingly periphrastic system. Its

aspectual function was to present an "internal perspective" on a situation, i.e., to perspectivize a situation as unfolding (Langacker 1991:208). This reanalysis can be said to have been completed with the development of the progressive passive in the late eighteenth century (Warner 1995), though inevitably further changes are occurring in PDE, for example, the extension of the progressive to perspectivize a situation as contingent (compare *I'm living at 6 Railway Cuttings* with *I live at 6 Railway Cuttings*, Comrie 1976:37). In other words, over a period of about a thousand years, a new discontinuous aspect marker *be – ing* came into being to mark a new category (progressive aspect).

In the next sections we focus on the development of the present participial Adjs, and of participial prepositions, conjunctions, and degree Advs (also known as "intensifiers," "emphasizers," "boosters," "uptoners," and "downtoners"). We will argue that the development of present participial Adjs is a case of lexicalization. On the other hand, the development of present participial Preps/Conjs and degree Advs is a case of grammaticalization. This grammaticalization of the individual items occurred over a short period of time while a division of labor between the extant categories Adj and Prep was crystallizing.

5.1.2 The development of the present participial adjective

As mentioned in the preceding Section, in OE V-forms in *-ende* were restricted to a small class of Vs, mainly intransitives and transitive Vs of communication. This suggests that although *-ende* was an inflection with respect to agreement, it was a derivational morpheme with respect to distribution. In EME, before the progressive was well developed, the V-*ing* form was used much as the V-*ende* form that it replaced had been in OE: predicatively after verbs of perception and mental action, including stative verbs like *know* and *like*. It was also used attributively in the construction (Det)__N, and appositively as a free adjunct, which is a non-finite construction without overt subject (Mustanoja 1960:552), e.g., *thilke fooles sittynge hire aboute* 'those same fools [who are] sitting about her' (Chaucer, *TC* 4.715 [Mustanoja 1960:555]). Free adjuncts are very productive but are restricted to these three constructions; members do not behave like Adjs in that they do not allow *very*, *seem*, or derivation. They are hybrids: in essence V-forms with marginal Adj-form properties.

In OE and EME V-*ende*/V-*ing* forms were used with a very different class of Vs than are PDE present participial Adjs. However, during the later ME and especially the EModE periods a large number of verbs were borrowed from OFr. or later from Lat., and some of these developed into present participial Adjs. These borrowed Vs included transitive Psych-Vs (*please*) or transitive Vs which later came to have Psych-V polysemies, e.g., *surprise* 'attack suddenly.' In most cases the *OED* exemplifies present participial Adj

uses from fifty to two hundred years later than the original verbal use. *Revolt* is an interesting case in point. The *OED* cites the verb in the meaning 'rise against' in 1548 and in the meaning 'affect with disgust' in 1751. An adjectival *-ing* form meaning 'rebellious' is dated 1593, and the first meaning 'repulsive' in 1806. A clearly present participial Adj form appears in (8), though such a form was probably used earlier since derived *revoltingly* is cited in 1835.

(8) There is, to us, something so **revolting** in the very idea.
(1849 Maitland, *Essays on Subject connected with the Reformation of England* 67 [*OED*])

Among the characteristics of the set of present participials related to Psych-Vs in PDE is that they are usually used intransitively with *to*-NP instead of Object (see 3b). C. Allen (1995:279) notes that in ME *please* was often used with a Prep in translations from Fr. and Lat., and cites (9a). This same V is found in a distinctly present participial Adj form with Prep in Spenser (9b):

(9) a. . . . þat hii **plesen** **to** þe
. . . that they please to thee
(c1350 *MPPsalter* 18.15 [C. Allen 1995:279; citing *MED*])

b. Sweete and **pleasing unto** living sense
(1590–96 Spenser, *Fairie Queene* II.xii.42)

While the influence of Fr. and Lat. is doubtless important, the fact that earlier present participial forms had largely been intransitive probably provided a model for the new forms. Over the course of EModE as the progressive was developing, the original present participial construction for the most part disappeared and came to be restricted primarily to Psych-Vs. This restriction was part of the general loss of erstwhile transitive Adjs (e.g., *(un)worthy*, *(un)becoming*, *next*), which

(a) came to have PrepP as complements (*next*, *unworthy*),
(b) were recategorized as prepositions (*like*) (Denison 2001:132), or
(c) were replaced.[5]

Example (10) illustrates how long earlier constructions without Preps continued to be used:

[5] Of the nine examples of "dative with adjective" cited in Mustanoja (1960:103) in addition to *(un)worthy*, *near*, *next*, *like*, only one is still current, *unlike* (cf. *She is very unlike me*).

(10) and any such feeling on her part was mean, ignoble, and **unbecoming**
['inappropriate to'] the spirit with which she wished to think she was endowed
(1860–1861 Trollope, *Framley* xxxv.343 [Denison 2001:132[6]])

While some of the early present participial Adjs may simply be examples of word formation by derivation, many, certainly among those derived from Psych-Vs, appear to have been lexicalized. That is, they have acquired meanings not directly derivable from either their source Vs or the present participial Adj construction, which signifies that the quality expressed by the stem is perceived as typical of the N. For example, *know* is a member of the subset of Psych-Vs which select experiencer subjects. A *knowing look* is a look construed as conveying not only that the looker typically knows something, but also that what he or she knows is probably secretive. Furthermore, the polysemies of the present participial Adj usually come to be considerably more restricted than those of the V; for example, the most salient meanings of *surprising*, *entertaining*, *revolting*, *stirring* are the Psych-related meanings, not the more concrete ones, such as the meaning of *surprise* as 'attack unexpectedly.'

5.1.3 The development of present participial prepositions and conjunctions

We turn now to the development of present participial Preps and Conjs (also known as "converbs,"[7] see Kortmann and König 1992). Participial V-forms have been used throughout the history of English as non-finite adjuncts with adverbial functions (see Mitchell 1985, II:914–930; Visser 1963–1973, II:1073–1075). This term includes what Kortmann (1991) calls "free adjuncts" without overt subject (11a), and "absolutes" with overt subjects (11b):

(11) a. *Inflating her lungs*, Mary screamed.
b. *The coach being crowded*, Fred had to stand.
(Kortmann 1991:5; italics original)

Both constructions are often considered to be borrowings from Latin and to be characteristic of literate style. The association of absolutes with Latinate

[6] Denison treats (10) as an example of a prepositional use of *unbecoming* (see discussion of *during* below) and compares it to prepositional uses of *worthy* and *next*, which are also transitive. A fuller study of transitive Adjs in general is needed to support this analysis. Despite the lateness of the date, parallels with (10) suggest that transitive *unbecoming* may be a relic of the earlier situation and not a prepositional use.

[7] Nedjalkov says "the definition of converb involves two features, non-finiteness and adverbiality" (1998:421).

language is far from recent. The Prologue to the *Wycliffite Bible* treats absolutes as a problem to be resolved in translation and recommends use of adverbial subordinate clauses for them, e.g.:

(12) *the maister reading, I stonde*, mai be resoluid thus, *while the maister redith, I stonde*
'the master reading, I stand may be resolved in this way: while the master reads, I stand'
(c. 1390 *Wycliffite Bible*, Prologue [Aertsen 1992:673; italics original])

These constructions do not appear to be entirely foreign in origin, but they are strongly associated with formal styles.

Kortmann and König (1992) propose that the participials in free adjuncts and absolutes were reanalyzed in syntactically different ways, but in both cases the outcome was a preposition and the change was an instance of grammaticalization. They compare (13a), which has a free adjunct that co-indexes with the matrix, with (13b), which has a free adjunct that does not. *Savynge* in (13a) is a V-*ing* form, while in (13b) it is a present participial Prep (note the significant difference in meaning):

(13) a. My lady yaf me al hooly/ The noble yifte of hir mercy,/
my lady gave me all holy the noble gift of her mercy
Savynge hir worship by al weyes
preserving her virtue in every way
(c. 1370 Chaucer, *BD* 1269–1271 [Kortmann and König 1992:679])

b. A doghter hadde they bitwixe hem two/ Of twenty yeer . . . /
Savynge a child that was of half yeer age
'They had a daughter between the two of them, who was twenty years old, . . . except a child that was half a year old'
(c. 1386 Chaucer, *CT.Rv.* 3969–3971 [Kortmann and König 1992:679])

Absolutes are more complex. In the case of free adjuncts with no clear controller, i.e., "dangling participles" that do not co-index with an argument of the matrix, "the direct object is simply reanalyzed as internal argument of a preposition" (Kortmann and König 1992:676). This may be manifested by oblique case in the pronoun, which may undergo a word order change; compare (14a) where *failing* occurs in an absolute construction, and (14b) where it is a preposition (*he failing* > *him failing* > *failing him*):

(14) a. in these fear'd ['mixed with fear'][8] hopes, I barely
gratify your love; they **failing**, I must die much your debtor
(1611 Shakespeare, *Cymbeline* II.iv.6
[Kortmann and König 1992:677, citing Visser 1963–1973, II:1219])

[8] Kortmann and König, following Visser, have "sear'd"; the correct form is "fear'd."

b. In default of these, the heritage goes to the son of the ... aunt. Or, **failing** him, it passes ...
(1810 Colebrooke, *2 Treat. Hindu Law Inher.* 225 [Kortmann and König 1992:676, citing *OED*])

The primary semantic domains of present participial Preps are "communicative, textual, or discourse-structuring functions" (Kortmann and König 1992:689), e.g., topic markers (*considering*, *regarding*, *excepting*, *notwithstanding*), and inclusives and exclusives (*including*, *failing* 'in the absence of'). Another semantic domain is temporals, e.g., *during*.[9] A subset of these Preps was used with a clause as complement, i.e., as conjunctions:[10]

(15) a. ye sholde enforce yow to haue pacience,/ **considerynge** that the tribulaciouns of this world but litel while endure
'you should force yourselves to have patience, considering that the tribulations of this world endure only a little while'
(c. 1386 Chaucer, *CT.Mel.* 2696–07 [Visser 1963–1973, II:1218])

b. The place [is] death, **considering** who thou art, If any of my kinsmen find thee.
(1592 Shakespeare, *Romeo and Juliet* II.ii.64 [*OED*])

Reinterpretation of V-*ing* forms in free adjuncts and absolutes as present participial Preps continued through the EModE period (for examples see Visser 1963–1973, II:1218–1222). Among these V-*ing* forms was *being* in the sense of 'since.' This has now become non-standard:

(16) And **being** we are, as I perceive, going some considerable way together, I will give you an account of the whole of the matter
(1678 Bunyan, *Pilgrim's Progress* 283 [Rissanen 1999:321])

Although such reinterpretation appears to have become rare in ModE, it has nevertheless occurred recently with respect to *following* (Olofsson 1990), despite being proscribed in many grammars until the late twentieth century. Prepositional *following* is found as early as 1851.

(17) **Following** his ordination, the Reverend Mr. Henry Edward intends to go to Rome
(1851 *Tablet* [Olofsson 1990:25, citing Addendum to Shorter *OED*])

[9] Mustanoja says *during* came to be used as a preposition in the fourteenth century "presumably as a calque on OF[r.] *durant*: – *I have the power durynge al my lyf* (Ch. CT D WB 158)" (1960:376). While *during* may be formed on analogy with *durant*, there seems to be no reason to assume it is a calque since only the ending is Anglicized.

[10] On the development of Conj and Prep uses of *considering* as well as of an Adv use (e.g., *I think you're pretty safe, considering*), see Kawabata (2003).

Prepositional *following* appears to be favored with Ns "denoting an event (*collision*, *reaction*; *drowning*, *killing*; *experiment*, *treatment*; *victory*, *recovery* . . .)" (Olofsson 1990:33), most of which are deverbal. Unlike *after*, it "does not take nouns that denote a *point* in time" (33, italics original) such as *yesterday*, but only Ns that denote a period of time, such as *month*, *decade*, *Sunday*. It is also unlike *after* in not being used in extended senses such as are found in *seek after*, *look after* (33).

Of those present participial Preps that have survived in PDE, some split from their source forms and no longer form a pair with a V, e.g., *during*, *excepting*, or became restricted to specific contexts, e.g., *excepting*, which now collocates primarily with *not*. Many of them derive from borrowed Vs and occur primarily in formal speech styles (e.g., *excepting*, *concerning*). The grammaticalization of individual items that occurred, including restructuring of word order such as illustrated by (14b), was fairly short-lived. One possible reason is that the new Preps did not match the shorter, less semantically transparent, morphological template of extant "core" Preps that mark case relations (*of*, *to*, *with*, etc.).

Visser lists present participial Preps in a section entitled "conversion or functional shift" (1963–1973, II:1217; see also Huddleston and Pullum 2002:610). Again the question is whether conversion has occurred. The answer seems to be an unequivocal no. If conversion was involved, major areas of parsing ambiguity would not be expected. However, there are many potentially ambiguous examples such as occur in typical cases of reanalysis, e.g.:

(18) a. Who discourst his voyage long, **according** his request
(1590–1596 Spenser, *Fairie Queene* I, 12, 15 [Visser 1963–1973, II:1218])

b. The day **following** his intervention
(Olofsson 1990:23)

Example (18a) could be interpreted as 'She, agreeing to his desire . . . ' (V-*ing* form) or 'In agreement with his desire, she . . . ' (PrP Prep). Olofsson says of example (18b) that it could be adjectival ('the day which followed his intervention') or prepositional ('the day after his intervention') (1990:23). The restriction to certain classes of nominals outlined above further supports the likelihood that grammaticalization, not conversion, occurred.

5.1.4 The development of present participial degree adverbs

Finally, it should be mentioned that from EME times several V-*ing* forms could be used as degree Advs that modify Adjs (cf. *very*, *exceedingly*), e.g.:

(19) a. wafres, **pipyng** hoot out of the gleede
'cakes, piping hot out of the ember'
(c.1386 Chaucer, *CT.Mil.*3379 [Visser 1963–1973, II:1128])

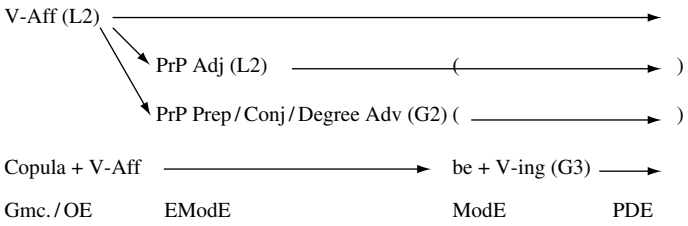


Figure 5.1 Schematic of developments of V-*ende/-ing* forms

b. A preste þat trowid he was **passand** gude synger, not-withstondyng he was not so.
'A priest who thought he was a very good singer, although he was not'
(1450 *Alph. Tales* 85, 27 [Visser 1963–1973, II:1129])

Many present participial degree Advs were figurative like *piping* ('emitting a hissing/piping sound because of the heat'), *passing* ('(sur)passing/exceeding expectation'). Most were replaced by the end of the nineteenth century, while *piping hot* has become a fixed expression.

Like present participial Preps these present participial Advs appear to be reanalyzed V-*ing* forms that underwent grammaticalization.

5.1.5 Summary

In sum, the developments discussed in this section have raised questions regarding the assumption that *-ing* is an inflection in the case of present participial Adjs, and regarding the claim that present participial Adjs and present participial Preps/Conjs are converted.

We postulate an original hybrid derived V-*ende* (later *-ing*) form with verbal and adjectival properties, in other words, a form that can be said to have L2 status. The development of the progressive is a prototypical instance of grammaticalization (toward G3) since it involves (re)constitution of the aspectual system[11] and extensive expansion of type frequency. In the process the copula *be* became an auxiliary, and *-ing* an inflection that occurs obligatorily, but discontinuously, with *be*.

In non-copula contexts, the V-*ing* form split into two further constructions. One was development of the present participial Adj in later ME and EModE; this is word formation affecting Psych-Vs, with some subsequent lexicalization as present participial Adjs came to be idiomatized and, presumably, in at least some cases, felt to be univerbated items independent of their verbal sources (e.g., *knowing*), a shift from L2 toward L3. The other was

[11] IE aspect marking had been lost in Germanic.

the development to present participial Preps/Conjs, or degree Advs, a short-lived case of grammaticalization (to G2). Ignoring all the complex contributions to the development of *be – ing*, the changes can be schematized as in Figure 5.1. (Aff is used as a cover term for *-ende* and *-ing*; parentheses around arrows show the change is no longer productive although the forms still exist):

5.2 Multi-word verbs

As pointed out in 3.1, opinion is divided concerning the lexical or grammatical status of multi-word verbs. What is encompassed by this class and how the different forms are to be subclassified are complex matters. Cowie and Mackin (1993 [1975]:xxix–lvii) distinguish transitive and intransitive combinations of V + Prt, i.e., "phrasal verbs" (see 20a), transitive and intransitive combinations of V + Prep, i.e., "prepositional verbs" (see 20b), and transitive and intransitive combinations of V + Prt + Prep, i.e., "phrasal-prepositional verbs" (see 20c) (for a similar classification, see Quirk, Greenbaum, Leech, and Svartvik 1985:1150–1168; Claridge 2000:26–39 discusses earlier classifications):[12]

(20) a. *nod off* 'fall asleep'
fall down 'collapse'
shoot up 'increase sharply'
come through 'recover'
cut NP *out* 'delete'
grind NP *down* 'oppress'
write NP *up* 'give a full written account'
pick NP *up* 'collect'

b. *frown on* 'disapprove'
look after 'attend'
get at 'reach'
look into 'examine'
put NP *to* 'ask'
take NP *for* 'regard'
hold NP *against* 'allow to count against'
cheat NP *of* 'prevent from enjoying'

c. *send away for* 'post money to obtain'
face up to 'accept honestly or courageously'
look forward to 'anticipate'
put NP *up to* 'suggest'
take NP *out on* 'make the scapegoat'
let NP *in on* 'bring into the know'

[12] For a detailed examination of the syntactic behavior of transitive prepositional verbs, see Akimoto (1989).

A more finely grained typology is provided in Huddleston and Pullum (2002:272–291) because diprepositional complements (e.g., *think of* NP *as* NP, *look to* NP *for* NP) and predicate complements (e.g., *count as* Adj, *regard* NP *as* Adj, *think of* NP *as* Adj, *show* NP *up as* Adj, etc.) are included as well.

We will focus here on phrasal verbs and prepositional verbs (types (20a) and (20b)). Because both phrasal verbs and prepositional verbs consist of more or less idiomatic combinations of verb plus particle (post-verbal particle or preposition, respectively), they are frequently treated together, yet they are distinctly different. They can be differentiated as follows. Phrasal verbs, but not prepositional verbs, allow:

(a) movement of the post-verbal particle to a position after the object (*write **up** the report*/*write the report **up***; contrast *look **after** the children*/**look the children **after***),
(b) stressing of the post-verbal particle (*I wrote it **úp***; contrast **The dog ran **áfter** the ball*).

Prepositional verbs, but not phrasal verbs, allow:

(a) an unstressed pronoun in postposition (*frown on **her***; contrast **pick up **her***),
(b) fronting of the preposition and its object (*the candidate **on whom** he frowned*; contrast **the woman **up whom** he picked*), and
(c) placing of an adverbial modifier between the verb and preposition (*look **immediately** after the problem*; contrast **pick **immediately** up the reporter*)

(see Quirk et al. 1985:1156–1157, 1166–1167; Huddleston and Pullum 2002:281–283).

We will argue that the particles of phrasal verbs are grammaticalized and prepositional verbs are lexicalized.

5.2.1 Phrasal verbs

Phrasal verbs in PDE consist of a small set of post-verbal particles, *up*, *down*, *off*, *out*, *over*, *through*, *away*, *on*, and *along*, which typically collocate with native, monosyllabic verbs (Huddleston and Pullum 2002:281, 284). The standard view is that in the shift from Object-Verb to Verb-Object order, post-verbal particles come to replace verbal prefixes, which fall into disuse by the end of the ME period (see, e.g., O'Dowd 1998:152). However, prior to the word order change, already in OE, verbs with adverbial particles (rather confusingly called "separable prefixes") are well established, though the collocations are not as fully idiomaticized as in PDE. Also unlike in PDE, the particle can either precede or follow the verb in OE. Brinton (1988:216–225) finds that the particles most commonly combining with the verb in OE are *onweg*/*aweg*, *ofdune*/*adune*, *forð*, *of* 'from', *up*, and *ut*.

Historically, phrasal verbs seem to be well established by early ME (on the history of phrasal verbs, see Hiltunen 1983, 1994; Brinton 1988:185–234; Claridge 2000:84–89), with verb plus particle combinations beginning to assume metaphorical meanings (see Hiltunen 1983:148–149; Brinton 1988:225–234). In the Lampeter Corpus of texts dated 1640–1740 (see *ICAME* 1999), the phrasal verb is substantially more numerous and varied than the prepositional verb (669 types as opposed to 199 types; see Claridge 2000:280–287). Postposition of the particle is regular (Hiltunen 1994:132). In EModE the most common particles employed in this form (*away*, *back*, *down*, *off*, *out*, *up*) closely approximate those found in both OE and PDE, with the exception that in ModE *forth* has virtually disappeared and *out* has become considerably more common; the verb combining with the particle is typically native (namely, *take*/*go away*; *bring*/*come*/*go back*; *come*/*go*/*set*/*sit*/*lay down*; *cut*/*take off*; *fall*/*find*/*go out*, and *bring*/*come*/*go*/*lift*/*take up*) (Hiltunen 1994:132–133). Thus, the phrasal verb exhibits a direct line of development from OE to the present and is more appropriately understood as the continuation of OE verbs accompanied by adverbial particles in pre- or postposition than as a replacement for lost prefixed verbs. Further evidence of the independence of phrasal verbs from prefixed verbs is the fact that the particles of phrasal verbs do not, for the most part, correspond in form to the most frequent verbal prefixes in OE (i.e., *a-*, *be-*, *for-*, *ge-*, *in-*, *of-*, *ofer-*, *on-*, *to-*). Some of the particles may function less often as prefixes (e.g., *for*ð-, *of-*, *ut-*, *up-*), while others are never prefixal (e.g., *onweg*, *ofdune*).

The verbal particles are also semantically and functionally distinguished from the verbal prefixes. The meaning of the particles is originally spatial; however, very early they come to be grammaticalized as markers of verbal aspect. Because of isomorphism between physical movement and event movement, the meaning of the particles can, in certain ambiguous contexts, undergo a metonymic shift from directional to telic meaning, that is, they perspectivize the situation as 'having an intrinsic endpoint or goal' (see Comrie 1976:44–48; Brinton 1988:25–26).[13] While Denison (1985a) finds no purely completive *up* in OE – only aspectual meaning mixed with spatial and metaphorical extensions of spatial meaning – Brinton (1988:215–234) finds telic meaning combined with spatial meaning in many of the other particles, in "bridging contexts," as in (21):

(21) a. **Aslat** þa þa tunas eall ymb þa burg **onweg**
tore then the enclosures all around the town away
'[He] then tore away all the enclosures around the town'
(c. 890 *Bede* 3 14.202.2 [Brinton 1988:222])

[13] There is considerable overlap in terminology in the aspectual domain, with the terms "telic," "completive," "resultative," and "bounded" often used interchangeably. Here, "telic" is used in the narrow sense given.

b. þa **wearp** se broðor þæt glæsene fæt **ut**
then threw the brother that glass vat out
'then the brother cast out that glass vat'
(c. 1000 *ÆCHom* II ii.104.425 [Brinton 1988:223])

In contrast to the verbal particles, the verbal prefixes in Old English show a much greater range of lexical meaning and more consistently express concrete spatial meanings, e.g., *be-* 'round, over,' *forð-* or *to-* 'motion towards,' *ofer-* 'over,' *ðurh-* 'through,' and *ymb(e)-* 'around.' Aspectual meaning, though possible (this has been most forcefully argued for *ge-* in OE, which has been described as a purely "perfective" particle), does not appear to be the primary meaning of the prefixes (on the perfective meaning of the prefixes in OE, see Brinton 1988:202ff.).

In ModE the particles are widely (albeit not universally) recognized as denoting aspectual meaning, both telic aktionsart (*up*, *down*, *out*, *off*, *through*, *over*, *away*) and iterative/durative aspect (Brinton 1988:243–246 includes a compilation of earlier scholarly opinion on the aspectual meanings of the particles). The aspectual meanings of completion/perfectivity (*break up*, *catch up*, *cut down*) or of repetition/duration (*beaver away*, *carry on*) expressed by the particles are noted in Huddleston and Pullum (2002:284). Quirk et al. point out that the verb in the phrasal verb construction keeps its meaning, while "it is the particle which establishes a family resemblance" (1985:1162), such as 'persistent action' (*chatter away*) or 'completion' (*drink up*, *point out*), but also 'endurance' (*hold out*). In her study of prepositions and particles in English, O'Dowd (1998) finds that in her spoken corpus of Modern English, *up*, *out*, *over*, *down*, *off*, and *around* and are specialized as particles with non-literal "situating," especially aspectual, meaning; most do not have or need not have NP complements (e.g., *eat up*). By contrast, *in*, *on*, *about*, *by*, and *through* are specialized as prepositions with relatively literal spatio-directional meaning; most are "linking" elements with NP complements. For example, 98 percent of uses of *up* and *out* have situating meaning, while 81 per cent of uses of *in* and 83 percent of *on* have spatio-directional linking meaning (O'Dowd 1998:32). Despite this gradience in grammaticality noted by O'Dowd, we conclude that particles are grammatical formatives expressing aspectual meaning and that originally relatively free phrasal combinations with a limited set of prepositions that invited more abstract aspectual meanings were grammaticalized by ME as G1 periphrases.[14]

[14] O'Dowd (1998:Chap. 8) hypothesizes a similar conclusion based on principles of grammaticalization and some evidence from IE languages.

5.2.2 Prepositional verbs

In contrast to phrasal verbs, prepositional verbs consist of a larger set of prepositions (e.g., *after*, *against*, *around*, *at*, *by*, *for*, *from*, *in*, *into*, *of*, *on*, *to*, *upon*, *with*), which collocate with Romance as well as native verbs. The verb selects the accompanying preposition and seemingly forms a semantic and syntactic unit with it.[15] It is frequently noted (see, e.g., Quirk et al. 1985:1156) that there are two possible structural analyses of prepositional verbs. The first (22a) underscores the unity of the PrepP and its separation from the verb and thus might be understood as "grammatical." The second (22b) underscores the unity of the V + Prep and emphasizes the semantic resemblance between the prepositional verb (*look after*) and the simplex (*tend*); it might be understood as a "lexical" analysis:

(22) a. She [looked] [after her son]
b. She [looked after] [her son]

Certain linguistic behavior (such as fronting [*her son after whom she looked*]) supports the bracketing in (22a) since it points to the existence of a phrase boundary between the V and the PrepP. However, other behavior (such as the prepositional passive [*her son was looked after*]) is evidence for the bracketing in (22b) since it maintains the contiguity of the verb and preposition apart from the NP (Quirk et al. 1985:1163–1166). Quirk et al. are undecided as to the correctness of one analysis over another, noting only that the lexical analysis is "to varying degrees, a suitable alternative" to the grammatical analysis (1985:1163). Biber, Johansson, Leech, Conrad, and Finegan (1999:59) also support the bracketing given in (22b), primarily for the semantic reason that prepositional verbs are often synonymous with simple verbs (e.g., *look like*/*resemble*, *deal with*/*handle*). In contrast, according to Huddleston and Pullum there is "no support" in syntactic behavior for the bracketing given in (22b) in which the preposition forms a constituent with the verb (2002:277; see also Huddleston 1984:200–206). At the same time, they are less certain of the appropriateness of analysis (22a) for prepositional verbs, such as *come across* or *get* NP *through*, in which they see the preposition as "fixed" or "**fossilised**, i.e., it blocks the application of the syntactic processes that can normally apply to such combinations" (277; boldface original).

[15] Denison observes that "[t]he syntax of the particle is particularly strange. The particle is a preposition as far as lexical category is concerned . . . , it has an evident transformational relationship with a preposition in the conventional pre-NP position, and in PE [=$PDE_{LJB\ \&\ ECT}$] at least it has the unstress typical of prepositions. On the other hand, its constituency seems to be with the verb, and it has no explicit NP object; these are the properties more typical of adverbs. If there is a gradience between preposition and adverb, the particle of the prepositional passive lies somewhere in the middle" (1985b:199).

A view of the history of these structures perhaps sheds light on this controversy. Claridge observes that "the appearance of the prepositional verbs on the linguistic stage seems to have been rather sudden," occurring in late OE / early ME (2000:89; on the history of prepositional verbs, see further Visser 1963–1973, I:387–407). The rise of prepositional verbs is concurrent with the loss of verbal prefixes, which over the OE period had weakened, overextended, and lost information content. Prepositional verbs – rather than phrasal verbs – thus represent the functional replacement for prefixed verbs, which already in OE could be highly idiomaticized. Metaphorical meanings appear very early in the prepositional verb. Denison (1985b:191ff.) argues that prepositional verbs undergo structural reanalysis, from a structure like that in (22a) to a structure like that in (22b). Factors contributing to this reanalysis, among others, include preposition stranding resulting from various syntactic operations, the obsolescence of the OE prefixal system, what Denison (1985b) terms "lexicalization" (the impossibility of substituting a synonymous V and/or Prep, thus *look up* cannot be substituted by *watch up*), reanalysis of coordinate structures (see (23a)), the fixing of SVO order, and a bridging context such as (23b) where non-spatial *of* precedes a subordinate clause with which it may or may not form a constituent. Denison (1985b) argues that the nature of the first *of* in (23b) is indeterminate or gradient: it is either a stranded particle or it is a preposition governing a clausal complement:

(23) a. To **lufe** and **lok on** þe
'To love (you) and look on you'
(a. 1450 Rolle, *EDormio* (Cmb) 72.359 [Denison 1985b:191])

b. ȝet is meast dred **of** hwen þe sweoke of helle eggeð
yet is most dread of when the deceiver of hell eggs
to a þing
to a thing
'Yet there is most fear of (it) when the deceiver of hell incites one to something'
(c. 1230 *Ancr.* [Corp.-C] 60a.19 [Denison 1985b:200])

As Claridge notes, these are all conditions "which would cause closer as well as more frequent contact of verbs and prepositions, and secondly conditions which would favour the re-analysis as one verbal unit of two previously separate words" (2000:89). To Denison's (1985b) list of factors she adds the loss of case endings (and concomitant increased use of prepositions) and calquing on Fr., Lat., or ON (90). Thus, a division of labor emerges between prepositional verbs, which come to replace prefixed verbs for a variety of reasons, and phrasal verbs, which continue the V and Prt constructions of OE, and become aspectual.

The prepositional passive is frequently seen as the strongest indicator of the unity of prepositional verbs.[16] The earliest examples of this construction date from c. 1400 (Denison 1985b):

(24) children – unarrayde, unkepide, and noght **tente to** as þam aughte for to be
'children, undressed, neglected, and not tended to as they ought to be'
(c.1440 Rolle, *Prose Treatises* [EETS] IX, 28 [Visser 1963–1973, III-2:2123])

Visser believes that end-position of the preposition is a "time-honoured usage" and that movement away from this construction (e.g., *whom I delite in* > *in whom I delite*) is the result of stylists' attacks on end-position in the eighteenth century (1963–1973, I:402–403). An additional criterion of unity relates to whether the NP prepositional object is questioned with a *wh*-pronoun (*She tended to her mother? **Who** did she tend to?*) or whether the PrepP is questioned with a *wh*-adverbial (*She went to the store. **Where** did she go?*) (see Quirk et al. 1985:1165). In addition to syntactic factors contributing to the unity of verb + preposition, it can be observed that, unlike the particles of phrasal verbs, the prepositions do not contribute in any systematic (or grammatical) way to the meaning of the unit. Rather, if they have any identifiable meaning (some do not, e.g., *refer to*, *accuse of*), they contribute lexically to the meaning of the combination in idiosyncratic ways often related to the lexical meaning of the preposition: *take* vs. *take after* 'resemble,' *stand* vs. *stand by* 'support,' *come* vs. *come between* 'interfere,' or *hold* vs. *hold against* 'allow to affect one's judgments.'

We conclude that prepositional verbs represent a lexicalized unit equivalent to L1 on our scale. We observe, however, that there exists a gradient from less lexicalized to more lexicalized prepositional verbs, differentiated by their degrees of fusion and idiomaticization. Using the syntactic criteria mentioned above (prepositional passive, *wh*-question formation), Quirk et al. (1985:1166) suggest that there exists a gradient from free combinations (*She left before noon*) to semiprepositional verbs (*She died of pneumonia*, *His job also comes into the picture*) to more fully prepositional verbs (*The police have asked for details*, *The queen slept in this bed*, *White wine goes with poultry*). A somewhat different gradient is suggested in Huddleston and Pullum (2002:275–277), where free combinations such as *fly to*, which permit fronting

[16] The prepositional passive is not restricted to lexicalized prepositional verbs but may also occur with free combinations such as *sit on*, *sleep in*, or *walk over* (see Huddleston and Pullum 2002:276). Denison contemplates but rejects the possibility of considering forms such as *walk over* as in *every inch of the carpet was walked over* (Dickens) as "nonce lexicalizations," choosing instead to consider the function of the passive more broadly (1985b:193–194).

but not prepositional passive, are distinguished from prepositional verbs; among prepositional verbs, "mobile specified" prepositional verbs such as *refer to*, which allow prepositional passive and fronting, can be differentiated from "fixed specified" prepositional such as *come across*, which allow prepositional passive but not fronting. Though these two gradients differ in detail, they both suggest a range from less to more lexicalized prepositional verbs.

5.2.3 "Inseparable" prefixes

Before leaving this topic, we should note that some of the locative particles would seem to function as "inseparable" prefixes in ModE, especially in substantival, participial, and adjectival forms, such as *outlook*, *oncoming*, and *overanxious* (see Marchand 1969 [1960]:109). However, Marchand argues for treating the verbal forms as compounds: "because either constituent is an independent morpheme. I am not overlooking the fact that particles used in conjunction with verbs display senses other than the ones found in the independent words which brings them nearer to prefixes. But as they do not have their dependent morphological status, a term such as semi-prefix would have had to be introduced. To avoid this I have classed them as verbal compounds" (112). He notes that apart from some remnants from Old and Middle English, such as *understand*, *withstand*, or *gainsay*, the only particles which productively combine to form verbal compounds in ModE are *out-* (e.g., *outrank*, *outstay*, *outplay*), *over-* (e.g., *overwhelm*, *overstay*, *overrule*), and to a lesser extent *under-* (e.g., *undercut*, *undercharge*, *undervalue*) (96–100). He considers *up-* to be a "doubtful bookish preparticle" (111) and sees most verbal compounds with *up-* (e.g., *update*, *uproot*, *upend*) as translations from Latin or modern analogical formations (120–121). Verbal compounds with *back-* (e.g., *backform*, *backstab*, *backfire*), *down-* (e.g., *downgrade*, *download*, *downsize*), and *off-* (e.g., *offset*, *offload*) are also found, though many involve conversion or back formation. The meaning of the locative particles when they combine with verbal elements is not predictable from the meaning of the Adv when it occurs independently, e.g., *over-* typically has the sense 'excessively' (see Marchand 1969 [1960]:98–99) from the seventeenth century *under-* has meant 'below a fixed norm' rather than simply 'below,' cf. *undercut* vs. *undersign* (see Nevalainen 1999:420). Furthermore, the particles contribute to the lexical meaning of the verb (either concretely or metaphorically) but have no grammatical consequences. We therefore consider these formations to be the result of lexicalization processes.[17]

[17] The importance of the function of the particle is emphasized by comparison with "inseparable complex verbs" that derive from "separable complex verbs" in Dutch (also in German, Finno-Ugric, etc.), e.g., Modern Du. *doorléven* 'live completely through.' Blom and Booij show that the inseparable prefixes are aspectual and instances of grammaticalization; by contrast, separable prefixes belong to

5.3 Composite predicates

As discussed in Section 3.1, composite predicates such as:

(25) give a response
make a promise
have a try
take a look
do a study

have been variously interpreted as lexicalized or grammaticalized elements. A brief look at the development of these constructions is informative in understanding their status (for details and data, see Brinton and Akimoto 1999; Visser 1963–1973, I:138–141). Historically, composite predicates have always been available in English, though they were less frequent, less idiomaticized (more transparent semantically), and less unified and fixed in OE than in ModE. Composite predicates formed with *don* 'do,' *habban* 'have,' *niman* 'take,' *sellan* 'give,' and later *macian* 'make' were the most type frequent in OE, with *niman* being replaced by *take* and *sellan* by *give* (both borrowings from Norse) in early ME. These are the same "light verbs" recognized by Algeo (1995:214–216) and Huddleston and Pullum (2002:290–296) as most common in contemporary composite predicates. In other words, the same semantic set has been used throughout the history of English. Borrowings from French (e.g., *bataille*, *deliberacioun*, *force*, *promyse*, *querele*) vastly increased the deverbal nominal objects of composite predicates in ME. Far more important was the fact that with the development of the article system during this period, the construction came to be associated with individuated, countable situations and hence a difference in meaning arose between *bathe* '(cause to) wash' (on any occasion, unless the context indicated otherwise) and *have a bath* 'prepare a bath' (on a particular occasion). Composite predicates began to be associated with telic aspect, that is, the situation is perspectivized as an accomplishment or an achievement as in *take a walk* vs. *walk* (see Brinton 1996a). This division of labor between the simple verb and the phrasal version became more marked during the course of the EModE period.[18] The aspectual function is particularly noticeable with composite predicates introduced by *have* whereas main verb *have* is typically stative, "light verb" *have* is often not, especially if the deverbal noun originates in a process verb, e.g., *have a dance*. A further grammatical distinction arose between certain *have* and *give*/*take* composite predicates, in which the former expresses stative meaning and the latter

"constructional idioms" with idiomatic meanings, cf. *aanval* 'attack,' *bijval* 'applaud,' *genval* 'disappoint,' all with the root *vallen* 'fall' (2003:63).

[18] However, Claridge (2000:245–246) finds telic meaning in complex predicates to be rare in the Lampeter Corpus (1640–1740).

dynamic meaning (as in *have a cold*/*take cold*, *have a fancy for*/*take a fancy for*, *have a wish*/*make a wish*) (see Huddleston and Pullum 2002:295).[19] During the ModE period the construction was extended to deverbal phrasal verbs, e.g., *make a clean-up*, *have a warm-up*.

If one thinks about an individual composite predicate construction as a unit (e.g., *lose sight of* = 'forget,' *give an answer* = 'answer'), and takes lexicalization to be "adoption into the lexicon," then the development of composite predicates would appear to be a case of lexicalization; the same conclusion is reached if one thinks about idiomatization and fusion as criterial to lexicalization. Although there is some syntactic freedom (e.g., many composite predicates can be used in the passive, cf. *an answer was given*, or take adjectival modifiers, cf. *raise a serious objection to*), the constructions are for the most part fixed. Indeed that is the approach taken in Nevalainen (1999) and Traugott (1999). However, this approach misses important distinctions among composite predicates which would suggest that some composite predicates are indeed lexicalizations, while others are not.

In a composite predicate such as *lose sight of*, the verb *lose* is nonproductive, i.e., does not form other patterns of the type *lose N (of)*. Except for tense and voice alternation, this form is fixed: there is now no *lose the sight of*, or *lose exceptional sights of* (Akimoto 1989:355). The meaning of such composite predicates is often idiomaticized. Complex predicates of this type have thus "fallen out of the productive rules of the grammar" (see Anttila 1989 [1972]:151). We can think of these as instances of L1 (phrasal constructions) that over time become more fossilized and less compositional, and in which the NP is decategorialized in that it loses many of its nominal properties. Other examples of lexicalized composite predicates would include *curry favor with*, *raise an objection to*, *cast doubt on*, *lay claim to*, *pay attention to* (see Huddleston and Pullum 2002:289).

However, composite predicates that have "light" verbs, i.e., those with *make*, *take*, *give*, *have*, and *do*, participate in a different construction that emerged primarily in the later ME period. This construction is different from those like *lose sight of* in that the light verbs are highly productive in forming new composite predicates (see, e.g., the lists of *make*- and *take*-complex predicates cited by Nunberg, Sag, and Wasow 1994:532–534). Furthermore, the verb has a clearly identifiable grammatical (i.e., aspectual or dynamic/stative) function. In these cases, we would say that the verb has grammaticalized. We can think of composite predicates of this type as instances of G1 (phrasal constructions) in which the NP is decategorialized, and the verb has taken on a grammatical function, and might conceivably one day come to be a grammatical (derivational) prefix not unlike *be*- in

[19] Dialect variation may also account for some differences, speakers of British English preferring *have* and speakers of American English preferring *take* (see Algeo 1995:215–216).

bedevil (this would of course involve a typological shift in word formation, given the high frequency of composite predicates with light verbs). Possibly the construction is one of several other changes that are leading to the obscuring of differences between form classes in English (in this case between V and N) that Bauer (1983:226–227) speculated was underway in contemporary English (see 2.1.3). Where Bauer foresees "closed sets such as pronouns and a single open set of lexical items that can be used as required" (227), we might foresee closed sets such as light verbs and lexical items that can be used as required.

It should be noted that the two types of composite predicates – lexical and grammatical – meet the criteria in Himmelmann (2004) in respect to host contraction/expansion as well. In the case of *lose sight of*, *lose* can be said to become fixed to *sight of*, and *sight of* to *lose*. The possibilities of type expansion have been reduced (an indicator of lexicalization). However, in the case of the light verbs, the possibilities of type expansion have increased, and the semantics-pragmatics have been enriched to include aspectual meanings (indicators of grammaticalization).

5.4 Adverbs formed with *-ly*

Both cross-linguistically and language-internally, adverbs have been considered problematic in terms of their definition and of their status as major or minor classes. Adverbs like those ending in *-ly* have been interpreted in different ways with respect to derivational vs. inflectional, open vs. closed class status. After reviewing some cross-linguistic analyses of Advs, we will conclude, like Nevalainen (1997), that *-ly* is becoming grammaticalized, and that individual adverbs formed with it have a tendency to grammaticalize.

"Adverbs" (word-level expressions) and adverbials (phrase-level expressions) belong to "an elusive category" (van der Auwera 1998:3). Like Adjs, they are modifiers, but Advs modify a larger range of elements, most especially Vs (*she ran quickly*), but also Adjs (*she is very intelligent*), Ns (*only Jim was bored*), Preps (*exactly on target*), Advs (*unnecessarily softly*), sentences (*obviously*/*logically*/*frankly you are right*), and discourses (*so, in fact* [in the sense 'more precisely']). Cross-linguistically they are positionally freer than Adjs (Bhat 2000:60). They are also less well attested than N, V, Adj and are often derived. Indeed, Talmy says that Advs "seem in all languages to be derived, as from nouns, verbs, or adjectives ... rather than to comprise in their own right an open class of intrinsically adverbial roots," and concludes that they are therefore closed class, i.e., minor, grammatical category members (2000, I:23). This, however, simplifies the facts. As mentioned in Section 1.2.2, Ramat and Ricca (1994, 1998), while acknowledging the tendency for Advs to be derived, emphasize that there is a range from relatively open class, derived Advs like *fortunately* to relatively closed class, monomorphemic Advs like *now*, *just*. Quirk et al. (1985:52,

Section 7.4.6ff.) also emphasize that Advs and adverbials cover a wide "spectrum of types," but nevertheless treat most Advs with the exception of *not* as open class items (their examples are *steadily* [manner] and also *completely*, *really* [degree]) (1985:67).

Ramat and Ricca (1994) conceptualize the range of Advs not as a single continuum, but rather as a synchronically highly articulated radial set that has formal overlaps (often via derivation) with all the major classes (N, V, Adj) and with minor ones, specifically prepositions and conjunctions. According to them, prototypical Advs, which are primarily negative (*not*), time-related (*now*), and focusing (*even*, *just*), are formally monomorphemic or at least opaque, though many are derived historically by processes of univerbation (1994:292–293).[20] In English and most European Verb-Object languages, prototype Advs occur in the "auxiliary range," i.e., between the subject and the main verb, as well as in many other positions; cross-linguistically their distribution is relatively free (Ernst 2002:357), and a subset is morphologically "light", i.e., short (e.g., *not*, *hardly*, *just*) (171–175). It appears then that Advs may range, depending on the language, from contentful, lexical, typologically non-prototypical Advs that are members of an open class, to grammatical, typologically prototypical, closed class Advs. It also appears that cross-linguistically prototypical Advs have undergone more changes than less prototypical ones, either with respect to form (cf. *not* < OE *nawiht* 'no thing') or meaning (cf. focalizer *even* < OE *efene* 'evenly').

The adverbial morpheme *-ly* has its origins in the OE adjectival ending *-lic* 'having the form/body of' + adverbial *-e* derivation,[21] i.e., ultimately in a compound. While often associated with the class of manner Advs, it is actually attested in PDE across all classes, as evidenced by "degree" Advs such as *extremely* which locate the head on a gradable scale, "sentence" Advs such as *frankly*, *legally*, "setting" Advs such as *recently*, "text" Advs like *consequently*, and even "focalizers" such as *exclusively* (Ramat and Ricca 1994:307–308), a point to which we will return below.

The most important theoretical issues relevant to this book that are raised by *-ly* are:

[20] According to Ernst (2002), of these, only negatives appear extensively cross-linguistically. Cinque hypothesizes that a rich hierarchy of functional projections exists in Universal Grammar, "possibly, on the order of 40" (1999:141), and in a relatively rigid order, whether or not they are morphophonologically realized. Cinque is clearly using the term "adverb" to refer not only to word-level but also to phrase-level expressions. Nevertheless, the widespread absence of a distinct class of Advs in many languages casts significant doubt on this hypothesis.

[21] Kastovsky (1992:397) points out that there are a few instances of Adverbial *-lice* that are in fact derived from Adjs without *-lic*, e.g., *blind* (Adj), **blindlic* 'blind,' *blindlice* 'blindly' (Adv).

(1) What is the status of *-ly*; is it a derivational or inflectional morpheme?
(2) What is its relation to zero-derived Advs (cf. *slowly* vs. *slow*)?
(3) Does the history of Advs in *-ly* suggest that the category is becoming increasingly grammaticalized?
(4) Does the history of Advs in *-ly* point to a tendency to increase categorial differences in English?

In PDE it appears that *-ly* is used very productively to form a manner Adv from any Adj.[22] This has led some researchers to claim that it is now (primarily) an inflection (e.g., Marchand 1969 [1960]:158) and was one in OE (Campbell 1959:275). However, its distribution is somewhat restricted; for example, it does not usually occur with Adjs that already have suppletive Advs (*good–well*, **goodly*) or end in *-ly* (*friendly*, **friendlily*). Furthermore, there is often an unstable semantic relationship between the meaning of the Adj base and the Adv (e.g., clause-initial *actually* does not mean 'in an actual manner' but rather 'in actual fact'; *frankly* does not mean 'in a frank manner' but rather 'frankly speaking' [cf. Huddleston 1984:332]). Therefore, adverbial *-ly* has often been analyzed as a derivational morpheme (Jackendoff 2002:155), though sometimes with a hedge; e.g., Quirk et al. say that it "could almost be regarded as inflexional" (1985:1556), and Baayen and Renouf say that the productivity of de-adjectival *-ly* Advs is "nearly 'syntactic'" (1996:93). However, Kastovsky says "[t]he formation of Advs is a borderline case between word formation and inflexional morphology" and therefore treats OE adverbial suffixes separately from either type of word formation (1992:396).

Nevalainen argues in detail that adverbial *-ly* is derivational throughout its history, not only for the reasons given above, but because it changes grammatical category, and because derivational analysis unifies the three main kinds of Adv-formation in English: suffixation (*-ly*), zero-formation, and compounding (*-wise*, *-ward(s)*, *-fashion*, *-style*) (1997:148–149; see also Nevalainen and Rissanen 2002:379). Throughout its history until recently, adverbial *-ly* has also been in competition with another derivational morpheme. At first this was OE *-e* (which has also been treated as both an inflection and a derivational morpheme). The loss of unstressed final *-e* by the end of the ME period led to zero-derivation of former *-e* adverbs. Zero-derivation in turn became recessive by the beginning of ModE, at least in Standard English (Nevalainen and Rissanen 2002), and is almost non-productive in PDE.

[22] Baayen and Renouf (1996) show that adverbial *-ly* is highly productive in the London *Times* and the Kobold/CELLED database, contrary to prior claims such as that of Cannon (1987), based on Merriam-Webster's *Third New International Dictionary of the English Language* (1981 [1961]), that *-ly* is no longer productive. The claim that derivation of manner Advs is currently productive is a claim about productivity of neologisms in terms of type-frequency.

Zero-derived forms are often said to be more concrete (Donner 1991), as in the case of *the sun shines bright* but *she smiles brightly*. However, many are, in fact, quite abstract. A particular instance is the set of degree Advs such as *just*, *pretty*, *right*, *very* (< MFr. *verrai* 'true'), a set which developed especially rapidly in the EModE period. Where zero-derived and *-ly*-derived pairs exist, they may be distinguished distributionally: zero-derived Advs are favored as modifiers of Adjs and adverbial heads, *-ly*-derived Advs as modifiers of verbs and participles (Mustanoja 1960:319–320). Nevalainen (1997:1267) cites:

(26) a. Moises and Aaron felden **lowe** to erthe, bifor al the multitude of the sones of Israel
'Moses and Aaron fell low to earth, before all the multitude of the sons of Israel'
(c.1380 Wycliffe. *Old Testament, Numbers* 14.5 [*HCET*])

b. we **lowly** beseche þe kyng of kynges
'we humbly beseech thee, king of kings'
(1418 *London Letters* 73 [*HCET*])

In many cases the distinction between Adj and zero-derived Adv is indeterminate in ME and EModE. In ME, especially, there are large numbers of borderline cases or "quasi-adverbials" (Peters 1994; Nevalainen 1997:154, 168), e.g., *liven chaste* 'live chaste(ly)', *bare hym so meke and softe* 'bore himself so meek(ly) and soft(ly)'. This is occasionally still true in PDE, cf. *They married young* (Quirk et al. 1985:737), which means both 'They were young when they married' and 'They married at a young age/soon.' The gradual disappearance of zero-formation of Advs has resulted in fewer such ambiguities.

Even though Nevalainen regards *-ly* as still derivational, she considers its development to be a case of grammaticalization (1997:182) and furthermore to be currently "in the transition area between derivation and inflection" (Nevalainen 2004:25). This is because *-ly* has become so extensively generalized with respect to its host subclasses that it can occur with stative Adjs, such as *green* (Nevalainen and Rissanen 2002 cite Killie 2000),[23] and is favored even for new degree Advs (e.g., *highly*, *truly*). The extension of *-ly* has contributed to the increase of Advs in English. Ramat and Ricca (1994) note that not only do *-ly*-derived manner Advs reveal the highest type frequency in PDE but also *-ly*-derived focalizers (even discounting *only*, which is now opaque) have the highest token frequency in the corpora they studied (318). In other words, *-ly* is an unexpectedly significant factor even in a subclass of Advs that prototypically does not show transparent

[23] Note, however, that Baayen and Renouf (1996) did not find any Advs of this form in their London *Times* database. Manner Advs derived from stative color Adjs appear to be somewhat stylistically restricted.

morphology (293). Another factor that has contributed to the extension of the Adv category is the increase in polysemy. Almost all degree Advs in English are derived from Advs with more literal functions, e.g., local-dimensional (*highly*), quantitative (*much*), qualitative (*terribly*), emphasizing (*really*, *even*) (Peters 1994);[24] the qualitative source came to be particularly productive in the EModE period. Sentential Advs in *-ly*, including modal (OE *witodlice* 'truly'), subject disjuncts (*cleverly*), speech act disjuncts (*frankly*), and evaluative (*luckily*) forms, mostly derive from manner Advs; among them the evaluative type has increased most rapidly in the twentieth century (Swan 1988). The degree Advs and sentential Advs are not only semantically but also distributionally more grammatical than the uses from which they derive because they are more versatile. "Overall, there appears to be a long-term diachronic trend towards increased grammaticalization in adverbialization in the English language" (Nevalainen 1997:182), so striking that Swan (1997) is cited by Nevalainen and Rissanen (2002:379) as calling English an adverbial language if there ever was one.

In sum, while adverbial *-ly* has remained derivational throughout its history, it has over time become more productive and regular (less idiosyncratic), and has gradually supplanted other means for forming Advs, such as *-e* or zero-derivation. It is moving toward inflectional status (G3). Furthermore, individual Advs with *-ly* that originate in open class manner Advs are being assigned increasingly grammatical functions (G2 status).

The spread of the adverbial *-ly* suffix at the expense of the zero-suffix has had the effect of progressively sharpening the distinction between Advs and Adjs. Further sharpening derives from the fact that adjectival *-ly* (*friendly*, *lonely*, *masterly*) is distributionally different and far more limited than adverbial *-ly*. Neologisms are restricted almost entirely to derivation from concrete nouns (by contrast Advs in *-ly* are largely derived from Adjs). Baayen and Renouf found 5,196 word types with *-ly* in their study of the London *Times*. Of these 1,362 were hapax legomena, i.e., occurred only once. Of these 1,362 as many as 1,278 were adverbial (e.g., *breathcatchingly*, *onely*), and only 84 adjectival (e.g., *lizardly*) (1996:82). As Dalton-Puffer concludes, "adjectival LY is lexical-derivational whereas adverbial LY is quasi-syntactic" (1996:175).

5.5 Discourse markers

Many phrasal discourse markers, because they show some degree of fusion (fossilization and routinization), have been seen as lexicalized (see 3.1). However, we will argue that they are grammaticalized. Included among this

[24] Adamson (2000:55) suggests that some degree Advs (she calls them intensifiers), e.g., *lovely* as in *lovely little house*, derive directly from affective uses of Adjs and proposes a cline Descriptive Adj > Affective Adj > Intensifier; the observation that degree Advs are derived holds in this case too.

class of discourse markers are certain types of parenthetical "comment clauses," namely, those resembling matrix clauses (e.g., *I believe*) and those resembling adverbial finite clauses (e.g., *as you know*) (see Quirk et al. 1985:1112).[25] Quirk et al. observe that functionally these comment clauses may serve as hedges, express the speaker's certainty (or concession), express an emotional attitude toward the content of the adjoined clause, or be used to claim the hearer's attention (1985:1114–1115); i.e., they have the functions generally identified with "discourse markers." Phrasal discourse markers also encompass complex adverbials showing some degree of fusion, such as *indeed*, *in fact*, *instead*, *as far as*/*insofar as*, *besides*, *after all*, and *anyway*, and various kinds of "conversational routines," such as *thank you* and *I'm sorry* (see Aijmer 1996a).

Historically, it is possible to identify a number of sources for phrasal discourse markers: subject + verb (+ object) matrix clauses (27a), imperative + subject (+ object) matrix clauses (27b), adjunct adverbial/relative clauses (27c), or adverbial prepositional phrases (Prep + N (+ Prep)) and nominal complexes (27d):

(27) a. first-person: *I say* (> *say*), *I mean*, *I think*, *I guess*, *I suppose*, *I reckon*, *I pray you* (> *pray*), *I pray thee* (> *prithee*), *I thank you* (> *thank you*, *thanks*), *I'm afraid*, *I'm sorry*, etc. as well as the impersonal ME *me thinks/thinketh* (> *methinks/methinketh*)
second-person: *you see*, *you know* (> *y'know*), *you realize*
third-person: *God forbid*, ME *God woot* (> *Goddot*, *Goddoth*, *Goddote*)

b. *look ye* (> *look'ee*; also *lookahere*), *look to it* (*look to't*, *lookit*), *hark ye* (> *hark'ee*), *mind you*, *mark you*, *say to me/us* (> *say*)

c. *if you please* (> *please*), *as it seems*, *as far as*, *insofar as (concerns/touches, regards)*

d. *indeed*, *in fact*, *instead*, *besides*, *after all*, ME *for the nones*, OE *on an(e)* (> ME *anon*), *anyway*

As has been discussed in a series of studies,[26] these forms undergo many of the changes characteristic of grammaticalization. Thompson and Mulac (1991:323–325) argue that the change from matrix clause (*I think that*

[25] Three other types of comment clauses (*to*-infinitives [e.g., *to be frank*], *-ing* clauses [e.g., *speaking honestly*], and *-ed* clauses [e.g., *stated simply*]) serve as "style adjuncts" (Quirk et al. 1985:1117) and will not be considered further here. A fourth type (those like nominal relative clauses [e.g., *what annoys me*]) is uncommon and will also not be included in the following discussion.

[26] For example, see Traugott (1995b), Brinton (1996b; In press), Palander-Collin (1999), Traugott and Dasher (2002), and several of the contributions in Jucker (1995). See also Moore (Forthcoming) on the borrowed discourse marker, *videlicet*, in ME legal language and Jucker (Online) for a bibliography of historical pragmatics.

S, *I guess that* S) to epistemic parenthetical (*I think*, *I guess*) involves changes conforming to Hopper's (1991) principles of grammaticalization (see 1.4.1).[27]

Taking *look ye/you (here)* as an example, we can see that, in addition to exhibiting Hopper's (1991) principles of divergence, persistence, and layering, it is decategorialized from a full complement-taking verb construction to an invariable particle-like form (a shift from major (open) > minor (closed) class). It also exhibits morphological bonding and phonological attrition (*look'ee*, *lookahere*), it loses concrete perceptual meaning (desemanticization), it shifts from propositional to pragmatic function, it comes to encode features of speaker attitude (subjectification), and it conventionalizes invited inferences (see Brinton 2001). In (28a) *lok* is an imperative matrix clause with the meaning 'attend to' followed by a dependent nominal clause. In (28b), *look you* is syntactically parenthetical and desemanticized (i.e., it does not retain its literal meaning), and it gives a pragmatic instruction to the hearer to 'be careful'; *look'ee* in (28c) shows further fusion and conveys the speaker attitude of impatience:

(28) a. But **lok** thou dele nought withl
'but look you deal not therewith'
'see to it that you do not deal with it'
(c. 1386 Gower, *Confessio Amantis* 1.1225 [Brinton 2001:182])

b. **Look you**, she loved her kinsman Tybalt dearly,/And so did I
(1594–1596 Shakespeare, *Romeo and Juliet* III.iv.3–4 [Brinton 2001:184])

c. **Look'ee** Serjeant, no Coaxing, no Wheedling, d'ye see
(1706 Farquhar, *Recruiting Officer* i.i [*OED*] [Brinton 2001:185])

Since they come to have scope over stretches of discourse beyond the sentence, discourse markers do not exhibit the scope reduction, or "condensation," that Lehmann first identified with grammaticalization (1995 [1982]:143ff.). However, the notion of scope reduction has been challenged in grammaticalization generally (Tabor and Traugott 1998) and in the case of the grammaticalization of discourse markers specifically (Traugott 1995b; Brinton 1996b:274). Lehmann's notion of "fixation," or loss of syntactic variability (1995 [1982]:158–160), is also problematic with respect to the grammaticalization of discourse markers since they are often quite moveable (Traugott 1995b; Brinton 1996b).[28]

[27] However, Brinton (1996b:239–254) argues that Old and Middle English evidence suggests that a development from adverbial/relative structures such as *as I think*, *since I guess* is more plausible than demotion from matrix to subordinate clause.

[28] Taking quite a different approach to a similar form, *guarda* 'look!', in Italian, Waltereit (2002) argues that the imperative is reanalyzed as a discourse marker. However, because in his view discourse markers are not part of the structure of

Given such a strong correspondence, why, then, has the development of discourse markers as a case of grammaticalization been called into question? There would appear to be two reasons:

(1) discourse markers such as *I think* or *look'ee* do not fit categories that are considered to be part of grammar proper: they typically occupy an extra-sentential position, they have non-truth-conditional meaning, and they function pragmatically; i.e., they are seen as "extra-grammatical," or
(2) they show fusion and semantic demotivation.

Reason (1) has led certain scholars, especially in Sweden, to argue that phrasal discourse markers undergo PRAGMATICALIZATION rather than grammaticalization (e.g., Erman and Kotsinas 1993; Aijmer 1996b).[29] However, increasingly it has come to be recognized that many core grammatical (functional) categories such as tense, aspect, and mood convey non-truth-conditional meaning (consider the pragmatic meaning of the past tense in *What was your name*? or of the progressive in *Are you wanting to go now*?) and that discourse-related categories such as Topic and Focus are part of grammar. Though discourse markers have primarily pragmatic meaning and carry scope over more than the sentence, they are indubitably "part of the grammar," or part of the structure of the sentence (i.e., they are not extra-grammatical) (Traugott and Dasher 2002:158–159). In this view, phrasal discourse markers can be regarded as belonging to grammar rather than lying outside it. Further support for treatment of these items as grammatical items comes from comparison with items with similar functions standardly viewed as grammatical, such as sentence adverbials (e.g., *frankly*, *honestly*) or second position clitics (e.g., connectives like Lat. *que* 'and' or Gk. *ge* 'indeed,' *te* 'and, also,' *nu(n)* 'now, indeed' in Gk.; see, e.g., Anderson 1993).[30]

sentence grammar, this change cannot be accounted for by grammaticalization. Reanalysis occurs "in contexts of floor-seeking" where speakers use the form "illegitimately" in order "to begin a turn before the next transition place" (989, 999, 1002–1003), that is, as a means of interruption and as a justification for self-selection. While arguing that language change does not follow from the inherent meaning of forms, Waltereit nonetheless relates the discourse functions of *guarda* to its original meaning as a verb of perception.

[29] Pragmaticalization is also seen to account for the development of pragmatic meanings more generally, as in Sw. *bara/ba'* 'naked' > 'merely,' Eng. *mere* 'pure' > 'nothing more than' (both focalizers).

[30] In their synchronic study of French discourse particles and back-channelers, e.g., Quebec Fr. *là*, Port. *né* Vincent, Votre, and LaForest (1993) speak of POSTGRAMMATICALIZATION. Despite the diachronic sound of the term, no prior historically more grammatical stage seems to be posited by the authors. Rather, it seems to mean pragmaticalization, and migration to an "extra-grammatical" level.

Reason (2) has lead some scholars to suggest that phrasal discourse markers have been lexicalized (see 3.1). For example, Wischer (2000:363) argues that the development of *methinks* qualifies as lexicalization because "a once productively formed impersonal construction has been fossilized, 'conventionalized,' partly demotivated (since impersonal constructions have become unproductive), and therefore changed into a symbol, a formula, and as such it has to be stored as a whole entity in the lexicon." However, as we have seen (see 4.4), grammaticalization and lexicalization share the features of fusion and semantic demotivation as strong correlations. Furthermore, grammatical items, no less than lexical items, constitute part of the "inventory" of items in a language. Most importantly, the development of (phrasal) discourse markers is characterized by decategorialization – a characteristic of grammaticalization and not lexicalization – and discourse markers, unlike the results of lexicalization generally, do not belong to a major lexical category (N, V, Adj). If they bear a clear resemblance to any other category, it is to that of Adv, many members of which, as we have seen, must be understood as grammaticalized. Crucially, discourse markers introduce sentences and discourse chunks very productively – the ultimate in host-extension.

We thus conclude that the development of discourse markers, both phrasal and non-phrasal, is best understood as a process of grammaticalization.

6

Conclusion and research questions

6.0 Introduction

This chapter will provide a summary of the book (6.1) followed by a discussion of questions which have emerged from the work and which require further study (6.2), including:

(1) Are there changes in lexicalization and grammaticalization that do not occur or are unlikely to occur?
(2) What constraints are there on transition from one category to another in lexicalization and grammaticalization?
(3) How do large-scale typological shifts affect transition in lexicalization and grammaticalization?
(4) What correlations are there between discourse types and lexicalization and grammaticalization?
(5) What correlations are there between specific lexicalization and grammaticalization patterns and language contact?

6.1 Summary

The foregoing study has been an examination of the process of diachronic lexicalization, most especially in light of contemporary work on diachronic grammaticalization, where it has been common to label phenomena which run counter to the expected directionality of grammaticalization as either "lexicalization" or "degrammaticalization." To a lesser extent, this study has also examined lexicalization in the context of work in morphology, where lexicalization is often equated with routine processes of word formation. Most of the examples have been from English.

The current state of work in lexicalization suggests a number of conflicting interpretations that deserve attention:

(1) Lexical and grammatical have been seen as discrete categories or as part of a gradient/continuum.
(2) Lexicalization is on the one hand viewed from its starting point, as "syntacticogenetic" (i.e., as "falling outside the productive rules of grammar") or on the other from its endpoint, as "lexicotelic" (i.e., as "adoption into the inventory").
(3) Lexicalization has been identified with a diverse and often opposing set of phenomena, including those involving processes of fusion leading to decreased compositionality and those involving processes of separation leading to increased autonomy.
(4) Lexicalization has at times been equated with other processes, variously of a formal nature (such as fossilization), of a semantic nature (such as idiomaticization), of a pragmatic nature (such as routinization), and of a sociolinguistic nature (such as institutionalization), while at other times it has been seen as distinct from these processes.
(5) Lexicalization and grammaticalization may be considered converse, and hence lexicalization is identified with *de*grammaticalization. In contrast, lexicalization and grammaticalization may be considered either complementary (at least in part) or "orthogonal" (Lehmann 2002), and hence lexicalization is distinguished from degrammaticalization.

Chapter 1 establishes some preliminaries to the discussion of lexicalization and grammaticalization. We begin with a brief summary of contemporary approaches to grammar and to language change and then turn to views of the lexicon. Much theoretical debate has focused on the nature of lexical items, whether they are to be conceived of holistically or compositionally. Debate has also centered on the distinction between lexical and grammatical categories and their relation to word classes, especially as there are significant differences cross-linguistically in these domains and as it remains unclear where to draw boundaries, or even if clear binary distinctions can be made, between lexical and grammatical. While we do not presume to resolve these debates, we adopt – along with others – a continuum model of the lexical/grammatical split and understand there to be a gradient from more to less grammatical, more to less productive/type frequent. The latter half of the chapter provides a brief introduction to lexicalization, first as a synchronic process (concerned with the formal encoding of conceptual structures) and then as a diachronic process, which is the focus of the ensuing study. Likewise, we provide an introduction to synchronic and diachronic perspectives on grammaticalization, focusing on factors which have been recognized as essential in the diachronic process, including decategorialization, gradualness, fusion/coalescence, typological generality, metaphorization/metonymization, subjectification, bleaching, frequency. We come back to all of these factors in Chapter 4 as part of our unified treatment of grammaticalization and lexicalization.

Chapter 2 explores various definitions of, and viewpoints on lexicalization, which in its broadest sense has been understood as "institutionalized adoption into the lexicon." We begin with a survey of routine processes of word formation – compounding, derivation, conversion (especially the shift from minor > major word class), clipping, ellipsis, blending, back formation, initialism, acronym creation, metalinguistic citation, and coinage – which, because they create new lexemes (neologisms), are often conceived of as lexicalization. While word formation processes are typically treated as a synchronic phenomenon, they result in additions to the inventory (the lexicon) and hence may also be seen as having a diachronic dimension. We then move to the more obviously diachronic aspects of lexicalization in which new lexemes may arise via processes of either fusion or separation. After looking at how nonce formations become neologisms, typically via institutionalization (here again there exists some debate about whether institutionalization is coincidental with or a precursor of lexicalization), we survey processes of morphophonemic fusion and coalescence that have been labeled lexicalization, including the change from syntagm > lexeme and from complex > simple lexeme (including the amalgamation of compounds, the rise of derivational affixes from roots, phonogenesis, and phonologization). Processes of semantic fusion, such as idiomaticization and semantic demotivation, are included. Processes of separation, such as clitic > word or affix > word, may also result in the rise of new lexemes and hence have been said to constitute lexicalization; however, such changes have also been categorized as "decliticization" (ceasing to be a clitic) and hence associated with degrammaticalization (as opposed to cliticization [coming to be a clitic] which has been associated with grammaticalization).

Chapter 3 moves on to a discussion of examples in the literature which have been variously identified with processes of either lexicalization or grammaticalization, including the development of derivational affixes from independent words and the fixing of phrasal forms (complex prepositions, multi-word verbs, composite predicates, and discourse markers). What lies behind these different interpretations is the process of coalescence, which has been seen as central to both lexicalization and grammaticalization. In grammaticalization studies, fusion and coalescence have been subsumed under the more encompassing principle of "unidirectionality." The view of certain scholars who have pursued the implications of unidirectionality for lexicalization (Lehmann 2002; Himmelmann 2004) is that although lexicalization is morphologically unidirectional, it may involve narrowing of semantic-pragmatic contexts or of an item's range of collocational possibilities, neither of which are characteristic of the unidirectionality found in grammaticalization. Other aspects of unidirectionality which seem to be shared by lexicalization and grammaticalization are layering (Hopper 1991), the development of new grammaticalized and lexicalized forms and replacement of older forms (after a period of competition), as well as the revival of extant but rarely used forms.

The chapter then turns to another view of the relationship of grammaticalization and lexicalization, namely, that they are converses or mirror images of each other and that lexicalization is, or is a part of, degrammaticalization. While it is fully recognized that exact reversals of grammaticalization are not possible (that is, of a specific form following the exact path in reverse), the shift from more grammatical > less grammatical > lexical (what is often understood as lexicalization) is interpreted as the reverse of the shift from lexical > less grammatical > more grammatical (what is generally accepted as grammaticalization). For Ramat (1992; 2001), lexicalization **is** degrammaticalization (along with demorphologization). However, increasingly scholars are arguing that lexicalization and degrammaticalization must be uncoupled, either as in the case of Lehmann (2002), whose starting point is that lexicalization necessarily involves coalescence (hence its reverse involves "bestowing structure" [2002:14], i.e., folk etymology), or in the case of Norde (2001, 2002) or Haspelmath (2004), whose starting point is that degrammaticalization as the reverse of grammaticalization is necessarily gradual (while lexicalization is not, in their view, gradual). The chapter ends with a brief excursus on the status of derivation vis-à-vis inflection, which has likewise proved contentious in respect to lexicalization and grammaticalization.

In Chapter 4, we begin with the observation that the "lexicon" – better understood as an "inventory" – is not homogeneous but contains some forms which are productive or not, some that are contentful or not. Synchronically, the inventory is structured as a continuum of type frequency from nonproductive to semiproductive to productive, which correlates with a continuum from lexical to grammatical (open∼closed, free∼obligatory, contentful∼functional). Degrees of grammaticality range from semiproductive to productive and are based on extent of structural fusion (G1–G3); degrees of lexicality range from semiproductive to nonproductive and are based on degrees of compositionality (L1–L3). Diachronically, both lexicalization and grammaticalization concern adoption of items into the inventory, and modifications of items within the structure of the inventory, with respect to changes in such factors as function, productivity, fusion, coalescence, compositionality, and contentfulness. Change occurs in context of language use and involves institutionalization.

We define lexicalization as:

> the change whereby in certain linguistic contexts speakers use a syntactic construction or word formation as a new contentful form with formal and semantic properties that are not completely derivable or predictable from the constituents of the construction or the word formation pattern. Over time there may be further loss of internal constituency and the item may become more lexical.

The input of lexicalization is anything stored in the inventory and the output is a new contentful (lexical) form of any complexity. Lexicalization is a

gradual change involving formal reduction and loss of semantic compositionality (idiomaticization). Once formed, lexical items may undergo further phonological and structural change to become more "lexical." A number of the examples cited in the literature fail this definition of lexicalization, including most cases of conversion and clipping, while the coalescence of certain syntactic phrases, the creation of semantic affixes, and phonogenesis would meet the definition.

We define grammaticalization as:

> the change whereby in certain linguistic contexts speakers use parts of a construction with a grammatical function. Over time the resulting grammatical item may become more grammatical by acquiring more grammatical functions and expanding its host-classes.

We conclude that both lexicalization and grammaticalization are gradual processes typically involving semantic demotivation (loss of semantic compositionality) and coalescence (loss of phonological segments). Lexicalization, by our definition, obligatorily requires fusion (loss of morpheme boundaries) whereas grammaticalization frequently, but not always, involves fusion. Both lexicalization and grammaticalization are characterized by unidirectionality, although grammaticalization is more consistently so than lexicalization (see above). We reject the attempted equation of metaphor with grammaticalization and of metonymy with lexicalization (see Moreno Cabrero 1998) and argue that metaphor and metonymy potentially participate in both processes. However, only grammaticalization, and not lexicalization, centrally involves decategorialization, bleaching, subjectification, increased frequency and productivity, and typological generalization.

We thus do not conceptualize lexicalization as a mirror image of grammaticalization in any sense. We understand "delexicalization" as a process that leads to an increase in compositionality and undoes coalescence, moving from L3 > L2 > L1; examples of such would be folk etymology (following Lehmann 2002) and certain types of back formation. We understand "degrammaticalization" as a process that leads to an increase in autonomy of form and often decreased productivity, moving from G3 > G2 > G1; examples of such would be decliticization as well as certain instances of shifts from inflection to clitic (see Haspelmath 2004).

Chapter 5 takes up some of the problematic examples of semiproductive constructions which lie on the border between lexicalization and grammaticalization and considers their historical development vis-à-vis grammaticalization and lexicalization (as defined in Chapter 4). First, while the details of development are quite complex, it would appear that many present participial Adjs, especially those containing Psych-Vs (e.g., *the disturbing news*) are lexicalized forms (L2) with new and not entirely predictable meanings; they are not conversions from the inflected V-*ing* form (cf. *The noise is disturbing us*), as suggested in contemporary

grammars. In contrast, present participial prepositions (e.g., *concerning the news*) are grammaticalizations (G2) involving reanalysis of free adjuncts and absolutes; present participial degree adverbs (*piping hot*) are likewise grammaticalizations involving reanalysis of verbal V-*ing* forms.

Second, we briefly contrast the development of phrasal verbs (e.g., *think out*) and prepositional verbs (e.g., *think about*) in English. We understand phrasal verbs as the grammaticalized descendants (G1) of verb + adverbial particle combinations in OE, in which the particle now serves a grammatical function within the aspectual system of ModE. While the analysis of prepositional verbs as lexicalized or grammaticalized is the subject of dispute in contemporary accounts, we argue that historically these are the descendents of prefixed verbs in OE and exhibit a range from less to more lexicalized (L1 > L3) in ModE.

Third, we see composite predicates as falling into two types, those formed with the "light" verbs *have*, *do*, *make*, *take*, *give* + deverbal noun (e.g., *take a bribe*) and those formed with more specialized verbs + deverbal noun (e.g., *raise an objection*). The former are productive and form an integral part of the aspectual system of ModE and hence are grammaticalized (G1). The latter are less fully productive, more fixed, and less compositional; we consider these to be lexicalized phrasal constructions (L1).

Fourth, noting that the category of adverb has been interpreted as both an open and a closed class, we undertake to determine the lexical or grammatical status of adverbial *-ly* in English. Although it has remained derivational, this affix has become increasingly productive and regular during its history (has been grammaticalized) and is on the road towards becoming an inflection (G3).

Finally, we look at a number of phrasal discourse markers, such as *I think*, *mind you*, *if you please*, and *after all*. While these forms exhibit (partial) fusion and semantic demotivation (features of both lexicalization and grammaticalization, as discussed in Chapter 4), they are best understood as grammaticalizations because in their development they undergo many of the changes characteristic of grammaticalization (chief among them decategorialization) and they come to belong to a functional rather than to a lexical word class. Moreover, we argue that they may be understood as grammaticalization rather than pragmaticalization as it is now recognized that many core grammatical categories convey non-truth-conditional meaning.

Note, in conclusion, that the results reached in this chapter rest crucially on the assumption of a continuum from L1 > L3 and G1 > G3, as shown in Table 4.2, though in the process of change, units do not always achieve full L3 or G3 status. Nor do units necessarily begin the lexicalization or grammaticalization process at the L1 and G1 stage, respectively; they may enter at the L2 or G2 stage.

6.2 Research questions

We here outline some research questions that arise from the issues discussed in the preceding chapters. Most obviously, the language base for studies of grammaticalization and most especially lexicalization needs to be enormously expanded so that the claims in Chapter 4 regarding similarities and differences between them can be tested, and so that our knowledge of types of language change can be refined. There has been substantial work on grammaticalization in languages with long historical records, such as Chinese (e.g., Sun 1996) and Japanese (see, e.g., papers in the series *Japanese and Korean Linguistics*). There has also been work on grammaticalization in languages with little recorded history, such as Korean (e.g., Rhee 1996; and articles in *Japanese and Korean Linguistics*) or very little, including African languages (see, e.g., Heine and Reh 1984; Heine, Güldemann, Kilian-Katz, Lessau, Roberg, Schladt, and Stolz 1993), Australian languages (see, e.g., Evans 1995), and Native American languages, (see, e.g., Campbell and Mithun 1979; Mithun 1999; and on the problem of reconstructing grammars in general, Gildea 1999). However, to our knowledge little has been done in these languages and language groups on lexicalization (but see Mithun 2001). Like all historical research, questions regarding lexicalization and grammaticalization should be embedded in overarching questions formulated in terms of the problems for historical linguistics posed in Weinreich, Labov, and Herzog (1968) and cited in Section 1.1.2. They concern "constraints," "transition," "actuation," and "actualization" problems. These problems have guided much of the work to date on grammaticalization, but less attention has been paid to them in systematic ways where lexicalization is concerned. We suggest a few questions in the hopes that the asymmetry between work on lexicalization and grammaticalization can be rectified.

6.2.1 Possible and impossible changes

The grammaticalization literature has a long history of studies attempting to specify constraints on possible changes cross-linguistically. The various hypotheses about unidirectional changes from "free" forms or open class items > G1 > G2 > G3, and the attempts to challenge them are designed to answer this constraints question and also to give insights into foundations of cognition. Much of the work has been conceptualized in terms of the semantics of the source and "target," e.g., spatial > temporal marker > concessive marker, but not vice versa; spatial > case marker, but not vice versa, or of the particular lexical types that give rise to particular grammatical markers, e.g., terms for GIVE > benefactive or dative, but not to instrumental or accusative. A detailed summary of possible changes, and by inference unlikely ones,

can be found in Heine and Kuteva (2002). Other hypotheses relevant to the data discussed in this book have been made about constraints on structural sources such as particular linguistic subcategories, e.g., serial verbs > core (case-related) Preps, present participial Adjs > communicative, textual, discourse-structuring Preps/Conjs, but not serial verbs > discourse-structuring preps/conjunctions (Kortmann and König 1992:688–693).

Comparable work needs to be done in areas relevant to diachronic lexicalization. We have hypothesized a change from "free" forms > L1 > L2 > L3. Like the equivalent unidirectionality hypothesis in grammaticalization it is a strong, empirically testable hypothesis, and should be challenged, as should the flowchart in Figure 4.1. and the schema of correlations in Table 4.2. This can only be done by looking, as has been done for grammaticalization, at particular instances as well as classes of lexicalization changing in context. However, as has been pointed out many times, lexicalization operates more idiosyncratically than grammaticalization, and so the task appears more difficult. Possible questions include whether there are changes in addition to folk etymology and back formation that are cases of antilexicalization.

A set of questions concerning constraints on possible changes that we have touched on several times concerns whether it is historically feasible to determine whether a particular change in affix is from derivation to inflection or vice versa. Kuryłowicz (1975 [1965]) cites derivation > inflection as in instances of "less to more grammatical," but also provides putative examples of the reverse. As discussed in Section 3.4, Bybee (1985) and Bybee, Perkins, and Pagliuca (1994) regard the "derivation cline" as primarily synchronic and typological. However, a distinction between derivation and inflection is not specified in the "grammaticalization" cline, which is characterized as "phrases or words – non-bound grams – inflection (bound grams)" (Bybee, Perkins, and Pagliuca 1994:40). As we have mentioned, others have likewise rejected or avoided a distinction between derivational and inflectional affixes in historical schemas. For example, in discussing antigrammaticalization and putative counterexamples to unidirectionality in grammaticalization, Haspelmath treats changes between the two as "internal to the morphology" because in his view "[i]nflectional patterns do not show stronger internal dependencies than derivational patterns" (2004:32) (see 3.4), and Himmelmann (2004) (see 2.1.2) suggests derivation may be a process separate from both lexicalization and grammaticalization. Nevertheless, Nevalainen (2004) suggests that the synchronic derivation cline, which she cites in Bybee's (1985:12) version (lexical – derivational – inflectional – free grammatical – syntactic),[1] could be used to characterize diachronic shifts. However, since inflection does not normally become "free

[1] Contrast the Bybee, Perkins, and Pagliuca (1994) version cited in Section 3.4, which is in the reverse order.

grammatical" diachronically (see 3.3.1) this raises further questions, rather than resolving the issue. It seems best to retain a version of Bybee's and Bybee, Perkins, and Pagliuca's grammaticalization cline, but restated, as in (1), since not all bound grams are inflections:

(1) phrases or words > non-bound grams > bound grams

Part of the problem is that there is so little agreement about how to distinguish the two with respect to particular historical data. This is no doubt in part because the traditional distinction is fuzzy at best between category changing derivational markers such as *-ness* and adverbial *-ly*, which are thought of as grammatical, and non-category changing markers such as *un-*, which are thought of as semantic and lexical (see 2.1.2). Recently the criterion of type frequency has been used more successfully for distinguishing them. For example, in the case of adverbial *-ly* it has become so productive that it is considered to be almost inflectional by Nevalainen (1997, 2004), whereas adjectival *-ly* is so relatively unproductive that it is derivational (see 5.4). Dalton-Puffer (1996:175) suggests that as early as the ME period adjectival *-ly* is "derivational-lexical" because it has limited productivity, but adverbial *-ly* is "derivational-grammatical" because it is highly generalized and very productive. According to this terminology, adverbial *-ly* is still derivational, as Nevalainen claims, thus emphasizing its lexical and idiosyncratic characteristics. A different set of distinctions proposed by Booij (2002) would put it in the inflectional category, thus emphasizing its grammatical characteristics. Booij distinguishes "derivation," which is lexical, "inherent inflection," which may change class category membership, but crucially is not obligatory (e.g., participle and infinitive affixes), and "contextual inflection," which is a default category (e.g., agreement). In either case the distinctions are on a continuum, not absolute. Historically, Dalton-Puffer's and Booij's categories can be schematized as in (2) (where "lex" stands for "lexical," "deriv" for "derivational," "gram" for "grammatical," and "infl" for "inflectional").

(2) a. lex > deriv-lex > deriv-gram > infl
b. lex > deriv > inherent infl > contextual infl

Using these two sets of distinctions, we can differentiate between adverbial *-ly* (< OE *lice*) and adjectival *-ly* (< OE *lic*) as in (3):

(3) a. lex > deriv-lex > deriv-gram > infl
b. lex > deriv > inherent infl > contextual infl
OE *lice* > ME *-ly* > PDE *-ly*
OE *lic* > ME *-ly*

The obvious question, then, is whether criteria for distinguishing points on the continuum between derivational and inflectional morphology from a historical perspective can be satisfactorily established, and whether, for any

particular example, (2a) or (2b) is a better description, and why. The next step would be to test whether the historical transition path is from derivational-grammatical/inherent inflectional to inflection, as would be predicted on grounds of gradualness, and never directly from derivational-lexical/derivational to inflectional. Furthermore, is the default transition always from left to right as in (2), as would be predicted by the hypothesis of unidirectionality, or do generalizable changes from inflection to derivation also occur?

6.2.2 Transitions from one category to another

While the constraints problem focuses on larger schematic changes of the type A > B and tends to assume that A and B are categories, the transition problem is concerned with understanding the tiny local steps between A and B that the arrow ">" encompasses. As discussed in Sections 1.2.3, 1.2.4, and 4.1, we assume a model of grammar that pays attention to gradience and gradualness, in other words, to the "numerous imperceptible changes with no radical break at any point" (Haspelmath 2004:28). Indeed, we have defined both lexicalization and grammaticalization as changes that are gradual (4.2.1). Gradience is a primarily synchronic phenomenon, gradualness a primarily diachronic one, and the assumption is that gradience is both the result of and the reason for change. Questions necessarily arise concerning how the gradualness of change maps on to the relative stability of gradience.

Denison (2001) distinguishes two kinds of gradience and discusses changes in English during the narrow time span from 1800 on, distinguishing some that are gradual from some that are not. This paper is suggestive of what a research program on constraints on transition from one category to another in lexicalization and grammaticalization might look like. The two types of gradience are "subsective" and "intersective" (terms Denison attributes to Bas Aarts; see further Aarts 2004):

(a) Subsective gradience is found when X and Y are in a gradient relationship within the same form class. This is a question of prototype vs. marginal members of a category (e.g., *house* is a more prototypical N than *home* with respect to determiners and quantifiers; *house* is also less subject to idiomatic use).
(b) Intersective gradience is found when X and Y are in a gradient relationship between classes; see the notion of "category squish" (from Ross 1972 on).

The concept of subsective gradience (but not the term) has been very important in work on decategorialization and grammaticalization. One structural diagnostic of early grammaticalization is the development of non-prototypical uses of a relatively prototypical item or construction,

e.g., the beginnings of structural auxiliarihood are identified when a verb has ceased to have prototypical morphology and distribution,[2] as in the case of the core modals in English, which ceased to have infinitive forms, object complements, etc. In other words, the gradualness of structural changes leading to modal status begins in the subsective gradience of the source V. Subsective gradience could pertain to lexicalization with respect to the question whether there is any correlation between the likelihood of being affected by fusion, coalescence, and core or peripheral status of the source item(s).

The notion of intersective gradience (but again not the term) has also been important in historical syntax with respect to reanalysis and grammaticalization (see Harris and Campbell 1995) and the development of blends as in the case of the gerundive and *sort of* (Tabor 1994) since it concerns constraints on ways in which a member of one category can be reinterpreted as a member of another. As lexicalization may involve blends (e.g., *infomercial*), the question remains to what extent the blends are similar or different; for example, can they be generalized in lexicalization to classes of lexical items as can *sort/kind/type of* (by hypothesis it would not involve morphosyntactic blends such as are illustrated by gerundives, precisely because they are morphosyntactic)?

Writing from a strictly synchronic perspective, Aarts (2004) argues that categories in an intersective gradience relation are not on a continuous cline. Rather, in many cases where gradience between categories has been postulated, the categories are too fine-grained (e.g., distinctions between prepositions, adverbs, and conjunctions, and even the distinction between Adj and Adv, are called into question (30)). On his view, as an idealization, while there may be convergence between classes (sharing of distributional properties, as in the case of gerunds) there is always a clear cut-off point (36). "Hybridity" or structural overlap between constructions is recognized, as in the case of *I dislike Brown's painting his daughter*. Aarts says that despite the verbal ending and NP complement, the fact that *painting* can potentially be modified by an Adv such as *deftly* but not an Adj such as *deft* "tips the balance" in favor of categorizing it as a verb (34). One problem for the historical linguist is that such potentials arise only in the aftermath of extensive indeterminacy found (synchronically) in historical data (see discussion of *be going to* in 1.4.2) and presumably synchronically in small subsections of the grammar of PDE, or any other present-day language. They can therefore hardly be used as diagnostics prior to their development. Another problem is that Aarts' model seems to assume monotonic continua

[2] Morphosyntactic changes indicating structural auxiliarihood are not to be equated with the beginning of auxiliary status. They must be preceded by semantic-pragmatic changes (e.g., Brinton 1988; Hopper and Traugott 1993; Himmelmann 2004).

in convergence (he adopts "the null hypothesis that all the properties are equal" (38)). However, the historical data seem to suggest otherwise. The most appropriate metaphor would appear to be a "cline" with "focal points where phenomena may cluster" (Hopper and Traugott 1993:7).

In the grammaticalization literature, the "cluster" approach has been used to identify points at which decategorialized items become recategorized into an extant category, e.g., of *concerning* into the extant category Prep, or into a new category for the language, e.g., the development of syntactic auxiliaries in English, of a complementizer phrase in IE (Kiparsky 1995) or a determiner phrase in Latin (Vincent 1996).[3] It is a truism that grammaticalized items split from their sources (a phenomenon Hopper [1991] has called "divergence") and coexist at least for a time with their sources. Because the study of grammaticalization concerns development of or recruitment into relatively closed categories with few members, a certain amount of idiosyncratic behavior is often expected, at least by functionalist researchers. However, the trend seems to be for changing items to take on more and more characteristics of their new class (Y), e.g., quasi-modals. Krug suggests that quasi-modals, which he prefers to call "emerging modals," that allow drastic phonological reduction (*wanna*, *gotta*, *hafta*, *be gonna* but not *attemma* < *attempt to*) are forming a new "focal point" toward which other quasi-modals are gravitating syntactically and in part phonologically (cf. *ought to*, *dare to*, *need to*) (2000:215–217). The change may take several centuries as in the case of *be – ing*, but the latter did eventually come to pattern with auxiliaries in collocating with the passive *be – en* (see 5.1 and 6.2.4 below). Since change is not deterministic, individual changes may also abort and not "go to completion," e.g., ME *ginnen* 'begin to' (see Brinton 1988 on this and several other aspectual verbs that fell into disuse). However, for those that do not abort, continual indeterminacy does not seem to be an expected outcome (provided the researcher can view a long enough time span).

Since Denison poses questions about open classes, he focuses on possible ways in which changes may not just lead to clusters/gravitational points, but may actually lead to "sharpening" of major open categories (2001:132). Examples mentioned in Chapter 5 are the way in which the category Adj came to be distinguished, at least in part, from the category Prep as transitive Adjs including present participials came to be used as intransitives (5.1) and from Advs as *-ly* became a salient adverbial derivational morpheme (5.4). In fact, Aarts (2004:35) says he has "a suspicion that languages do not tolerate truly hybrid structures." We can modify this proposal to the research question "Do speakers not tolerate truly hybrid structures and eliminate them over time?" It may be that the answer will be negative,

[3] The latter kind of change where a new category comes into being is relatively rare.

somewhat like the answer to the question whether "pernicious ambiguity" or "homonymiphobia" in polysemies can be tolerated (for discussion, using the example of *sanction* 'penalty, approval,' see Traugott and Dasher 2002:53). But it raises the important research question of what linguistic resources of a language, lexical and grammatical, are used over time to distinguish categories.

6.2.3 Typological shifts

A full understanding of many changes subsumed under both lexicalization and grammaticalization necessitates situating them in the larger typological shifts that a particular language is undergoing.

We have cited the examples of developments leading to such pairs in English as *sit*/*set*, *drink*/*drench*, *lie*/*lay*. They have been conceptualized as instances of lexicalization by Bybee (1985), morphologization by Anderson (1992), and as the last stage of grammaticalization ("phonologization") by Hopper (1994) (see 2.3.3). Whatever term is used, original morphology and original allophonic variation resulting from its presence is lost due to phonological changes, and the forms split. Such changes may seem idiosyncratic when cited as instances of particular kinds of change, and as topics of terminological debate. However, when situated in the word formation typology of early Germanic, the examples can be seen as representatives of a major and productive word formation strategy, which was coming to be obsolete by the end of the OE period. Very briefly, early Germanic word formation was "stem-based"; that is, there was no major class form without an inflectional ending: Ns and Adjs were inflected for number, case, and gender, Vs for tense, aspect, and mood (Wurzel 1984; Kastovsky 1992).[4] For example, in OE, vowel alternation that had its source in *i*-mutation could signal causative in a pair such as *full* 'full' ~ *fyllan* 'to fill,' *dom* 'judgment, jurisdiction' ~ *deman* 'to judge' (Kastovsky 1994:151) while vowel alternation that had its source in ablaut[5] could be far more freely used to form deverbal nouns such as *brecan* 'to break,' *bræc* 'breaking' ~ *gebrecness* 'breach' ~ *brecþa* 'broken condition' ~ *broc* 'fragment' ~ *(ge)bryce* 'breach, breaking' (147–148). Although the derivational relationships were originally transparent, they became more and more opaque over time for a variety of reasons. Particularly importantly, various types of vowel alternations that in earlier times had arisen due to the erosion of conditioning factors, such as umlaut and ablaut, ceased to be productive by the beginning of ME (Kastovsky 1992, 1994). Also in the Germanic languages "weak" Ns and Vs

[4] Note this is a narrower sense of "stem" than has been used above.

[5] "Ablaut" is a vowel alternation found especially in verbs (cf. PDE *sing-sang-sung*) that originated in IE and had come to have allomorphological function very early on.

without vowel alternations started to compete with and eventually become more productive than "strong" ones with vowel alternations (cf. OE weak N *guma* 'man' – *guman* 'man + PL' vs. strong N *fot* 'foot' – *fet* 'foot + PL'; weak V *ic lufie* 'I love + PRES' – *ic lufode* 'I love + PAST' vs. strong V *ic ride* 'I ride + PRES' vs. *ic rad* 'I ride + PAST'). Furthermore, meanings not derivable entirely from the constituent elements of the word came to be institutionalized (consider the various forms cited for 'breach'). As a result of these and other changes including loss of unstressed vowels, most Ns and Vs came to be "word-based," i.e., some forms were not inflected (typically subject for the N, infinitive for the V). From the ME period on, new native vocabulary was not formed with vowel alternations, but with derivational morphology attached to the word base.

However, extensive borrowing led to new lexical items with alternation, most especially words with Latinate sources such as *science* ~ *scientific*. By contrast, speakers of German, in which sound changes did not eliminate unstressed vowels, maintained the stem-based word formation type that was typical of OE, e.g., *rot* 'red' ~ *röten* 'become red,' *Bach* 'stream' ~ *Bächlein* 'little stream,' *werfen* 'to throw' ~ *Wurf* 'a throw' ~ *Würfel* 'die, cube' ~ *Zerwürfnis* 'quarrel' ~ *Werfer* 'thrower' (Kastovsky 1994:143). As in English, borrowing came to be widespread in the Renaissance, and additional stem formations were added and accommodated to extant patterns; see, e.g., the umlaut in *nervös* 'nervous' ~ *Nervosität* 'nervousness' (143). We might ask why when stem-based word formation with vowel alternation such as *full* ~*fill* became unproductive and was replaced by word-based word formation, new borrowed stem formations such as *science* ~ *scientific* were not eliminated by analogy with the new word-based formations (Kastovsky 1994:155). Another question might be whether stem-based word formation favors lexicalization more or less than word-based word formation (the first would appear to be more subject to coalescence due to phonological and morphological changes, and therefore perhaps more subject to lexicalization).

The shift in word formation strategies surely accounts to a considerable extent for the rapid growth of new vocabulary in the ME period; other factors include the influence of other languages, especially Scandinavian and French (see 6.2.5 below), and most especially a morphosyntactic typological shift that also occurred around this time. Throughout the OE period we see increasing use of prepositions in addition to case, of pre-verbal periphrastic aspect, modality, and voice markers in addition to post-verbal tense inflection; also a shift from object-V > V-object word order, Adv-V > V-Adv, and other factors consistent with a typological shift to verb-object and "right-branching" order (see Dryer 1992 for discussion of the cross-linguistic correlates of object-V and V-object order, and references therein). This typological shift to V-object order and periphrastic expression is the context in which phrasal Vs were grammaticalized (5.2.1), while prepositional Vs and "inseparable" prefixes were lexicalized (see 5.2.2,

5.2.3, respectively). As we have argued, prepositional Vs such as *look after*, which are highly idiomatic, replaced OE prefixes as in *utweopan/weorpan ut* 'out-throw/throw out.' Although locally right-branching, constructions like *look after* are not productive in English, perhaps in part because they are not fully right-branching. Contrast aspectual constructions such as *eat up*, which allow the object to occur between V and particle. These are highly productive (cf. *I googled your husband up*), and pattern with V-object-Adv constructions such as *She planted the rose carefully/in a green pot*. The shift to periphrastic expression is also the context in which the "newer" auxiliaries like *ought to*, *need to*, *dare to* developed; they could not have the properties of the "core" modals (*will*, *shall*, *must*, *may*, *can*) because they developed syntactic auxiliary characteristics after the word order change, while the "core" modals had developed modal properties earlier (Bybee and Dahl 1989:60).[6]

In this book we have focused on the development of the meaning and function of form–meaning pairs, keeping form constant, barring phonological changes.[7] However, as mentioned in Section 1.3.1, synchronic lexicalization is often understood to concern the question of what semantic elements are conflated in a single morpheme or word. In this approach, form is often in focus. Conflation resulting from phonological change (e.g., *set* < *sit* + PAST + CAUSE) has been much discussed in the historical literature (see 2.3.3 on phonologization). However, little historical work appears to have been done on the question of how and why lexical typologies change with respect to the characteristics that Talmy (2000) has called "satellite-framing" (a motion event is conflated with manner of motion in a single lexeme, and the path of motion is expressed in a satellite to the verb such as a particle, or adverbial) and "verb-framing" (a motion event is conflated with path in a single lexeme, and manner is expressed in a satellite to the verb). Clearly, borrowing from one language type to another is an important factor. In Section 1.3.1 we mentioned that while English is strongly satellite-framing (cf. *creep*, *swim*), there is a substantial number of borrowed verb-framed forms (cf. *enter*, *exit*). We also mentioned that Latin tended to be satellite-framing, but Romance languages became verb-framing; however, recently Italian has become more satellite-framing. Strategies for expressing complex concepts are by no means restricted to verbs, however. Blank (2003), citing Koch (1999), discusses tendencies for cognitively and culturally salient concepts to be expressed by simple nouns. For example, the concept "PEAR TREE is verbalized on the basis of PEAR ... and

[6] It should be noted that *be*, *have*, *need*, and *ought* all occur in OE with *to*-infinitives and modal meaning; however, they show few structural properties of auxiliaries until later ME and EModE and are quite rare. According to Visser (1963–1973:1434) *dare to* does not appear before the beginning of the seventeenth century.

[7] This approach is often referred to as "semasiological."

BEECHNUT on the basis of BEECH" (Blank 2003:56). The reason for this tendency is hypothesized to be that pear trees are saliently valued for their fruit, while beeches are saliently valued as trees (56); if change occurs it is predicted to occur in the direction of greater experiential saliency. The problem with this hypothesis is why earlier forms would be less salient than later ones. Blank points out that sometimes there may be major shifts in conceptualization. We are reminded once again that it is not possible to project external factors from the present back onto the past indiscriminately. How and why changes occur with respect to simple vs. complex expression of concepts is a topic of great interest on the intersection of linguistics, cognitive psychology, and culture studies.

The basic research question that a typological perspective suggests is to what extent any particular instance of lexicalization or grammaticalization is harmonic with and hence presumably analogical to an ongoing larger change, to what extent sets of such changes may trigger it.

6.2.4 Discourse types

One of the most difficult questions to answer in linguistics is how change is initiated and why it occurs in one language at a particular time, but not in another language or at another time. Weinreich, Labov, and Herzog have called this the "actuation" question (1968:102). We have stressed that change in general and the particular changes we have discussed here originate in speaker–hearer interaction, and that the acquisition of new uses by speakers of all ages, not exclusively children, is a major factor in change (1.1.2 and 4.1). Any claims about the role of children or adults in change must eventually be tested by long-term studies of several generations of speakers. However, since present-day societal mores and practices regarding the role of children and present-day demographics of literacy (including visual literacy demanded by computers) are so fundamentally different from that of prior ages, any evidence from contemporary and future generations will not be easy to generalize to past eras.

Furthermore, since our data for language change prior to the twentieth century is always written, and often "standard," there is the problem of assessing to what extent changes in the textual record reflect changes in daily speech. One of the best ways to control for this is to compare records of the past with written records from the present, though even then the emergence of new genres and styles inevitably makes exact comparisons impossible.[8] Corpora that include a variety of discourse types that are

[8] See articles in Lindquist and Mair (2004), especially Nevalainen (2004), for discussion of some of the problems as well as advantages of using corpora.

assumed to be relatively representative of spoken language, namely, drama, letters, trials, conversation in fiction, etc. (see, e.g., the *Helsinki Corpus of English Texts, Diachronic Part and Corpus of Early English Correspondence* [*ICAME* 1999; see Rissanen, Kytö, and Palander-Collin 1993]) provide invaluable resources to supplement historical dictionaries such as the *OED*, *Dictionnaire Robert*, and Wartburg (1966), which tend to cite more formal, often literary works.

A number of terms have been developed to refer to different types of discourse to be found in historical corpora (for a useful survey including reasons for changes in terminology over the last twenty years, see Moessner 2001; see also Saukkonen 2003). Currently the following distinctions between "text type," "text style," "genre"/"register," and "text tradition" are widely used (see articles in Raumolin-Brunberg, Navala, Nurmi, and Rissanen 2002). "Text type" refers to texts identified with fairly fine-grained linguistic correlations, ranging from contractions (e.g., *'ll*), to choice of pronouns, subordination, nominalization, hedges, discourse particles, etc. The relevant linguistic criteria were initially developed by Biber (1988), based on patterns emerging from synchronic text types. They were applied to historical texts by Biber and Finegan (1992),[9] and then modified as further text types were developed (e.g., Taavitsainen 2001 on ME recipes). Biber and Finegan (1992) conducted a major study of the linguistic features that characterize different text types over time, such as essays, fiction, and personal letters, using multifactorial analysis of three "dimensions": (a) literate – oral, (b) informational – interactive, and (c) elaborated – situation-dependent. They concluded that there has been a drift from the seventeenth century on toward more "oral" characteristics, despite a slight shift back toward more formal styles in the eighteenth century (1992:695). While "text type" focuses on internal linguistic characteristics of a text, "genre" (often identified with "register") is a term for a larger category that highlights external criteria, such as subject-matter and purpose. Genres change significantly over time; e.g., in ME most recipes gave instructions on how to make medicines, rather than food. The third term, "text tradition," concerns transmission and ways in which new texts build on earlier ones. Finally, "text type style" and "genre type style" are terms used for the "set of linguistic features which characterizes a particular text type or genre" (Moessner 2001:135). It has long been suggested that grammaticalization of connectives and other items with discourse-structure functions may be connected with styles that are less interactive (see Givón 1979). In Section 5.5 we mentioned that free adjuncts and absolute constructions with present participial Adjs have been associated with formal styles (see Aertsen 1992).

[9] They use the term "genre" and "register" in this paper.

Olofsson points out that the participial Prep *following* occurs predominantly in "informative writings" and non-fiction generally (1990:28).

Among external features correlated with text types that have received much attention are gender and community. Wright discusses the development of the "experiential (modal) progressive" in intimate, private letters written by women during the eighteenth century; an example given by Wright (1994:471) is *Daughter, daughter, don't call names. You* ***are*** *always* ***abusing*** *my pleasures* (1757, Lady Mary Wortley Montague). In Wright (1995) she shows how this construction came to be construed as a part of a literary style in the novels of Aphra Behn.

Conceptualizing community partly in terms of values held for certain kinds of text traditions, Cuenca shows how different values regarding academic writing in English and Spanish are correlated with significantly different degrees of grammaticalization of reformulation markers, i.e., markers signaling that what is to follow is a re-elaboration of "an idea in order to be more specific and 'facilitate the hearer's understanding of the original'" (2003:1071; citing Blakemore 1993:107). Reformulation is conceptually a paraphrase introduced by 'that is to say.' English academic writers, Cuenca (2003) argues, value a rhetorical style that is simple and linear (for the development of this style, see Thomas and Turner 1994; Adamson 1999 on the parallel "drive to perspicuity" in literary language). Spanish academic writers, by contrast, value "content and *variatio*" (Cuenca 2003:1089).[10] There were significantly fewer reformulation markers in the English than the Spanish texts Cuenca studied (with Catalan showing intermediate numbers). Furthermore, if a connective was used, the English texts showed a preference for relatively simple, fixed (univerbated) connectives, the Spanish for more complex connectives, and especially for free phrases. Simple connectives include English *or*, *namely*, and *i.e.* (borrowed from Lat. *id est* 'that is'; here an acronym has become grammaticalized), Sp. *o* 'or.' More complex forms that are fixed are *that is (to say)*, *in other words*, Sp. *es decir* 'is to say'. But Spanish writers favor variable, relatively free, expressions, e.g., *en/con otras palabras* 'in/with other words,' *dicho con otras palabras* 'said with other words,' *formulado en otros términos* 'formulated in other terms,' etc. (Cuenca 2003:1081). Forms of *that is* are especially interesting because (like the complementizer *that*), they originate in an anaphoric deictic that points back (1078); when they became connectives they came to be forward-pointing ("cataphoric").

Community may also be defined in terms of social group (membership, status, educational opportunities in it, etc.) (see Nevalainen and Raumolin-

[10] Cuenca associates the distinction with contrasts made by Kaplan (1966) and Clyne (1994) between "formal-oriented cultures" that favor "writer-responsible style" and "content-oriented cultures" that favor "reader-responsible styles" (2003:1085).

Brunberg 2003; Bergs 2005) and networks of interaction such as coalitions or patron-client interactions, where norms of politeness, paths of influence, and shaping of identity can be correlated with linguistic choices (see, e.g., Milroy 1987 [1980]; Fitzmaurice 2000). In a study of the literary circle of Southey and Coleridge as evidenced primarily by their letters of the 1790s, Pratt and Denison (2000) suggest that the discursive practices of these writers contributed to the spread of the newly developing passive construction (see, e.g., *the speech ... was being debated* [1772 Harris, Sr., *Lett. 1st Earl Malmesbury* I, 430] which was replacing *the speech was debating*). While the progressive passive may have been a regionalism of South West England, the authors suggest that the Southey–Coleridge circle seized on it "in a kind of radical experimentation," a possibly unconscious act of linguistic symbolism through which they established their group identity (417).

Given the enormous increase of available corpora not only in English but in other languages, the possibilities are expanding exponentially for studying aspects of lexicalization as well as grammaticalization in the context of the text types now available, and of tracing the actualization or spread of certain changes through them. One could, for example, compare the spread of prepositional and phrasal verbs across text types and genres. Further, there is the opportunity to identify additional linguistic correlations with specific text types and genres.

6.2.5 Language contact

Degree and type of language contact are closely connected with the text type and community. Formal styles and certain text types may be directly tied to language contact. For example, the ME "romance" was of French origin; many legal documents were in Latin or French in the ME period (see Görlach 1999 on correlations between genre type and language type, specifically English dialects, Latin, French, and Standard English from 700 to 1700). Until at least the Renaissance, many texts in Europe were close or loose translations of Latin or, sometimes, French texts. Certainly Latin and French were used as prose models throughout ME and EModE, and this presumably had an influence on productivity as well as idiomaticization of certain forms and form classes. When large numbers of words or phrases are borrowed (as was the case in the EModE periods, see Nevalainen 1999) or a particular class is borrowed (e.g., most Psych-Vs were borrowed in English), the question naturally arises to what extent the morphosyntax and semantics/pragmatics of the source language carry over into and affect native developments. In particular, did the prepositional uses of the Psych-V in French contribute to the division of labor between participial Adjs and Preps? In trying to answer such questions, as well as other questions related to a particular structure, it is important to investigate the whole set of relevant constructions or of a particular grammatical class. Many

researchers from Mossé (1938) on have asked whether Celtic influenced the development of *be going to*, but little systematic comparison of the various relevant constructions has been conducted; in English they include gerunds, gerundives, and participial Adjs, in Celtic they might be a similar or a different set. Without such study the question remains unanswered.

Discourse markers are a category that is frequently borrowed. If borrowed, they may serve particular discourse functions; for example, Solomon shows that in Yucatan Mayan narrative native connectives like *ká* '(and) then' are used to signal succession within narrative episodes, but borrowed connectives like Sp. *entonses* 'then, so' to signal succession between narrative episodes and orientation apart from action (1995:287).

Pidgins and creoles have often been cited as a highly interesting domain for the study of borrowed lexicon and the development of word formation strategies (see, e.g., Mühlhäusler 1979, 1997), but to date not much has been done on lexicalization as defined here. Evidence for the (perceived as well as actual) coalescence and fixing of various phrases is frequently cited as characteristic of pidgins, e.g., Tok Pisin *atink* 'maybe' < *I think*, *baimbai* 'later' < *by and by*. Especially notable are changes in class category, as in Beach-la-Mar *hariap* 'quickly' < *hurry up*, *tasol* 'but' < *that's all*, or change in semantics, as in Tok Pisin *bagarap* 'tired, ruined' < *buggered up* (Mühlhäusler 1997:154–155). It would appear that when pidgins stabilize and compounds are developed (e.g., Tok Pisin and Samoan Plantation Pidgin English *nuboi* 'freshly indentured laborer' < *new boy*; *olboi* 'laborer having served a three-year term' < *old boy*, *waitman* 'European' < *whiteman*; [Mühlhäusler 1997:178]), the specific meanings of these items are derived not by relatively arbitrary semantic reanalysis of the compounds with English phrases, but language-internally by idiomatization, i.e., lexicalization (Hancock 1980). In creole studies most recent attention has been paid not to the lexicon but to grammaticalization (an early study focused on Tok Pisin *baimbai* > *bai* 'future' [Sankoff and Laberge 1980 [1976]]) and the extent to which it may have started prior to contact or during the creolization process, i.e., was external or internal (see, e.g., Baker and Syea 1996; DeGraff 1999). Examples pertaining to lexicalization include the incorporation in French Creoles of the definite article and sometimes the partitive genitive determiner into the N stem (a type of phonogenesis from the perspective of the lexifier, French) as in Haitian Creole (e.g., *lank* 'anchor' < Fr. *l'ancre* 'the anchor', *dlo* 'water' < Fr. *de l'eau* 'of the water'). Clearly, here as elsewhere, the field is wide open for further exploration.

References

Aarts, Bas. 2004. Modelling linguistic gradience. *Studies in Language* 28:1–49.

Adamson, Sylvia. 1999. Literary language. In Lass, ed., 539–653.

Adamson, Sylvia. 2000. A lovely little example: Word order options and category shift in the premodifying string. In Fischer, Rosenbach, and Stein, eds., 39–66.

Aertsen, Henk. 1992. Chaucer's *Boece*: A syntactic and lexical analysis. In Rissanen, Ihalainen, Nevalainen, and Taavitsainen, eds., 671–687.

Aijmer, Karin. 1996a. *Conversational Routines in English: Convention and Creativity*. (Studies in Language and Linguistics.) London and New York: Longman.

Aijmer, Karin. 1996b. *I think* – an English modal particle. In Toril Swan and Olaf Jansen Westvik, eds., *Modality in Germanic Languages: Historical and Comparative Perspectives*, 1–47. (Trends in Linguistics, Studies and Monographs, 99.) Berlin: Mouton de Gruyter.

Aijmer, Karin. 2004. The semantic path from modality to aspect: *Be able to* in a cross-linguistic perspective. In Lindquist and Mair, eds., 57–78.

Akimoto, Minoji. 1989. *A Study of Verb-Nominal Structures in English*. Tokyo: Shinozaki Shorin.

Algeo, John. 1995. Having a look at the expanded predicate. In Bas Aarts and Charles F. Meyer, eds., *The Verb in Contemporary English: Theory and Description*, 203–217. Cambridge, UK: Cambridge University Press.

Allen, Andrew S. 1995. Regrammaticalization and degrammaticalization of the inchoative suffix. In Andersen, ed., 1–8.

Allen, Cynthia L. 1995. On doing as you please. In Andreas H. Jucker, ed., *Historical Pragmatics: Pragmatic Developments in the History of English*, 275–308. (Pragmatics & Beyond, New Series, 35.) Amsterdam and Philadelphia: John Benjamins.

Allen, Cynthia L. 1997. The origins of the "group genitive" in English. *Transactions of the Philological Society* 95:111–131.

Allen, Cynthia L. 2003. Deflexion and the development of the genitive in English. *English Language and Linguistics* 7:1–28.

The American Heritage Dictionary of the English Language. 2000. Boston and New York: Houghton Mifflin, 4th edn.

Andersen, Henning. 1987. From auxiliary to desinence. In Martin Harris and Paolo Ramat, eds., *Historical Development of Auxiliaries*, 21–51. (Trends in Linguistics; Studies and Monographs, 35.) Berlin, New York, and Amsterdam: Mouton de Gruyter.

Andersen, Henning, ed. 1995. *Historical Linguistics 1993: Selected Papers from the 11th International Conference on Historical Linguistics, Los Angeles, 16–20 August 1993*, 33–47. (Current Issues in Linguistic Theory, 124.) Amsterdam and Philadelphia: John Benjamins.

Andersen, Henning. 2001. Actualization and the (uni)directionality of change. In Henning Andersen, ed., *Actualization: Linguistic Change in Progress*, 225–248. (Current Issues in Linguistic Theory, 219.) Amsterdam and Philadelphia: John Benjamins.

Anderson, Stephen R. 1985a. Typological distinctions in word formation. In Shopen, ed., Vol. III, 3–56.

Anderson, Stephen R. 1985b. Inflectional morphology. In Shopen, ed., Vol. III, 150–201.

Anderson, Stephen R. 1992. *A-Morphous Morphology*. (Cambridge Studies in Linguistics, 62.) Cambridge, UK: Cambridge University Press.

Anderson, Stephen R. 1993. Wackernagel's revenge: Clitics, morphology, and the syntax of second position. *Language* 69:68–98.

Anttila, Raimo. 1989 [1972]. *Historical and Comparative Linguistics*. (Current Issues in Linguistic Theory, 6.) Amsterdam and Philadelphia: John Benjamins, 2nd edn.

Arnovick, Leslie K. 1999. *Diachronic Pragmatics: Seven Case Studies in English Illocutionary Development*. (Pragmatics & Beyond, New Series, 68.) Amsterdam and Philadelphia: John Benjamins.

Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. (Linguistic Inquiry Monograph, 1.) Cambridge, MA: MIT Press.

Aronoff, Mark. 1980. Contextuals. *Language* 56:744–758.

Asher, R. E. and J. M. Y. Simpson, eds. 1994. *The Encyclopedia of Language and Linguistics*. Oxford: Pergamon Press, 4 vols.

Aston, Guy and Lou Burnard. 1998. *The BNC Handbook Exploring the British National Corpus with SARA*. (Edinburgh Textbooks in Empirical Linguistics.). Edinburgh: Edinburgh University Press.

Axmaker, Shelly, Annie Jaisser, and Helen Singmaster, eds. 1988. *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society. General Session and Parasession on Grammaticalization*. Berkeley, CA: Berkeley Linguistics Society.

Baayen, R. Harald. 2003. Probabalistic approaches to morphology. In Rens Bod, Jennifer Hay, and Stefanie Jannedy, eds., *Probabilistic Linguistics*, 229–287. Cambridge, MA: MIT Press, A Bradford Book.

Baayen, R. Harald and Antoinette Renouf. 1996. Chronicling the *Times*: Productive lexical innovations in an English newspaper. *Language* 72:69–96.

Baker, Philip and Anand Syea, eds. 1996. *Changing Meanings, Changing Functions: Papers Relating to Grammaticalization in Contact Languages*. (Westminster Creolistics Series, 2.) London: University of Westminster Press.

Barcelona, Antonio, ed. 2000. *Metaphor and Metonymy at the Crossroads: A Cognitive Perspective*. (Topics in English Linguistics, 30.) Berlin and New York: Mouton de Gruyter.

Bauer, Laurie. 1978. On lexicalization (neither a lexicalist nor a transformationalist be). *Archivum Linguisticum* 9:3–14.

Bauer, Laurie. 1983. *English Word Formation*. (Cambridge Textbooks in Linguistics.) Cambridge, UK: Cambridge University Press.

Bauer, Laurie. 1988. *Introducing Linguistic Morphology*. Edinburgh: Edinburgh University Press.

Bauer, Laurie. 1992. Lexicalization and level ordering. *Linguistics* 30:561–568.

Bauer, Laurie. 1994. Productivity. In Asher and Simpson, eds., Vol. VI, 3354–3357.

Benveniste, Emile. 1971a [1958]. Delocutive verbs. In Benveniste 1971c, 239–246. (Orig. publ. in A. G. Hatcher and K. L. Selig, eds., *Studia Philologica et Litteraria in Honorem L. Spitzer*, 57–63, Bern, 1958.)

Benveniste, Émile. 1971b [1958]. Subjectivity in language. In Benveniste 1971c [1966], 223–230. (Originally publ. in *Journal de psychologie* 55: 267ff., 1958.)

Benveniste, Emile. 1971c [1966]. *Problems in General Linguistics*, trans. Mary Elizabeth Meek. (Miami Linguistics Series, 8.) Coral Gables, FL: University of Miami Press. (Orig. publ. as *Problèmes de linguistique générale*. Paris: Editions Gallimard, 1966.)

Bergs, Alexander. 2005. *Social Networks and Historical Sociolinguistics: Studies in Morphosyntactic Variation in the Paston Letters (1421–1503)*. (Topics in English Linguistics, 51.) Berlin and New York: Mouton de Gruyter.

Berman, Ruth A. and Dan I. Slobin. 1994. *Relating Events in Narrative: A Crosslinguistic Developmental Study*. Hillsdale, NJ: Lawrence Erlbaum.

Beths, Frank. 1999. The history of *dare* and the status of unidirectionality. *Linguistics* 37:1069–1110.

Bhat, D. N. S. 2000. Word classes and sentential functions. In Vogel and Comrie, eds., 47–63.

Biber, Douglas. 1988. *Variation Across Speech and Writing*. Cambridge, UK: Cambridge University Press.

Biber, Douglas and Edward Finegan. 1992. The linguistic evolution of five written and speech-based English genres from the 17th to the 20th centuries. In Rissanen et al., eds., 688–704.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad, and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Harlow: Longman.

Bierwisch, Manfred. 1970. On classifying semantic features. In Manfred Bierwisch and Karl Erich Heidolph, eds., *Progress in Linguistics: A Collection of Papers*, 27–50. (Janua Linguarum, Series Maior, 43.) The Hague: Mouton.

Blakemore, Diane. 1987. *Semantic Constraints on Relevance*. Oxford: Blackwell.

Blakemore, Diane. 1993. The relevance of reformulations. *Language and Literature* 2:101–120.

Blank, Andreas. 2001. Pathways of lexicalization. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher, and Wolfgang Raible, eds., *Language Typology and Language Universals*, Vol. II, 1596–1608. (Handbücher zur Sprach- und Kommunikationswissenschaft, 20.2.) Berlin and New York: Walter de Gruyter.

Blank, Andreas. 2003. Words and concepts in time: Towards diachronic cognitive onomasiology. In Regine Eckardt, Klaus von Heusinger, and Christoph

Schwarze, eds., *Words in Time: Diachronic Semantics from Different Points of View*, 38–65. (Trends in Linguistics. Studies and Monographs, 143.). Berlin and New York: Mouton de Gruyter.

Blank, Andreas and Peter Koch, eds. 1999. *Historical Semantics and Cognition*. (Cognitive Linguistics Research, 13.) Berlin and New York: Mouton de Gruyter.

Blom, Corrien and Geert Booij. 2003. The diachrony of complex predicates in Dutch: A case study in grammaticalization. *Acta Linguistica Hungarica* 50:61–91.

Bloomfield, Leonard. 1933. *Language*. New York: Holt, Rinehart and Winston.

Boersma, Paul and Bruce Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32:45–86.

Bolinger, Dwight. 1976. Meaning and memory. *Forum Linguisticum* 1:1–14.

Booij, Gert. 2002. *The Morphology of Dutch*. Oxford and New York: Oxford University Press.

Bresnan, Joan. 2001. *Lexical-Functional Syntax*. (Blackwell Textbooks in Linguistics, 16.) Oxford: Blackwell.

Bresnan, Joan, Shipra Dingare, and Christopher D. Manning. 2002. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In *Proceedings of the LFG01 Conference, Hong Kong*, 13–82. Stanford: CSLI Publications and Chicago: University of Chicago Press.

Brinton, Laurel J. 1988. *The Development of English Aspectual Systems: Aspectualizers and Post-verbal Particles*. (Cambridge Studies in Linguistics, 49.) Cambridge, UK: Cambridge University Press.

Brinton, Laurel J. 1996a. Attitudes toward increasing segmentalization: Complex and phrasal verbs in English. *Journal of English Linguistics* 24:186–205.

Brinton, Laurel J. 1996b. *Pragmatic Markers in English: Grammaticalization and Discourse Functions*. (Topics in English Linguistics, 19.) Berlin and New York: Mouton de Gruyter.

Brinton, Laurel J. 2001. From matrix clause to pragmatic marker: The history of *look*-forms. *Journal of Historical Pragmatics* 2:177–199.

Brinton, Laurel J., ed. 2001. *Historical Linguistics 1999. Selected Papers from the 14th International Conference on Historical Linguistics, Vancouver, 9–13 August 1999*. (Current Issues in Linguistic Theory, 215.) Amsterdam and Philadelphia: John Benjamins.

Brinton, Laurel J. 2002. Grammaticalization versus lexicalization reconsidered: On the "late" use of temporal adverbs. In Teresa Fanago, María José López-Couso, and Javier Pérez-Guerra, eds., *English Historical Syntax and Morphology: Selected Papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000*, 67–97. (Current Issues in Linguistic Theory, 223.) Amsterdam and Philadelphia: John Benjamins.

Brinton, Laurel J. 2004. Subject clitics in English: A case of degrammaticalization? In Lindquist and Mair, eds., 227–256.

Brinton, Laurel J. In press. The development of *I mean*: Implications for the study of historical pragmatics. In Susan Fitzmaurice and Irma Taavitsainen, eds., *Methods of Historical Pragmatics: Recovering Speaker Meaning and Reader Inference*. Berlin and New York: Mouton de Gruyter.

Brinton, Laurel J. and Minoji Akimoto, eds. 1999. *Collocational and Idiomatic Aspects of Composite Predicates in the History of English*. (Studies in Language Companion Series, 47.) Amsterdam and Philadelphia: John Benjamins.

Brinton, Laurel J. and Dieter Stein. 1995. Functional renewal. In Andersen, ed., 33–47.

Bruyn, Adrienne. 1996. On identifying instances of grammaticalization in Creole languages. In Baker and Syea, eds., 29–46.

Burridge, Kate. 1998. From modal auxiliary to lexical verb: The curious case of Pennsylvania German *wotte*. In Richard M. Hogg and Linda van Bergen, eds., *Historical Linguistics 1995, Selected Papers from the 12th International Conference on Historical Linguistics, Vol. II. Germanic Linguistics*, 19–31. (Current Issues in Linguistic Theory, 162.) Amsterdam and Philadelphia: John Benjamins.

Burridge, Kate. 2002. Changes within Pennsylvania German grammar as enactments of Anabaptist world view. In N. J. Enfield, ed., *Ethnosyntax: Explorations in Grammar and Culture*, 207–230. Oxford: Oxford University Press.

Bussmann, Hadumod. 1996. *Routledge Dictionary of Language and Linguistics*, trans. and eds. Gregory Trauth and Kerstin Kazzazi. London and New York: Routledge.

Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. (Typological Studies in Language, 9.) Amsterdam and Philadelphia: John Benjamins.

Bybee, Joan L. 1988. Morphology as lexical organization. In Hammond and Noonan, eds., 119–141.

Bybee, Joan L. 1994. The grammaticization of zero: Asymmetries in tense and aspect systems. In Pagliuca, ed., 235–252.

Bybee, Joan L. 2003. Mechanisms of change in grammaticization: The role of frequency. In Joseph and Janda, eds., 602–623.

Bybee, Joan L. and Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world. *Studies in Language* 13:51–103.

Bybee, Joan L. and Paul Hopper, eds. 2001. *Frequency and the Emergence of Linguistic Structure*. (Typological Studies in Language, 45.) Amsterdam and Philadelphia: John Benjamins.

Bybee, Joan L. and William Pagliuca. 1987. The evolution of future meaning. In Anna Giacalone Ramat, Onofrio Carruba, and Giuliano Bernini, eds., *Papers from the 7th International Conference on Historical Linguistics*, 109–122. (Current Issues in Linguistic Theory, 48.) Amsterdam and Philadelphia: John Benjamins.

Bybee, Joan L., Revere Perkins, and William Pagliuca. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.

Campbell, Alistair. 1959. *Old English Grammar*. Oxford: Clarendon Press.

Campbell, Lyle. 1991. Some grammaticalization changes in Estonian and their implications. In Traugott and Heine, eds., Vol. I, 285–299.

Campbell, Lyle. 2001a. What's wrong with grammaticalization? In Campbell, ed., 113–161.

Campbell, Lyle, ed. 2001b. Grammaticalization: A critical assessment. *Language Sciences* 23, Numbers 2–3.

Campbell, Lyle and Marianne Mithun, eds. 1979. *The Languages of Native America: Historical and Comparative Assessment*. Austin and London: University of Texas Press.

Cannon, Garland. 1987. *Historical Change and English Word-Formation*. New York: Peter Lang.

Childs, G. Tucker. 1994. African ideophones. In Leanne Hinton, Johanna Nichols, and John J. Ohala, eds., *Sound Symbolism*, 178–204. Cambridge, UK: Cambridge University Press.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1988. *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, MA and London: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. (Oxford Studies in Comparative Syntax.) New York and Oxford: Oxford University Press.

Claridge, Claudia. 2000. *Multi-Word Verbs in Early Modern English: A Corpus-Based Approach* (Language and Computers: Studies in Practical Linguistics, 32.) Amsterdam and Atlanta, GA: Rodopi.

Clark, Brady Z. 2004. Early English clause structure change and stochastic optimality theory setting. In Anne Curzan and Kimberley Emmons, eds., *Studies in the History of the English Language II: Unfolding Conversations*, 343–369. (Topics in English Linguistics, 45.) Berlin and New York: Mouton de Gruyter.

Clark, Eve V. and Herbert H. Clark. 1979. When nouns surface as verbs. *Language* 55:767–811.

Claudi, Ulrike. 1994. Word order change as category change: The Mande case. In Pagliuca, ed., 191–231.

Claudi, Ulrike and Bernd Heine. 1986. On the metaphorical basis of grammar. *Studies in Language* 10:297–335.

Clyne, Michael G. 1994. *Inter-Cultural Communication at Work: Cultural Values in Discourse*. Cambridge, UK: Cambridge University Press.

Comrie, Bernard. 1976. *Aspect: An Introduction to the Study of Verbal Aspect and Related Problems*. (Cambridge Textbooks in Linguistics, 2.) Cambridge, UK: Cambridge University Press.

Corbett, Greville D. 1987. The morphology/syntax interface. *Language* 63:299–345.

Corpus of Early English Correspondence. 1998. Compiled by Terttu Nevalainen, Helena Raumolin-Brunberg, et al. University of Helsinki. Available in *ICAME*.

Cowie, Anthony Paul and Ronald Mackin. 1993 [1975]. *Oxford Dictionary of Phrasal Verbs*. London: Oxford University Press. (Previously publ. as *Oxford Dictionary of Current Idiomatic English. Vol. I: Verbs with Prepositions & Particles*.)

Cowie, Claire. 1995. Grammaticalization and the snowball effect. *Language and Communication* 15:181–193.

Croft, William. 1991. *Syntactic Categories and Grammatical Relations: The Cognitive Organization of Information*. Chicago: University of Chicago Press.

Croft, William. 1995. Autonomy and functionalist linguistics. *Language* 71:490–532.
Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach*. Harlow, Essex: Pearson Education.
Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. New York: Oxford University Press.
Croft, William, Keith Denning, and Suzanne Kemmer, eds. 1990. *Studies in Typology and Diachrony: Papers Presented to Joseph H. Greenberg on his 75th Birthday*. (Typological Studies in Language, 20.) Amsterdam and Philadelphia: John Benjamins.
Cuenca, Maria-Josep. 2003. Two ways to reformulate: A contrastive analysis of reformulation markers. *Journal of Pragmatics* 35:1069–1093.
Culicover, Peter W. and Andrzej Nowak. 2003. *Dynamical Grammar: Minimalism, Acquisition, and Change*. (Foundations of Syntax, 2.) New York: Oxford University Press.
Dalton-Puffer, Christiane. 1996. *The French Influence on Middle English Morphology: A Corpus-Based Study of Derivation*. (Topics in English Linguistics, 20.) Berlin and New York: Mouton de Gruyter.
Darnell, Michael, Edith Moravcsik, Frederick Newmeyer, Michael Noonan, and Kathleen Wheatley, eds. 1999. *Functionalism and Formalism in Linguistics*. (Studies in Language, Companion Series, 42.) Amsterdam and Philadelphia: John Benjamins, 2 vols.
DeGraff, Michel, ed. 1999. *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, MA: MIT Press.
de Groodt, Sarah. 2003. Unidirectionality in grammaticalization: The development of concessive subordinating conjunctions with *ob-* in German. *Folia Linguistica Historica* XXIV:193–204.
Denison, David. 1985a. The origins of completive *up* in English. *Neuphilologische Mitteilungen* 86:37–61.
Denison, David. 1985b. Why Old English had no prepositional passive. *English Studies* 66:189–204.
Denison, David. 1993. *English Historical Syntax: Verbal Constructions*. (Longman Linguistics Library.) London and New York: Longman.
Denison, David. 2001. Gradience and linguistic change. In Brinton, ed., 119–144.
Detges, Ulrich and Richard Waltereit. 2002. Reanalysis vs. grammaticalization: A semantic-pragmatic account of functional change in grammar. *Zeitschrift für Sprachwissenschaft* 21:151–195.
Dictionary of Old English Corpus. 2000. Antonette di Paolo, ed. Toronto: Centre for Medieval Studies, University of Toronto. (http://ets.umdl.umich.edu/o/oec/.)
Dictionnaire Robert. 1992. *Dictionnaire historique de la langue Française*. Alain Rey, ed. Paris: Dictionnaires le Robert.
Dixon, Robert M. 1982. *Where have all the Adjectives Gone? And other Essays in Semantics and Syntax*. (Janua Linguarum, Series Maior, 107.) Berlin: Mouton.
Donner, Morton. 1991. Adverb form in Middle English. *English Studies* 72:1–11.
Dowty, David R. 1991. Thematic proto-roles and argument selection. *Language* 67:547–619.
Doyle, Aidan. 2002. Yesterday's affixes as today's clitics: A case study in grammaticalization. In Wischer and Diewald, eds., 67–81.

Dressler, Wolfgang U. 1989. Prototypical differences between inflection and derivation. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42:3–10.

Dryer, Matthew. 1992. The Greenbergian word order correlations. *Language* 68:81–138.

Eckert, Penelope. 2000. *Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten High*. (Language in Society, 27.) Malden, MA: Blackwell.

Enfield, N. J. 2003. *Linguistic Epidemiology: Semantics and Grammar of Language Contact in Mainland Southeast Asia*. (RoutledgeCurzon Asian Linguistics Series.) London and New York: Routledge.

Enger, Hans-Olaf. 2002. The story of Scandinavian *-s(t)* retold: Grammaticalising a clitic to a derivational affix. *Folia Linguistica Historica* XXIII:79–105.

Erman, Britt and Ulla-Britt Kotsinas. 1993. Pragmaticalization: The case of *ba'* and *you know*. *Studier i Modernspråkvetenskap*, 76–93. (Acta Universitatis Stockholmiensis. Stockholm Studies in Modern Philology. New Series, 10.) Stockholm: Almqvist and Wiksell.

Ernst, Thomas. 2002. *The Syntax of Adjuncts*. (Cambridge Studies in Linguistics, 96.) Cambridge, UK: Cambridge University Press.

Evans, Nicholas D. 1995. *A Grammar of Kayardild; with Historical-Comparative Notes on Tangkic*. (Mouton Grammar Library, 15.) Berlin and New York: Mouton de Gruyter.

Evans, Nicholas D. and David Wilkins. 2000. In the mind's ear: The semantic extensions of perception. *Language* 76:546–592.

Fillmore, Charles J. 1970. Types of lexical information. In F. Kiefer, ed., *Studies in Syntax and Semantics*, 109–137. (Foundations of Language, Supplementary Series, 10.) Dordrecht, Holland: D. Reidel.

Fillmore, Charles J., Paul Kay., Laura A. Michaelis, and Ivan A. Sag. 2003. *Construction Grammar*. Stanford: CSLI Publications and Chicago: University of Chicago Press.

Fillmore, Charles J., Paul Kay, and Mary Catherine O'Connor. 1988. Regularity and idiomaticity in grammatical constructions: The case of *let alone*. *Language* 64:501–538.

Fischer, Olga, Anette Rosenbach, and Dieter Stein, eds. 2000. *Pathways of Change: Grammaticalization in English*. (Studies in Language, Companion Series, 53.) Amsterdam and Philadelphia: John Benjamins.

Fitzmaurice, Susan. 2000. Coalitions and the investigation of social influence in linguistic history. *European Journal of English Studies* 4:265–276.

Fleischman, Suzanne. 1976. *Cultural and Linguistics Factors in Word Formation: An Integrated Approach to the Development of the Suffix* -age. (University of California Publications in Linguistics, 86.) Berkeley and Los Angeles: University of California Press.

Fleischman, Suzanne. 1982. *The Future in Thought and Language*. (Cambridge Studies in Linguistics, 36.) Cambridge, UK: Cambridge University Press.

Fraser, Bruce. 1996. Pragmatic markers. *Pragmatics: Quarterly Publication of the International Pragmatics Association* 6:167–190.

Gabelentz, Georg von der. 1901 [1891]. *Die Sprachwissenschaft, ihre Aufgaben, Methoden und bisherige Ergebnisse*. Leipzig: Weigel.

Gelderen, Elly van. 1996. The reanalysis of grammatical prepositions in Middle English. *Studia Linguistica* 50:106–124.
Gelderen, Elly van. 1997. *Verbal Agreement and the Grammar behind its "Breakdown": Minimalist Feature Checking*. (Linguistische Arbeiten, 364.) Tübingen: Niemeyer.
Giacalone Ramat, Anna. 1998. Testing the boundaries of grammaticalization. In Giacalone Ramat and Hopper, eds., 107–127.
Giacalone Ramat, Anna and Paul J. Hopper, eds. 1998. *The Limits of Grammaticalization*. (Typological Studies in Language, 37.) Amsterdam and Philadelphia: John Benjamins.
Gildea, Spike, ed. 1999. *Reconstructing Grammar: Comparative Linguistics and Grammaticalization*. (Typological Studies in English, 43.) Amsterdam and Philadelphia: John Benjamins.
Givón, Talmy. 1971. Historical syntax and synchronic morphology: An archaeologist's fieldtrip. In *Papers from the Seventh Regional Meeting, Chicago Linguistic Society*, 394–415. Chicago: Chicago Linguistic Society.
Givón, Talmy. 1975. Serial verbs and syntactic change: Niger-Congo. In Charles N. Li, ed., *Word Order and Word Order Change*, 47–112. New York: Academic Press.
Givón, Talmy. 1979. *On Understanding Grammar*. New York: Academic Press.
Givón, T. 1982. Tense-aspect-modality: The creole prototype and beyond. In Paul J. Hopper, ed., *Tense-Aspect: Between Semantics & Pragmatics*, 115–162. (Typological Studies in Language, 1.) Amsterdam and Philadelphia: John Benjamins.
Givón, T. 1991. The evolution of dependent clause morpho-syntax in Biblical Hebrew. In Traugott and Heine, eds., Vol. II, 257–310.
Goldberg, Adele. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
Görlach, Manfred. 1991. *Introduction to Early Modern English*. Cambridge: UK: Cambridge University Press.
Görlach, Manfred. 1999. Regional and social variation. In Lass, ed., 459–538.
Greenberg, Joseph H. 1991. The last stages of grammatical elements: Contractive and expansive desemanticization. In Traugott and Heine, eds., Vol. I, 301–314.
Greenberg, Joseph H., Charles A. Ferguson, and Edith Moravcsik, eds. 1978. *Universals of Human Language*. Stanford: Stanford University Press, 4 vols.
Gruber, Jeffrey S. 1976. *Lexical Structures in Syntax and Semantics*. Amsterdam: North-Holland.
Guilbert, Louis. 1975. *La créativité lexicale*. Paris: Larousse.
Hagège, Claude. 1993. *The Language Builder: An Essay on the Human Signature in Linguistic Morphogenesis*. (Current Issues in Linguistic Theory, 94.) Amsterdam and Philadelphia: John Benjamins.
Haiman, John. 1980. The iconicity of grammar. *Language* 56:515–540.
Haiman, John. 1983. Iconic and economic motivation. *Language* 59:781–819.
Haiman, John. 1994. Ritualization and the development of language. In Pagliuca, ed., 3–28.
Hale, Mark. 1998. Diachronic syntax. *Syntax* 1:1–18.
Hammond, Michael and Michael Noonan, eds. 1988. *Theoretical Morphology: Approaches in Modern Linguistics*. San Diego: Academic Press.

Hancock, Ian F. 1980. Lexical expansion in Creole languages. In Arnold Highfield and Albert Valdman, eds., *Historicity and Variation in Creole Studies*, 63–88. Ann Arbor: Karoma.

Harris, Alice C. and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. (Cambridge Studies in Linguistics, 74.) Cambridge, UK: Cambridge University Press.

Haspelmath, Martin. 1992. Grammaticalization theory and heads in morphology. In Mark Aronoff, ed., *Morphology Now*, 69–82. Albany: State University of New York Press.

Haspelmath, Martin. 1998. Does grammaticalization need reanalysis? *Studies in Language* 22:315–351.

Haspelmath, Martin. 1999a. Are there principles of grammatical change? *Journal of Linguistics* 35:579–595. (Review of Lightfoot 1999.)

Haspelmath, Martin. 1999b. Why is grammaticalization irreversible? *Linguistics* 37:1043–1068.

Haspelmath, Martin. 2000a. Why can't we talk to each other? *Lingua* 110:235–255.

Haspelmath, Martin. 2000b. The relevance of extravagance: A reply to Bart Geurts. *Linguistics* 38:789–798.

Haspelmath, Martin. 2002. *Understanding Morphology*. (Understanding Language Series.) London: Arnold and New York: Oxford University Press.

Haspelmath, Martin. 2004. On directionality in language change with particular reference to grammaticalization. In Olga Fischer, Muriel Norde, and Harry Perridon, eds., *Up and Down the Cline – The Nature of Grammaticalization*, 17–44. (Typological Studies in Language, 59.) Amsterdam and Philadelphia: John Benjamins.

HCET. Helsinki Corpus of English Texts. Diachronic Part. 1993. Compiled by Matti Rissanen, Merja Kytö, Minna Palander-Collin, et al. Available in *ICAME*.

Heine, Bernd. 2002. On the role of context in grammaticalization. In Wischer and Diewald, eds., 83–101.

Heine, Bernd. 2003a. Grammaticalization. In Joseph and Janda, eds., 575–601.

Heine, Bernd. 2003b. On degrammaticalization. In Barry J. Blake and Kate Burridge, eds., *Historical Linguistics 2001. Selected Papers from the 15th International Conference on Historical Linguistics, Melbourne, 13–17 August 2001*, 165–179. (Current Issues in Linguistic Theory, 237.) Amsterdam and Philadelphia: John Benjamins.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A Conceptual Framework*. Chicago: University of Chicago Press.

Heine, Bernd, Tom Güldemann, Christa Kilian-Katz, Donald A. Lessau, Heinz Roberg, Mathias Schladt, and Thomas Stolz. 1993. *Conceptual Shift: A Lexicon of Grammaticalization Processes in African Languages*. (Afrikanistische Arbeitspapiere, 34/35.) University of Cologne.

Heine, Bernd and Tania Kuteva. 2002. *World Lexicon of Grammaticalization*. Cambridge, UK: Cambridge University Press.

Heine, Bernd and Mechthild Reh. 1984. *Grammaticalization and Reanalysis in African Languages*. Hamburg: Helmut Buske.

Helftoft, Lars. 1996. Paradigmatic structure, word order and grammaticalization. In Elisabeth Engberg-Pedersen, Michael Fortescue, Peter Harder, Lars Helftoft,

and Lisbeth Falster Jakobsen, eds., *Content, Expression and Structure: Studies in Danish Functional Grammar*, 469–494. (Studies in Language Companion Series, 29.) Amsterdam and Philadelphia: John Benjamins.

Hiltunen, Risto. 1983. *The Decline of the Prefixes and the Beginnings of the English Phrasal Verb: The Evidence from some Old and Early Middle English Texts*. (Annales Universitatis Turkensis, Ser. B., 160.) Turku: Turun Yliopisto.

Hiltunen, Risto. 1994. On phrasal verbs in Early Modern English. In Kastovsky, ed., 129–140.

Himmelmann, Nikolaus P. 2004. Lexicalization and grammaticalization: Opposite or orthogonal? In Walter Bisang, Nikolaus P. Himmelmann, and Björn Wiemer, eds., *What Makes Grammaticalization? A Look from its Fringes and its Components*, 21–42. (Trends in Linguistics, Studies and Monographs, 158.) Berlin and New York: Mouton de Gruyter.

Hock, Hans Henrich and Brian D. Joseph. 1996. *Language History, Language Change, and Language Relationship: An Introduction to Historical and Comparative Linguistics*. (Trends in Linguistics, Studies and Monographs, 93.) Berlin and New York: Mouton de Gruyter.

Hopper, Paul J. 1988. Emergent grammar. In Michael Tomasello, ed., *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure*, 155–173. Mahwah, NJ: Lawrence Erlbaum.

Hopper, Paul J. 1990. Where do words come from? In Croft, Denning, and Kemmer, eds., 151–160.

Hopper, Paul J. 1991. On some principles of grammaticization. In Traugott and Heine, eds., Vol. I, 17–35.

Hopper, Paul J. 1994. Phonogenesis. In Pagliuca, ed., 29–45.

Hopper, Paul J. and Sandra A. Thompson. 1985. The iconicity of "noun" and "verb". In John Haiman, ed., *Iconicity in Syntax*, 151–183. (Typological Studies in Language, 6.) Amsterdam and Philadelphia: John Benjamins.

Hopper, Paul J. and Elizabeth Closs Traugott. 1993. *Grammaticalization*. (Cambridge Textbooks in Linguistics.) Cambridge, UK: Cambridge University Press.

Hopper, Paul J. and Elizabeth Closs Traugott. 2003. *Grammaticalization*. (Cambridge Textbooks in Linguistics.) Cambridge, UK: Cambridge University Press, 2nd revised edn.

Horn, Laurence R. 2001 [1989]. *A Natural History of Negation*. (The David Hunter Series.) Stanford, CA: CSLI Publications, 2nd edn.

Huddleston, Rodney. 1984. *Introduction to the Grammar of English*. (Cambridge Textbooks in Linguistics.) Cambridge, UK: Cambridge University Press.

Huddleston, Rodney and Geoffrey K. Pullum. 2002. *The Cambridge Grammar of the English Language*. Cambridge, UK: Cambridge University Press.

ICAME. International Computer Archives of Modern English. 1999. Compiled by Knut Holland, Anne Lindebjerg, and Jorn Thunestvedt. Bergen: Norwegian Computing Center for the Humanities, CD-ROM, 2nd edn.

Jackendoff, Ray. 1983. *Semantics and Cognition*. (Current Studies in Linguistics, 8.) Cambridge, MA: MIT Press.

Jackendoff, Ray. 1990. *Semantic Structures*. (Current Studies in Linguistics, 18.) Cambridge, MA: MIT Press.

Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. (Linguistic Inquiry Monograph, 28.) Cambridge, MA: MIT Press.

Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.

Jakobson, Roman. 1971 [1959]. Boas' view of grammatical meaning. *Selected Writings*, Vol. II, *Word and Language*, 489–496. The Hague: Mouton.

Janda, Richard D. 1980. On the decline of declensional systems: The overall loss of OE nominal case inflections and the ME reanalysis of *-es* as *his*. In Traugott, Labrum, and Shepherd, eds., 243–252.

Janda, Richard D. 1981. A case of liberation from morphology into syntax: The fate of the English genitive-marked *-(e)s*. In Brenda B. Johns and David R. Strong, eds., *Syntactic Change*, 59–114. (Natural Language Studies.) Ann Arbor: Department of Linguistics, University of Michigan.

Janda, Richard D. 1995. From agreement affix to subject "clitic" – and bound root: ***-m****os* > ***-n****os* vs. *(-)****n****os(-)* and ***n****os-otros* in New Mexican and other regional Spanish dialects. In Audra Dainora, Rachel Hemphill, Barbara Luka, Barbara Need, and Sheri Pargman, eds., *CLS Parasession on Clitics: Papers from the Thirty-First Regional Meeting, Chicago Linguistic Society*, 118–139. Chicago: Chicago Linguistic Society.

Janda, Richard D. 2001. Beyond "pathways" and "unidirectionality": On the discontinuity of transmission and the counterability of grammaticalization. In Campbell, ed., 265–340.

Janda, Richard D. and Brian D. Joseph. 2003. On language, change, and language change – Or, of history, linguistics, and historical linguistics. In Joseph and Janda, eds., 3–180.

Japanese and Korean Linguistics. Stanford University: Center for the Study of Language and Information.

Jeffers, Robert J. and Arnold M. Zwicky. 1980. The evolution of clitics. In Traugott, Labrum, and Shepherd, eds., 221–231.

Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen: A. F. Høst.

Jespersen, Otto. 1961 [1909–1941]. *A Modern English Grammar on Historical Principles*. London: George Allen and Unwin, and Copenhagen: Ejnar Munksgaard, 7 vols.

Jones, Michael. 1993. *Sardinian Syntax*. London: Routledge.

Joseph, Brian D., ed. 1986. *Studies in Language Change*. (Working Papers in Linguistics, 34.) Columbus: Department of Linguistics, Ohio State University.

Joseph, Brian D. 2003. Morphologization from syntax. In Joseph and Janda, eds., 472–492.

Joseph, Brian D. and Richard D. Janda. 1988. The how and why of diachronic morphologization and demorphologization. In Hammond and Noonan, eds., 193–210.

Joseph, Brian D. and Richard D. Janda, eds. 2003. *The Handbook of Historical Linguistics*. (Blackwell Handbooks in Linguistics.) Malden, MA: Blackwell.

Jucker, Andreas H., ed. 1995. *Historical Pragmatics: Pragmatic Developments in the History of English*. (Pragmatics & Beyond, New Series, 35.) Amsterdam and Philadelphia: John Benjamins.

Jucker, Andreas H. Online. Bibliography of Historical Pragmatics. http://www.es.unizh.ch/ahjucker/HistPrag.htm.

Kaplan, Robert B. 1966. Cultural thought patterns in intercultural education. *Language Learning* 16:1–60.

Kastovsky, Dieter. 1982. *Wortbildung und Semantik*. (Studienreihe Englisch, 14.) Düsseldorf: Pädagogischer Verlag Schwann-Bagel GmbH.

Kastovsky, Dieter. 1986. The problem of productivity in word formation. *Linguistics* 24:585–600.

Kastovsky, Dieter. 1992. Semantics and vocabulary. In Richard M. Hogg, ed., *The Cambridge History of the English Language*, Vol. I, *The Beginnings to 1066*, 290–408. Cambridge, UK: Cambridge University Press.

Kastovsky, Dieter, ed. 1994. *Studies in Early Modern English*. (Topics in English Linguistics, 13.) Berlin and New York: Mouton de Gruyter.

Kastovsky, Dieter. 1994. Typological differences between English and German morphology and their causes. In Toril Swan, Endre Mørck, and Olaf Jansen Westvik, eds., *Language Change and Language Structures: Older Germanic Languages in a Comparative Perspective*, 135–157. (Trends in Linguistics, Studies and Monographs, 75.) Berlin and New York: Mouton de Gruyter.

Katz, Jerrold J. and Jerry A. Fodor. 1963. The structure of a semantic theory. *Language* 39:170–210.

Kawabata, Tomohiro. 2003. On the development of *considering*: The prepositional, conjunctive and adverbial usages. In *Studies in Modern English* (*The Twentieth Anniversary Publication of the Modern English Association*), 139–152. Tokyo: Eichosha.

Keller, Rudi. 1994. *On Language Change: The Invisible Hand of Language*, trans. Brigitte Nerlich. London and New York: Routledge.

Kemenade, Ans van. 1999. Functional categories, morphosyntactic change, grammaticalization. *Linguistics* 37: 997–1010.

Killie, Kristin. 2000. Stative adverbs in English: A study of adverbial productivity and orientation. Unpublished Ph.D. dissertation, University of Tromsø.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In Emmon Bach and Robert Harms, eds., *Universals in Linguistic Theory*, 171–202. New York: Holt, Rinehart and Winston.

Kiparsky, Paul. 1992. Analogy. In William Bright, ed., *International Encyclopedia of Linguistics*. Vol. I, 56–61. New York: Oxford University Press.

Kiparsky, Paul. 1995. Indo-European origins of Germanic syntax. In Adrian Battye and Ian Roberts, eds., *Clause Structure and Language Change*, 140–169. Oxford: Oxford University Press.

Klausenburger, Jurgen. 2002. Grammaticalization within a theory of morphocentricity. In Wischer and Diewald, eds., 31–43.

Koch, Harold. 1996. Reconstruction in morphology. In Mark Durie and Malcolm Ross, eds., *The Comparative Method Reviewed: Regularity and Irregularity in Language Change*, 218–263. New York, Oxford: Oxford University Press.

Koch, Peter. 1999. TREE and FRUIT. A cognitive-onomasiological approach. *Studi di Linguistica Italiana Teorica ed Applicata* 28:331–347.

Kortmann, Bernd. 1991. *Free Adjuncts and Absolutes in English: Problems of Control and Interpretation*. London and New York: Routledge.

Kortmann, Bernd and Ekkehard König. 1992. Categorial reanalysis: The case of deverbal prepositions. *Linguistics* 30:671–697.

Kövekses, Zoltán and Günter Radden. 1998. Metonymy: Developing a cognitive linguistic view. *Cognitive Linguistics* 9:37–77.

Kroch, Anthony. 2001. Syntactic change. In Mark Baltin and Chris Collins, eds., *The Handbook of Contemporary Syntactic Theory*, 699–729. (Blackwell Handbooks in Linguistics.) Malden, MA: Blackwell.

Kroch, Anthony, John Myhill, and Susan Pintzuk. 1982. Understanding *do*. In K. Tuite, Robinson Schneider, and Robert Chametzky, eds., *Papers from the Eighteenth Regional Meeting, Chicago Linguistic Society*, 282–294. Chicago: Chicago Linguistic Society.

Krug, Manfred 1998. British English is developing a new discourse marker, *innit*? A study in lexicalisation based on social, regional and stylistic variation. *Arbeiten aus Anglistik und Amerikanistik* 23:145–197.

Krug, Manfred G. 2000. *Emerging English Modals: A Corpus-based Study of Grammaticalization*. (Topics in English Linguistics, 32.) Berlin and New York: Mouton de Gruyter.

Kuryłowicz, Jerzy. 1975 [1965]. The evolution of grammatical categories. *Esquisses linguistiques* 2:38–54. (Originally publ. in *Diogenes* 1965:55–71.)

Kytö, Merja. 1996. *Manual to the Diachronic Part of the Helsinki Corpus of English Texts: Coding Conventions and Lists of Source Texts*. Helsinki: Department of English, University of Helsinki, 3rd edn.

Labov, William. 1972. Negative attraction and negative concord. Chapter IV, *Language in the Inner City: Studies in Black English Vernacular*. Philadelphia: University of Pennsylvania Press.

Lampeter Corpus of Early Modern English Tracts (1640–1740). 1999. Available in *ICAME*.

Langacker, Ronald W. 1977. Syntactic reanalysis. In Li, ed., 57–139.

Langacker, Ronald W. 1987. *Foundations of Cognitive Grammar*, Vol. I, *Theoretical Perspectives*. Stanford: Stanford University Press.

Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1:5–38.

Langacker, Ronald W. 1991. *Foundations of Cognitive Grammar*, Vol. II, *Descriptive Application*. Stanford: Stanford University Press.

Langacker, Ronald W. 1999. Losing control: Grammaticalization, subjectification, and transparency. In Blank and Koch, eds., 147–175.

Lass, Roger. 1990. How to do things with junk: Exaptation in language evolution. *Journal of Linguistics* 26:79–102.

Lass, Roger, ed. 1999. *The Cambridge History of the English Language: Vol. III 1476–1776*. Cambridge, UK: Cambridge University Press.

Lazzeroni, Romano. 1998. Divagazioni sulla degrammaticalizzazione. In Giuliano Bernini, Pierluigi Cuzzolini, and Piera Molinelli, eds., *Ars Linguistica, Studi Offerti a Paolo Ramat*, 275–283. Rome: Bulzoni.

Lee, Hanjung. 2001. Optimization in argument expression and interpretation: A unified approach. Unpublished Ph.D. dissertation, Stanford University.

Leech, Geoffrey. 1981 [1974]. *Semantics*. Harmondsworth: Penguin, 2nd edn.

Lehmann, Christian. 1985. Grammaticalization: Synchronic variation and diachronic change. *Lingua e stile* 20:303–318.

Lehmann, Christian. 1989. Grammatikalisierung und Lexikalisierung. *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 42:11–19.

Lehmann, Christian. 1990. Towards lexical typology. In Croft, Denning, and Kemmer, eds., 161–185.

Lehmann, Christian. 1992. Word order change by grammaticalization. In Marinel Gerritsen and Dieter Stein, eds., *Internal and External Factors in Syntactic Change*, 395–416. (Trends in Linguistics, Studies and Monographs, 61.) Berlin and New York: Mouton de Gruyter.

Lehmann, Christian. 1993. Theoretical implications of grammaticalization phenomena. In William A. Foley, ed., *The Role of Theory in Language Description*, 315–340. (Trends in Linguistics, Studies and Monographs, 69.) Berlin: Mouton de Gruyter.

Lehmann, Christian. 1995 [1982]. *Thoughts on Grammaticalization*. (LINCOM Studies in Theoretical Linguistics, 1.) München and Newcastle: LINCOM EUROPA.

Lehmann, Christian. 1995. Synsemantika. In Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, and Theo Vennemann, eds., *Syntax: Ein internationales Handbuch zeitgenössischer Forschung*, 1251–1266. (Handbücher der Sprach- und Kommunikationswissenschaft, 9/2.) Berlin: Walter de Gruyter.

Lehmann, Christian. 2002. New reflections on grammaticalization and lexicalization. In Wischer and Diewald, eds., 1–18.

Lessau, Donald A. 1994. *A Dictionary of Grammaticalization* (Bochum-Essener Beiträge zur Sprachwandelforschung, 21.) Bochum: Universitätsverlag Dr. N. Brockmeyer, 3 vols.

Levin, Beth. 1993. *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Levin, Beth and Malka Rappaport Hovav. 1995. *Unaccusativity: At the Syntax-Lexical Semantics Interface*. Cambridge, MA: MIT Press.

Levinson, Stephen. 2000. *Presumptive Meaning: The Theory of Generalized Conversational Implicature*. Cambridge, MA: MIT Press.

Li, Charles N., ed. 1977. *Mechanisms of Syntactic Change*. Austin: University of Texas Press.

Lichtenberk, Frantisek. 1991. On the gradualness of grammaticalization. In Traugott and Heine, eds., Vol. I, 37–80.

Lightfoot, David W. 1979. *Principles of Diachronic Syntax*. (Cambridge Studies in Linguistics, 23.) Cambridge, UK: Cambridge University Press.

Lightfoot, David W. 1999. *The Development of Language: Acquisition, Change, and Evolution*. (Maryland Lectures in Language and Cognition, 1.) Malden, MA and Oxford: Blackwell.

Lindquist, Hans and Christian Mair, eds. 2004. *Corpus Approaches to Grammaticalization in English*. (Studies in Corpus Linguistics, 13.) Amsterdam and Philadelphia: John Benjamins.

Lindström, Therese Å. M. 2004. The history of the concept of grammaticalisation. Unpublished Ph.D. dissertation, University of Sheffield.

Lipka, Leonhard. 1992. Lexicalization and institutionalization in English and German or: *Piefke, Wendehals, smog, perestroika, AIDS*, etc. *Zeitschrift für Anglistik und Amerikanistik* 40:101–111.

Lipka, Leonhard. 1994. Lexicalization and institutionalization. In Asher and Simpson, eds., Vol. IV, 2164–2167.

Lipka, Leonhard. 2002 [1990]. *English Lexicology: Lexical Structure, Word Semantics & Word-Formation*. (Narr Studienbücher.) Tübingen: Max Niemeyer Verlag, 3rd revised edn. of *An Outline of English Lexicology*.

Luraghi, Silvia. 1998. On the directionality of grammaticalization. *Sprachtypologische Universal-Forschungen* (STUF) 51:355–365.

Lyons, John. 1977. *Semantics*. Cambridge, UK: Cambridge University Press, 2 vols.

McArthur, Tom. 1992. *The Oxford Companion to the English Language*. Oxford and New York: Oxford University Press.

McCawley, James D. 1968. Lexical insertion in a transformational grammar without deep structure. In Bill J. Darden, Charles-James N. Bailey, and Alice Davison, eds., *Papers from the Fourth Regional Meeting, Chicago Linguistic Society*, 71–80. Chicago: Chicago Linguistic Society.

McElhinney, Bonnie. 1992. The interaction of phonology, syntax, and semantics in language change: The history of modal contraction in English. In Costas P. Canakis, Grace P. Chan, and Jeannette Marshall Denton, eds., *Papers from the Twenty-Eighth Regional Meeting, Chicago Linguistic Society*, Vol. I:367–381. Chicago: Chicago Linguistic Society.

McKercher, David. 2001. Children's acquisition of the meaning of with: A case study of polysemy in child language development. Unpublished Ph.D. dissertation, Stanford University.

Mair, Christian. 1994. Is *see* becoming a conjunction? The study of grammaticalisation as a meeting ground for corpus linguistics and grammatical theory. In Udo Fries, Gunnel Tottie, and Peter Schneider, eds., *Creating and Using English Language Corpora: Papers from the Fourteenth International Conference on English Language Research on Computerized Corpora, Zurich 1993*, 127–137. Amsterdam: Rodopi.

Mair, Christian. 2004. Corpus linguistics and grammaticalisation theory: Statistics, frequencies, and beyond. In Lindquist and Mair, eds., 121–150.

Marchand, Hans. 1969 [1960]. *The Categories and Types of Present-Day English Word-Formation: A Synchronic-Diachronic Approach*. (Handbücher für das Studium der Anglistik.) Munich: Beck'sche Verlags Buchhandlung, 2nd edn.

Matsumoto, Yo. 1988. From bound grammatical markers to free discourse markers: History of some Japanese connectives. In Axmaker, Jaisser, and Singmaster, eds., 340–351.

Matthews, P. H. 1974. *Morphology: An Introduction to the Theory of Word-Structure*. (Cambridge Textbooks in Linguistics.) Cambridge, UK: Cambridge University Press.

Matthews, P. H. 1997. *The Concise Oxford Dictionary of Linguistics*. Oxford and New York: Oxford University Press.

MED. The Middle English Dictionary. 1956–2001. Ann Arbor: University of Michigan Press. (See also http://www.hti.umich.edu/dict/med/.)

Meillet, Antoine. 1958 [1912]. L'évolution des formes grammaticales. In Meillet 1958, 130–148. (Originally publ. in *Scientia* [*Rivista di Scienza*] 12, No. 26, 6, 1912.)

Meillet, Antoine. 1958 [1915–16]. Le renouvellement des conjonctions. In Meillet 1958, 159–174. (Originally publ. in *Annuaire de l'École Pratique des Hautes Études*.)

Meillet, Antoine. 1958. *Linguistique historique et linguistique générale*. (Collection linguistique publiée par la Société de Linguistique de Paris, 8.) Paris: Champion.
Miller, D. Gary. 2002. *Nonfinite Structures in Theory and Change*. Oxford and New York: Oxford University Press.
Milroy, James. 1992. *Linguistic Variation and Change: On the Historical Sociolinguistics of English*. Oxford: Blackwell.
Milroy, James. 2003. On the role of the speaker in language change. In Raymond Hickey, ed., *Motives for Language Change*, 143–157. Cambridge, UK: Cambridge University Press.
Milroy, Lesley. 1987 [1980]. *Language and Social Networks*. Oxford: Blackwell.
Mitchell, Bruce. 1985. *Old English Syntax*. Oxford: Clarendon Press, 2 vols.
Mithun, Marianne. 1999. The reordering of morphemes. In Gildea, ed., 231–255.
Mithun, Marianne. 2000. Noun and verb in Iroquoian languages: Multicategorisation from multiple criteria. In Vogel and Comrie, eds., 397–420.
Mithun, Marianne. 2001. Lexical forces shaping the evolution of grammar. In Brinton, ed., 241–252.
Moessner, Lilo. 2001. Genre, text type, style, register: A terminological maze. *European Journal of English Studies* 5:131–138.
Moreno Cabrera, Juan C. 1998. On the relationship between grammaticalization and lexicalization. In Giacalone Ramat and Hopper, eds., 209–227.
Moore, Colette. Forthcoming. The use of *videlicet* in Early Modern slander depositions: A case of genre-specific grammaticalization. *Journal of Historical Pragmatics* 7.2.
Morita, Junya. 1995. Lexicalization by way of context-dependent nonce-word formation. *English Studies* 76:468–473.
Mossé, Ferdinand. 1938. *Histoire de la forme périphrastique être + participe présent en germanique*. (Collection Linguistique, Société de linguistique de Paris, 42–43.) Paris: C. Klincksieck, 2 vols.
Motsch, Wolfgang. 2003. Derivational morphology. In William J. Frawley, ed., *International Encyclopedia of Linguistics*, Vol. I, 427–429. New York: Oxford University Press, 2nd edn.
Mühlhäusler, Peter. 1979. *Growth and Structure of the Lexicon of New Guinea Pidgin*. (Pacific Linguistics: Series C, 52.) Canberra: Australian National University.
Mühlhäusler, Peter. 1997. *Pidgin and Creole Linguistics*. (Westminster Creolistics Series, 3.) London: University of Westminster Press.
Mustanoja, Tauno F. 1960. *A Middle English Syntax*. (Mémoires de la Société Néophilologique, 23.) Helsinki: Société Néophilologique.
Napoli, Donna Jo. 1993. *Syntax: Theory and Problems*. New York and Oxford: Oxford University Press.
Nedjalkov, Igor' V. 1998. Converbs in the languages of Europe. In van der Auwera, with Ó Baoill, eds., 421–455.
Nevalainen, Terttu. 1997. The processes of adverb derivation in Late Middle and Early Modern English. In Matti Rissanen, Merja Kytö, and Kirsi Heikkonen, eds., *Grammaticalization at Work: Studies of Long-Term Developments in English*, 145–189. (Topics in English Linguistics, 24.) Berlin and New York: Mouton de Gruyter.

Nevalainen, Terttu. 1999. Early Modern English lexis and semantics. In Lass, ed., 332–458.
Nevalainen, Terttu. 2004. Three perspectives on grammaticalization: Lexico-grammar, corpora and historical linguistics. In Lindquist and Mair, eds., 1–31.
Nevalainen, Terttu and Helena Raumolin-Brunberg. 2003. *Historical Sociolinguistics: Language Change in Tudor and Stuart England*. (Longman Linguistic Library.) Harlow and London: Pearson Education.
Nevalainen, Terttu and Matti Rissanen. 2002. Fairly pretty or pretty fair? On the development and grammaticalization of English downtoners. *Language Sciences* 24:359–380.
Nevis, Joel A. 1986a. Decliticization and deaffixation in Saame: Abessive *taga*. In Joseph, ed., 1–9.
Nevis, Joel A. 1986b. Decliticization in Old Estonian. In Joseph, ed., 10–27.
Newmeyer, Frederick J. 1998. *Language Form and Language Function*. (Language, Speech, and Communication.) Cambridge, MA and London: MIT Press.
Nickel, Gerhard. 1966. *Die expanded Form im Altenglischen. Vorkommen, Funktion und Herkunft der Umschreibung* beon/wesan + *Partizip Präsens*. Neumünster: Karl Wachholtz.
Norde, Muriel. 2001. Deflexion as a counterdirectional factor in grammatical change. In Campbell, ed., 231–264.
Norde, Muriel. 2002. The final stages of grammaticalization: Affixhood and beyond. In Wischer and Diewald, eds., 45–65.
Norrick, Neal R. 1979. The lexicalization of pragmatic functions. *Linguistics* 17:671–685.
Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70:491–538.
O'Dowd, Elizabeth M. 1998. *Prepositions and Particles in English: A Discourse-Functional Account*. Oxford and New York: Oxford University Press.
OED. Oxford English Dictionary. 3rd edn. (Online http://dictionary.oed.com/.)
Olofsson, Arne. 1990. A participle caught in the act. On the prepositional use of *following*. *Studia Neophilologica* 62:23–35.
Páez Urdaneta, Iraset. 1982. Conversational "*pues*" in Spanish: A process of degrammaticalization? In Anders Ahlqvist, ed., *Papers from the Fifth International Conference on Historical Linguistics*, 332–340. (Current Issues in Linguistic Theory, 21.) Amsterdam and Philadelphia: John Benjamins.
Pagliuca, William, ed. 1994. *Perspectives on Grammaticalization*. (Current Issues in Linguistic Theory, 109.) Amsterdam and Philadelphia: John Benjamins.
Palander-Collin, Minna. 1999. *Grammaticalization and Social Embedding: I THINK and METHINKS in Middle and Early Modern English*. (Mémoires de la Société Néophilologique de Helsinki, 55.) Helsinki: Société Néophilologique.
Palmer, F. R. 1988. *The English Verb*. (Longman Linguistics Library.) London and New York: Longman, 2nd edn.
Paradis, Carita. 1997. *Degree Modifiers of Adjectives in Spoken British English*. (Lund Studies in English, 92.) Lund: Lund University Press.
Pawley, Andrew. 1986. Lexicalization. In Deborah Tannen and James E. Alatis, eds., *Languages and Linguistics: The Interdependence of Theory, Data, and*

Application, 98–120. (GURT 1985.) Washington, DC: Georgetown University Press.

Pawley, Andrew and Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In Jack C. Richards and Richard W. Schmidt, eds., *Language and Communication*, 191–226. London and New York: Longman.

Pérez, Aveline. 1990. Time in motion: Grammaticalisation of the *be going to* construction in English. *La Trobe University Working Papers in Linguistics* 3:49–64.

Peters, Hans. 1994. Degree adverbs in Early Modern English. In Kastovsky, ed., 269–288.

Pinker, Steven. 1999. *Words and Rules: The Ingredients of Language*. New York: Basic Books.

Plag, Ingo. 1999. *Morphological Productivity: Structural Constraints in English Derivation*. (Topics in English Linguistics, 28.) Berlin and New York: Mouton de Gruyter.

Plank, Frans. 1989. From cases to adpositions. In Nicola Pantaleo, ed., *Aspects of English Diachronic Linguistics; Papers Read at the Second National Conference of History of English, Naples, 28–29 April 1989*, 19–61. (Biblioteca della Ricerca. Cultura Straniera, 48.) Fasano: Schena Editore.

Plank, Frans. 1994. Inflection and derivation. In Asher and Simpson, eds., Vol. III, 1671–1677.

Pollard, Carl and Ivan A. Sag. 1994. *Head-Driven Phrase Structure Grammar*. (Studies in Contemporary Linguistics.) Stanford: CSLI Publications and Chicago: University of Chicago Press.

Pratt, Lynda and David Denison. 2000. The language of the Southey-Coleridge circle. *Language Sciences* 22:401–422.

Pustejovsky, James. 1995. *The Generative Lexicon*. Cambridge, MA: MIT Press.

Pustet, Regina. 2003. *Copulas: Universals in the Categorization of the Lexicon*. (Oxford Studies in Typology and Linguistic Theory.) Oxford and New York: Oxford University Press.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London and New York: Longman.

Ramat, Paolo. 1992. Thoughts on degrammaticalization. *Linguistics* 30:549–560.

Ramat, Paolo. 2001. Degrammaticalization or transcategorization? In Chris Schaner-Wolles, John Rennison, and Friedrich Neubarth, eds., *Naturally! Linguistic Studies in Honour of Wolfgang Ulrich Dressler Presented on the Occasion of his 60th Birthday*, 393–401. Torino: Rosenbach and Sellier.

Ramat, Paolo and Davide Ricca. 1994. Prototypical adverbs: On the scalarity/radiality of the notion of ADVERB. *Rivista di Linguistica* 6:289–326.

Ramat, Paolo and Davide Ricca. 1998. Sentence adverbs in the languages of Europe. In van der Auwera, with Ó Baoill, eds., 187–275.

Rappaport Hovav, Malka and Beth Levin. 1998a. Morphology and lexical semantics. In Andrew Spencer and Arnold Zwicky, eds., *Handbook of Morphology*, 248–271. (Blackwell Handbooks in Linguistics.) Oxford: Blackwell.

Rappaport Hovav, Malka and Beth Levin. 1998b. Building verb meanings. In Miriam Butt and Wilhelm Geuder, eds., *The Projection of Arguments: Lexical*

and Compositional Factors, 97–134. (CSLI Lecture Notes, 83.) Stanford: Center for the Study of Language and Information Publications.

Raumolin-Brunberg, Helena, Minna Navala, Arja Nurmi, and Matti Rissanen, eds. 2002. *Variation Past and Present: VARIENG Studies on English for Terttu Nevalainen*. (Mémoires de la Société Néophilologique de Helsinki, 61.) Helsinki: Société Néophilologique.

Rhee, Seongha. 1996. *Semantics of Verbs and Grammaticalization: The Development in Korean from a Cross-Linguistic Perspective*. (Hankuk Dissertation Series.) Seoul: Hankuk Publisher.

Rickford, John R. 1999. *African American Vernacular English: Features, Evolution, Educational Implications*. (Language in Society, 26.) Oxford and Maldon, MA: Blackwell.

Rissanen, Matti. 1999. Syntax. In Lass, ed., 187–331.

Rissanen, Matti. 2004. Grammaticalisation from side to side: On the development of *beside(s)*. In Lindquist and Mair, eds., 151–170.

Rissanen, Matti, Ossi Ihalainen, Terttu Nevalainen, and Irma Taavitsainen, eds. 1992. *History of Englishes: New Methods and Interpretations in Historical Linguistics*. (Topics in English Linguistics, 10.) Berlin and New York: Mouton de Gruyter.

Rissanen, Matti, Merja Kytö, and Minna Palander-Collin, eds. 1993. *Early English in the Computer Age: Explorations through the Helsinki Corpus*. (Topics in English Linguistics, 11.) Berlin and New York: Mouton de Gruyter.

Roberts, Ian G. 1993a. A formal account of grammaticalisation in the history of Romance futures. *Folia Linguistica Historica* XIII:219–258.

Roberts, Ian. 1993b. *Verbs and Diachronic Syntax: A Comparative History of English and French*. (Studies in Natural Language and Linguistic Theory, 28.) Dordrecht: Kluwer.

Roberts, Ian G. and Anna Roussou. 2003. *Syntactic Change: A Minimalist Approach to Grammaticalization*. (Cambridge Studies in Linguistics, 100.) Cambridge, UK: Cambridge University Press.

Romaine, Suzanne and Deborah Lange. 1991. The use of *like* as a marker of reported speech and thought: A case of grammaticalization in progress. *American Speech* 66:227–279.

Rosch, Eleanor. 1978. Principles of categorization. In Eleanor Rosch and Barbara B. Lloyd, eds., *Cognition and Categorization*, 27–48. Hillsdale, NJ: Lawrence Erlbaum.

Rosenbach, Anette. 2002. *Genitive Variation in English: Conceptual Factors in Synchronic and Diachronic Studies*. (Topics in English Linguistics, 42.) Berlin and New York: Mouton de Gruyter.

Ross, John Robert. 1972. The category squish: Endstation Hauptwort. In Paul Peranteau, Judith N. Levi, and Gloria C. Phares, eds., *Papers from the Eighth Regional Meeting, Chicago Linguistic Society*, 316–328. Chicago: Chicago Linguistic Society.

Rubino, Carl. 1994. Against the notion of unidirectionality in lexeme genesis. *Linguistica Atlantica* 16:135–147.

Ryder, Mary Ellen. 1999. Complex *-er* nominals: Where grammaticalization and lexicalization meet? In Ellen Contini-Morava and Yishai Tobin, eds., *Between Grammar and Lexicon*, 291–332. (Current Issues in Linguistic Theory, 183.) Amsterdam and Philadelphia: John Benjamins.

Sag, Ivan and Carl Pollard. 1991. An integrated theory of complement control. *Language* 67:63–113.

Sankoff, Gillian and Suzanne Laberge. 1980 [1976]. On the acquisition of native speakers by a language. In Gillian Sankoff, ed., *The Social Life of Language*, 195–209. Philadelphia: University of Pennsylvania Press. (Originally publ. in *Kivung* 6:32–47.)

Sapir, Edward. 1920. *Language: An Introduction to the Study of Speech*. New York: Harcourt Brace Jovanovich.

Sasse, Hans-Jürgen. 1988. Der irokesische Sprachtyp. *Zeitschrift für Sprachwissenschaft* 7:173–213.

Saukkonen, Pauli. 2003. How to define and describe genres and styles. *Folia Linguistica Historica* XXXVII:399–414.

Saussure, Ferdinand de. 1986 [1916]. *Course in General Linguistics*, trans. Roy Harris. Chicago: Open Court.

Schiffrin, Deborah. 1987. *Discourse Markers*. (Studies in Interactional Sociolinguistics, 5.) Cambridge, UK: Cambridge University Press.

Shopen, Timothy, ed. 1985. *Language Typology and Syntactic Description*, Vol. III, *Grammatical Categories and the Lexicon*. Cambridge, UK: Cambridge University Press.

Skeat, Walter W. 1887. Obscure compounds. *Principles of English Etymology*. *First Series*, Vol. I, *The Native Element*. Oxford: Clarendon Press.

Slobin, Dan I. 2004. The many ways to search for a frog: Linguistic typology and the expression of motion events. In Sven Strömqvist and Ludo Verhoeven, eds., *Relating Events in Narrative: Typological and Contextual Perspectives*, Vol. II, 219–257. Mahwah, NJ: Lawrence Erlbaum.

Solomon, Julie. 1995. Local and global functions of a borrowed/native pair of discourse markers in a Yucatec Maya narrative. In Jocelyn Ahlers, Leela Bilmes, Joshua S. Guenter, Barbara A. Kaiser, and Ju Namkung, eds., *Proceedings of the 21st Annual Meeting of the Berkeley Linguistics Society, February 17–20, 1995*, 287–298. Berkeley, CA: Berkeley Linguistics Society.

Stein, Dieter and Susan Wright, eds. 1995. *Subjectivity and Subjectivisation: Linguistic Perspectives*, 151–172. Cambridge, UK: Cambridge University Press.

Sun, Chaofen. 1996. *Word Order Changes and Grammaticalization in the History of Chinese*. Stanford: Stanford University Press.

Swan, Toril. 1988. *Sentence Adverbials in English: A Synchronic and Diachronic Investigation*. (Tromsø-studier i Sprakvitenskap, 10.) Oslo: Novus Verlag.

Swan, Toril. 1997. From manner to subject modification: Adverbialization in English. *Nordic Journal of Linguistics* 20:179–195.

Sweetser, Eve E. 1988. Grammaticalization and semantic bleaching. In Axmaker, Jaisser, and Singmaster, eds., 389–405.

Sweetser, Eve E. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. (Cambridge Studies in Linguistics, 54.) Cambridge, UK: Cambridge University Press.

Taavitsainen, Irma. 2001. Middle English recipes: Genre characteristics, text type features, and underlying traditions of writing. *Journal of Historical Pragmatics* 2:85–113.

Tabor, Whitney. 1994. Syntactic innovation: A connectionist model. Unpublished Ph.D. dissertation, Stanford University.

Tabor, Whitney and Elizabeth Closs Traugott. 1998. Structural scope expansion and grammaticalization. In Giacalone Ramat and Hopper, eds., 229–272.

Talmy, Leonard. 1985. Lexicalization patterns: Semantic structure in lexical forms. In Shopen, ed., Vol. III, 57–149.

Talmy, Leonard. 2000. *Toward a Cognitive Semantics*. Cambridge, MA: MIT Press, 2 vols.

Taylor, John R. 1997 [1989]. *Linguistic Categorization: Prototypes in Linguistic Theory*. Oxford: Clarendon, 2nd edn.

Thomas, Francis-Noél and Mark Turner. 1994. *Clear and Simple as the Truth: Writing Classic Prose*. Princeton: Princeton University Press.

Thompson, Sandra A. and Anthony Mulac. 1991. A quantitative perspective on the grammaticization of epistemic parentheticals in English. In Traugott and Heine, eds., Vol. II, 313–329.

Timberlake, Alan. 1977. Reanalysis and actualization in syntactic change. In Li, ed., 141–177.

Traugott, Elizabeth Closs. 1980. Meaning change in the development of grammatical markers. *Language Sciences* 2:44–61.

Traugott, Elizabeth Closs. 1982. From propositional to textual and expressive meanings: Some semantic-pragmatic aspects of grammaticalization. In Winfred P. Lehmann and Yakov Malkiel, eds., *Perspectives on Historical Linguistics*, 245–271. (Current Issues in Linguistic Theory, 24.) Amsterdam and Philadelphia: John Benjamins.

Traugott, Elizabeth Closs. 1986. "Conventional" and "dead" metaphors revisited. In Wolfgang Paprotté and René Dirven, eds., *The Ubiquity of Metaphor: Metaphor in Language and Thought*, 17–56. (Current Issues in Linguistic Theory, 29.) Amsterdam and Philadelphia: John Benjamins.

Traugott, Elizabeth Closs. 1994. Grammaticalization and lexicalization. In Asher and Simpson, eds., Vol. III, 1481–1486.

Traugott, Elizabeth Closs. 1995a. Subjectification in grammaticalization. In Stein and Wright, eds., 31–54.

Traugott, Elizabeth Closs. 1995b. The role of discourse markers in a theory of grammaticalization. Paper presented at the 12th International Conference on Historical Linguistics, Manchester, August 1995. www.stanford.edu/~traugott/ect-papersonline.html

Traugott, Elizabeth Closs. 1999. A historical overview of complex predicate types. In Brinton and Akimoto, eds., 239–260.

Traugott, Elizabeth Closs. 2002. From etymology to historical pragmatics. In Donka Minkova and Robert Stockwell, eds., *Studies in the History of the English Language: A Millennial Perspective*, 19–49. (Topics in English Linguistics, 39.) Berlin and New York: Mouton de Gruyter.

Traugott, Elizabeth Closs. 2003. Constructions in grammaticalization. In Joseph and Janda, eds., 624–647.

Traugott, Elizabeth Closs. 2005. Lexicalization and grammaticalization. In Alan Cruse, Franz Hundsnurscher, Michael Job, and Peter Rolf Lutzeier, eds., *Lexikologie/Lexicology*, Vol. II, *1702–1712*. Berlin and New York: Mouton de Gruyter.

Traugott, Elizabeth Closs and Richard B. Dasher. 2002. *Regularity in Semantic Change*. (Cambridge Studies in Linguistics, 97.) Cambridge, UK: Cambridge University Press.

Traugott, Elizabeth Closs and Bernd Heine, eds. 1991. *Approaches to Grammaticalization*. (Typological Studies in Language, 19.) Amsterdam and Philadelphia: John Benjamins, 2 vols.

Traugott, Elizabeth and Ekkehard König. 1991. The semantics-pragmatics of grammaticalization revisited. In Traugott and Heine, eds., Vol. I, 189–218.

Traugott, Elizabeth Closs, Rebecca Labrum, and Susan Shepherd, eds. 1980. *Papers from the Fourth International Conference on Historical Linguistics*. (Current Issues in Linguistic Theory, 14.) Amsterdam and Philadelphia: John Benjamins.

Ullmann, Stephen. 1962. *Semantics: An Introduction to the Science of Meaning*. New York: Barnes and Noble.

van der Auwera, Johan. 1999. Dutch verbal prefixes: Meaning and form, grammaticalization and lexicalization. In Lunella Mereu, ed., *Boundaries of Morphology and Syntax*, 121–136. (Current Issues in Linguistic Theory, 180.) Amsterdam and Philadelphia: John Benjamins.

van der Auwera, Johan. 2001. On the typology of negative modals. In Jack Hoeksema, Hotze Rullmann, Victor Sánchez-Valencia, and Ton van der Wouden, eds., *Perspectives on Negation and Polarity Items*, 23–48. (Linguistik Aktuell, 40.) Amsterdam and Philadelphia: John Benjamins.

van der Auwera, Johan. 2002. More thoughts on degrammaticalization. In Wischer and Diewald, eds., 19–29.

van der Auwera, Johan, ed., in collaboration with Dónall P. Ó Baoill. 1998. *Adverbial Constructions in the Languages of Europe*. (Empirical Approaches to Language Typology; EUROTYP 20–3.) Berlin and New York: Mouton de Gruyter.

van der Wurff, Wim. 2002. The word *withal*: Some remarks on its historical development. In Jacek Fisiak, ed., *Studies in English Historical Linguistics and Philology: A Festschrift for Akio Oizumi*, 469–487. Bern: Peter Lang.

Viberg, Åke. 1983. The verbs of perception: A typological study. *Linguistics* 21:123–162.

Vincent, Diane, Sebastião Votre, and Marty LaForest. 1993. Grammaticalisation et post-grammaticalisation. *Langues et linguistique* 19:71–103.

Vincent, Nigel. 1996. The emergence of the D-system in Romance. In Ans van Kemenade and Nigel Vincent, eds., *Parameters of Morphosyntactic Change*, 149–169. Cambridge, UK: Cambridge University Press.

Visser, F.Th. 1963–1973. *An Historical Syntax of the English Language*. Leiden: E. J. Brill, 3 vols.

Vogel, Petra M. and Bernard Comrie, eds. 2000. *Approaches to the Typology of Word Classes* (Empirical Approaches to Language Typology, 23.) Berlin and New York: Mouton de Gruyter.

Voyles, Joseph B. 1973. Accounting for semantic change. *Lingua* 31:95–124.

Waltereit, Richard. 2002. Imperatives, interruption in conversation, and the rise of discourse markers: A study of Italian *guarda*. *Linguistics* 40:987–1010.

Warner, Anthony R. 1993. *English Auxiliaries: Structure and History*. (Cambridge Studies in Linguistics, 66.) Cambridge, UK: Cambridge University Press.

Warner, Anthony R. 1995. Predicting the progressive passive: Parametric change within a lexicalist framework. *Language* 71:533–557.

Wartburg, Walter von. 1996. *Französisches etymologisches Wörterbuch*. Basel: Zbinden.

Webster's Third New International Dictionary of the English Language. 1981 [1961]. Springfield, MA: Merriam-Webster.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In W. P. Lehmann and Yakov Malkiel, eds., *Directions for Historical Linguistics*, 97–195. Austin and London: University of Texas Press.

Wierzbicka, Anna. 1985. *Lexicography and Conceptual Analysis*. Ann Arbor: Karoma.

Williams, Edwin. 1981. On the notions "lexically related" and "head of a word." *Linguistic Inquiry* 12:245–274.

Wischer, Ilse. 2000. Grammaticalization versus lexicalization – "methinks" there is some confusion. In Fischer, Rosenbach, and Stein, eds., 355–370.

Wischer, Ilse and Gabriele Diewald, eds. 2002. *New Reflections on Grammaticalization – Proceedings from the International Symposium on Grammaticalization, 17–19 June 1999, Potsdam, Germany*. (Typological Studies in Language, 49.) Amsterdam and Philadelphia: John Benjamins.

Wright, Susan [Fitzmaurice]. 1994. The mystery of the modal progressive. In Kastovsky, ed., 467–485.

Wright, Susan [Fitzmaurice]. 1995. Subjectivity and experiential syntax. In Stein and Wright, eds., 115–172.

Wurzel, Wolfgang U. 1984. *Flexionsmorphologie und Natürlichkeit. Ein Beitrag zur morphologischen Theoriebildung*. (Studia Grammatica, 21.) Berlin: Akademie Verlag.

Ziegeler, Debra. 1996. A synchronic perspective on the grammaticalisation of *will* in hypothetical predicates. *Studies in Language* 20:411–442.

Ziegeler, Debra. 2003. Redefining unidirectionality: Insights from demodalisation. *Folia Linguistica Historica* XXIV:225–266.

Žirmunskij, V. M. 1966. The word and its boundaries. *Linguistics* 27:65–91. (Originally publ. in Russian in 1961.)

Index of names

Index of subjects

Index of words and forms