FREQUENCY OF USE AND THE ORGANIZATION OF LANGUAGE

Joan Bybee

Frequency of Use and the Organization of Language

This page intentionally left blank

Frequency of Use and the Organization of Language

Joan Bybee

OXFORD
UNIVERSITY PRESS

2007

OXFORD
UNIVERSITY PRESS

Oxford University Press, Inc., publishes works that further
Oxford University's objective of excellence
in research, scholarship, and education.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Library of Congress Cataloging-in-Publication Data
Bybee, Joan L.
Frequency of use and the organization of language / Joan Bybee.
p. cm.
Includes bibliographical references and index.
ISBN-13 978-0-19-530156-4; 978-0-19-530157-1 (pbk.)
ISBN 0-19-530156–0; 0-19-530157-9 (pbk.)
1. Frequency (Linguistics) 2. Linguistic change. I. Title.
P128.F73B935 2006
410.1'51—dc22 2006040009

1 3 5 7 9 8 6 4 2

Printed in the United States of America
on acid-free paper

ACKNOWLEDGMENTS

This volume reprints articles written over a period of almost three decades. Needless to say, I am grateful to a large community of scholars who have reacted to the hypotheses and findings of these articles and eventually to the theory that has emerged from them. Of course, I am most indebted to my co-authors, Mary Alexandra Brewer, Elly Pardo, Dan I. Slobin, Carol Lynn Moder, Jean E. Newman, Sandra Thompson, and Joanne Scheibman. Collaboration with these researchers allowed me to move beyond my own limitations and produce results that I could never have achieved alone. For helping with the preparation of the manuscript, I thank Vsevolod Kapatsinski and Clayton Beckner. I am also grateful to Peter Ohlin of Oxford University Press for suggesting this volume and to my family for their continued support.

I gratefully acknowledge permission from the publishers to reprint the following articles and book chapters, which constitute this volume.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In *Current progress in historical linguistics*, ed. William Christie, 95–105. Amsterdam: North Holland. Reprinted with kind permission from Elsevier.

Bybee, Joan L., and Mary Alexandra Brewer. 1980. Explanation in morphophonemics: Changes in Provençal and Spanish preterite forms. *Lingua* 52, 201–242. Reprinted with kind permission from Elsevier.

Bybee, Joan L., and Elly Pardo. 1981. On lexical and morphological conditioning of alternations: A nonce-probe experiment with Spanish verbs. *Linguistics* 19, 937–968. Reprinted with kind permission from Mouton de Gruyter.

Bybee, Joan L., and Dan I. Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58, 265–289. Reprinted with kind permission from the Linguistic Society of America.

Bybee, Joan L., and Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59, 251–270. Reprinted with kind permission from the Linguistic Society of America.

Bybee, Joan L., and Jean E. Newman. 1995. Are stem changes as natural as affixes? *Linguistics* 33, 633–654. Reprinted with kind permission from Mouton de Gruyter.

Bybee, Joan L. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10, 425–455. Reprinted with kind permission from Lawrence Erlbaum Associates.

Bybee, Joan. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language,* ed. Michael Barlow and Suzanne Kemmer, 65–85. Stanford, Calif.: CSLI. Reprinted with kind permission from CSLI Publications.

Bybee, Joan L. 2000. Lexicalization of sound change and alternating environments. In *Papers in Laboratory Phonology*, vol. 5: *Acquisition and the Lexicon,* ed. Michael Broe and Janet Pierrehumbert, 250–268. Cambridge: Cambridge University Press. Reprinted with kind permission from Cambridge University Press.

Bybee, Joan L. 2002. Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language variation and change* 14, 261–290. Reprinted with kind permission from Cambridge University Press.

Bybee, Joan, and Sandra Thompson. 1997. Three frequency effects in syntax. *Berkeley Linguistics Society* 23, 65–85. Reprinted with kind permission from the Berkeley Linguistics Society.

Bybee, Joan L. 1998. The emergent lexicon. In *CLS 34: The panels*, ed. M. Catherine Gruber, Derrick Higgins, Kenneth S. Olson, and Tamra Wysocki, 421–435. University of Chicago: Chicago Linguistic Society. Reprinted with kind permission from the Chicago Linguistic Society.

Bybee, Joan L., and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in American English. *Linguistics* 37(4), 575–596. Reprinted with kind permission from Mouton de Gruyter.

Bybee, Joan L. 2002. Sequentiality as the basis of constituent structure. In *The evolution of language out of pre-language*, ed. T. Givón and Bertram F. Malle, 109–132. Amsterdam: John Benjamins. Reprinted with kind permission from John Benjamins Publishing.

Bybee, Joan L. 2003. Mechanisms of change in grammaticization: The role of frequency. In *Handbook of historical linguistics,* ed. R. Janda and B. Joseph, 602–623. Oxford: Blackwell. Reprinted with kind permission from Blackwell Publishers Limited.

CONTENTS

PART I

BACKGROUND AND CURRENT CONTEXT

This page intentionally left blank

Introduction to Part I

Chapter 1 provides background on the effects of frequency of use on linguistic elements, explaining how such effects have been regarded in recent linguistic theory. This introduction lays out briefly the topics further developed in the chapters of this volume, which together demonstrate a major role for frequency of use in determining the nature of language.

The second chapter is my first paper on the topic of frequency of use; it was presented at a conference on historical linguistics in 1975. My interest in and investigations of frequency began rather tentatively, given the climate in which I was educated and in which I worked. The first research in which I investigated frequency effects was in the context of language change, and it involved two frequency effects that most historical linguists knew about and perhaps considered plausible, though of little interest given the prevailing view that language change is change in the grammar.

The strength of the paper is not in its methodology; at that time, linguists were unaccustomed to evaluating quantitative or experimental data and I was no exception. The evidence presented, however, seems to be suggestive enough to render the theoretical points valid. The paper addresses mechanisms of change that involve usage and the interaction of usage with cognitive representations. The hypotheses of this paper are further developed in the chapters presented here and in the work on lexical diffusion by Phillips (1984, 2001, and elsewhere). The phonological aspects of the schwa-deletion process discussed in this chapter are further explored in Hooper (1978).

References

Hooper, Joan B. 1978. Constraints on schwa deletion in American English. In *Recent developments in historical phonology*, ed. Jacek Fisiak, 183–207. The Hague: Mouton.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–342.

———. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In *Frequency and the emergence of linguistic structure*, ed. Joan L. Bybee and Paul Hopper, 123–126. Amsterdam: John Benjamins.

Chapter 1

Introduction

1. Background

A newcomer to the field of linguistics might be surprised to learn that for most of the twentieth century facts about the frequency of use of particular words, phrases, or constructions were considered irrelevant to the study of linguistic structure. To the uninitiated, it does not seem unreasonable at all to suppose that high-frequency words and expressions might have one set of properties and low-frequency words and expressions another. So how is it that so many professional linguists for so many decades, maybe even centuries, have missed (or perhaps avoided) this basic point?

One factor is that frequency effects tend to be observable at the level of the individual word or expression, while linguists have tended to focus their interest on the broader patterns of structure and the more abstract and generalized categories. While language is full of both broad generalizations and item-specific properties, linguists have been dazzled by the quest for general patterns. Of course, the abstract structures and categories of language *are* fascinating. But I would submit that what is even more fascinating is the way that these general structures arise from and interact with the more specific items of language use, producing a highly conventional set of general and specific structures that allow the expression of both conventional and novel ideas.

In terms of the history of the field, one can see not just the glamour but also the utility of generalization as the basis of the Neogrammarian doctrine. Many sound changes turn out to be completely regular in the sense that they affect all the words of a language with the relevant phonetic environment. This fact is interesting in its own right, but it is also the indispensable basis of the comparative method—the fact that systematic relations can be found between the sounds of the words of two languages establishes their genetic relationship. It has also become the basis of the generative position that sound change is grammar change. This latter position in particular also rules out irregularity in the spread of a sound change.

Similarly, on the level of grammar for both descriptive and theoretical purposes, a focus on general patterns is absolutely essential. In cross-linguistic studies, the discovery of typologies of word order, case marking, or morphological form would be impossible without a focus on the general patterns of a language. However, it is now time to examine the role that specific expressions and structures play in grammar and here frequency becomes important.

The other major theoretical factor working against an interest in frequency of use in language is the distinction, traditionally traced back to Ferdinand de Saussure (1916), between the knowledge that speakers have of the signs and structures of their language and the way language is used by actual speakers communicating with one another. American structuralists, including those of the generativist tradition, accept this distinction and assert furthermore that the only worthwhile object of study is the underlying knowledge of language (Chomsky 1965 and subsequent works). In this view, any focus on the frequency of use of the patterns or items of language is considered irrelevant. In fact, given Chomsky's disdainful review of Skinner's *Verbal Behavior*, any mention of frequency, repetition, or imitation has been dismissed by Chomsky's followers as irrelevant to the true nature of language.

A major challenge to this doctrine, arising in the 1970s and continuing with growing strength to the present, is a view of grammar as arising from the patterns of language use in actual discourse. This functionalist view fostered studies of grammar in the context of discourse and interaction, as well as the dynamic study of grammar as found in works on grammaticization. In this view, it is proposed that grammar arises diachronically because of the commonly used discourse patterns that humans need to communicate (Li 1975, 1977; Givón 1979; Hopper and Thompson 1980, 1984). The empirical support for this hypothesis is found in diachronic, typological, and discourse studies (e.g., DuBois 1987). The importance of the work of Joseph Greenberg in inspiring this movement cannot be overestimated. His detailed cross-linguistic investigations revealed common diachronic patterns across related and unrelated languages, allowing him to formulate both synchronic universals and universals of change. This typological work inspired others to seek to explain the implicational hierarchies he discovered by considering discourse patterns (e.g., Greenberg 1978; Givón 1979). Perhaps more germane to this work, however, is Greenberg's (1966) explicit investigations of the role of frequency in deriving the effects underlying the theory of markedness. While not proposing a causal mechanism, he demonstrates that unmarked members of categories throughout the grammar are more frequent than marked members. The argument that is gaining strength is that separating language from the way it is used removes a valuable source of explanation for why language has grammar and what form that grammar takes.

Despite the empirical bent of the functionalist movement and the acceptance of the notion that conventionalization through repetition creates grammar, there was still very little investigation into the nature of the effects of repetition or frequency on the cognitive representation of language. One reason for this gap has to do with the unfortunate separation in both method and theory between psycholinguists and linguists, such that linguists often have trouble making use of psycholinguistic findings. The other reason is that the more empirically minded functionalists, reacting negatively to the arrogance of generativists' proposing "mental representations" on very scant evidence, shied away from proposals about what might be going on in people's heads.

More recently the field of linguistics has begun to diversify in the range of evidence considered relevant and the range of views of grammar and how it relates to the rest of cognition. In particular there is a new appreciation of item-specific knowledge and how it interacts with and gives rise to more general knowledge. The role of experience in the formation of linguistic categories and representations is being investigated in what have come to be called usage-based models (Bybee 1985, this volume, chapter 8; Langacker 1987, 2000; Barlow and Kemmer 2000) and probabilistic models (Bod, Hay, and Jannedy 2003) An important force in raising the visibility of item-specific factors is the continuing work on connectionist modeling of language, where it is shown that it is possible to model the buildup of generalizations from specific instances without throwing away the knowledge of the specific instances (Rumelhart and McClelland 1986; Elman 1990; Bybee and McClelland 2005). This suggests a model of language in which the predictable patterns are not separate from representations, where lexicon and grammar are interwoven as are the specific and the general. Both type and token frequency in the input have been found to be factors important to the modeling of linguistic systems, such as the English past tense (see section 2 of this chapter).

A new movement in language acquisition focuses on specific word combinations used early by children (Lieven, Pine, and Baldwin 1997; Tomasello 2003), and how these specific combinations begin to pattern into constructions (Savage et al. 2003; Casenhiser and Goldberg 2005; Dąbrowska and Lieven 2005). In these works, both token and type frequency in the child's experience are shown to be important for acquiring the specific combinations and the generalized pattern. Rather than viewing acquisition as guided by innate categories and rules, these researchers demonstrate that children build up categories and sequential patterns using the utterances they have experienced in their own and in their caregivers' speech.

An important methodological breakthrough comes with the availability of large corpora recording natural language use, and the concordance software that allows for quantitative studies. Corpus studies often reveal quantitative patterns that are not available to introspection but that are likely to be important to the understanding of how speakers store and access units of language. Along with fairly traditional positional analysis, which for instance reveals that English *of* does not behave like other prepositions (Sinclair 1991), quantitative studies of phonetic reduction using large corpora reveal that the units of speech are grouped in production according to their specific histories of co-occurrence (Gregory et al. 1999; Jurafsky et al. 2001; this volume, chapters 11, 14, and 15).

Thus more recent theories are approaching a common ground in which it is hypothesized that specific instances of experience give rise to generalizations, and they can do so without being swallowed up themselves by the general pattern. This interaction of the specific and the general is a general property of human cognition that has been known in psychology from the late 1960s. Experiments in which subjects are exposed to patterns of dots, colored geometrical objects, and even line drawings of facial features demonstrate that subjects form categories based on similarity to a prototype that may never have appeared in the experiment, but they do so without losing the specific information about the exemplars they have experienced (Posner and Keele 1968; Medin and Schaffer 1978). These studies have shown that both similarity and frequency in experience determine the nature of categories. When the stimuli are very different and each is shown repeatedly, sensitivity to the individual patterns is relatively high; when the stimuli are very similar and each is shown only once, sensitivity to the individual patterns is relatively low and sensitivity to the prototype or central tendency is higher. The essential points are that performance with each item reflects both specific and general information and, in addition, the number of repetitions influences the nature of the category.

This type of system, often referred to as an emergent system, has the properties we find described as complex systems in nature. In complex systems, a small number of mechanisms operate in real time and with repetition lead to the emergence of what appears to be an organized structure, such as a sand dune. However, we know that a sand dune is not fixed in time and space but is ever altering and becoming. So we see that language is also always in a process of becoming—creating, losing, and re-creating structures that are never absolutely fixed, allowing for continued variation and change (Lindblom, MacNeilage, and Studdert-Kennedy 1984; Hopper 1987; Holland 1998). In such theories repetition of actions brings about the formation of structures; thus in language, too, we see that repetition is a necessary component of grammar formation (Haiman 1994). The reason frequency or repetition plays a role in grammar formation is that the mind is sensitive to repetition. This is a domain-general principle; that is, it does not apply just to language but to other cognitive domains as well.

What are the cognitive responses to repetition? There are multiple responses and their effects can change according to the extent of the repetition. The chapters of this volume represent my investigations into these effects over the last thirty years, from the first recognition of the potential explanatory power of frequency effects through a continuing effort to understand them and their interaction with processing factors, as well as phonetics, semantics, and grammar.

One might reasonably ask how it was that in the early 1970s I became interested in a topic that was so unpopular, one that indeed might brand me as a behaviorist. The source of my interest was, and remains, language change. When I was a student at UCLA, Theo Vennemann had briefly introduced the historical linguistic class I attended to the ideas of Hugo Schuchardt, who had observed that high-frequency words tend to undergo sound change before low-frequency words. Of course, this phenomenon is readily observable in the extreme reduction of high-frequency phrases. I, in turn, passed on this observation to my students at SUNY at Buffalo. Two of my students, Richard Mowrey and William Pagliuca, became very interested in phonetic reduction and frequency, which led me to consider the topic further. In teaching historical linguistics, I also noted that it was generally observed that analogy in the form of regularization affected low-frequency paradigms before high-frequency ones. This paradox seemed worth pointing out for the implications it has for the understanding the mechanisms involved in these two types of change (see this volume, chapter 2). Admittedly, at the time I did not realize the potential for both of these observations to totally change the way I looked at phonology and grammar.

When I undertook a cross-linguistic study of morphology, the importance of frequency was observable both within and across paradigms in all the languages I investigated, leading to a novel conception of morphological structure (see Bybee 1985 and in this volume part II). A view of morphological structure as affected by language use kept producing more and more interesting hypotheses and findings. Similarly, the investigation of frequency effects in phonology has led to many interesting questions and solutions to some old problems (see Bybee 2001b and in this volume part III). I feel the issues in morphosyntax may be more complex, but many interesting findings have already arisen in this usage-based conception of grammar (see part IV). This volume reports on these investigations, which together paint a picture of language as having a variegated and mutable shape. While there are many general patterns of great interest, the topology of the cognitive organization of language is neither flat, regular, nor permanent. High-frequency words and phrases grow strong with repetition and loom large, forming looser connections with other items, while low-frequency words and expressions are less prominent but gain stability by conforming to patterns used by other items. General patterns dominate networks where more specific patterns can be overpowered unless represented by high-frequency items. Words that have phonological similarities cluster together; constructions are connected if they have properties in common. Instances of constructions that grow to high frequency slowly disengage from the more general pattern to become independent constructions. Thus the phonetic and semantic substance of language is ever being shaped by the effects of usage.

In the next section I provide an overview of the frequency effects that I have investigated. At the beginning of each part of the book, I have provided a few paragraphs that put the ideas in these articles in perspective with respect to the environment in which they were written, as well as subsequent research.

2. What are frequency effects?

There are various ways of counting frequency, various units upon which to base the count, and various ways in which frequency has cognitive effects. First, I distinguish token and type frequency. Token frequency counts the number of times a unit appears in running text. Any specific unit, such as particular consonant [s], a syllable [ba], a word *dog* or *the*, a phrase *take a break*, or even a sentence such as *Your toast popped up* can have a token frequency. Type frequency is a very different sort of count. Only patterns of language have type frequency, because this refers to how many distinct items are represented by the pattern. Type frequency may apply to phonotactic sequences; it would be the count of how many words of the language begin with [sp] versus how many begin with [sf]. It may apply to morphological patterns, such as stem + affix combinations. For instance, the English past tense pattern exemplified by *know*, *knew*; *blow*, *blew* has a lower type frequency than the regular pattern of adding the *–ed* suffix. Syntactic patterns or constructions also have type frequencies: the ditransitive pattern in English, exemplified by *He gave me the change*, is used with only a small set of verbs, while the alternate pattern *He gave the change to me* is possible with a large class of verbs (Goldberg 1995). In this section I treat the three known effects of token frequency and the effect of type frequency, as well as their interaction.

2.1. Token frequency: The Conserving Effect

The first effect of token frequency to be discussed, which we can call the Conserving Effect, depends upon the fact that repetition strengthens memory representations for linguistic forms and makes them more accessible. *Accessibility* in this sense refers to the fact that in experiments where subjects are asked to say whether a string of letters or sounds is a word of their language, they respond much more quickly to high-frequency words than to low-frequency words. This greater accessibility suggests that each token of use strengthens the memory representation for a word or phrase (Bybee 1985; this volume, chapters 2, 3, 4, 5, and 8). The strength of representation of higher frequency forms explains why they resist reformation on the basis of analogy with other forms. For instance, for English irregular verbs, there is a general trend diachronically toward regularization, a trend also witnessed in child language development. However, the higher frequency verbs resist this trend; thus *keep*, *sleep*, *weep*, *leap*, and *creep* and other verbs of this shape acquired irregular past forms when the vowel was shortened in early Middle English, giving *kept*, *slept*, *wept*, *leapt*, and *crept*. Only the lower frequency verbs of this class have subsequently developed regularized pasts *weeped*, *leaped*, and *creeped* (still used alongside the irregulars). The mechanism behind this type of change (analogical reformation) is that a new past form is created by accessing the base/present form and adding the suffix *–ed* (in this case its allomorph [t]) to it. For those verbs of high frequency, the greater accessibility of the irregular past makes such a reformation unlikely. For this reason, the lower frequency paradigms tend to regularize before the higher frequency paradigms.

Another aspect of the Conserving Effect is the fact that within a paradigm it is the higher frequency form that serves as the basis for the reformation just illustrated (Mańczak 1980; Bybee 1985; this volume, chapters 3 and 4). Thus in our example, the base/present form *weep*, which is more frequent than *wept*, survives and serves as the base upon which the new past is created, in contrast to the hypothetical situation in which a new base is formed from *wept*, creating **wep* as the new base. The latter result would also be a regular paradigm, but this type of change is rare (see Tiersma 1982). Thus Mańczak (1980) has written that the most frequent form(s) of a paradigm are the most likely to resist change and to serve as the basis for change in other forms.

The Conserving Effect is also observable in morphosyntactic constructions (this volume, chapters 12 and 16). For instance, of the two forms of negation exemplified by *I know nothing about it* (negative incorporation) and *I don't know anything about it* (*not* negation), the former is older and more conservative. Tottie (1991) shows that where the two constructions are interchangeable, the older construction (negative incorporation) is used in certain high-frequency constructions, such as existential constructions, constructions with possessive *have* and copular *be*. That frequency is important to the retention of this construction is shown by the fact that it also occurs commonly with certain high-frequency verbs, such as *know*, *do*, *give*, and *make*.

The Conserving Effect also operates in the formation of new constructions, such as those involving the English auxiliary: certain properties of the auxiliaries, such as inverting with the subject, being followed by negation, and taking an infinitive without *to* are actually ancient grammatical properties that applied to all verbs. After some changes that occurred in the fifteenth through the seventeenth century, only the most frequent verbs (the modals, progressive *be*, perfect *have*, copular *be*, and possessive *have*) retained these characteristics (see this volume, chapters 12 and 16, and Bybee forthcoming).

Another interesting phenomenon is the interaction of phonological alternations with morphosyntactic constructions such as consonant liaison in French (Bybee 2001b; this volume, chapter 13). Obligatory liaison occurs primarily in grammatical morphemes (determiners, clitic pronouns, verb endings, plural on adjectives), all of which are very frequent. A corpus study by Ågren (1973) demonstrates that optional liaison is highly influenced by frequency of use, such that even forms of the same verb maintain liaison to a greater or lesser extent depending upon the frequency with which they occur in the liaison environment.

While this volume is only suggestive of the ways that morphosyntactic constructions might be affected by frequency, the articles contained here point to important areas for future research in usage-based theory. As we will see below, all of the frequency effects discussed here are highly relevant for the understanding of the cognitive representation of constructions.

2.2. Token frequency: The Reducing Effect

It is a common observation that oft-repeated phrases, such as greetings (*God be with you* > *goodbye, how are you* > *hi*) and titles, tend to reduce phonetically. The same observation applies to much-used grammatical items, such as auxiliaries, modals, negatives, and pronouns. As mentioned above, it has recently been fully documented that reductive sound change applies probabilistically across all frequency levels, affecting high-frequency items more quickly and radically than low-frequency items (this volume, chapters 2, 9, and 11). The reason for this trend is that repetition of neuromotor sequences leads to greater overlap and reduction of the component articulatory gestures. As articulation becomes more efficient, the output appears more and more to have been affected by assimilation and reduction.

What is the cause of such reduction? My view, as presented in the chapters of this book, is that phonetic reduction is directly tied to neuromotor processing: repeated sequences of neuromotor commands and actions tend to be processed as single units; at the same time, repeated sequences tend to become more efficient by the increased overlap and reduction of the gestures involved (Browman and Goldstein 1992; Haiman 1994; Mowrey and Pagliuca 1995; Boyland 1997). This domain-general process is responsible both for the fact that general reductive sound change occurs earlier in high-frequency words and that special reduction occurs in very high-frequency words and phrases. Thus frequency of use is one factor in explaining sound change.

As mentioned before, the lexical regularity of sound change is the cornerstone of the comparative method. It was argued for vociferously by the Neogrammarians of the late nineteenth century. Over the years, there have been many detractors from this strict position. Contemporaneously with the Neogrammarians, Schuchardt (1885; see Vennemann and Wilbur 1972) documented his observations that sound changes in progress saw frequent words scurrying ahead, while the less frequent ones dragged behind. Many philologists maintained open minds about regularity; the famous historian of the Spanish language Ramón Menéndez-Pidal (1950) observed that in sound change words were like leaves floating on running water: some were caught by the current and rushed ahead, while others circled in eddies and lagged behind, yet the general flow was unmistakably directional. In my first investigation of frequency effects, I attempted to document the role of high-token frequency in sound change (see chapter 2 of this collection). Taking off from this point, Phillips (1984, 2001, and other works) has studied many sound changes of various types in the light of frequency of use. The effect of frequency can be seen in many cases of more or less stable variation, as well as in changes that end up affecting all the words of a language, and turn out regular. In the latter case, the course of development of the change involves the gradual spread through the lexicon known as lexical diffusion.

The idea that high-frequency words grow shorter brings to mind the work of George Kingsley Zipf, which, however, proposes a very different explanation for the correlation of high frequency with shorter expression. In *The Psychobiology of Language*, published in 1935, Zipf documents across a number of languages the tendency for high-frequency words to be shorter than lower frequency words. Zipf takes the frequency distribution of short and long words to represent an equilibrium that languages strive to maintain. The mechanism he sees as responsible for this equilibrium is the process of "clipping" or abbreviation by which *laboratory* becomes *lab*, and so on. Mańczak's theory of irregular sound change due to frequency is similar to Zipf's:

> There is a synchronic law according to which the linguistic elements which are more often used are smaller than those which are less often used . . . If a word or morpheme becomes too short in relation to its frequency, it is replaced by a longer one. But if a linguistic element (i.e., a morpheme, word or group of words) becomes too long in relation to its frequency, it must be shortened, and then there are two possibilities: either mechanical shortening (Universität > Uni) or an irregular sound change due to frequency. (1978: 309–310)

My own view of Zipf's finding, as represented in the chapters of this volume, is that high-frequency words undergo reductive changes at a faster rate than low-frequency words. While "clipping" certainly occurs and produces this result as well, the major mechanism is gradual phonetic reduction brought about by the reduction and overlapping of articulatory gestures. In phrases of extreme high frequency this reduction may appear "irregular," as Mańczak has observed, but my view is that such reduction follows the regular phonetic patterns of the language but carried to a more extreme point than in words of lesser frequency. Thus when *I don't know* reduces, the [d] undergoes flapping and even deletion, the nasal consonant deletes, and the [o] vowel may be reduced to schwa. All of these phonetic changes affect other words of English as well. They are not "irregular"; they are just extreme instances of otherwise regular phonetic processes.

A related effect of the repetition of sequences of units is chunking, or the representation of the sequence as a single unit at a higher level (Haiman 1994; Boyland 1997; this volume, chapters 14 and 15). As *I don't know* is repeated frequently, it ceases to be produced by the concatenation of three items (as *don't* has long ago coalesced) and comes to be accessed as a single unit (this volume, chapter 14). This packaging facilitates both access and further phonetic reduction in the sequence. Chunking is particularly obvious in high-frequency word sequences, but Jurafsky et al. (2001) and other studies (Krug 1998; Brown 2004; Alba 2005) have shown that phonetic reduction between words is highly dependent upon how frequently the words co-occur, suggesting the probabilistic relations among words at all levels of frequency.

Phonetic reduction of high-frequency combinations of words is also one of the processes of change that make up the phenomenon of grammaticization (Bybee, Perkins, and Pagliuca 1994; this volume, chapter 16). As the words or phrases that enter into the grammaticization process increase in frequency, they also undergo extreme versions of phonetic reduction, as, for instance, when *be going to* reduces to [gənə] or *I'm gonna* reduces to [aimənə]. Other concomitant processes that lead to this reduction are the loss of stress on grammaticizing items and their increasing use in redundant contexts.

The Reducing Effect and the Conserving Effect at first blush seem contradictory (see this volume, chapter 2) since in one case high frequency encourages change and in the other high frequency discourages change. The contradiction is only apparent, however, because very different types of change are involved. On the one hand, the strengthening of memories makes complex units resist change by reformation on analogy with productive patterns, while on the other hand, the greater fluency and reduction of repeated units is a factor in phonetic change and semantic change. The two types of change involved here are very different. This is the point of chapter 2. (See also chapter 16 for these effects in grammaticization of morphosyntax.)

2.3. Token frequency: Autonomy

In previous work, I have identified two effects of token frequency, the two just discussed. However, I think it is worthwhile to take "autonomy" to be a third effect of token frequency even though it may be also considered an extreme instance of the Conserving Effect. In Bybee and Brewer 1980 (this volume, chapter 3) we introduce the notion of autonomy, as it applies to words (although it can also be extended to phrases), by saying it is "the extent to which a word is likely to be represented in the speaker's lexicon as a whole and separate unit" (see page 50 of this volume). Autonomy is thus probabilistic, varying for each word or phrase. Three factors are said to influence the degree of autonomy: semantic simplicity, morphophonemic regularity, and word frequency. Autonomy is based on MacWhinney's (1978) notion of lexical strength, which is a reflection of frequency in experience and discussed more thoroughly in Bybee (1985). Highly autonomous words have weaker connections to other related words—either words of the same paradigm or words of the same lexical class. The idea behind autonomy is that when words (and phrases) are highly frequent they can be accessed independently of related items and are thus not as interconnected in the network.

Autonomy can be observed to be operating in grammaticization. A common result of grammaticization is that the lexical item within a grammaticizing construction splits off from its origins: thus possessive *have* and perfect *have* in English are no longer the "same" item (Heine and Reh 1984; Hopper 1991). Besides the semantics, evidence for this is the fact that in American English possessive *have* takes the auxiliary *do* in questions and negatives and perfect *have* does not (*Do you have a driver's license?* vs. *Have you seen Betty?*). Thus the construction exemplified by *have seen*, *have gone*, *have eaten* is independent of other uses of *have*. Similarly, *be going to*, discussed in chapter 16, is autonomous from other uses of *go* as well as from other instances of the purpose construction that gave rise to it.

Bybee and Brewer (1980; this volume, chapter 3) argue that autonomy in inflectional morphology leads to resistance to change and it also allows forms to serve as the basis of reformation within a paradigm. The high frequency of autonomous forms weakens their connections to other forms, leading in extreme cases to suppletion, as when the Past form *went* split off from *wend* and became the past of *go* (Bybee 1985). Such cases of suppletion occur only in the highest frequency paradigms of a language.

In derivational morphology a related tendency is for high-frequency derived forms to split off semantically from their bases (Bybee 1985). Hay (2001) has refined this hypothesis by postulating that derived forms move away from their bases semantically if they are more frequent than their bases.

Bybee and Moder (1983; this volume, chapter 6) also argue that high-frequency inflected forms do not contribute to the productivity of morphological patterns. Thus an irregular verb such as *begin, began, begun* is of extreme high frequency and could well be autonomous from the semiproductive class to which it originally belonged. This could explain why the class it belonged to extended by adding verbs that end in velars (such as *string*, *strung*; *strike*, *struck*) rather than adding more members ending in alveolars. Moder (1992) tests this hypothesis experimentally, demonstrating that high-frequency class members are less successful at priming nonce responses than lower frequency class members.

The same principle has been shown to apply in phonotactic studies. Bailey and Hahn (2001) show that the type frequency (see section 2.4) of phonotactic patterns partially determines how acceptable native speakers will find them, but that patterns with very high token frequency and very low token frequency have less of an effect on acceptability judgments.

2.4. Type frequency

Type frequency is a property of patterns or constructions and refers to the number of distinct items that can occur in the open slot of a construction or the number of items that exemplify a pattern, such as a phonotactic sequence. Type frequency is a major factor determining the degree of productivity of a construction (Guillaume 1927/1973; MacWhinney 1978; Bybee 1985; this volume, chapters 4–8). Constructions that apply to a high number of distinct items also tend to be highly applicable to new items. In determining productivity, however, factors other than type frequency must also be taken into account: often the member items that occur with a construction must also belong to certain phonological or semantic categories. The verbs of the *string*, *strung* class must end in a nasal or a velar (chapters 5 and 6); the adjectives that can be used in the construction [X drives me (or someone) Adj], (as in *it drives me mad*, *it drives me crazy*) must suggest some degree of insanity, either literally or figuratively (Boas 2003). As pointed out in the previous section, not all members contribute equally to productivity: extreme high-frequency members of classes may not contribute at all, due to their autonomy.

The contribution of type frequency to productivity is the fact that when a construction is experienced with different items occupying a position, it enables the parsing of the construction. If *happiness* is learned by someone who knows no related words, there is no way to infer that it has two morphemes. If *happy* is also learned, then the learner could hypothesize that *–ness* is a suffix, but only if it occurs on other adjectives would its status as a suffix become established. Thus a certain degree of type frequency is needed to uncover the structure of words and phrases. In addition, a higher type frequency also gives a construction a stronger representation, making it more available or accessible for novel uses.

In addition, other factors important for productivity have been identified by Jennifer Hay and colleagues. In particular, Hay (2002) and Hay and Baayen (2002) present evidence that it is the type frequency of words within which the affix can be parsed out that determines productivity, not just the type frequency of the affix. The "parsibility" of a word depends on both phonological and semantic factors.

2.5. Frequency and category formation

The formation of grammar is in large part an issue of categorization. Every construction is characterized by the categories that fill its various slots. Every phoneme constitutes a category, often made complex by the existence of allophones. Semantic notions such as past or future tense also constitute categories, again made complex by meaning changes in specific contexts. Even the contexts of use of particular constructions constitute categories. Both type and token frequency are important to category formation.

The properties of the types included in a category establish its boundaries while the number of types relates to the degree of productivity of the construction referring to the category. In research into exemplar models (in which the category consists of the experienced exemplars), token frequency can be seen to influence the perception of the center of the category, as well as its boundaries (Nosofsky 1988). In phonetic categorization, high-frequency exemplars tend to be maintained while low-frequency ones are marginalized and lost (this volume, chapter 10). In semantic categorization, a similar phenomenon occurs; in a corpus and experimental study of the pairing of verbs that mean 'become' with adjectives in Spanish, it was found that the high-frequency pairs served as the center of some of the most productive categories (Bybee and Eddington 2006). Similarly, Casenhiser and Goldberg (2005) show that children and adults learn a new construction faster if they are exposed to one higher frequency token as well as several types that exemplify the construction.

3. How high, how low?

Throughout this introduction, I have used the terms *high* and *low frequency* in a fairly vague way. In the chapters collected here frequency is counted with a variety of cutoff points, depending upon the goals of the particular study. In addition, different sources for frequency information are consulted. There is no one method for doing frequency research. When I began to study frequency effects, the available counts were almost all based on written language. Fortunately, now there are large databases of spoken language as well as software that makes extracting frequency information from such databases fairly straightforward. Thus now it is possible to choose a database that is relevant to the particular hypotheses being tested.

When one is studying token frequency, there is an inherent problem in determining the point at which high should be distinguished from low. The problem is that there will be many tokens of high-frequency types and many types represented among the low-frequency tokens. Thus if one draws the line so that half the tokens are high frequency and half are low, the high-frequency group may have very few types in it. Conversely, if one puts half the types in the high group and half in the low group, the number of tokens in the high group will vastly outnumber those in the low group. Most studies take a compromise between these two position by looking for a natural gap in the frequency ranks that puts about 30 to 50% of the tokens in one group and 50 to 70% of the tokens in the other.

For the research I have done, the strategy of including more than two frequency groups—say high-, mid-, and low-frequency groups—has not improved the description of the data. But the possibility certainly remains that items in the medium range have their own properties and effects. Certainly it has been useful to refer to items of extreme high frequency, as these are most likely grammaticizing or exhibiting autonomy in one way or another. However, it has also been found that for types of phonetic reduction that are quite general (e.g., final [t] or [d] deletion in English) a continuous effect throughout the frequency levels can be observed (Gregory et al. 1999).

It is important not to overlook the fact that even a small number of repetitions in the speaker's experience has a cognitive effect. As pointed out in chapter 13, certain English words that occur only in fixed expressions, such as *dint*, *bated*, and *hale* in the expressions *by dint of*, *with bated breath*, and *hale and hearty*, are actually of rather low frequency (not listed at all among the 1 million words of Francis and Kučera 1982). Yet these expressions are part of the knowledge of English speakers. The pervasive use of prefabricated word sequences, which form a major part of any discourse, points to the same conclusion. Prefabs are word sequences such as *prominent role*, *beyond repair*, *to need help*, which are transparent semantically but used conventionally. Some counts show that more than half of the word choices in a spoken or written discourse are determined by membership in a prefab (Pawley and Syder 1983; Erman and Warren 2000; Wray 2002). Prefabs are not especially high in frequency, and yet it is clearly their repetition that has given them their social status of conventionalized and their cognitive status of easily accessed routines. In addition, certain phrases that are not particularly frequent, though conventionalized for particular situations, such as *who goes there?* and *how goes it?*, are conservative in using main verb inversion rather than *do*. We also find conservative behavior in phrases such as *far be it from me*, which includes a subjunctive form and an archaic word order. Perhaps related to these lower frequency repetitions, subordinate clauses often exhibit conservative behavior (Bybee 2001a).

In addition, constructions can take on special pragmatic and semantic values through repetition despite not being of especially high frequency. The special connotations of the construction discussed by Kay and Fillmore (1999), "what is X doing Y" of incongruity or disapproval, as in *What's that box doing up there?*, have to be acquired by repetition in the appropriate contexts. A few repetitions in the relevant context allow a new construction to be formed (Johnson 1997). I conclude, then, that repetition is necessary for the formation of grammar (Haiman 1994), but that different levels of repetition have different effects (Bybee forthcoming).

It is important to bear in mind, as always, that frequency interacts with other factors, such as phonological and semantic similarity, categorization, and semantic/pragmatic change. It is often difficult to discern which factors are the most important in determining linguistic behavior.

4. Frequency of what?

In order for frequency counts to be useful the researcher must know what to count. Often it is a more specific and longer string that must be counted rather than a more general and shorter one. For instance, the frequency of *don't* in itself does not account for its reduction, because it only reduces in certain phrases. Thus it is the frequency of phrases such as *I don't know*, *I don't think*, *I don't have*, and so on, that figures in its most extreme reduction. Whether the reduction can spread from these contexts into new ones is a question that remains to be answered.

Similarly, in using type frequency to study productivity, the question arises as to how similar or how specific the pattern is that counts toward productivity. For instance, in Spanish the alternations in the verbs *tener/tengo*, 'to have, 1s have', *poner/pongo* 'to put, 1s put', do not extend to nonce verbs such as **roner* (this volume, chapter 4). However, the derivatives of *tener* and *poner* all use the same alternations in their conjugations: *detener/detengo* 'to detain', *proponer/propongo* 'to propose'. Thus while the *–n/ngo* alternation does not seem productive, a longer, more specific string *tener/tengo, poner/pongo* appears to be. Thus the linguist's goal of always finding the most general patterns may not serve to reveal the significant patterns in the speaker's mental grammar.

However, sometimes the relevant comparison class is broader than the phenomenon under study. Diessel and Tomasello (2005) argue that the ease of acquisition of relative clauses does not always depend upon the frequency of the type of relative clause but also on the similarity of the word order of that clause to a simple clause.

5. Is frequency a cause or an effect?

Over the years many linguists have expressed the opinion that what is really interesting is the question of *why* certain words, expressions, or sounds are frequent.

My claim is that no matter the source of the frequency, the cognitive effects shape grammar. However, it is worthwhile to investigate the question of why certain items are frequent, even though we will find that items become frequent for many different reasons. First, there is the question of what people want to talk about. The answer is, "Themselves!" One of the most used words in English is *I*, and other first person pronouns (object and possessive) are also very frequent (Scheibman 2002).

Second, in addition to the content, the way speakers structure their discourse leads some elements to be more frequent than others. These discourse patterns are conventionalized into grammar, as summed up in DuBois's slogan: "grammars code best what speakers do most." One could question the use of the word *best* here, but presumably DuBois means that through repetition, reduction, and conventionalization grammars arise that provide economical means of expressing frequent discourse patterns. His example is the development of ergative cross-referencing out of the tendency to pronominalize or omit the agent and use full noun phrases for the absolutives (the subject of the intransitive and the object of the transitive verb).

There is also a relationship between frequency and generality or flexibility of meaning. Grammaticized elements, such as past or future markers (*will*, *be going to*), have very general and abstract meaning that can be applied to almost any verb type. Therefore, these markers tend to be of very high frequency. Polysemous items also can have a higher frequency. For instance, a phrase such as *I don't know*, which has its literal meaning in addition to a pragmatic use as a conversational hedge or turn organizer, will have a higher frequency than a phrase with only literal meaning (Scheibman 2000). Gentner (1981) demonstrates that verbs are more flexible in their interpretation and thus more polysemous than nouns are. This relative flexibility of meaning gives rise to the different frequency distributions of nouns versus verbs in English: there are more high-frequency verbs than nouns and more low-frequency nouns than verbs.

Thus the answer to the question of whether frequency is a cause or an effect is complex. On the one hand, frequency is just a tally, a pattern observable in texts, which is of course an effect. On the other hand, frequency or repetition of experiences has an impact on cognitive representations and in this way becomes a cause for the effects discussed in this book.

6. The parts of this volume

This book is divided into four parts. The second chapter of part I consists of an article that discusses two effects of token frequency; it represents my first foray into this territory and lays out the premises for much of the later work. Part II is made up of six chapters that discuss the way in which token and type frequency modulate morphological structure. Part III has three chapters that explore further how token frequency affects the spread of sound change. The five chapters of part IV apply the findings for phonetics and morphology to the level of morphosyntax. The organization of the book partly reflects chronology: part I's chapter 2 was originally published in 1976, the articles of part II appeared between 1980 and 1995, the articles of part III between 2000 and 2002, the articles of part IV between 1997 and 2003. Each part of the book begins with commentary on the chapters that follow.

References

Ågren, John. 1973. Etude sur quelques liasons facultatives dans le français de conversation radiophonique: Fréquence et facteurs. Uppsala: Acta Universitatis Upsaliensis. Studia Romanica Upsaliensa 10.

Alba, Matt. 2005. Hiatus resolution between words in New Mexican Spanish: A usage-based account. Ph.D. diss., University of New Mexico.

Bailey, Todd M., and Ulrike Hahn. 2001. Determinants of wordlikeness: Phonotactics or lexical neighborhoods? *Journal of Memory and Language* 44, 568–591.

Barlow, Michael, and Suzanne Kemmer. 2000. *Usage-based models of language.* Stanford, Calif.: CSLI.

Boas, Hans C. 2003. *A constructional approach to resultatives.* Stanford, Calif.: CSLI.

Bod, Rens, Jennifer Hay, and Stefanie Jannedy (eds.). 2003. *Probabilistic linguistics*. Cambridge, Mass.: MIT Press.

Boyland, Joyce Tang. 1997. Morphosyntactic change in progress: A psycholinguistic approach. Ph.D. diss., University of California at Berkeley.

Browman, Catherine P., and Louis M. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49, 155–180.

Brown, Esther. 2004. Reduction of syllable-initial /s/ in the Spanish of New Mexico and Southern Colorado: A usage-based approach. Ph.D. diss., University of New Mexico.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins.

———. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10, 425–455.

———. 2001a. Main clauses are innovative, subordinate clauses are conservative: Consequences for the nature of constructions. In *Complex sentences in grammar and discourse: Essays in honor of Sandra A. Thompson*, ed. Joan L. Bybee and Michael Noonan, 1–17. Amsterdam: John Benjamins.

———. 2001b. *Phonology and language use.* Cambridge: Cambridge University Press.

———. forthcoming. From usage to grammar: The mind's response to repetition. *Language* 82.

Bybee, Joan L., and Mary Alexander Brewer. 1980. Explanation in morphophonemics: Changes in Provençal and Spanish preterite forms. *Lingua* 52, 201–242.

Bybee, Joan L., and David Eddington. 2006. A usage-based exemplar model approach to Spanish verbs of "becoming." *Language* 82, 323–355.

Bybee, Joan L., and James McClelland. 2005. Alternatives to the combinatorial paradigm of linguistic theory based on domain general principles of human cognition. In *The role of linguistics in cognitive science*, ed. Nancy A. Ritter. Special Issue of The *Linguistic Review* 22(2–4), 381–410.

Bybee, Joan, Revere D. Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world.* Chicago: University of Chicago Press.

Casenhiser, Devin, and Adele Goldberg. 2005. Fast mapping between a phrasal form and meaning. Paper presented at the International Association for the Study of Child Language Development. Berlin.

Chomsky, Noam. 1965. *Aspects of the theory of syntax.* Cambridge, Mass.: MIT Press.

Dąbrowska, Ewa, and Elena Lieven. 2005. Towards a lexically specific grammar of children's question constructions. Paper presented at the International Association for the Study of Child Language Development. Berlin.

Diessel, Holger, and Michael Tomasello. 2005. A new look at the acquisition of relative clauses. *Language* 81.4, 882–906.

DuBois, John. 1987. The discourse basis of ergativity. *Language* 63, 805–855.

Elman, Jeffrey L. 1990. Finding structure in time. *Cognitive Science* 14, 179–211.

Erman, Britt, and Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text* 20, 29–62.

Francis, W. Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Gentner, Dedre. 1981. Some interesting differences between nouns and verbs. *Cognition and Brain Theory* 4, 161–178.

Givón, Talmy. 1979. *On understanding grammar.* New York: Academic Press.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure.* Chicago: University of Chicago Press.

Greenberg, Joseph. 1966. *Language universals: With special reference to feature hierarchies.* The Hague: Mouton.

———. 1978. How does a language acquire gender markers? In *Universals of Human Language*, vol. 3: *Word Structure*, ed. Joseph H. Greenberg, Charles A. Ferguson, and Edith A. Moravcsik, 47–82. Stanford, Calif.: Stanford University Press.

Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *CLS* 99, 151–166.

Guillaume, P. 1927/1973. The development of formal elements in the child's speech. In *Studies of child language development*, ed. C. A. Ferguson and Dan I. Slobin, 240–251. New York: Holt, Rinehart and Winston.

Haiman, John. 1994. Ritualization and the development of language. In *Perspectives on grammaticalization*, ed. William Pagliuca, 3–28. Amsterdam: John Benjamins.

Hay, Jennifer. 2001. Lexical frequency in morphology: Is everything relative? *Linguistics* 39(6), 1041–1070.

———. 2002. From speech perception to morphology: Affix ordering revisited. *Language* 78, 527–555.

Hay, Jennifer, and Harald Baayen. 2002. Parsing and productivity. In *Yearbook of morphology 2001*, ed. Geert Booij and Jaap van Marle, 203–235. Dordrecht: Klewer Academic.

Heine, Bernd, and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske Verlag.

Holland, John H. 1998. *Emergence: From chaos to order*. New York: Basic Books.

Hopper, Paul J. 1987. Emergent grammar. *BLS* 13, 139–157.

———. 1991. On some principles of grammaticization. In *Approaches to grammaticalization*, vol. 1, ed. Elizabeth C. Traugott and Bernd Heine, 17–35. Amsterdam: John Benjamins.

Hopper, Paul, and Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, 251–299.

———. 1984. The discourse basis for lexical categories in universal grammar. *Language* 60, 703–752.

Johnson, Christopher. 1997. The acquisition of the "What's X doing Y?" construction. *Proceedings of the 21st annual Boston University Conference on Language Development* 2, 343–353. Somerville, Mass.: Cascadilla Press.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency* *and the emergence of linguistic structure*, ed. Joan L. Bybee and Paul Hopper, 229–254. Amsterdam: John Benjamins.

Kay, Paul, and Charles Fillmore. 1999. Grammatical constructions and linguistic generalizations: The What's X doing Y? construction. *Language* 75, 1–33.

Krug, Manfred. 1998. String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26, 286–320.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical prerequisites*. Stanford, Calif.: Stanford University Press.

———. 2000. A dynamic usage-based model. In *Usage-based models of language*, ed. Michael Barlow and Suzanne Kemmer, 1–63. Stanford, Calif.: CSLI.

Li, Charles N., ed. 1975. *Word order and word order change*. Austin: University of Texas Press.

———, ed. 1977. *Mechanisms of syntactic change*. Austin: University of Texas Press.

Lieven, Elena V. M., Julian M. Pine, and Gillian Baldwin. 1997. Lexically-based learning and early grammatical development. *Journal of Child Language* 24, 187–220.

Lindblom, Björn, Peter MacNeilage, and Michael Studdert-Kennedy. 1984. Self-organizing processes and the explanation of phonological universals. In *Explanations for language universals*, ed. Brian Butterworth, Bernard Comrie, and Östen Dahl, 181–203. New York: Mouton.

MacWhinney, Brian. 1978. *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development, no. 174, vol. 43.

Mańczak, Witold. 1978. Irregular sound change due to frequency in German. In *Recent developments in historical phonology*, ed. Jacek Fisiak, 309–319. The Hague: Mouton.

———. 1980. Laws of analogy. *Historical morphology*, ed. Jacek Fisiak, 283–288. Berlin: Mouton.

Medin, Douglas L., and M. M. Schaffer. 1978. Context theory of classification learning. *Psychological Review* 85, 207–238.

Menéndez-Pidal, Ramón. (1950). *Orígenes del español: Estado lingüístico de la península ibérica hasta el siglo XI*. Madrid: Espasa-Calpe.

Moder, Carol Lynn. 1992. Productivity and categorization in morphological classes. Ph.D. diss., SUNY at Buffalo.

Mowrey, Richard, and William Pagliuca. 1995. The reductive character of articulatory evolution. *Rivista di Linguistica* 7(1), 37–124.

Nosofsky, R. M. 1988. Similarity, frequency and category representation. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14:54–65.

Pawley, Andrew, and Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In *Language and communication*, ed. Jack C. Richards and Richad W. Schmidt, 191–226. Longman: London.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–342.

———. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In *Frequency and the emergence of linguistic structure*, ed. Joan L. Bybee and Paul Hopper, 123–126. Amsterdam: John Benjamins.

Posner, Michael I., and Steven W. Keele. 1968. On the genesis of abstract ideas. *Journal of Experimental Psychology* 77, 353–363.

Rumelhart, David E., and James L. McClelland. 1986. On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In *Parallel distributed processing: Explorations in the microstructure of cognition*, ed. James L. McClelland, David E. Rumelhart, and the PDP Research Group, 216–271. Cambridge, Mass.: MIT Press.

Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.

Savage, Ceri, Elena V. M. Lieven, Anna Theakston, and Michael Tomasello. 2003. Testing the abstractness of young children's linguistic representations: Lexical and structural priming of syntactic constructions? *Developmental Science* 6, 557–567.

Scheibman, Joanne. 2000. *I dunno but*: A usage-based account of the phonological reduction of *don't* in American English conversation. *Journal of Pragmatics* 32(1), 105–124.

———. 2002. *Point of view and grammar: Structural patterns of subjectivity in American-English conversation.* Amsterdam: John Benjamins.

Schuchardt, Hugo. 1885. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: Robert Oppenheimer.

Sinclair, John. 1991. *Corpus, concordance, collocation*. Oxford: Oxford University Press.

Tiersma, Peter. 1982. Local and general markedness. *Language* 58, 832–849.

Tomasello, Michael. 2003. *Constructing a language.* Cambridge, Mass.: Harvard University Press.

Tottie, Gunnel. 1991. Lexical diffusion in syntactic change: Frequency as a determinant of linguistic conservatism in the development of negation in English. In *Historical English syntax*, ed. Dieter Kastovsky, 439–467. Berlin: Mouton de Gruyter.

Vennemann, Theo, and Terence Wilbur. 1972. *Schuchardt, the Neogrammarians and the transformational theory of phonological change*. Frankfurt: Athenäum.

Wray, Alison. 2002. *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Zipf, George K. 1935. *The psychobiology of language*. Boston: Houghton Mifflin.

Chapter 2

Word Frequency in Lexical Diffusion and the Source of Morphophonological Change

Two ideas that have been around for almost a century need to be reexamined in the light of our current theories of linguistic change. These ideas involve the relationship between word frequency and morphophonological change. I will divide morphophonological change in the traditional way into sound change, which I view as being largely, perhaps entirely, phonetically motivated, and analogy, which I consider to be primarily conceptually motivated. The particular point I want to discuss is that the relation between sound change and word frequency is just the reverse of the relation between analogy and word frequency. That is, as Schuchardt observed in 1885, sound changes affect the most frequent lexical items first (1885 [1972]). And Paul, around 1886, observed that analogical leveling tends to affect infrequent items first (1890 [1972]). Or, stated differently, infrequent items are the most resistant to phonetically motivated change, while frequent items are the most resistant to conceptually motivated change. I have gathered some data that document these two tendencies, and I will present these data before speculating on the meaning of this difference between the types of change for theories of the source of morphophonological change. In particular I will speculate on the role of the child versus the adult speaker in the initiation of linguistic change.

1. Phonetically conditioned change

First consider phonetically motivated change, sound change. It has often been observed that reductive changes affect frequent words and phrases, thus *God be with* *you* becomes *good-bye*, *vuestra merced* becomes *usted*. In English, auxiliaries and object pronouns contract. Leslau (1969) has observed similar reductions of very frequent verbs in Ethiopian languages. Furthermore, Schuchardt observed that general sound changes affect very frequent items before they affect infrequent items. Some recent investigations support Schuchardt's hypothesis. For example, Fidelholtz (1975) has shown that vowel reduction in initial syllables in English correlates significantly with frequency: the more frequent the word is, the more likely it is to have a reduced vowel.

I have gathered some data on schwa deletion in English that also support Schuchardt's hypothesis. Schwa deletion is in an extremely variable state in American English. Unstressed schwas tend to be deleted generally except for a few phonologically constrained environments that I will mention later. Yet there are some words for which a schwa-less pronunciation is clearly the norm and others for which a schwa-ful pronunciation is preferred; words of this sort troubled Zwicky (1972): compare *nursery* and *cursory*, *scenery* and *chicanery*, *celery* and *artillery*,[1] *memory* and *armory*. I was interested in finding out if the frequency of a word had any correlation to the progress of schwa deletion with regard to that word, so I devised a test, which I tried out on a few subjects.

To get a corpus of words with phonologically similar environments, I extracted from Dolby and Resnikoff, *The English Word Speculum*, (1967) all words ending in stressed $V\ C_o\ V\ r\ y$.[2] I checked the frequency of these words in Kučera and Francis (1967).[3] I took out the words not listed in the Kučera and Francis frequency count and was left with a list of 112 words. To find out which of these words undergo schwa deletion, I could not consult a dictionary as Fidelholtz did for his vowel reduction data, because dictionary pronunciations are extremely conservative and usually give a pronunciation with the schwa. The ideal way to find out which words undergo schwa deletion would be to listen to speakers' pronunciations in spontaneous conversation, but this is not practical, for the amount of data one can gather in this way is far out of proportion to the amount of time it would take. Instead, I asked speakers rather directly how these words are normally pronounced: I gave eight speakers a list of the 112 words and asked them to classify the words into one of three categories: Usually delete, Sometimes delete, and Rarely delete. All of the subjects but one were graduate students in linguistics, so they were able to understand the directions. None of them knew my hypothesis. This was a difficult task, and in some cases the decisions were rather arbitrary. A speaker's judgment on any one word may be different today from what it was yesterday.[4] Since the conclusions are all stated in rather general terms, that is, in terms of the average frequencies of the words in each category, I make no claims about any particular word. It should also be remembered that my test shows only the speaker's conception of how these words are pronounced; it does not reveal any phonetic reality.

Table 2.1 gives the results for each individual speaker. There are four columns; the first three represent the three categories the subjects used to categorize the words. The first number in each column is the average frequency of the words the subject put into this category; the number in parentheses is the number of words in the category. The fourth column represents adjustments made on the Rarely category. As I suspected, there are certain phonological conditions that systematically block schwa

TABLE 2.1. Average Frequency of Words Categorized for the Application of Schwa-deletion

	Usually	*Sometimes*	*Rarely*	*Rarely without Phonologically Constrained Words*
A	38.75 (16)	13.95 (38)	14.17 (53)	6.40 (f, c, a)
B	[a]	23.87 (46)	12.33 (66)	5.81 (f, c, a)
C	22.00 (67)	5.54 (24)	15.00 (20)	8.70 (c)
D	24.03 (37)	20.44 (33)	8.76 (42)	5.81 (c)
E	29.57 (35)	3.90 (11)	12.67 (66)	5.48 (f, c, a)
F	10.38 (8)	31.44 (25)	13.11 (79)	11.77 (f, c)
	20.75 (4)			
G	27.24 (45)	21.27(11)	9.04 (49)	7.28 (f, c)
H	35.04 (28)	7.5 (10)	11.57 (74)	9.93 (f, c)
Average of all subjects	26.72	15.99		6.18

Average frequency of all 112 words: 17.07

[a]Subject B did not put any words in the Usually category.

deletion. When words that do not have schwa deletion for phonological reasons are taken out of the Rarely category, as we might expect, the average frequency of words in this category falls to the level that we see in the fourth column.

There are three phonological environments that block schwa deletion before *r*. One is the presence of a preceding *flap*, derived historically from a *t* (and possible from a *d*, but there are so few words in the corpus with *-dVry* that no conclusions can be drawn).[5] The flap appears where a /t/ was preceded by a vowel or /r/, for example, *artery*, *flattery*, but not where a consonant precedes /t/, for example, *factory*, *adultery*, *elementary*. In the latter forms, deletion may occur, but in the former forms, the eight words with flaps, *artery*, *battery*, *buttery*, *flattery*, *lottery*, *pottery*, *rotary*, and *watery*, deletion does not usually occur.

Deletion is also constrained if a syllable-initial consonant cluster precedes the schwa. In *burglary* a deletion would leave a totally unacceptable consonant cluster of *glr* for syllable-initial position. Similarly, deletion is constrained in *library*, *contrary*, *mercury*, and *penury*, although [laybri] and [mɚkri] are possible pronunciations. Finally, a preceding affricate, as in *forgery* or *butchery*, constrains deletion for many speakers.

The individual speakers displayed different phonological constraints, so the fourth column was computed separately for each subject. I considered the subject to be exercising a phonological constraint if all or all but one of the words of a phonological type appeared in the Rarely category. When this occurred, all the words of that type were removed and the average was recomputed. However, if a subject put two or more words of a phonological type in the Sometimes or Usually category, I did not take the words of this type out of the Rarely category. The lowercase initials following the fourth column indicate which constraints were considered applicable for that subject: *f* stands for flap, *c* for cluster, *a* for affricate. Thus subjects A, B, and

E put all flap words, affricate words, and cluster words in the Rarely category; for these subjects all words of these categories were removed to obtain the fourth column average. However, subjects C and D had two or more flap words and two or more affricate words in the Sometimes and Usually categories, so flap and affricate words were averaged in to get the fourth column score.

Now if we examine the figures we see that the overall pattern is what Schuchardt's hypothesis predicts, although there are a few problems of detail, which I will return to. To make the general pattern more apparent, compare the highest and lowest categories for each speaker. The highest category has at worst an average frequency three times greater than the lowest category and at best an average frequency six times greater than the lowest (cf. subject A versus subject F).

The major problems in table 2.1 show up in the middle category, where for three subjects the average frequency of the Sometimes category is higher than the average frequency of the Rarely category. I attribute this largely to the difficulty of assigning words to three arbitrary categories. In fact, given the difficulty of the task, it is amazing that the results are so consistent. Notice also that for subjects E and H the Sometimes category has so few members that their average is hardly comparable to the other averages.

The only other incongruity in table 2.1 occurs in the Usually column for subject F. The average frequency for his Usually category is much lower than his Sometimes category. However, his Usually category contained only eight words, and of these five were words ending in *-mentary*. It appears that while he is a very conservative deletor in general, for him certain words or formatives are restructured, and *-mentary* is one of them. If the five *-mentary* words are counted as a single item the frequency average for this category doubles to 20.75. Still the average for this category is lower than for his Sometimes category, but with the five *-mentary* words counted as one item there are only four items in this category, which makes the average considerably less reliable and not comparable to the other averages. For this reason, in table 2.2 I used subject F's Sometimes category, where for the other subjects I used the Usually category.

I will only mention two other obvious problems with the test. First, the frequency count used is a frequency count based on written texts. Though a variety of types of texts were used, written texts will never reflect spoken language accurately. But the

TABLE 2.2. High and Low Categories

	High	*Low*
A	38.75	6.40
B	23.87	5.81
C	22.00	8.70
D	24.03	5.81
E	29.57	5.48
F	31.44	11.77
G	27.24	7.28
H	35.04	9.93
Total average	28.99	6.18

primary difference in word frequency between written and spoken language will be that frequent words in written language will be even *more* frequent in spoken language, while infrequent words in written language will be even more infrequent or totally non-existent in spoken language. Thus the difference in frequency found in written texts will be increased, probably in geometric proportions, in spoken language. This means that the frequency differences displayed in tables 2.1 and 2.2 would be greatly increased if I had used a frequency count of spoken language.

Another problem may be the fact that only eight subjects were used. However, the findings are quite consistent, and it seems to me in the case used here, of schwa deletion, the hypothesis is really non-controversial. This is because it is a reductive type change and I do not think anyone would deny that reduction in frequent items is to be expected. It would be more interesting (although more difficult) to run more extensive tests on sound changes in progress that are not so obviously reductive, for example, assimilations and vowel shifts. I would venture to guess that all sound changes, even those that we do not consider to be reductive, take place first in frequent words.

Let me make clear my interpretation of the status of schwa deletion in English. I claim that schwa deletion, especially poststress schwa deletion, represents a "sound change" in progress. It is potentially totally general, but at present it has not yet reached all the lexical items of English. At present schwa deletion may appear sporadic. Similarly, vowel reduction in English appeared sporadic to Fidelholtz, who referred to it as "an exceptional rule" and concluded that "frequent words can do exceptional things." I do not think that frequent words do "exceptional things"; rather, they are merely ahead of other words in undergoing phonetic change. My view is more like that of Schuchardt, who says "every sound change in some phase is sporadic" (1885 [1972]: 63), except that I would say that every sound change in some phase *appears* to be sporadic until we know what all the variables are. I will return in section 3 to a discussion of the implications of word frequency as a variable affecting sound change.

2. Analogical leveling

Analogical leveling appears to affect less frequent forms; frequent paradigms stubbornly cling to their suppletion. (I am leaving analogical extension out of the discussion.) While our general experience supports this idea, I wanted to check the hypothesis a little more rigorously before basing any claims on it. In particular I looked at classes of Old English strong verbs that have some members that survive as strong verbs in Modern English and some members that have become weak verbs, that is, that have undergone analogical leveling in Modern English. Using *Sweet's Anglo-Saxon Primer* (Davis 1965), I found three appropriate classes of verbs.[6] The frequency counts for these verbs (again from Kučera and Francis) are listed in table 2.3. To obtain these frequency counts, I totaled all entries that could be verb forms, for example, *drive*, *drives*, *driven*, *drove*, and *driving*. In some cases, however, the frequency score will include non-verbs that are spelled identically, since the frequency count is done totally on the basis of spelling, disregarding the words' function in the sentence. The verbs whose frequency score may include non-verbs and thus overrepresent the frequency of the verb, are marked with asterisks.

TABLE 2.3. Frequency of Leveled versus Unleveled Old English Strong Verbs

Strong Verbs			*Strong Verbs That Have Become Weak*	
Class I				
*drive	208		bide	1
*rise	280		reap	5
*ride	150		*slit	8
write	599		*sneak	11
*bite	128	Partially leveled *shine 35		
Average frequency	273.00		Average frequency	6.25
Class II				
choose	177		rue	6
*fly	119		seethe	0
*shoot	187		*smoke	59
lose	274		*float	23
flee	40		shove	16
Average frequency	159.40		Average frequency	32.50
Class VII				
*fall	338		*wax	19
*hold	498		weep	31
know	1227		*beat	96
grow	257		hew	1
blow	81		*leap	42
			mow	1
Average frequency	473.80		sow	3
			*flow	95
			*row	53
			Average frequency	37.89

In the left-hand column are strong verbs that remained strong; in the right-hand column are strong verbs that lost the vocalic alternation. I used all the verbs that Sweet listed under each class, except, of course, the verbs that do not survive in Modern English. The results overwhelmingly bear out the hypothesis; in fact, the frequency differences are so great that I am certain an expanded corpus would produce the same results.

Let me comment briefly on a few details: (1) In Class I, *shine* is not classified, because as an intransitive verb it has, presumably, the past tense and past participle *shone*; but as a transitive verb, it is weak and has the past tense and past participle form *shined*. (2) In Class VII, *weep* and *leap* are considered to be leveled, despite the lax vowels occurring in *wept* and *leapt*, because this lax vowel is a result of an early Middle English vowel shortening (Moore 1968, 67) and is not the same as the ablaut found in the Old English strong verbs.

A problem with the results displayed in table 2.3 is that the frequency count used was based on Modern English, but the analogical leveling took place sometime during the last ten centuries. However, since the results show such a striking difference in frequency between leveled and nonleveled forms, I do not think a more accurate frequency count would alter the general picture. A way to avoid this problem would be to study modern leveling. One case I have investigated involves the six verbs *creep*, *keep*, *leap*, *leave*, *sleep*, and *weep*, all of which have a past form with a lax vowel (due to the Middle English laxing mentioned earlier). Of these verbs, three, *creep*, *leap*, and *weep*, all may have, at least marginally, a past forms with a tense vowel, *creeped*, *leaped*, and *weeped*. The other three verbs are in no way threatened by leveling; past forms **keeped*, **leaved*, **sleeped* are clearly out of the question. Now consider the frequency differences among these verbs, in table 2.4. Again the hypothesis that less frequent forms are leveled first is supported.

3. Implications for the source of change

Of course I have not examined enough data to produce strong statistical support for the two tendencies upon which I will base my theoretical speculations, but this small body of data satisfies me that the intuitions of linguists over the years have been essentially correct: phonetic change tends to affect frequent words first, while analogical leveling tends to affect infrequent words first. The implication of this difference is that phonetic change and analogical change must be treated differently in both diachronic theory and synchronic theory.

With regard to theories of diachronic change, I would like to explore the implications of the reverse tendencies for theories concerning the source of linguistic change. I will argue that since the two types of change diffuse through the lexicon in opposite manners, we must posit two different sources for these two types of change.

One hypothesis about the source of linguistic change, which I will call the imperfect learning hypothesis, holds that a language is changed in the process of being transmitted from one generation to the next. This view has been proposed by Halle (1964), Kiparsky (1968), King (1969), Stampe (1969, 1973), and others. One point of view, stated in Kiparsky (1970), is that analogical change comes about as the result of overgeneralization on the part of children acquiring their language. This claim is quite plausible: the similarities between children's overgeneralization and analogical levelings are quite striking and familiar. The tendency I have discussed here concerning the lexical diffusion of analogical change fits in quite well with the imperfect learning hypothesis. If children are going to get away with any overgeneralization, it is more likely to be in infrequent forms, where it will not be noticed so much. If they are going to learn any suppletive paradigms, these will

TABLE 2.4. Modern English Leveling

Not Subject to Leveling		*Subject to Leveling*	
keep	531	*creep	37
*leave	792	*leap	42
*sleep	132	weep	31

surely be the more frequent paradigms; for frequent forms there is a greater availability of the model, more opportunity to practice, and greater pressure to conform. (Notice that I am not claiming that children acquire the most frequent words first.) Actually, the imperfect learning hypothesis need not be restricted to young children. Adult speakers may very well play a role in analogical leveling. For infrequent paradigms adults may not be exactly sure of all the forms, or the mere infrequency of a suppletive paradigm makes an analogical formation more acceptable. For instance, as I mentioned earlier, *creeped* is not standard, but I would not flinch if I heard it, and I might even produce it myself, although I know *crept* is "correct." However, *keeped* would definitely cause a negative reaction, because the form *kept* is so solidly established, due to its frequency. Thus I conclude that the imperfect learning hypothesis is quite a reasonable hypothesis about analogical change, but it need not be applied only to children.

The imperfect learning hypothesis has also been proposed to explain phonetic changes, chiefly by David Stampe (1969, 1973). Again we find certain correspondences between the substitutions that children make and known historical sound changes, although the correspondence is by no means perfect (Dressler 1974). Stampe claims that children possess a set of innate phonological processes that they must learn to suppress in order to acquire a language. The failure to suppress a natural process may result in a sound change.

If the reasoning I have followed in explaining how imperfect learning allows analogical change is correct, that is, if it is correct that the forms that are most likely to be imperfectly learned are the infrequent forms, then phonetic changes such as schwa deletion and any other change that follows the same pattern of lexical diffusion cannot be attributed to imperfect learning. A child's command of the phonetic shape for a frequent word will certainly be closer to the adult model than for an infrequent word, since as I mentioned earlier, for frequent words the model is more available and the pressure to conform is greater. Thus we cannot postulate that children initiate phonetic changes such as vowel reduction or schwa deletion.

What, then, is the source of a change such as schwa deletion? I suggest that the source and cause of such changes should be sought in adult speech, especially the casual style of speech. It is no accident that "new rules" appear to start out as optional or variable rules in the most casual styles of speech. These phonetic changes (or rules) are naturally occurring in casual or fast speech and are due to a tendency to minimize effort; thus vowels move closer to the neutral position, as in Fidelholtz's vowel reduction, and schwas grow shorter and shorter until they are simply skipped. Or these phonetic changes come about as changes in the timing of articulatory gestures—thus we get assimilations as in *didya*, *wouldya*, and so on.

Zipf (1935 [1965]) would claim that these phonetic changes originate because certain sounds or sequences of sounds are frequent. I would suggest, however, that these phonetic changes are motivated solely by phonetics, not frequency, but that they affect frequent items first because it is the frequent words that make up the substance of casual speech. That is, the more casual the style, the greater the chance that the lexical items used will be the more frequent ones. If schwa deletion in its initial stages is restricted to casual speech, then only the words used in casual speech will have variable forms.

After the initial stages, the language acquisition process begins to play a role, for it is probably in the transmission of the language that restructuring takes place. A very frequent word such as *every* may be variable for an adult speaker, but a new learner may take the schwa-less pronunciation to be the norm or base form. The younger speaker will have a schwa-less underlying form, and the process will be complete for that word. The change in underlying forms will progress similarly over several generations.

This is only a very general outline of how changes originating in casual speech may become permanent changes in lexical items. A number of problems remain. These concern in particular a synchronic grammar, such as that of present-day American English, which must represent an ongoing change such as schwa deletion. For example, if schwa deletion is a "rule," as we normally think of rules, how can it be made to affect different words in different ways? I do not think the relative frequency of words is a part of native speaker competence, so I would not propose to make the rule sensitive to word frequency. We could propose that schwa deletion is not a rule in the ordinary sense of "rule," but then we must discover what sort of phenomenon it is.

Let me now draw some conclusions for diachronic linguistics. I have suggested earlier that I expect that sound changes that we do not ordinarily think of as reductions, for example, vowel shifts, would also affect frequent items before infrequent items.[7] If it turns out that they do, then we can make the general claim that all sound change initiates in language use or casual speech rather than in the acquisition process. However, if it turned out that vowel shifts and some other phonetic changes affect infrequent forms before frequent forms, then we would have an interesting indication that phonetic changes arise from different sources, and furthermore, if my hypotheses are correct, a way of determining which types of changes are traceable to which source. Thus it appears that lexical diffusion, studied in terms of word frequency, may turn up some interesting evidence concerning the source of morphophonological change.

Appendix: Word list

Accessory	Factory	Plenary
Adultery	Feathery	Plenipotentiary
Anniversary	Fernery	Pottery
Armory	Fishery	Powdery
Artery	Flattery	Quackery
Artillery	Forgery	Raillery
Battery	Gallery	Recovery
Beggary	Glossary	Refectory
Boundary	Granary	Refractory
Bravery	Haberdashery	Robbery
Brewery	Hickory	Rosary
Burglary	History	Rotary
Butchery	Illusory	Salary
Buttery	Injury	Satisfactory

Calvary
Cannery
Celery
Century
Chancery
Chicanery
Complementary
Complimentary
Compulsory
Contradictory
Contrary
Creamery
Cursory
Debauchery
Delivery
Desultory
Directory
Discovery
Dispensary
Documentary
Drapery
Drudgery
Elementary
Embroidery
Introductory
Ivory
Leathery
Library
Lottery
Luxury
Machinery
Mastery
Memory
Mercury
Misery
Mockery
Mystery
Nursery
Papery
Parliamentary
Penury
Peppery
Peremptory
Perfumery
Perfunctory
Periphery
Perjury
Savory
Scenery
Sedimentary
Sensory
Shivery
Shrubbery
Silvery
Slavery
Slippery
Snobbery
Sorcery
Spidery
Splintery
Summary
Supervisory
Supplementary
Surgery
Treachery
Treasury
Unsatisfactory
Vagary
Victory
Watery

Notes

I am very grateful to Richard Mowrey for many stimulating discussions on the topic of this paper.

1. Kučera and Francis (1967) give a frequency of 4 for *celery* and 11 for *artillery*; this is due to the type of written texts they use. I am certain that for speakers other than militarists *celery* is the more frequent of the pair in spoken language. This pair illustrates clearly a general problem with relying on frequency counts of written material. Since I have computed the results of the test in terms of average frequencies for groups of words, this problem does not interfere greatly with the validity of my conclusions.

2. The "double-standard" reverse word list was used. It is a list of all left-justified, boldface words, with standard meanings occurring in both the *Shorter Oxford English Dictionary* and the *Merriam Webster New International Dictionary*, third edition.

3. For nouns total the frequency counts for singulars and plurals to obtain the figure I used.

4. One of the problems that made the task difficult is that particular words may have deletion or not, depending on how they are used. Some examples: (i) [mɚkri] is a possible shortening of *mercury* only when it refers to the make of an automobile, but never for the element or planet; (ii) Richard Mowrey (personal communication) observed the following: *parliamentary* reduces in *a parliamentary procedure*, but not in *Their form of government is parliamentary*; *rosary* for this speaker usually has a schwa, but in the name Rosary Hill (a college in Buffalo) the schwa is deleted. These examples suggest that the degree of stress may condition the presence or absence of the schwa in words that have variable form.

5. The phonological constraints on schwa deletion are due to syllable structure constraints. These are explained in detail in Hooper (1976). The deletion under discussion here

creates a new syllable-initial cluster; e.g., *hickory* gives [hIkri]. Under the theory of syllable structure developed in Vennemann and Wilbur (1972) and Hooper (1976), a consonant in initial position must be stronger than the consonant in second position. The flap in English is an extremely weak consonant; I suggest that it is too weak to cluster with an /r/. In order to have a cluster with an /r/, the alveolar consonant must be pronounced as a stop, either [t^h] or [d]. The data show, however, that speakers prefer to keep the schwa, rather than pronouncing a full stop here.

Observe further that the phonological constraints operating here are not language-specific constraints, for many violations of English syllable structure are allowed; e.g., nasal plus /r/ is allowed, *memory*, *scenery*; /s/ plus /r/, *nursery*. Neither of these clusters occurs ordinarily as a syllable-initial cluster in English.

6. Sweet's other classes of strong verbs were not appropriate, either because they had too few surviving members or because leveling occurred according to phonological subtype.

7. Labov (1966) finds that the degree of vowel shift is greater for some speakers in casual styles than in more formal styles, but this fact in itself does not prove that the vowel shifts originate in casual speech.

References

Davis, Norman. 1965. *Sweet's Anglo-Saxon primer*. Oxford, Clarendon Press.

Dolby, J. L., and H. L. Resnikoff. 1967. *The English word speculum*. The Hague: Mouton.

Dressler, Wolfgang. 1974. Diachronic puzzles for natural phonology, *Papers from the Parasession on Natural Phonology*, 95–102. Chicago: Chicago Linguistic Society.

Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. In annual *Papers from the eleventh meeting of the Chicago Linguistic Society*, 200–213. Chicago: Chicago Linguistic Society.

Halle, Morris. 1964. Phonology in generative grammar. In *The structure of language*, ed. J. A. Fodor and J. J. Katz, 334–352. Englewood Cliffs, N.J.: Prentice-Hall.

Hooper, Joan B. 1976. *An introduction to natural generative phonology*. New York: Academic Press.

King, Robert D. 1969. *Historical linguistics and generative grammar*. Englewood Cliffs, N.J.: Prentice-Hall.

Kiparsky, Paul. 1968. Linguistic universals and linguistic change. In *Universals in linguistic theory*, ed. Emmond Bach and Robert Harms, 170–202. New York: Holt, Rinehart and Winston.

———. 1970. Historical linguistics. In *New horizons in linguistics*, ed. John Lyons, 302–315. London: Penguin.

Kucera, Henry, and Nelson Francis. 1967. *Computational analysis of present-day American English*. Providence, R.I.: Brown University Press.

Labov, William. 1966. *The social stratification of English in New York City*. Arlington: Center for Applied Linguistics.

Leslau, Wolf. 1969. Frequency as a determinant of linguistic changes in the Ethiopian languages," *Word* 25, 180–189.

Moore, Samuel. 1968. *Historical outlines of English sounds and inflections*. Revised by Albert H. Marckwardt. Ann Arbor, Mich.: George Wahr.

Paul, Hermann. 1970 [1890]. *Principles of the history of language*. Trans. by H. A. Strong. College Park, Md.: McGrath.

Schuchardt, Hugo. 1885 [1972]. On sound laws: Against the Neogrammarians, in *Schuchardt, the Neogrammarians, and the transformational theory of phonological change* (*Linguistische Forschungen* 26), ed. Theo Venneman and Terence H. Wilbur, 29–72. Frankfurt am Main. Athenaum.

Stampe, David. 1969. On the acquisition of phonetic representation, *Papers from the fifth regional meeting of the Chicago Linguistic Society*, 443–454.

———. 1973. *A dissertation on natural phonology*. Diss., University of Chicago.

Zipf, George K. 1935 [1965]. *The psycho-biology of language*. Cambridge, Mass.: MIT Press.

Zwicky, Arnold. 1972. Note on a phonological hierarchy in English. In *Linguistic change and generative theory*, ed. R. P. Stockwell and R. K. S. Macaulay, 275–301. Bloomington: Indiana University Press.

PART II

FREQUENCY AS A DETERMINANT OF MORPHOLOGICAL STRUCTURE

This page intentionally left blank

Introduction to Part II

The chapters of part II resulted from a period of intense focus on the impact of meaning and use on morphological structure. Data from language change, child language, and experimentation are considered in trying to determine the nature of the cognitive organization of morphology, especially inflectional morphology. While the effects of both type and token frequency are explored in these works, frequency effects are by no means the only factors considered in these chapters. Phonological similarity and degree of semantic relatedness are also considered to have an effect on the organization of paradigms.

Chapter 3 reprints Bybee and Brewer (1980), which focuses on analogical changes in the preterite verbal paradigms of dialects of Castilian and Provençal. Though the changes work with different phonological material, indicating their independence from one another, the principles underlying the changes are the same. These changes point to the relative autonomy of certain forms, particularly the third singular and the first singular, which both resist change and serve as the basis of new analogical formations. Autonomy is said to depend upon semantic simplicity, morphophonemic irregularity, and word frequency. Because of the accepted dogma of the day—that frequency has nothing to do with linguistic structure—we were somewhat hesitant to hypothesize frequency as the major factor, a position I would be more likely to choose today.

While chapter 3 deals with token frequency, chapter 4 (Bybee and Pardo 1981) reports on a nonce-probe experiment with Spanish-speaking adults that investigates a variety of issues, including the internal organization of paradigms, as in chapter 3, but also the role of type and token frequency in determining productivity in morphophonological alternations. It is found that (1) certain widespread vowel alternations, especially diphthongization, are not readily extended to nonce forms; (2) when alternations are extended, it is under conditions in which both alternates are made available in the experiment and one is in a form that is closely related semantically to the target form; (3) morphophonological classes do not necessarily make use of the broadest phonological generalizations but often have more specific and local phonological reference (see also Bybee 2001:130ff); (4) type frequency and phonological similarity relate to the degree of productivity; and (5) paradigms with very high token frequency do not serve as analogical models, thus detracting from productivity. Subsequent studies of Spanish diphthongization have confirmed the lexical specificity of diphthongization, identifying both morphological and phonological contexts that favor or disfavor diphthongs (Eddington 1996, 1998; Albright, Andrade, and Hayes 2001).

Chapter 5 (Bybee and Slobin 1982) deals with the English past tense in children and adults, taking into account the type frequency of the irregular patterns, the token frequency of members of the irregular classes, the phonological similarity between the base and the past, and the phonological similarity among past forms themselves. Because of its focus on the differences in type and token frequency between regular and irregular past forms, Rumelhart and McClelland (1986) chose to simulate the results of this study with a connectionist model, setting off a long-standing debate about the role of structure versus frequency in the processing of regular versus irregular forms (Pinker 1991; McClelland and Patterson 2002 and references there).

One particular class of irregular verbs in English, exemplified by *string/strung*, shows some productivity, offering an opportunity to understand further the factors that govern productivity. This class is studied in chapter 6, which reprints Bybee and Moder (1983). Here it is shown that the phonological shape of the verb, including not just the final consonants, but also the initial ones, defines the class. This chapter argues for product-oriented schemas on the basis of the fact that the vowel of the base is the least important factor in predicting class membership, for example, verbs such as *strike/struck* are recent additions to the class. The results reported in this paper have been replicated by Prasada and Pinker (1993).

In chapter 7, Bybee and Newman (1995) contrast explanations traced to grammaticization with explanations based on synchronic processing preferences. We argue that the rarity of stem changes versus affixes in the languages of the world is due to the way each mode of expression develops diachronically and not to a postulated extra processing cost for stem changes. The acquisition difficulty associated with stem changes is due to their low type frequency in the languages in which they have been studied. We report here on an experiment in which subjects were taught an artificial language in which stem changes have type frequency comparable to affixes. In this experiment, subjects learned the stem changes as easily as the affixes.

Chapter 8 is a response to some claims that arose in the Great Past Tense Debate, which, as mentioned before, was triggered by Rumelhart and McClelland's connectionist simulation of the acquisition of the English past tense. An important topic in the debate centers on the role of type frequency. The dual-mechanism camp (Pinker 1991; Marcus et al. 1992 and subsequent papers) argue that productivity is structurally determined; somehow the child learns which is the default rule, possibly by having heard it in a default environment. In this theory, type frequency is of no consequence. In the connectionist account and my own network model, type frequency is a strong determining factor for productivity. In this chapter, I argue that there are degrees of productivity that correspond to the interaction of phonological similarity with type frequency. The chapter also argues that high-frequency regular past tense forms are represented in the mental lexicon, for which further evidence is presented in chapter 9.

References

Albright, Adam, Argelia Andrade, and Bruce Hayes. 2001. Segmental environments of Spanish diphthongization. In *Papers in phonology 5*. UCLA Working Papers in Linguistics 7, ed. A. Albright and T. Cho, 117–151. Los Angeles: UCLA Department of Linguistics.

Bybee, Joan L. 2001. *Phonology and language us*e. Cambridge: Cambridge University Press.

Eddington, David. 1996. Diphthongization in Spanish derivational morphology: An empirical investigation. *Hispanic Linguistics* 8(1), 1–13.

———. 1998. Spanish diphthongization as a non-derivational phenomenon. *Rivista di Linguistica* 10(2), 335–354.

Marcus, G. F., S. Pinker, M. Ullman, M. Hollander, T. J. Rosen, F. Xu, and H. Clahsen. 1992. *Overregularization in language acquisition*. Monographs of the Society for Research in Child Development 57. Chicago: University of Chicago Press.

McClelland, James L., and Karalyn Patterson. 2002. Rules or connections in past-tense inflections: What does the evidence rule out? *Trends in Cognitive Sciences* 6, 465–472.

Pinker, Steven. 1991. Rules of language. *Science* 253, 530–535.

Prasada, Sandeep, and Pinker, Steven. 1993. Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes* 8, 1–15.

Rumelhart, David E., and James L. McClelland. 1986. On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In *Parallel distributed processing: Explorations in the microstructure of cognition*, ed. James L. McClelland, David E. Rumelhart, and the PDP Research Group, 216–271. Cambridge, Mass.: MIT Press.

This page intentionally left blank

Chapter 3

WITH MARY ALEXANDRA BREWER

Explanation in Morphophonemics

Changes in Provençal and Spanish Preterite Forms

Morphophonemic alternations (alternations with some morphosyntactic or lexical conditioning) would seem to be among the most arbitrary and conventional aspects of natural language. Their treatment in most theories characterizes them as synchronically arbitrary—Bloomfield (1933) and later Chomsky and Halle (1968) describe morphophonemics with historical reconstructions, while others are satisfied with long lists of alternants and statements about distribution. More recently, morphophonemic alternations have been excluded from phonology, as not being explainable by phonological principles but belonging to some other domain (Andersen 1969; Vennemann 1972; Stampe 1973). The question now arises, what is this other domain, and are morphophonemic alternations synchronically arbitrary, or are they somehow motivated by the communication process and its participants?

In this essay we take the point of view that morphophonemic alternations are not arbitrary but rather tend to form diagrammatic relations as described in Andersen (1980). Andersen's proposal (based on some brief suggestions in Jakobson 1965) is that the relations among signantia (the phonological forms of morphemes) or the relations among variants of signantia reflect parallel or diagram relations among signata (the semantic values of morphemes). In particular, we present evidence, as others have, that there is a basic-derived relationship among the forms of a stem paradigm that diagrams the basic-derived relationship in the semantic domain (section 1). More important, we investigate the question of what determines particular basic-derived relationships and the question of why diagrams exist.

Our proposal is that words may be classified as to degree of autonomy—the extent to which a word is likely to be represented in the speaker's lexicon as a whole and independent unit (Zager 1980). The factors that determine the degree of autonomy of a word are discussed in section 4. The significance of autonomy is that highly autonomous words will be learned as whole units, even if they are (to the linguist) morphologically complex, and less autonomous words will be analyzed as derived from the autonomous words of the same paradigm. Thus autonomy is one of the determinants of the basic-derived relationship. The other determinants are the morphological categories and the hierarchical relations among those categories (section 1).

Because there may be more than one autonomous word in a stem paradigm, there may be more than one basic form for the paradigm. Furthermore, derived forms are not all related to basic forms in the same way. In sections 6–9 we discuss a series of morphophonemic changes in Provençal dialects and a related but different series of changes in Spanish dialects that help us understand the nature of the relations in a small area of verbal morphology—the person/number forms within a tense.

1. Diagrams

1.1. The basic-derived diagram

One example of a diagram that Andersen (1969) cites is the morphophonemic reflection of the semantic basic-derived relationship. Vennemann discusses such diagrams saying: "Usually in natural languages, a semantic derivation of secondary conceptual categories from primitive ones, tertiary from secondary ones, etc., is reflected by a parallel syntactic or morpho-phonological derivation" (1972:240). Consider Vennemann's representation of this principle (fig. 3.1). Here *X* is the signans corresponding to the signatum *A* and *y* corresponds to the secondary category. The manifestation of *y* may be some modification of *X*, that is, the application of a morphophonemic rule, or it may be the addition of an affix to *X* or both.

Vennemann proposes that this diagram represents a language acquisition strategy and that it explains the frequent observation that when morphophonemic alternations are eliminated (in so-called analogical leveling), it is the alternant found in the primitive or basic category that survives, for example, the third person, the present, the indicative, the singular (Mańczak 1963). To take a brief example, we analyze the past tense of English *creep*, *crept*, as consisting of *creep* plus the past-tense suffix, and a vowel change of [iy] to [ɛ], rather than analyzing the present as derived by a vowel change of [ɛ] to [iy]. The alternate past-tense form *creeped* [kriypt] is described

	Primitive category	Secondary category
Semantic level	A	A + b
Level of overt manifestation	X	X + y

FIGURE 3.1

FIGURE 3.2

as arising by the elimination of the vowel change, which allows the basic form to emerge. In section 3 we will discuss more changes motivated by this type of diagram, and in section 4 we will try to explain why this type of diagram exists.

1.2. The relevance diagram

Another diagrammatic relation is discussed in Hooper (1979), where it is shown that morphophonemic alternations in verb stems diagram the hierarchical relation in which person/number distinctions are subordinate to tense/aspect distinctions. The hypothesis was suggested by child language data in which it was evident that the children made verb stem substitutions of 3s for 1s within the same tense, while the relation of 3s stems in two different tenses was more remote.[1] That is, the children under study, who were acquiring Brazilian Portuguese, appeared to learn the 3s preterite form separately from the 3s present form (Simões and Stoel-Gammon 1979). The arrows in figure 3.2 represent the direction of substitution. This pattern of substitution together with the loss of or failure to acquire morphophonemic alternations would lead to a situation in which verb stems within tenses would be more uniform than verb stems across tenses. The result is the following testable hypothesis:

> *Hypothesis*: Alternations in verb stems will correspond to tense/aspect distinctions across person distinctions, but not to person distinctions across tense/aspect lines.

For instance, there are many cases of special stem forms for present, past, or future but no reported cases of a special stem for, for example, a 1s that is used throughout the tenses. This hypothesis was tested on a stratified probability sample of fifty languages. Eighteen of these languages inflected for both person and tense/aspect, and in none of these were verb stem alternations found that corresponded to person distinctions across tense/aspect lines (Hooper 1979).

If morphophonemic alternations in verb stems coincide with tense/aspect distinctions, then they outline the major category distinctions in verbal paradigms. The greater importance of tense/aspect distinctions to verbs (over person/number distinctions) is evident in the fact that the presence of person/number inflections in a language implies the presence of inflections for tense, aspect, or mode (Greenberg 1963:112). Tense/aspect distinctions are more important to verbs than person/number distinctions because of their meaning. Tense and aspect have meanings that refer directly to the event described by the verb, tense placing the action in time relation to the moment of speech and aspect referring to the internal structure of the action. Person and number, however, refer to the arguments of the verb and, in their meaning have more to do with nominals than verbs. Thus, as regards the meaning of the *verb*, tense and aspect make

Present		Preterite
1s	⟶	1s
3s	⟶	3s

FIGURE 3.3

larger meaning changes than person and number. The more direct relation of tense and aspect to the verb is reflected in the possibility of these categories conditioning morphophonemic modification of the verb stem itself. The peripheral status of person and number as verbal categories is reflected in their inability to affect verb stem morphophonemics across tense and aspect lines. The greater meaning change made by tense and aspect is reflected in a greater change in form. The result is that in both meaning and overt form persons within a tense or aspect will show greater similarity and be more closely related to one another than to the same persons in another tense of aspect. In the discussion that follows we will be studying the basic-derived diagrams among the very closely related person forms within a tense.

It is important to note at this point that the morphemes whose function is to directly signal person/number form different patterns of relations than the verb stems of persons within a tense. Language acquisition data as well as historical data show many examples of person/number suffixes of one tense replacing those of another tense in the same person, as shown in figure 3.3 (Hooper 1979). In our discussion we will have to ignore changes of this type.

Another type of morphophonemic change that we will not be discussing here is change in verb stems motivated from outside the paradigm. An example from Spanish is the loss of diphthongized stem variants, as in the verb *temblar* 'to shake', which has etymologically and in some dialects a diphthong when the stem is stressed, *tiémblo* (1s), *tiémblas* (2s), and so forth, but in other dialects has a simple vowel throughout. This type of change is a change in the class membership of the lexical item. An original diphthongizing verb moves into the class of regular verbs. This type of change is discussed in Bybee (1980).

2. Romance preterites

The verb forms we will be discussing are descended from Latin perfect forms. The perfect presents certain analytic problems in Latin, especially with regard to the person-number suffixes, for example, the occurrence of *-sti* in 2s, where other tenses have only *s*, the occurrence of *-runt* in 3p, where other tenses have only *-Vnt*, but there is a clearly separable perfect marker, at least in the weak perfects of formal Latin, in the form of *-v-*:

(1)

First conjugation[2]		*Fourth conjugation*	
cantāvī	cantāvimus	dormīvī	dormīvimus
cantāvistī	cantāvistis	dormīvistī	dormīvistis
cantāvit	cantāvērunt	dormīvit	dormīvērunt

The phonetically conditioned loss of the *-v-* ([w]) and at times the vowel following it led to a situation in which no one segment could be identified as the perfect marker. The following forms are what Menéndez-Pidal calls "latín popular" (1941[1968]:310):

(2)	cantā́i	cantā́mus	dormī́i	dormī́mus
	cantā́sti	cantā́stis	dormī́stī	dormī́stis
	cantā́ut[3]	cantā́runt	*dormī́ut	dormiérunt

Regular sound changes acting on these forms yielded the modern Castilian (Standard Spanish) forms:

(3)	canté	cantámos	dormí	dormímos
	cantáste	cantásteis[4]	dormíste	dormísteis
	cantó	cantáron	durmió	durmiéron

A slightly more complex history, which we will not discuss, gave the following Old Provençal forms (Anglade 1921:272, 294):[5]

(4)	*canta* 'to sing'		*venre* 'to sell'	
	cantéi	cantém	vendéi	vendém
	cantést	cantétz	vendést	vendétz
	cantét	cantéren	vendét	vendéron

(Also 3p *cantéron*, 1s *vendiei*, 2s *vendiést*.)

The paradigms in (3) and (4) present various analytic problems for the native speaker or learner of the language. As we mentioned earlier, it is difficult to isolate a single segment or group of segments that can be said to signal perfect (or preterite, as it becomes in the Romance languages), although there is a characteristic pattern of stressing the vowel following the stem. The speaker is faced with a number of questions: For example, in the Spanish data, is the *-é* of 1s a person marker or a preterite marker or both? Is it a special 1s marker for preterite or a special preterite marker for 1s? And, in these languages, as earlier in Latin, speakers are faced with the problem of assigning a role to the *-r-* of 3p.

The problems in Old Provençal are even more complex.[6] There was a good deal of inconsistency with regard to the suffix vowels, both within and between conjugations, and a considerable amount of homophony between forms of the preterite and forms of other tenses. Even the characteristic stress pattern of the preterite was disturbed by such factors as variation between final and penultimate stress on the 1p and 2p forms (Ronjat 1937:186, 192), and the existence of both weak and strong conjugations, as exemplified in (4) by *canta* and *venre* and in (5) by *saber* and *prenre*, respectively (Anglade 1921:300, 305):

(5)	*saber* 'to know'		*prenre* 'to take'	
	saup	saubém	pris	prezím
	saubíst	saubétz	prezíst	prezítz
	saup	sáupron	pres	préiron

The strong verbs (some infinitives in *-re* and *-er* and all *-ér* infinitives) were distinguished from the weak verbs (all *-a*[*r*] and *-ir* infinitives, and some in *-re* and *-er*) by exhibiting stem stress in 1s, 3s, and 3p and by their lack of person/number inflection in 1s and 3s (Anglade 1921:262). Finally, some verbs (including *prenre*) manifested as many as three stem allomorphs in the preterite.

The way speakers handle these complex analytical situations is revealed by the changes that take place in these paradigms in the various dialects of Spanish and Provençal. The location of the dialects surveyed is indicated on the map (see fig. 3.4). We have attempted to make our study exhaustive enough to provide the full picture of the range of possible changes that can result from the situations exemplified earlier.[7] An analysis of these changes provides evidence for some general principles governing diagrammatic relations. The first to be illustrated is the basic-derived relation described in the preceding section.

3. Third singular as a basic form

3.1. Provençal preterites

Watkins (1962) presents several clear examples of all the person forms of a tense being re-formed with the 3s taken as the base. Such re-formations are often preceded by a reanalysis of an original form composed of *stem + tense/aspect + person/number* as *stem + tense/aspect* + Ø. The following Provençal forms show just this sort of development:

(6) *Charente* (Meyer-Lubke 1890/1902[1925]:352)

cantí	cantétem
cantétei	cantétei
cantét	cantéten

(7) *Clermont-Ferrand* (Ronjat 1937:193)

cantéte	cantétem
cantétes	cantétetz
canté	cantéton

The original 3s suffix *-t* has been extended to the other persons of the preterite (except 1s in Charente—a matter we return to in section 7). This change has two steps: the interpretation of *cantét* as containing no specific mark for person; the *-t* instead is analyzed as signaling preterite. Second, the choice of 3s as a base that can be used to create the other forms by the addition of person/number inflections drawn from other tenses. This second step is further confirmed by the fact that the rebuilt forms have stress on what was the final, stressed syllable of the 3s. It does not matter that the final *-t* of 3s has been deleted in some dialects (e.g., Clermont-Ferrand in [7] earlier.[8] In Old Provençal the *-t* still appeared before a vowel-initial word.

In some dialects, final *-t* has been replaced by a velar *-c* ([k]) in the 3s (Ronjat 1937:192; Anglade 1921:272; Meyer-Lubke 1890/1902[1925]:353). In other dialects, such as that spoken in Foix, the velar appears throughout the preterite (voiced to [g] intervocalically) (Ronjat 1937:192):

FIGURE 3.4. Locations of dialects referred to in the text. Dialect areas indicated on map: points indicate cities or towns; dotted circles indicate regions of varying size.

Provence

1. Toulouse
2. Lomagne
3. Foix
4. Couserans
5. Montignac
6. Limoges (Limousin)
7. Bas-Limousin
8. Rochechouart
9. St. Mathieu
10. Confolens
11. Montembœuf
12. Montbron
13. Aurillac
14. Aquitaine
15. Charente
16. Marche
17. Clermont-Ferrand
18. Ambert
19. Gilhoc
20. Loriol
21. Chabrillan
22. Châtillon-en-Diois
23. Gap
24. Nice
25. Forcalquier
26. Marseille
27. Nîmes
28. Alais (Alès)
29. Gévaudan

Spain

30. Babia, Laciana, and Sisterna
31. Sanabria
32. Aliste
33. La Ribera
34. Lena
35. Cabrales
36. La Pas
37. Valle de Aragüés
38. Bielsa

(8) cantégui cantéguen
cantégues cantéguets
cantéc cantéguen

This development is apparently parallel to that of the preterites with *-t-*, with the 3s form in *-c* serving as the basis for a re-formation of the entire tense. Speakers may have taken the velar to be a marker of preterite because of the existence in the same dialects of a small class of highly frequent strong verbs with a velar stop characterizing the entire preterite. In these verbs, the velar arose because of phonetic developments involving the strengthening of Latin *-ui* to [g^w], and a concomitant absorption of the preceding stop:

(9) *voler* 'to wish, to want' (Grandgent 1905:144)
vǫlc, vuẹlc, volguí volguem
volguíst volguetz
vǫlc vǫlgron

This class included such verbs as *aber* 'to have', *poder* 'to be able', and *tener* 'to have, to hold'. Furthermore, it is significant that while the velar began to spread through the weak paradigms as other forms were rebuilt off the 3s form, the 1s of the strong verbs often added the weak inflection *-i*, accompanied by final stress, thus making the 1s and 3s forms distinct: for example, 1s *deguí*, 3s *dec*. Note that it is the 1s form that was remade; the 3s form itself was under no circumstances remade (Grandgent 1905:138).

As we noted earlier, the 3p form of the preterite was particularly problematic for speakers of Provençal, as the *-r-* that occurred there could not be assigned a specific semantic value. This *-r-* was restricted to 3p and so could not be taken as a general preterite marker. It did not occur in any other 3p forms and thus could not be a 3p marker. While there was an *-r-* in the future and conditional, it occurred in all persons and so could not be related to the *-r-* of 3p preterite. The problem the speakers faced, then, was that of determining the status of this *r*. As Blaylock noted (1975:434), with regard to another verbal augment, the Latin *-sk-*, "where an empty [i.e., apparently meaningless] morph . . . separates a neatly silhouetted base from an equally well-delineated inflectional ending, both grammarian and lexicologist can claim a legitimate interest, while the native speaker strives to clarify the status of the bothersome element . . ."

The changes in the various dialects show how speakers resolved the question of this "bothersome" element. In Lomagne, speakers apparently concluded that the *-r-* served no function whatever, and they dispensed with it entirely (Ronjat 1937:272):

(10) cantéi cantém
cantés cantéts
cantéc cantén

The stress was regularized, so that it is now consistently final throughout the preterite. The problem of potential homophony with the present tense, which shares many of the preterite inflectional endings, did not arise, since in the present the stress was retracted in 1p and 2p, making it uniformly penultimate.

The other possibility for dealing with the *-r-* involved a segmentation of the 3p such that the *-r-* was interpreted as linking the stem and the 3p inflection, *parlé + r + on.* Once its role as a linking consonant had been established, the *-r-* could be used in the re-formation of other persons from a 3s base that ended in a vowel. In the examples in (11), it can be seen that the addition of a vowel-initial person/number inflection to a vowel-final stem, such as the 3s, would require either a vowel merger or the "interfixing" of a consonant to separate the two vowels. The latter solution yields a transparent sequence of morphemes and is adopted by the following dialects (Ronjat 1937:187, 225, 226):[9]

(11)	*Literary*	*Provençal*	*Nîmes*	
	parlére	parlérian	cantére	cantén
	parléres	parlérias	cantéres	cantés
	parlé	parléron	canté	cantéroun
	Forcalquier			
	parlí	parlérian		
	parléres	parlérias		
	parlé	parléroun		

For example, 2s is re-formed by using the 3s as a base, *parlé*, adding the linking consonant *-r-* and a 2s suffix, *-es*: *parlé + r + es*. Subsequently, the *-r-* may be interpreted as a marker of preterite, as more forms are remade in this way. Observe that 3p does not serve as a base but only provides the *-r-* for use as a linking consonant, and that role depends entirely upon an analysis of 3p as consisting of the 3s as a base, plus *r*, plus the 3p suffix *-on.* Thus in all the Provençal preterite changes involving a consonant, whether it is *t*, *c*, or *r*, it is clear that the 3s serves as the basic form. In some dialects, not all persons of the preterite were re-formed. In section 7, we will discuss the different patterns of re-formation.

3.2. Spanish preterites

Another example that shows the 3s form to be the basis of the other persons of a tense comes from a Spanish dialect in the valley of Bielsa in the Pyrenees. Consider the preterite forms in Bielsa (Badia Margarit 1950:128–131):

(12)	*First conjugation*		*Second and third conjugations*	
	llevé	llevémos	metié	metiémos
	llevóres	llevéz	metióres	metiéz
	llevó	llevóren	metió	metióren

These paradigms exhibit several interesting peculiarities to which we will return later. In particular, note that the 2s has been remade on the basis of the 3s, using the *r* from the 3p, just as in the Provençal dialects discussed earlier.

What interests us for the moment, however, is the relation between the first conjugation suffixes and those of the second and third conjugations. It is immediately obvious from (12) that all three conjugations use the same suffixes: *-é*, *óres*, *ó*, *émos*, *éz*, *ónen*, but that in the second and third conjugations each suffix is preceded by *i*. This vowel is actually pronounced as the glide [y] when prevocalic. A comparison with the Castilian second and third conjugation forms reveals that the only unchanged form is the 3s:[10]

(13)	metí	metímos
	metíste	metísteis
	metió	metiéron

Furthermore, if the forms in (13) are compared with the first conjugation forms in (12), it is clear that the only form that could have served as the basis for the remodeling of the Bielsa second and third conjugations is precisely the 3s form.

On the basis of the form *llevó*, speakers could quite correctly decide that *o* is the marker of preterite, 3s. If this analysis is applied to *metió* [metyó], speakers might further conclude that the stem for the preterite is *meti-* [mety-]. When the 1st conjugation suffixes are attached to this preterite stem, the current Bielsa forms result. Notice that in keeping with the hypotheses discussed in section 1, the stem for the preterite tense is used to build all the persons of the same tense, but the person markers come from the same person in another paradigm.

This change, or series of changes, probably took place over an extended period of time, with some forms changing before others and with a period in which both conservative and innovating forms competed. The point we would like to stress here is that only the 3s form could be the source of this restructuring, for it is only in the 3s that the correspondence between the suffixes (*-o*) in the three conjugations could lead to the reanalysis of the stem as ending in a glide in the second and third conjugations.

These examples, then, demonstrate the existence of a basic-derived relationship among the forms of a paradigm, with the 3s often taken as the most basic.

4. Autonomy

4.1. Autonomy defined

Two factors influence the basic-derived relationships that speakers construct. (1) Morphological categories such as tense, person, and number establish the axes upon which such relationships stand. We will adopt a rather traditional view about such categories and not say more about them here. (2) The directionality of the basic-derived relationship depends on the degree of autonomy: the more autonomous form serves as the basic form, and the less autonomous forms are derived.

The degree of autonomy is a property of individual words, not of morphemes, nor of members of categories as whole sets. It is the extent to which a word is likely to be represented in the speaker's lexicon as a whole and separate unit. Our hypothesis is formulated as a matter of degree for two reasons: to allow for the possibility that individual speakers may make slightly different decisions concerning lexical storage and because there are (at least) three determinants of autonomy in morphologically related forms, all of which may be present in a given word in varying proportions. A hypothesis formulated in such a way makes predictions of statistical tendencies in diachronic change, language acquisition and psycholinguistic experimentation. It cannot, nor it is intended to, generate a unique grammar for a body of linguistic data.

Since a fully autonomous word is a word with its own lexical representation, any word that is not morphologically related by productive rule to another word will be autonomous. That is, a word is autonomous if it is not semantically and phonologically predictable from some other word by general rule (Hooper 1976a). Thus English *archive*, *sibling*, and *flagrant* (to take some random examples) are all autonomous, but a plural such as *siblings* need not be, for it can be derived by a productive rule. This position on lexical storage is not controversial and will not be discussed further. We turn now to a discussion of autonomy in words related by productive morphological rules. The discussion is phrased to apply primarily to inflectionally related words, but it is also applicable to words related by productive derivational rules.

4.2. Autonomy within paradigms

Among the inflectionally related words of a paradigm, some will be more autonomous than others; that is, some will be more amenable to lexical storage, and some to derivation by rule. We will consider the following three factors as determinants of autonomy: semantic simplicity, word frequency in discourse, and morphophonemic irregularity. We begin with a consideration of relative semantic simplicity.

Our hypothesis is that words with fewer semantic components to segment can more readily be stored as whole units, while words with more semantic parts are more likely to be broken down into separate morphological units. That is, more units of meaning look for more units of expression. This notion of semantic complexity depends upon the theory of markedness, in which one member of a category is "unmarked" semantically and thus simpler. For example, Spanish 3s present indicative *cánta* is simpler than 1p present indicative *cantámos* (even though it takes the same number of traditional terms to describe them—third, singular, present, and indicative for *cánta*, first, plural, present, and indicative for *cantámos*) because the 3s is unmarked for person and number. The 3s does not contain any information concerning person and number, unless one infers that the lack of person/number marking signals 3s. With regard to our hypothesis, it should be clear that there is less segmenting to do for *cánta* than for *cantámos* and less for *cantámos* than for *cantábamos* (1p imperfect indicative). Of these three *cánta* would be the most autonomous.[11]

Words that are highly frequent in discourse are more likely to have separate lexical entries than less frequent words. It is well known that frequency is an extremely important determinant of rote learning. A highly frequent model does not require decomposition by rule. Numerous diachronic changes attest to the nonsegmentation of highly frequent sequences, as when idioms become fixed, or verbal phrases such as *going to* and *got to* become *gonna* and *gotta.* Less frequent words and phrases maintain the separate identities of their parts and thus seem not to be autonomous.

Word frequency is an important diagnostic for establishing the degree of autonomy, because it is a property of words that can be measured with more confidence than some other properties that determine autonomy. It must be remembered, however, that word frequency is a result of the meaning of the word, the real world situation it refers to, and the view of the world taken by speakers.

Jakobson's (1957) definition of *unmarked* and *marked* clearly indicates a wider range of usage and thus a higher frequency for the unmarked member of a category:

> The general meaning of a marked category states the presence of a certain (whether positive or negative) property A; the general meaning of a corresponding unmarked category states nothing about the presence of A and is used chiefly but not exclusively to indicate the absence of A. (Jakobson 1957:136)

Furthermore, Greenberg observes that frequency can be based on "the situation in the world with which the users of the language must deal" (1966:66). Thus in English *author* means writer of either sex or male writer, but, because of the real world situation in which nurses are usually females, *nurse* refers to nurses in general or to female nurses. Similarly, the most frequent case of a noun may vary according to the referent—the word for hammer may be most frequent in the instrumental, place names in the locative (Mańczak 1958), and so on. In addition to the semantic range of the word and the situations in the real world, frequency may also be conditioned by the psychological processes that determine how we perceive reality and how we use language to refer to it. For instance, the more frequent use of singulars may be due to a tendency to focus attention on single referents, and frequency of 1s forms may be attributed to an egocentric bias. (Zubin 1979 applies these psychological principles to the German case system.)

It is uncontroversial that a certain amount of morphophonemic irregularity or opacity forces autonomy. (For our purposes here, *irregularity* will refer to the presence of a morphophonemic alternation that is restricted to certain morphemes and *opacity* will refer to portmanteau expression or the absence of a transparent segmentation of morphemes.) In cases of total suppletion, such as English *go*, *went*, separate lexical entries are clearly needed. With less striking cases of opacity or irregularity, such as the English strong verbs, it is not so easy to decide on the basis of internal evidence alone whether a derivation by rules, which would be quite restricted and complex, is preferable to separate lexical storage. We hypothesize, however, that any amount of irregularity or opacity can produce a tendency toward autonomy and that a more precise picture of the relation between these properties and autonomy can be established experimentally. (Cf. section 4.5.)

4.3. The autonomy hypothesis

The autonomy hypothesis is, then, that words will have their own lexical representation if they have one or more of the following properties: low semantic complexity, high frequency, and morphophonemic irregularity or opacity. There is considerable diachronic evidence for this hypothesis. It has long been observed, and more recently documented by Mańczak (1958, 1963, 1980), that certain forms in paradigms are more *resistant* to morphophonemic change than others. Mańczak reports that these are the more frequent words of the paradigms, usually singulars, third persons, present tense, and indicative mood. Since the morphophonemic changes in question result from the replacement of old forms with new forms that result from the application of morphological and morphophonemic rules, the autonomous forms, being registered lexically and not produced by rule, are less likely to change, and the nonautonomous forms, those produced by rule, are most likely to change.

It has also been observed at least since Paul (1880[1970]), and documented recently in Hooper (1976b) for English strong verbs, that very frequent paradigms retain their morphophonemic irregularity much longer than infrequent paradigms. This fact is also accounted for by the autonomy hypothesis and, furthermore, reveals the interplay between morphophonemic irregularity and frequency in determining autonomy. Specifically, in order for a word to maintain a morphophonemic irregularity diachronically, it must be frequent enough to be learned and registered as an autonomous form. If it is not, it will be replaced by a form derived by rule. The acquisition of autonomous and nonautonomous forms is discussed in the next section.

The phenomenon of split paradigms also presents evidence for the autonomy hypothesis. Split paradigms are paradigms that are based on more than one stem, and each of these stems has an etymologically different source. English examples are *go/went* and *be/am*, *is*, *are/was*, *were* (the latter having three etymologically separate stems). In order to create a split paradigm, such as Modern English *go*, the past forms of *wend* had to become independent of the present-tense forms and gradually replace the past forms of *go.* This process requires a split in both the paradigms of *go* and of *wend.* These splits could only occur if the past forms had separate lexical entries from the present forms. Such splits would be impossible if the past forms were derived by rule from the present forms. In the case of *go*, the split of the present from the past is aided by the fact that the prior past forms of *go* were also from a different stem (Old English *gā* [present] and *ēode* [past]) and would have been autonomous for this reason. A connection between split paradigms and morphophonemic irregularity is argued for by Bolozky (1980), who observes that a large number of split paradigms in Modern Hebrew have arisen in paradigms that previously displayed other morphophonemic irregularities. However, morphophonemic irregularity is not a necessary motivation for paradigm split. Paradigms do not split simply because their forms are so different as to appear unrelated, for while *went* was not a totally regular form, its morphological connection to *wend* could hardly be considered opaque. Rather, in this case, we must look to frequency of use and realize that a highly frequent word may be autonomous, even if it is transparently derivable from some semantically simpler related word. In this regard, it should also be mentioned that split paradigms occur most often in the very frequent lexical items in any language. (Rudes 1980 discusses many such examples.)

Paradigm splitting and resistance to morphophonemic change both support the autonomy hypothesis and, furthermore, give evidence for the roles of frequency and morphophonemic irregularity in determining autonomy. We turn now to an examination of the language acquisition process with respect to our hypothesis.

4.4. Autonomy in acquistion

The acquisition of morphology and morphophonemics takes place through a combination of rote learning and rule formation. MacWhinney (1978) gives a detailed account of the process, which is usually thought of as comprising at least three stages. In the earliest stage, morphologically complex words are learned by rote, as amalgams, but little segmentation or analysis is applied to them. At this point, many irregular forms are produced correctly, but children are not able to apply rules to nonce forms. In the second stage the child begins to formulate rules, and in this stage many overgeneralizations occur. Sometimes irregular forms that were earlier produced correctly are now overgeneralized. Thus a child, who made the past tense of *break* as *broke* earlier, may now begin to produce *breaked.* The child, then, has two ways of producing a past tense form: one is to use a stored irregular form; the other is to apply a regular rule. The child's task is to learn which method works for which verb. In this stage, the child seems to prefer rule application. However, irregular and regular forms of the same verb are often used side by side, and some verbs do not seem to be so prone to overgeneralization as others. Slobin (1971) studied overgeneralizations of English past tense based on a sample of the speech of twenty-four children between the ages of 1½ and 4 years and found many more correct past-tense formations than regularized ones. Furthermore, there are some verbs that are never regularized in this sample. Slobin also computed the input frequency of these verbs based on sixty-four hours of speech to children in this age range by mothers and investigators in natural dialog. Z. Greenberg (1975) found a positive correlation between correct forms and input frequency in Slobin's data. At this point in the acquisition process, according to MacWhinney, the child is producing forms and monitoring his or her productions: comparing the forms produced by rule with rote forms. The success of rote forms against rule-generated forms depends upon what MacWhinney refers to as their "lexical strength" (1978:9). Lexical strength apparently depends upon frequency in input–availability of the model and frequency in production by the child, since the more often a word is produced, whether correctly or incorrectly by adult standards, the more right it will seem to the child. In the final stage, the child must sort out the irregular forms and learn them. This process seems to proceed one verb at a time and resembles rote learning more than minor rule acquisition.

It is very difficult to account for the perseverance of irregular forms in a language, such as the English strong verbs, without according an important role to word frequency, especially considering that only the most frequently used irregular forms remain over periods of time. But the importance of *token* frequency suggests rote learning, since rule formation requires high *type* frequency (MacWhinney 1978).

Our autonomy hypothesis would further predict that there is a much higher degree of portmanteau expression in highly frequent forms, just because they can be learned unanalyzed and do not need to be transparently segmentable. For the same reasons, highly frequent *lexical* items will often maintain portmanteau expression. There is evidence that these properties are indeed the result of the way the forms are acquired. Simões and Stoel-Gammon (1979), reporting on the acquisition of verbal forms for four children acquiring Brazilian Portuguese, observe that several verbs in the children's speech were not subject to the regularizing processes that affected other verbs; that is, 3s was not used for 1s, but rather 1s was used correctly from the beginning. Simões and Stoel-Gammon point out that these verbs have two characteristics in common: they were very frequent in the children's speech, and they were monosyllabic. The majority of Portuguese verbs are polysyllabic, with the stem consisting of one syllable or more and the inflections adding extra syllables, for example, *falo* 'speak (1s)', *fala* (3s). The verbs that were acquired as exceptions, for example, *vou* 'go (1s)', *va* (3s), *dou* 'give (1s)', *da* (3s), have only a single consonant as a stem and are thus difficult to segment. It seems that both their frequency and their opacity contribute to the likelihood of their being learned by rote.

4.5. Psycholinguistic evidence

Psycholinguistic research has concentrated a good deal of attention on studying the rule-formation process in developing children but has devoted considerably less energy to determining the interplay between rote learning and rule formation. The major problem here is that if a child correctly produces a word and also shows evidence of having a productive rule that would generate the word, we still do not know if the word was produced by rote or by rule. For example, Berko (1958) found that children aged 4 to 7 years could for the most part produce a correct plural for the word *glass*, but only about a third of them supplied the plural of the nonce form *tass* as *tasses* (with the *əz* allomorph). This would indicate that they do not yet have the rule that produces -*əz* and that *glasses* is a rote form. At a later stage, when there is evidence of a rule producing -*əz*, that is, when *tass* is pluralized "correctly," is *glasses* still produced by rote, or is it now produced by rule? This we do not know.

One attempt to approach this question uses a reaction time experiment. MacKay (1976) asked subjects to produce the past-tense forms of English verbs and measured the time it took the subject to begin the response. Differences were found, for example, between verbs such as *talk/talked* and *teach/taught*, but it is not clear how this data should be interpreted. MacKay assumes that a longer reaction time will be required as more rules have to be applied to produce the past tense form. But there is no evidence that this is so. It may be that a longer time is required to look up an irregular past tense in the lexicon.

Perhaps a more promising experimental paradigm is used in Stanners et al. (1979). In these experiments, subjects were asked to indicate whether a sequence of letters presented visually constituted an English word or not, and their reaction time was measured. If the same word is presented twice in the same test, the decision time for the second presentation is considerably reduced, even if several minutes intervene between the first and second presentation. The authors found that approximately the same results obtained if the first presentation was a form of a verb inflected with *-s*, *-ed*, or *-ing* and the second presentation was the base verb (e.g., *lends*, *lend*, *pouring*, *pour*). However, if the first presentation was an irregular past tense, for example, *dug* or *shook*, and the second the base verb, *dig* or *shake*, the priming effect still occurred but was not nearly so great. They suggest that regularly inflected forms do not have separate lexical storage, so that activating any one of them activates the base verb, but that irregular past tenses have their own representations. These representations are connected to that of the base verb, however, which accounts for the existence of the (smaller) priming effect of irregulars. While these results are quite consistent with our hypothesis, it should be remembered that the stimuli in these experiments were orthographic, so that the relevance to spoken language remains to be established. This does illustrate, however, one of the ways our hypothesis might be tested rather directly.

4.6. Relative autonomy and the basic-derived relation

Our second hypothesis with regard to autonomy is that autonomous forms are the basic members in the basic-derived relationship. Given two words formed from the same stem, the more autonomous one is more likely to serve as the basis from which the less autonomous one is derived. This follows from our first hypothesis, because in order for a word to be nonautonomous, it must be derived from some word with a lexical representation.

We are proposing that the larger phenomenon of autonomy underlies Vennemann's basic-derived diagram discussed in section 1.1. Notice that our criterion of semantic simplicity for autonomy in a sense restates Vennemann's hypothesis. We have added a consideration of word frequency, however, as an important basis for establishing lexical representations and as an important diagnostic for autonomy. The interaction of word frequency with semantic simplicity predicts that all paradigms need not have the same structure, a situation not countenanced in Vennemann's hypothesis. The consideration of morphophonemic irregularity as a determinant of autonomy ensures the same result. Let us consider briefly how morphophonemic irregularity affects the basic-derived diagram.

Many Spanish verbs have an unchanging stem throughout the paradigm; for example, *cantár* 'to sing' has *cant-* in every one of its forms. It may be sufficient to set up only one lexical representation for such a verb (although the discussion in sections 7 and 8 should be taken into consideration). Some verbs, however, use a modified stem for the preterite; for example, *ponér* 'to put or place' has *pon-* for present indicative, imperfect, future, and conditional but *pus-* for preterite and past subjunctive. This paradigm, then, will consist of two autonomous forms and two sets of basic-derived relations. In one, a form of the present indicative, possibly the 3s, *póne*, will serve as the basis for forming present indicative, imperfect, and so forth, while a preterite indicative, such as 3s *púso*, might serve as the basis for other preterites and the past subjunctive. Actually, this paradigm may even have a third base, the 1s present indicative, *póngo*, which could serve as a base for the present subjunctive forms (*ponga*, *pongas*, etc.), if the velar is analyzed as belonging to the stem rather than to the suffix.

As we mentioned earlier, the major evidence for the basic-derived relation is morphophonemic change in which the allomorph used to express the semantically basic member replaces the allomorph of the derived member. The basic members are usually third person, singular, present tense, and indicative (Mańczak 1958, 1963, 1980). This list has the same membership as the list of forms that are the most resistant to change, which argues for a close relation of the basic-derived diagram with autonomy. Notice further that with the autonomy hypothesis we can predict which *lexemes* are more likely to change with the same hypothesis that predicts the *direction* of morphophonemic change.

5. Relative autonomy in person forms

Before turning to a more detailed discussion of changes in Provençal and Spanish preterite forms, it is necessary to discuss briefly the expected relative autonomy of person/number forms. We will refer only to very general principles that apply to many languages, but these principles are quite sufficient for the data to be discussed.

A general observation about person forms is that third person differentiates itself from first and second by being a non-person, the form used with all nouns (Benveniste 1946). First and second persons are of course extremely restricted in the sense that they refer only to the speaker and hearer, respectively. Considering also that singulars are unmarked with respect to plurals, the 3s form emerges as Jakobson's (1957) unmarked form and Kuryłowicz's "base" form (1947).

The various word frequency counts mentioned in Greenberg (1966) show third person and singular verbal forms to be the most frequent. Juilland and Chang-Rodriguez (1964) is a frequency count of Spanish words that distinguishes tense, mood and person forms of verbs. Unfortunately, no totals are given by morphological category membership, so it was necessary to take a small sample of verbs to obtain totals. For this purpose forty-nine highly frequent verbs were chosen, and the frequency of the person forms in the present and preterite were tabulated.[12] Juilland and Chang-Rodriguez (1964) used texts written in peninsular Spanish between the years 1920 and 1940 and distinguished five different styles. Written discourse would not be relevant for the study of person inflections because of the lowered frequency of first and second person reference as compared to spoken discourse. For this reason we have used only the figures for the category "Plays," in which the frequency of person forms should be comparable to spontaneous spoken discourse.

Another frequency count (Rodriguez Bou 1952) has some advantages over this one: it records actual spontaneous conversations of children in grades 1 through 6, and it is based on considerably more tokens. One disadvantage for our purposes here is that it is based on Puerto Rican Spanish, where the 2p is always expressed with the 3p form. Another disadvantage is that homonyms are not distinguished. Thus the 2s familiar imperative is included in the 3s present indicative, and the relation of lp present and 1p preterite had to be established on the basis of verbs that use different forms for the two, since many verbs use the same form. Despite these problems, the two frequency counts show very similar results:

(14) *Juilland and Chang-Rodriguez*

Present (n = 3570)				Preterite (n = 405)			
1s	23%	1p	7%	1s	31%	1p	4%
2s	16%	2p	1%	2s	7%	2p	0%
3s	44%	3p	9%	3s	47%	3p	10%

Rodriguez Bou

Present (n = 14,332)				Preterite (n = 10,414)			
1s	24%	1p	4%	1s	22%	1p	4%
2s	11%	2p	—	2s	4%	2p	—
3s	41%	3p	20%	3s	51%	3p	19%

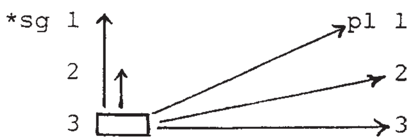
These figures show 3s to be the most frequent, as predicted by its status as least marked form. The competition offered 3s by the 1s form is striking, especially in the preterite, and clearly sets off 1s and 3s from the other forms. The figures suggest relative autonomy for 1s forms, and this suggestion is supported by other facts as well. J. Greenberg (1966:44) mentions the occurrence of zero marking for 1s in Dutch and German and suggests that between first and second person first is the least marked. (This is contrary to Jakobson's [1957] claim that first person is marked with respect to second.) It is easy enough to speculate on why 1s forms are so frequent even though their reference is so restricted and even though it is difficult to make a case for the semantic simplicity of 1s. The high frequency of 1s reflects the egocentric bias of language users, which is reflected in language in many other ways (Zubin 1979).[13]

Examining now the other forms, we find a frequency ranking after 3s and 1s of 1p and 3p (although the order is reversed in the preterite), followed by lp, then 2p. We will use this frequency ranking as a rough guide to autonomy in our discussion of the relations among person forms. This frequency count, although made on Spanish, is applicable to Provençal, as long as we bear in mind that Spanish uses the third person verbal forms for second person in formal contexts, thereby detracting from the frequency of second person in favor of third, while in Provençal the 2p verb form is used for 2s in formal contexts. Thus for Provençal dialects the frequency of 2p will be greater and the frequency of third person will be somewhat less. This will not affect our main points, however, since we are confident that even in Provençal 3s is by far the most frequent verb form and 1s is next in order.

6. Paradigm organization

The formal organization of a paradigm diagrams the semantic organization so that forms that are more similar semantically will be more similar in morphophonemic shape. Thus two forms sharing many semantic properties will be related in that the less autonomous of these will be derived from the more autonomous base form. Given that the six persons within the preterite (of Spanish and Provençal) are all closely related and 3s is the most autonomous of these, we might expect a set of basic-derived relations such as those indicated by the arrows in (15). However, we have found no evidence for paradigms organized as in (15):

(15)

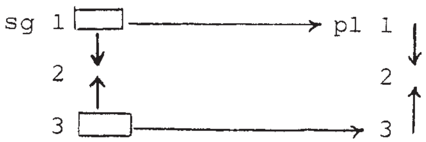


Rather, the person forms seem to be linked to 3s via some semantically related form with intermediate autonomy. Thus we have found evidence for a set of relations as shown in (16):

(16)


In (16) 3s is the basis of the paradigm and 1s, 2s, and 3p are derived directly from it. But 1p and 2p are derived from 3p, rather than directly from 3s. In this pattern the number distinction cuts across the person distinctions and generalizations apply within the singular and plural.

A second organizational type we have found evidence for is built around two base forms, 3s and 1s. Here the relative autonomy of 1s not only exempts it from being derived from 3s but also allows it to serve as the basis of other closely related forms. Evidence will be presented below for the relations diagrammed in (17):

(17)


This diagram represents the possibility that 1p may be derived directly from 1s. The relatively nonautonomous 2s and 2p are in an ambiguous situation and may be derived from either of the more autonomous forms.

One logical possibility for which we have not found any dynamic evidence is the derivation of 2p from what should be the more autonomous 2s. This should be possible because both forms are second persons. The reason that this alternative is not attested in our data is that the 2s form, while more autonomous than 2p, is not autonomous enough to be formally independent of 1s and 3s.

Given the two organizational possibilities shown in (16) and (17), or combinations of them, we are able to explain a wide array of morphophonemic changes that have taken place in preterite forms in Provençal and Spanish. It should be noted that our account of these changes assumes that the preterite forms (especially 3s) are autonomous vis-à-vis the present-tense forms. There are several reasons for regarding this as highly plausible. First, preterite forms in these languages are the most used of past forms and are second only to present-tense forms in frequency.[14] Second, there is a certain amount of morphophonemic opacity in these preterite forms. In order to derive a 3s preterite from a 3s present, it is necessary to go through two steps: a change in the final vowel and a change in the placement of stress. The result is a form that cannot be clearly segmented into parts corresponding to the components of meaning, as we mentioned earlier. Finally, there is evidence from children learning Brazilian

Portuguese, whose verbal morphology and morphophonemics are quite similar, that 3s preterite forms are learned as autonomous. Portuguese distinguishes three conjugation classes, as illustrated by the following forms:

Infinitive:		*falar* 'to speak'	*bater* 'to hit'	*abrir* 'to open'
Present:	1s	fálo	báto	ábro
	3s	fála	báte	ábre
Preterite:	1s	faléi	batí	abrí
	3s	falóu	batéu	abríu

These are the first four forms used by children. The 3s present is used first; then the 3s preterite is added. Note that the three conjugation classes are fully distinguished only in 3s preterite and only two classes are distinguished in 3s present. Simões and Stoel-Gammon (1979) report, however, that all four children they studied used the 3s preterite forms with correct suffixes from the very beginning. However, for 1s preterite forms the first conjugation suffix *-ei* was used with all verbs, including second and third conjugation verbs. This suggests an important difference between the way 3s and 1s preterite forms are acquired. In particular, it seems that 3s forms are learned as whole units, while 1s forms are produced by rule in this case.

7. Provençal preterites

Earlier we discussed the Provençal strategy for creating new preterite forms using the 3s, a linking consonant, and a person/number suffix. In some dialects, all persons except 3s are re-formed. In other dialects, only some of the persons are re-formed. The particular changes attested in the dialects form interesting patterns that can be explained by referring to degree of autonomy and morphological category membership.

All the types of preterite found in Provençal are shown in (18); the subsequent corresponding table in (19) shows the form patterns that are found in Provençal dialects. The "x" marks the forms with linking consonants. In the case of *r*, this consonant is etymological in 3p.

(18)

A. *Bas-Limousin* (Ronjat 1937:252)[15]

parlí	parlén
parléra	parlá
parlé	parléron

B. *Alais* (Ronjat 1937:263)[16]

parlére	parlén
parléras	parlés
parlé	parléron

C. *Forcalquier* (Ronjat 1937:226)

parlí	parlérian
parléres	parlérias
parlé	parléroun

Charente (Meyer-Lübke 1890/1902 [1925]:352)[17]

cantí	cantétem
cantétei	cantétei
cantét	cantéten

D. *Provençal litt.* (Ronjat 1937:224)

parlére	parlérian
parléres	parlérias
parlé	parléron

Clermont-Ferrand (Ronjat 1937:193)[18]

cantéte	cantétem
cantétes	cantétetz
canté	cantéton

(19)

A.	sg 1	pl 1	B.	sg 1x	pl 1
	2x	2		2x	2
	3	3x		3	3x
C.	sg 1	pl 1x	D.	sg 1x	pl 1x
	2x	2x		2x	2x
	3	3x		3	3x

We have already observed that the 3s form does not change, and we have claimed that this form serves as the basis for the other changes. The other striking regularity in the paradigm is that 2s is *always* changed, even if no other changes have taken place, as in A. This is because 2s is the least autonomous of the singular forms. We hypothesize that the 1s and 2s forms are the most likely to be re-formed because they are more closely related to 3s than the plurals. Between 1s and 2s, the 2s form is more likely to change than 1s because 2s is always less autonomous than 1s. These two factors together, the degree of autonomy and the morphological category membership, explain why 2s is most readily re-formed.

The 1s also undergoes change in some dialects, as we might expect, since it is a singular. However, 1s is the only form (other than 3s) that can still resist a change, even when all other forms have undergone it. Here, then, we have evidence for the relatively high autonomy of 1s.

In evaluating the changes in plural forms, it is important to remember that the *-r-* originates in the 3p. With this in mind, it can be seen that either all the plurals take the linking consonant or none of them do. This would depend upon whether the consonant is taken to be a general marker of preterite. If it is, none of the plural forms is autonomous enough to resist re-formation, so they all adopt the consonant and the new set of suffixes. We do not interpret types A and B as having 1p and 2p forms that resist change, but rather it appears that the re-formation has been restricted to the singular. Note in this regard that preterites using *-t-* or other consonants show a uniform re-formation of the plural.

8. Spanish preterites

Changes in Spanish preterites involve the vowel following the verb root, which is usually referred to as the theme vowel. This vowel was originally uniform throughout the tense, but due to a number of regular sound changes alternations in this vowel were introduced. The standard Castilian preterites are repeated here as examples of forms unaffected by the type of change we will discuss:

(20)

First conjugation		*Second and third conjugations*	
canté	cantámos	dormí	dormímos
cantáste	cantásteis	dormíste	dormísteis
cantó	cantáron	durmió	durmiéron

In many dialects, the first conjugation forms have *e* or *o* rather than *a* as the theme vowel in the 2s and the plural, giving evidence for re-formation of these person forms on the basis of 1s and 3s. The exact patterns of re-formation will be our concern in this section.

The examples in (21) illustrate all the attested patterns for theme vowels in first and second person of first conjugation. We will discuss the third person changes separately below.

(21)	A. *La Ribera*[19]		B. *Valle de Aragués*[20]	
	canté	cant*é*mos	canté	cant*é*mos
	cantátes	cantátis	cantástes	cant*é*is
	cantó	cantáron	cantó	cantóron
	C. *Asturias*		D. *Lena*[21]	
	canté	cantámus	eanté	cant*é*mos
	cant*é*ste	cantásti	cant*é*ste	cant*é*steis
	cantóu	cantánu	cantó	cantáron

The italicized vowels are the changed theme vowels that we will be discussing. There are also a number of changes in the person/number inflections, but we will have to ignore them and concentrate on the theme vowels.

The paradigm in A illustrates a single change—the re-formation of the lp form using 1s as a base. A form such as original *cantámos* is replaced by *cantémos*, which is made up of 1s *canté* plus the lp marker *-mos*, which is the unique and constant marker of 1p. This change is apparently the first one to occur, with ample documentation in León and other areas in the Middle Ages (Krüger 1954:72). It is also the most widespread geographically. This we attribute to the high autonomy of 1s and the close relation of 1p to 1s.

The B-paradigm type has undergone the same change as A and one further change—the 2p form has also been remade using the *e* vowel. The change of 2p is in some way dependent on the prior change of 1p. The change of 2p does not, however, depend on a change in 2s, as seen in B. This suggests that 2p is re-formed because of its relationship to 1p.

Only one case of the C-type paradigm was found, and this occurs where neighboring dialects usually have the D type, which is extremely common. Both types show that 2s may be re-formed on the basis of 1s. If all three changes occur, as in D, the result is a system in which all first and second persons have the same theme vowel, distinguishing them from third person. This would create a nice opposition of "participant in the speech event" versus "nonparticipant" if it were not for the fact that the third person verb form is used with *usted*(*es*) for second person formal. Thus the distinction between participant and nonparticipant applies only to informal speech situations. It may be that this constitutes the real generalization, but we are hesitant to make any claims in this regard. In any case, it is clear that 1s is serving as the morphophonemic base for both first and second persons.

It should be further noted that these changes across persons demonstrate that the primary function of the *é* of 1s is not to mark 1s but rather to mark preterite. If it were primarily a person marker, it could not serve as the basis for the re-formation of the second persons. This interpretation of the *é* as a marker of the more important verbal categories, tense/aspect, rather than person/number, is another reflection of the hierarchy discussed in section 1.2.

The patterns of theme vowel distribution illustrated in (21) represent the full range of attested patterns. The changes that led to these patterns give evidence for relations among first and second persons as illustrated in (22):

$$\text{(22)}\quad \begin{array}{llcll} \text{sg} & \boxed{1} & \longrightarrow & \text{pl} & 1 \\ & \downarrow & & & \downarrow \\ & 2 & & & 2 \end{array}$$

These examples clearly indicate that 1s can serve as the basis for other morphologically related forms.

In some dialects, another parallel series of changes overlays those we have just discussed. These changes occur in the dialects of León that have a diphthong in 1s, that is, *ei* (< *ai* < *avi*). Consider first Aliste, which has the diphthongs and has undergone the changes we have just discussed (Baz 1967):

(23) *Aliste*

sg	1	-éi	pl	1	-émos
	2	-éstes		2	-éstes
	3	-óu		3	-ón

Some dialects, such as Sanabria and Sisterna, farther to the north, have undergone a second re-formation of participant forms, using the 1s diphthong as the basis:[22]

(24) *Sanabria* (Krüger 1954:54) *Sisterna* (Alvarez Fernández 1960)

sg	1	-éi	pl	1	-éimos	sg	1	-éi	pl	1	-éimus
	2	-éste		2	-éstes		2	-éiste		2	-ésti, -éisti
	3	-óu		3	-ónen		3	-óu		3	-ánun

As in the cases discussed above, the most commonly attested change is that affecting the 1p form. The difference between this change and the preceding one is that 2s seems to be affected more commonly than 2p (Krüger 1954). This can be seen in the forms in Sisterna, where 2s has changed completely, but 2p still has two competing forms. More evidence of this sort would lead us to postulate a relation between 2s and 2p.

With the 1s serving as the basis for first and second persons, the 3s, in most dialects, has only 3p within its domain. Here, quite predictably, we find the *ó* of 3s showing up in 3p. There are a variety of 3p forms attested. The simplest occurs in Aliste (see [23] above) where the stressed *ó* of 3s is simply followed by the *-n* that marks 3p in all other tenses and moods. Most dialects, however, maintain the bisyllabic nature of the 3p suffix, altering it from *-áron* to one of the following: *-óron*, *-ónon*, *-órun*, *-óren*, *-órin* (Krüger 1954:76). Some of the variation in the unstressed vowel is phonological. The point that interests us, however, is the *ó* in the stressed syllable, which shows that 3p has been re-formed from 3s.[23]

To complete the picture of theme vowel changes, it should be mentioned that there are also regions in Aragón where the preterite theme vowel is consistently *ó* (Alvar 1953:236–237):

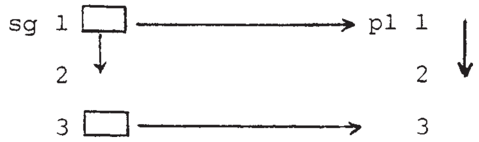
(25) cantó cantómos
cantós cantóz, cantóis
cantó cantóron

In this case all the person forms are rebuilt on the 3s base. There are also paradigms such as the following (Alvar 1953:236):

(26) canté cantémos
cantés cantéz
canté cantéron

This paradigm appears to have built on the 1s form, which would seem to be a possibility, given what we have learned about the role of 1s. However, Menéndez-Pidal (1941[1968]), Alvar (1953), and Zamora Vicente (1970) all subscribe to the idea that the *e* comes from the Latin *-avit*, which, instead of reducing to *-avt* > *-awt* > *-o*, reduced to *-avit* > *-ait* > *-e*. Malkiel (1976:449) points out, however, that this "pattern of contraction" is unproven and that the morphological explanation is quite plausible, given the cases where 3s *o* serves as preterite marker. Further, we may add, given the fact that 1s *e* serves as the base for first and second persons in so many dialects, we are not surprised to find a few cases where it serves as the basis of the entire tense paradigm.

All of these changes show that the theme vowel in the preterite is coming to be a signal of preterite that shows variation corresponding to the person distinctions. In the Provençal data we saw that 3s was the basis of all the re-formations and that the autonomy of 1s was apparent in the resistance that form exhibited to changes on the basis of 3s. In the Spanish data we see 1s not just resisting change but serving as the basis for the changes in other forms. The relations evidenced in the forms in (21) can be diagrammed as follows:

(27)


The precise directions of change in (25) and (26) cannot be discerned. It is clear from (25) that 3s may serve as the base for 1s, as in some dialects of Provençal, and (26) opens the possibility of 1s serving as the base for 3s.

The fact that 1s resists change in Provençal but actually serves as the basis for change in Spanish dialects is an important argument for the autonomy hypothesis. One could argue that 1s is not serving as a basis for remodeling 1p, 2s, and 2p but rather that a rule rewriting the preterite marker as /e/ in 1s has been generalized to apply to all participant categories (Long 1979). However, the resistance of 1s to change in Provençal demonstrates that a 1s preterite form can have its own lexical representation, since, as we argued above, this is the most straightforward way to account for resistance to change. If a form that resists change has its own lexical representation, then a form that not only resists change but also serves as the basis of change must also have its own lexical representation.

The competing autonomy of 1s and 3s opens up several possibilities for divisions of the six persons. In particular the second persons are in a somewhat ambivalent position—they may pattern with the first persons as participants and in opposition to nonparticipating third person, or they may be derived from the third person, creating a division between ego and non-ego forms. Thus in Provençal, 2s is derived from 3s, while 2p seems to always follow 1p.

Although we do not completely understand the factors that influence the way speakers organize and perceive the system, for example, whether second person relates to first or to third, we have been able to narrow considerably the range of possible organizations of the six person/number forms by employing the notions of autonomy and morphological relatedness. The following three generalizations hold for the many dialects we have considered:

1. 3s can serve as the basis of all other persons.
2. 1s can serve as the basis of first person and second person forms (and possibly all persons).
3. Singulars can serve as the basis for plurals.

As a final illustration of the organizational possibilities of person and number, we turn to a brief discussion of the verbal paradigms of Bielsa, where changes according to person have taken place not only in the preterite but also in the imperfect and conditional.

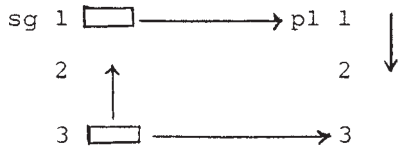
9. The organization of persons in Bielsa

In section 3.2 we discussed the preterites in Bielsa to show that the second and third conjugation suffixes were taken directly from 1st conjugation via the 3s form. The preterite forms are repeated here:

(28)

First conjugation		*Second and third conjugations*	
llevé	llevémos	metié	metiémos
llevóres	llevéz	metióres	metiéz
llevó	llevóren	metió	metióren

This is the only Spanish dialect we have encountered where we find *r* used as a linking consonant in the 2s, as it is in Provençal. Neighboring Catalan, however, has a regular preterite with *r* in the 2s and throughout the plural. As in the many cases discussed above, it is apparent that the 2s is formed directly off of 3s with the *r* or *re*, plus the 2s suffix *-s*. The 3p may also be analyzed as derived from 3s. In fact the reformation of 2s depends upon the analysis of 3p as *llevó* + *r* + *en*. However, 1p and 2p have been re-formed from 1s. (The *-z* of 2p is the unique and constant marker of 2p throughout the verbal paradigm.) The lines drawn among the forms in (28) show the division of the paradigm due to the relations illustrated in (29):

(29) 

What is interesting about this dialect is that several other tenses exhibit similar relations among the persons. Note first that this set of relations is apparent in the future tense of standard dialects, as well as in Bielsa:

(30) *Future*

Bielsa		*Castilian*	
llevaré	llevarémos	levaré	llevarémos
llevarás	llevaréz	llevarás	llevaréis
llevará	llevarán	llevará	llevarán

The only difference between Bielsa and Castilian is in the 2p marker. The conditional forms in Bielsa exhibit the same pattern, but this is due to changes undergone in Bielsa:

(31) *Conditional*

Bielsa		*Castilian*	
llevarí	llevaríos	llevaría	llevaríamos
llevarías	llevaríz	llevarías	lievaríais
llevaría	llevarían	llevaría	llevarían

In Castilian *ía* always follows the base *llevar* to mark conditional and the person/number suffixes follow *ía.* In Bielsa only *í* is present in 1s, 1p, and 2p, distinguishing these forms from 2s, 3s, and 3p. We do not know the mechanism for this change, but the result suggests that once 1s is changed, 1p and 2p were rebuilt with 1s as a base.

The imperfect indicative has also undergone a number of changes. Observe, however, that the consistent *b* throughout the Bielsa paradigms is a conservative feature:

(32) *Imperfect indicative*

Bielsa: first conjugation		*Castilian: first conjugation*	
llevábe	llevános[24]	llevába	llievábamos
llevábas	llevábez	llevábas	llevábais
llevába	llevában	llevába	llevában

Second conjugation		*Second and third conjugations*	
metébe	meténos	metía	metíamos
metébas	metébez	metías	metíais
metéba	metèban	metía	metían

Third conjugation

partíbe	partínos
partíbas	partíbez
partíba	partíban

In Castillian 1s and 3s are both unmarked for person, but in Bielsa the 1s marker is *e*. The 2p form also has this *e* plus the 2p marker -*z*. The 3s, 2s, and 3p all contain *ba*. The latter two can therefore be considered to be formed from the 3s. The 1p form has peculiarities all its own, which are discussed briefly in note 24. The imperfect indicative, then, divides up in much the same way as the preterite, future, and conditional.

These forms illustrate quite nicely the ambivalent position of the least autonomous forms, the second persons, and the nonnecessity of a symmetrical system. The pairing of 1p and 2p has precedent in the verbal system of Romance, in present indicative and subjunctive, where stress on the suffix vowel sets these forms apart from the others, which have stress on the stem. This stress difference gives rise to vowel alternations in both French and Spanish.

10. Conclusion

With these forms from Bielsa, we can illustrate the hypothesis concerning lexical storage that emerges from the autonomy hypothesis. The data considered here suggest that there could be as many as two lexically stored words for some tenses for the verbs of Bielsa, since there appear to be two autonomous words for each tense. Thus the preterite in Bielsa could consist of the lexical forms *llevé* (1s) and *llevó* (3s), plus a set of rules that add suffixes to derive the 1p and 2p from the 1s, and 2s and 3p from the 3s. Similarly, the future will have the forms *llevaré* and *llevará*, the conditional *llevarí* and *llevaría*, the imperfect *llevábe* and *llevába*.[25]

We have argued that the autonomy hypothesis helps to explain why 1s can resist change in some Provençal dialects and serve as the basis of change in some Spanish dialects. We have similarly presented the autonomy hypothesis as the best explanation for the role of 3s in the re-formation of the preterite in Provençal. An alternative account would claim, perhaps, that the 3s form is not being used as the base for other person forms, but rather that the rules that produce 3s are spreading to other persons. The difficulty with this account lies in explaining why a rule that spells out 3s as *t* or some other consonant would ever apply to a person other than 3s. Similarly, if one claimed that the rule that supplies *r* in 3p spreads to other persons, then one would have to explain why it always spreads to 2s and only sometimes to the rest of the plural and 1s. However, the notions of autonomy and degree of relatedness provide good explanations for these changes.

While we have argued from several examples that a speaker's analysis consists of an autonomous word or a paradigm plus rules to derive other closely related words, we have not shown that speakers *never* analyze a paradigm into a stem plus affixes, where the stem does not appear in the language as an autonomous word. In fact, it is conceivable that in more transparently agglutinative languages this type of analysis might be possible. Moreover, in the same languages we have been discussing here, there are cases that appear to show segmentation of stem and suffix. In section 3.2 we discussed the second and third conjugation preterite in Bielsa, saying that the creation of a new 1s preterite form *metié* [metyé] depended on the analysis of 3s *metió* as [mety + ó]. We have observed also that the existence of autonomous preterite forms does not mean that speakers do not know how to make a preterite given some other form of the paradigm. That is, there must be rules that link autonomous forms to one another, at least in morphophonemically regular cases. The question, then, is what is the nature of these rules and do they presuppose the segmentation and storage of the stem as a separate unit? This would certainly be one analysis, but another possible analysis maintains the notion of an autonomous word. Under this analysis, an autonomous base form may be modified, not just by affixation but also by changing features or segments. Thus *metió* serves as the base for *metié* in the sense that *metié* can be derived from *metió* by changing one feature of the final vowel, but without ever extracting and storing a separate stem. We would like to leave these and other questions concerning specific analyses open until further debate and more data can clarify the issues.

While the notions of autonomy and degree of relatedness do not in themselves settle all problems that might arise in trying to write a grammar, they do make some very specific predictions about the likelihood of forms to change and the direction of change, based on facts about frequency, morphophonemic irregularity, semantic simplicity, and semantic relatedness. Furthermore, these hypotheses make predictions that can be tested cross-linguistically, concerning the distribution of stem alternants and the location of morphophonemic irregularity and portmanteau expression in a morphological system.

We hope to have shown, furthermore, that morphological analysis does not need to depend entirely on considerations of economy, symmetry, and distributional facts. Indeed, these might be of very little importance in a speaker's analysis. Instead, if we look at dynamic data, we find that facts about child language and historical change support the hypothesis that the speaker's analysis is guided by certain psychological principles of acquisition and storage. These are the considerations that should be input to the development of morphological and morphophonemic theory.

The theory we have been developing here is a natural theory in the sense used by Donegan and Stampe, who describe natural phonology as follows:

> This is a natural theory . . . in that it presents language . . . as a natural reflection of the needs, capacities and world of its users, rather than as a merely *conventional* institution. It is a natural theory also in the sense that it is intended to *explain* its subject matter, to show that it follows naturally from the nature of things; it is not a conventional theory . . . in that it is not intended to *describe* its subject matter exhaustively and exclusively. (1979:127)

Morphology and morphophonemics have been subject to a great deal of description—from traditional grammars on the Latin model to *The Sound Pattern of English*. Morphological systems have not in general been subject to explanation, nor have they been treated as natural objects whose properties follow from their functions, whose outward form diagrams their content. We. hope to have shown here that it is useful to approach morphology and morphophonemics in just this way.

Notes

We are grateful to Henning Andersen, Edith Bavin, Suzanne Fleischman, Blair Rudes, Dan Slobin, Joe Stemberger, Peter Tiersma, Yakov Malkiel, David Zager, and David Zubin for discussing this research with us at various stages of its development and, moreover, for asking many difficult questions. We alone are responsible for the way we have decided to answer these questions. This research was supported in part by a Social Science Research Council Post-Doctoral Training Fellowship to the first author.

1. We employ the following abbreviations:

1s: first person singular	1p: first person plural
2s: second person singular	2p: second person plural
3s: third person singular	3p: third person plural

Person forms of verbs will be presented in the order given here.

2. Stress is indicated throughout by an acute accent and vowel length by an overbar.

3. In 3s the loss of *i* rather than *v* leads to the suffix *-aut* (or *avt*), which is found, along with *-ait*, in inscriptions. Menéndez-Pidal (1941[1968]:309) says that *-aut* prevailed in Latin. Non–first conjugation forms such as *posiut* are also found in Spanish inscriptions.

4. The *-steis* suffix of 2p for *-stes* comes from the addition of 2p *-is*, which occurs elsewhere, to the preterite in place of the simple *-s* (Menéndez-Pidal 1941[1968]:280).

5. The first conjugation suffixes were replaced by those of third conjugation, so that the inflections in (4) do not derive from those in (2) at all.

6. We use the term *Old Provençal* to refer to the "langue d'oc" and *Provençal* to refer to its descendants, as distinct from the descendants of the "langue d'oil" (French) to the north. Bec (1963) points out that Old Provençal was a koiné used for literary and administrative purposes and does not represent a homogeneous regional dialect. This is of little concern since our purpose here is merely to point out the analytic difficulties presented by the preterite at this time. Our theoretical points will concern the developments in the modern dialects, which are well documented.

7. The Provençal dialects shown on the map are scattered over the entire Provençal area and are representative of all the dialects to which we found any reference. The Spanish dialects are concentrated in the northern part of Spain because this is the only area where the changes to be discussed take place, with the possible exception of the change of 1p preterite *cantámos* to *cantémos*, which is quite widespread.

8. According to Ronjat, this is a quite recent phonetic development, not yet completed in some areas. In Marche and in some subdialects of Limousin, for example, the 3s *-t* is still pronounced, but only when followed by a pronominal subject.

9. The 1s and 2s inflections were taken from the present tense and the 1p and 2p inflections from the imperfect.

10. The Castilian forms are probably not identical with the forms that preceded the current Bielsa forms, but they do show the etymological vowels following the stem.

11. There are several reasons that we have chosen not to phrase the discussion entirely in terms of markedness. First, we do not consider markedness to constitute explanation in itself; rather, we propose to break the notion of markedness into its constituent parts to determine the causal relations among them. We find a variety of causes for autonomy, and these may be somewhat obscured by adopting a unitary notion of markedness. Further, we do not

believe that markedness relations are the same type in phonology and morphology, and we suspect that even within the realm of morphology there may be a variety of types of relations, which might remain uninvestigated if all relations were assumed to be of a single type. Finally, we feel that it is more useful, in the sphere of morphophonemics, to develop the notion of autonomy, which, unlike markedness, is a property of *words* that is independent of morphological relations.

12. We chose the same verbs that Gili Gaya used to obtain the frequencies reported in n. 14, with the exception of *ser*, which was excluded because it constituted nearly half of the tokens. With the other count, the *Recuento*, only thirty-three verbs could be used, because the others had noun homonyms for some crucial form, which distorted the frequency.

13. Note that children acquiring languages with person/number inflection use 1s and 3s verbal forms before any of the others (Mikeš 1967; Pačesova 1976; Simões and Stoel-Gammon 1979).

14. Gili Gaya (1960) chose fifty of the most frequent verbs used by schoolchildren and tabulated the frequency, according to the *Recuento*, of these verbs in various verbal forms. The five most frequent were:

Present indicative	64,372
Preterite indicative	36,247
Infinitive	29,585
Imperfect indicative	12,359
Imperative	10,423

The relative prominence of the preterite is quite apparent here.

15. Bas-Limousin is the only dialect we have found that displays this particular pattern.

16. Other dialects with this pattern are Gévaudan (Ronjat 1937) and Nîmes (Ronjat 1937).

17. This pattern is also found in Couserans Limousin, Bassiere-Badil, Rochechouart, Saint-Mathieu, Ambert (Ronjat 1937); in Confolens, Montbron, Montembœuf (Chabaneau 1874/1875); in Gilhoc (Meyer-Lübke 1890/1902). This is also the pattern of regular preterite formation in Catalan, with *r* occurring in all persons, except 1s and 3s.

18. This is the most common pattern for the preterite, and is found in Forcalquier, Marseille, Gap, Montignac, Aurillac, Clermont-Ferrand, Toulouse, Chabrillan, Loriol, Châtillon-en-Diois, and throughout Aquitaine (Ronjat 1937); in Niçois (Compan 1965); in Rhodanien (the modern "standard" Provençal, as reported by Ford [1921(1966)]); and in Provençal *littéraire* (Ford 1921[1966]).

19. Preterites with *e* in 1s and 1p only are documented in Asturias (Garcia de Diego 1946:168), the two provinces of Castille (Menéndez-Pidal 1941:311–312), and in the Pasiego dialect (Penny 1969), as well as in La Ribera (Llorente Maldonado de Guevara 1947).

20. Preterites with *e* in 1s, 1p, and 2p are found in the valley of Aragüés (González-Guzman 1953), in Babia and Laciana (Alvárez 1949), and in some of the Sanabria dialects studied by Krüger (1954).

21. Preterites with *e* in 1s, 1p, 2s, and 2p are found throughout Asturias, León and Aragón (Zamora Vicente 1970) specifically in Sisterna (Alvarez Fernández 1960), Cabrales (Cañedo 1963), Lena (Neira Martinez 1955), Aliste (Baz 1967), and La Ribera (Llorente Maldonado de Guevara 1947).

22. Krüger (1954) argues convincingly that the *-ei-* diphthongs in 2s, 1p, and 2p are not etymologically derived from *-avi-* as the 1s diphthong is. The argument is that this assumption leaves unexplained the *-e-* that occurs in 2p in Sisterna and in other forms in related dialects. Since in these dialects *ei* has quite consistently remained without monophthongization, the older *e* in 1p, 2s, and 2p must be analogical, and therefore the newer *ei* must also be analogical.

23. A much rarer form is *-eron* (Krüger 1954:52). This is usually attributed to the second and third conjugation 3p suffix *-ieron.*

24. A form similar to the Castillian 1p, *llevábamos*, probably preceded this Bielsa form *llevános.* The *b* has been deleted, possibly because of the phonetic environment: the 1p form is the only one is which two syllables follow the *b.* The change of *-mos* to *-nos* probably occurred before the deletion of the *b.* We suspect this because the appearance of *nos* rather than *-mos* is a widespread dialectal phenomenon in 1p forms with antepenultimate stress, e.g., *llevábanos*, and present subjunctive forms in dialects with stem stress, *llévenos.* This develops because speakers consider the verbal suffix *-mos* and the clitic pronoun *nos* to be allomorphs of a single 1p morpheme. The clitic *-nos* occurs attached to the end of the verb when the verb is an infinitive (*hablárnos* 'to speak to us'), a gerund (*hablándonos* 'speaking to us'), or imperative reflexive (*sentémonos* let's sit down', in which the *s* of *mos* is deleted). In the latter two cases, the addition of *-nos* produces antepenultimate stress. The form *mos* occurs primarily with penultimate stress (except in the imperfect indicative and subjunctive). Apparently, speakers form the abductive generalization that *-nos* appears preverbally and postverbally with antepenultimate stress, while *-mos* occurs only postverbally, with penultimate stress. Deductive application of this rule led to these changes in the imperfect forms.

25. This model is similar in some respects to the model of suppletive representation proposed by Hudson (1975) and applied in Hooper (1976a). What is better captured in the current model is the subordinate or derived status of certain forms, which is dependent upon morphological category membership. The arguments given in Hooper (1976a) in support of Hudson's model, i.e., that historical changes show some paradigms to have more than one underlying form, also apply to the current model.

References

Alvar, M. 1953. El dialecto aragonés. Madrid: Editorial Gredos.

Álvarez, G. 1949. El habla de Babia y Laciana. *Revista de filología española* (Madrid) 49.

Álvarez Fernández, J. 1960. El habla de Sisterna. *Revista de filología española* 74.

Andersen, H. 1969. A study in diachronic morphophonemics: The Ukrainian prefixes. *Language* 45, 807–830.

———. 1980. Morphological change: Towards a typology. In *Historical morphology*, ed. J. Fisiak. The Hague: Mouton.

Anglade, J. 1921. *Grammaire de l'ancien provençal*. Paris: Klincksieck.

Badia Margarit, A. 1950. *El habla del valle de Bielsa*. Barcelona: Consejo superior de estudios científicos (Instituto de estudios pirenáicos).

Baz, J. M. 1967. El habla de la tierra de Aliste. *Revista de filología española* (Madrid) 82.

Bec, P. 1963. *La langue occitan*. Paris: Presses universitaires de France.

Benveniste, E. 1946. Structure des relations de personne dans le verbe. *Bulletin de Ia société linguistique de Paris* 43, fasc. 1, no. 126, 1–12.

Berko, J. 1958. The child's learning of English morphology. *Word* 14, 150–177.

Blaylock, C. 1975. The Romance development of the Latin verbal augment *-sk-*. *Romance Philology* 28, 434–444.

Bloomfield, L. 1933. *Language*. New York: Holt, Rinehart and Winston.

Bolozky, S. 1980. On the autonomy of the tense paradigm: The Modern Hebrew evidence. *Lingua* 51(2–3), 115–123.

Bybee, J. 1980. Morphophonemic change from inside and outside the paradigm. *Lingua* 50, 45–59.

Cañedo, J. A. F. 1963. El habla y la cultura popular de Cabrales. *Revista de filología española* (Madrid) 76.

Chabaneau, C. 1874/1875. Grammaire limousine. *Revue des langues romanes* 6, 171–205; 7, 145–178.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Compan, A. 1965. *Grammaire niçoise*. Nice: Éditions Tiranty.

Donegan, P., and D. Stampe. 1979. The study of natural phonology. In *Current approaches to phonological theory*, ed. D. Dinnsen, 126–173. Bloomington: Indiana University Press.

Ford, H. E. 1921 (1966). *Modern Provençal phonology and morphology*. New York: AMS Press.

Garcia de Diego, V. 1946. *Manual de dialectología española*. Madrid: Instituto de cultura hispánica.

Gili Gaya, S. 1960. Funciones gramaticales en el habla infantil. *Publicaciones pedagógicas*, serie II, no. 24. Río Piedras: Universidad de Puerto Rico.

Gonzalez-Guzmán, P. 1953. *El habla viva de Valle de Aragüés*. Zaragosa: Monografías del instituto de estudios pirenáicos.

Grandgent, C. H. 1905. *An outline of the phonology and morphology of Old Provençal*. New York: D. C. Heath.

Greenberg, J. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of language*, ed. J. Greenberg, 73–113. Cambridge, Mass.: MIT Press.

———. 1966. *Language universal*s. The Hague: Mouton.

Greenberg, Z. 1975. Overgeneralization of past tense inflection. MS., University of California at Berkeley.

Hooper, Joan Bybee. 1976a. *Introduction to natural generative phonology*. New York: Academic Press.

———. 1976b. Word frequency in lexical diffusion and the source of morphophonological change. In *Current progress in historical linguistics*, ed. William Christie, 95–105. Amsterdam: North Holland. Reprinted in this volume as chapter 2.

———. 1979. Child morphology and morphophonemic change. *Linguistics* 17, 21–50. (Also in *Historical morphology*, ed. Jacek Fisiak, 157–187. The Hague: Mouton, 1980.)

Hudson, G. 1975. Suppletion in the representation of alternations. Doct. diss., University of California at Los Angeles.

Jakobson, R. 1957. Shifters, verbal categories and the Russian verb. In Roman Jakobson, *Selected writings*, vol. 2, 130–147. The Hague: Mouton.

———. 1965. Quest for the essence of language. In: Roman Jakobson, *Selected writings*, vol. 2, 345–359. The Hague: Mouton.

Juilland, A., and F. Chang-Rodriguez. 1964. *Frequency dictionary of Spanish words*. The Hague: Mouton.

Krüger, F. 1954. El perfecto de los verbos *-ar* en los dialectos de Sanabria y sus zonas colindantes. *Revista de filología española* 38, 45–82.

Kuryłowicz, Jerzy. 1947. La nature des procès dits analoguique. *Acta Linguistica* 5, 121–138.

Llorente Maldonado de Guevara, A. 1947. *Estudio sobre el habla de la Ribera*. Salamanca: Tesis y estudios salamantinos.

Long, M. 1979. On Romance preterites and the nature of phonological change. In *Essays in contemporary Romance linguistics*, ed. F. H. Nuessel, Jr., 230–246. Rowley, Mass.: Newbury House.

MacKay, Donald G. 1976. On the retrieval and lexical structure of verbs. *Journal of Verbal Learning and Verbal Behavior* 15, 169–182.

MacWhinney, Brian. 1978. *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development, no. 174, vol. 43.

Malkiel, Y. 1976. From falling to rising diphthongs: The case of Old Spanish _ió <*éu. Romance Philology_ 29, 435–500.

Mańczak, W. 1958. Tendances générales des changements analogiques. *Lingua* 7, 298–325, 387–420.

———. 1963. Tendances gênérales du développement morphologique. *Lingua* 12, 19–38.

———. 1980. Laws of analogy. In *Historical morphology*, ed. J. Fisiak, 283–288. The Hague: Mouton.

Menéndez-Pidal, R. 1941 (1968). *Manual de gramática histórica española*. Madrid: Espasa-Calpe.

Meyer-Lübke, W. 1890/1902 (1925). *Grammaire des langues romanes*, vol. 2. New York: Stechert.

Mikeš, M. 1967. Acquisition des catégories grammaticales dans le langage de l'enfant. *Enfance* 20, 289–298.

Neira Martínez, J. 1955. *El habla de Lena.* Oviedo: Instituto de estudios asturianos.

Pačesova, J. 1976. Some notes on developmental universals in Czech-speaking children. *Neurolinguistics* 5, 192–198.

Paul, H. 1880 (1970). *Principles of the history of language*. Trans. by H. A. Strong. College Park: McGrath.

Penny, R. J. 1969. *El habla Pasiega*. London: Tamesis Books.

Rodriguez Bou, J. 1952. *Recuento de vocabulario español*, vol. 2. Río Piedras: Universidad de Puerto Rico.

Ronjat, J. 1937. *Grammaire historique des parlers provençaux modernes*, vol. 3. Montpellier: Société des langues romanes

Rudes, B. A. 1980. On the nature of verbal suppletion. *Linquistics* 18, 655–676.

Simões, M. C. P., and C. Stoel-Gammon. 1979. The acquisition of inflections in Portuguese: A study of the development of person markers on verbs. *Journal of Child Language* 6, 53–67.

Slobin, Dan I. 1971. On the learning of morphological rules: A reply to Palermo and Eberhart In *The ontogenesis of language*, ed. Dan I. Slobin, 215–223. New York: Academic Press.

Stampe, David. 1973. A dissertation on natural phonology. Doct. diss., University of Chicago.

Stanners, R. F., J. J. Neiser, W. P. Hernon, and R. Hall. 1979. Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18, 399–412.

Vennemann, Theo. 1972. Rule inversion. *Lingua* 29, 209–242.

Watkins, C. 1962. *The Indo-European origins of the Celtic verb*. Dublin: Dublin Institute for Advanced Studies.

Zager, David. 1980. A real-time process model of morphological change. Ph.D. diss., SUNY at Buffalo.

Zamora Vicente, A. 1970. *Dialectología española*. Madrid: Editorial Gredos.

Zubin, D. A. 1979. Discourse function of morphology: The focus system in German. In *Syntax and semantics*, vol. 12: *Discourse and syntax*, ed. T. Givón, 469–504. New York: Academic Press.

Chapter 4

WITH ELLY PARDO

On Lexical and Morphological Conditioning of Alternations

A Nonce-Probe Experiment with Spanish Verbs

1. Background

Over the last six decades or so, American descriptive linguistics has produced a number of models for representing morphology and morphophonemics. These range from totally static item-and-arrangement and word-and-paradigm models to highly process-oriented generative models. In these descriptive models certain problems recur: problems of ambiguous segmentation, problems of formulating morphophonemic rules, deciding on conditioning and directionality. Despite a good deal of ingenuity and elbow grease, arguments on the basis of simplicity, elegance, and capturing generalizations, fail to solve these descriptive problems and do not even approach the question of how real speakers might actually internalize the morphological and morphophonemic systems of their language.

Research into morphological and morphophonemic change has contributed some ideas and has even made it possible to test some hypotheses (Bybee 1980); but the problem with using changes occurring in a natural environment is that if a particular type of change is not reported one is never sure if that change is ruled out in principle or if it just did not happen to occur or has not occurred yet. By examining a large number of changes in various languages some generalizations or "laws of analogy" have been proposed (by Kuryłowicz 1949 and Mańczak 1958, 1963, 1980), and these contribute some general understanding of the way speakers internalize morphological and morphophonemic systems.

A more systematic method of gathering data on this subject is by observing how speakers manipulate morphological forms in certain experimental tasks. In particular, the nonce-probe task, as designed by Berko (1958), has been particularly successful. It has been used, however, primarily with children to ascertain the order of acquisition of various morphological and morphophonemic rules. These tasks have been done numerous times on English, which unfortunately has a rather impoverished morphology, and much less often on inflectional languages, such as Spanish (Kernan and Blount 1966) and Hungarian (MacWhinney 1978). What all of these studies are lacking is an in-depth probe of the adult system. Either the adult system is not tested in these studies—that is, the authors assume they know what it is—or it is tested very superficially.

This chapter reports on a set of nonce-probe tasks designed to ask some theoretical questions about the way speakers have internalized their morphological system. The language chosen for the study is Spanish, which has a moderately well-developed inflectional system for verbs, with a few interesting subsystems of verbs with morphophonemic alternations. The nonce verbs used in the task were, for the most part, patterned on real Spanish verbs with an alternation or irregularity of some sort. The idea was to see how the subjects would handle the alternation when they encountered it in a nonce form.

1.1. The diphthong/mid-vowel alternation

A well-known morphophonemic alternation in Spanish verbs is the diphthong/mid-vowel alternation illustrated in (1):

(1) Spanish:

infinitive	*comenzár*	‘to begin’	*contár*	‘to tell’
Pres. Ind.	comiénzo	comenzámos	cuénto	contámos
	comiénzas	comenzáis	cuéntas	contáis
	comiénza	comiénzan	cuénta	cuéntan
Pret. Ind.	comencé	comenzámos	conté	contámos
	comenzáste	comenzásteis	contáste	contásteis
	comenzó	comenzáron	contó	contáron

This alternation is easily predicted on the basis of stress—the diphthong occurs in a stressed syllable and the mid-vowel in an unstressed syllable. The alternation is lexically arbitrary in the sense that there are many mid-vowels that do not alternate, for example, *flotar* ‘to float’ and *aprender* ‘to learn’, which have mid-vowels throughout their paradigms. However, non-alternating diphthongs in verb stems are very rare. They occur only in a few verbs formed from nouns or adjectives with stressed diphthongs, for example, *muébles* ‘furniture’, *amueblár* ‘to furnish’; *viéjo* ‘old’, *aviejár* ‘to grow old’. Despite these few examples, the occurrence of a diphthong in a stressed verb stem is an excellent indication that the verb will have a mid-vowel in unstressed syllables.

This alternation has been described in generative phonology by Harris (1969) by a rule of diphthongization and has been discussed in the generative literature as an important *phonological* rule of Spanish. The rule can be considered a phonological rule despite the fact that it is lexically restricted (i.e., requires a diacritic feature if it is stated as a diphthongization rule applying to certain mid-vowels), because no distinction is made in generative phonology between a rule with morphological conditioning and a rule with phonological conditioning. Such a distinction has, however, been made in natural generative phonology by Vennemann (1971), Hooper (1976), and others at about the same time (Stampe 1973; Skousen 1975). In these works it is emphasized that rules with morphological conditioning are of a fundamentally different nature from true phonological processes: they have exceptions, they tend to be unproductive, they are lexically bound, they make larger structural changes, they may be phonologically "unnatural," and according to Stampe they are even learned in a different way. We know that the truly phonetically conditioned processes of a language will usually be applied by speakers to nonce forms, much as they are applied to borrowed words and neologisms. Thus a Spanish speaker would render a nonce word *tado* as [taðo], using an interdental fricative for intervocalic *d*, while an English speaker would render the same form with an intervocalic flap [taɾo]. While we can make such predictions with some confidence, we know much less about whether, and under what conditions, we can induce speakers to apply morphologically conditioned or morphophonemic rules to nonce forms. That is, is it possible to get speakers of Spanish to use the diphthong/mid-vowel alternation on a nonce verb?

Kernan and Blount (1966) conducted a series of nonce-probe tasks on Spanish-speaking children and adults in a rural area of Jalisco. The researchers' main goal was to investigate the order of acquisition of certain morphological categories in Spanish. The test was given to adults only to arrive at a kind of local norm, to determine what system the children were aiming to attain. One of the verbs on the test was designed to elicit the diphthong/mid-vowel alternation. It was a present-tense verb form *suécha*. Kernan and Blount used this verb form in three items and elicited preterite, imperfect, and past participle forms of this verb. Given what we have said about the distribution of diphthongs in verbs in Spanish, we would expect speakers to assume that a 3s present form *suécha* would have a 3s preterite *sochó*, 3s imperfect *sochába*, and a past participle *sochádo*, since non-alternating diphthongs in verbs are so rare. However, Kernan and Blount's adult subjects all supplied the forms *suechó*, *suechába*, and *suechádo* for these items. The children also used the diphthong in these forms.

The unanimous selection of diphthongs in unstressed stem forms is a rather surprising result. It suggests that speakers are not using the types of rules and representations that our descriptions of the language would predict. It suggests in particular that (1) the diphthong/mid-vowel alternation is lexically bound to the extent that it cannot be extended to new verbs; (2) stress is not a good predictor of the occurrence of diphthongs. Indeed, the existing forms such as *amueblar* and *aviejar* give the same indications and thus should perhaps not be treated as exceptional. Thus we need to re-examine our notion of morphophonemic "rule" and determine to what extent they are indeed rulelike.

1.2. The current experiment

Similar items occurred on the nonce-probe task we constructed (examples are given in appendix 2). This test was administered to twenty-two subjects between the ages of 25 and 35. There were eleven women and eleven men. All had grown to adulthood speaking Spanish as a first language in a Spanish-speaking country outside the United States. We will refer to this group of subjects as the non-Chicano group. We also administered the test to fourteen adult Chicano subjects who lived in the San José and Palo Alto areas of California.

The subjects were told that they were trying out an experiment designed for children. They were shown Maurice Sendak illustrations from a book about Osito (Little Bear) and his friends. They were told that Osito and his friends used some words that were not ordinary Spanish words and that they would use these words in sentences. While the subject looked at the pictures, the experimenter read the sentences. The full texts of the tests are given in the appendices. Their responses were given orally and written down by the experimenter.[1]

In this test the sentences introduced the verbs *bierca* and *duenta* in their 3s present forms. The subjects were asked to complete sentences that required that the verb be conjugated in the 3s preterite. Given Kernan and Blount's results, we were expecting the majority of responses to be *biercó* and *duentó*, respectively. Table 4.1 gives a summary of the results obtained.

Although our results present more variety than Kernan and Blount's, they are still firmly slanted in the same direction. Our subjects did not make the inference that seems logical from the point of view of a descriptive grammar of Spanish, that a stressed diphthong implies an unstressed mid-vowel. Out of a total of seventy-two responses, only three were of the "grammatical" form *bercó* and *dontó*. Some of the variety obtained was due to the use of an *-íó* suffix rather than *-ó*. This is not the result of the assignment of these verbs to second or third conjugation but rather the use of a verb-forming device common for adapting loans and forming neologisms. The suffix for the infinitive is spelled *-ear*, as in *telefonear* 'to telephone' or Chicano *troquear* 'to truck', but the /e/ raises in unstressed position to /i/ or /y/, producing *ió* in the 3s preterite. The use of this type of suffix seems to reduce greatly the possibility of using *ie* in the stem, since there are no Spanish words with two diphthongs beginning with the same glide in successive syllables, for example, **bierquió*. There seems to be no problem with *ue* occurring before *ió*, as the eight forms in *duentió* attest. It should also be noted that the Chicano subjects used the *-ió* suffix much more than the non-Chicano subjects.[2]

The results show, then, that in this kind of experimental situation subjects are extremely reluctant to introduce an alternation into a nonce stem. They prefer to take the stem as given and apply only suffixation processes to it, even if this means using a diphthong in an unstressed syllable. Before concluding that speakers do not extend lexically bound rules to new forms, even if the rule already applies to a large number of items, we need to find out how the subjects respond if the nonce verb is given as having a diphthong/mid-vowel alternation.[3] We hypothesized that even under these conditions, some subjects would still use a diphthong in unstressed syllables. This hypothesis was correct. We presented four nonce verbs, *megar*, *nelar*, *ponzar*, and *monar*, with two forms, one a 3s present form, for example, *miéga*, and the other an infinitive, for example, *megár* or 3s preterite *megó*, and asked the subjects to supply a 3s preterite, *miegó* or *megó* (if the infinitive was presented), or a 1s preterite *miegué* or *megué* (if the 3s preterite was presented). (See items 6, 12, 19, and 21 in appendix

TABLE 4.1. Nonce Preterite Forms as Responses to a Diphthong in the Present

	bierea			*duenta*		
Non-Chicanos	biercó	16	73%	duentó	19	86%
	bercó	1	4.5%	dontó	1	4.5%
	bircó	1	4.5%	dontió	1	4.5%
	berquió	3	13.6%	duentió	I	4.5%
	birquió	1	4.5%			
Chicanos	biercó	9	64%	duentó	6	43%
	bercó	1	7%	duentió	7	50%
	berquió	2	14%	dontió	1	7%
	birquió	2	7%			

2.) The order of presentation of the two allomorphs was reversed in the two tests. Table 4.2 shows the results on these test items.

Table 4.2 shows that Chicanos and non-Chicanos behaved very much alike overall in using allomorphs with diphthongs or mid-vowels. As we predicted, a significant number of responses contain a diphthong, even though all the clues necessary for identifying the verb as having the diphthong/mid-vowel alternation were supplied by the experimenter. Another way to view the results is by subject: nine out of twenty-two non-Chicano subjects treated these verbs like real Spanish alternating verbs, while the other thirteen subjects used at least one diphthong in the unstressed position. Of the Chicano subjects, only four used only the mid-vowel in the unstressed position, following the pattern of real verbs, while the other ten used a diphthong in at least one of the four items. Overall, only 36% of the subjects used the right "rule" consistently.

These results show that a morphophonemic rule that is lexically restricted does not extend as easily to new lexical items as a phonological process does, even if it applies—as this one does—to approximately fifty common verbs. This indicates quite a different status for the two kinds of "rules." In fact, it casts doubt on the hypothesis that morphophonemic alternations are governed by rules at all.

There are good structural reasons for considering the alternations of *e/ie* and *o/ue* to be the result of a single rule. First, the distribution of the diphthongs and mid-vowels is precisely the same in the two cases, and second, the structural change in the two cases (no matter whether it is *e* → *ie*, *o* → *ue* or *ie* → *e*, *ue* → *o*) is analogous and may be easily abbreviated into a single statement. In Hooper (1976 and 1979b), the question has been raised as to whether such considerations are really relevant in morphophonemic rules. One way to test this is to determine whether or not the front

TABLE 4.2. Percentage of Diphthong and Vowel Allomorphs Used by Subjects when Experimenter Presents Both Allomorphs

Non-Chicanos	Diphthongs	23%	(N = 20)
	Mid-Vowels	77%	(N = 67)
Chicanos	Diphthongs	26%	(N = 14)
	Mid-Vowels	74%	(N = 40)

TABLE 4.3. Percentage of Front and Back Diphthongs and Mid-Vowels Used when Both Allomorphs Are Presented by the Experimenter

		o/ue		*e/ie*	
Non-Chicanos	Diphthongs	19%	(N = 8)	27%	(N = 12)
	Mid-Vowels	81%	(N = 35)	73%	(N = 32)
Chicanos	Diphthongs	8%	(N = 2)	43%	(N = 12)
	Mid-Vowels	92%	(N = 24)	57%	(N = 16)

and back alternations patterned the same way under these experimental conditions. These results are presented in table 4.3.

The non-Chicano subjects supplied slightly more front diphthongs than back ones, and the Chicano subjects supplied considerably more front diphthongs than back ones. In fact, the Chicanos' choice between diphthong and mid-vowel in the front series was almost random, suggesting that they did not pattern their responses on existing verbs at all. That is, they did not apply a rule, nor did they utilize "analogy" in the sense of Ohala (1974) and MacWhinney (1978). However, their responses on the back series almost categorically follow the pattern of alternation in real verbs. This striking difference is in part due to a phonological distaste for *ue* in unstressed syllables that does not affect *ie*. But something more must be involved, because the same subjects gave *duentó* and *duentió* in item 4 (see table 4.1), rather than falling back on *dontó*, and so forth. These results raise the possibility that the Chicano speakers may not have a unified generalization concerning the pairs of alternants *e/ie* and *o/ue*, but rather that they treat these alternations as unrelated.

2. Vowel alternation in third conjugation

All the verbs used in the items described above were first conjugation verbs, because first conjugation is the productive conjugation class, the class to which new verbs are always assigned. Third conjugation verbs have an infinitive in *-ir*. There are no third conjugation verbs with a non-alternating mid-vowel in the stem. Thus any verb with a mid-vowel in the infinitive, such as *pedir* 'to ask for' or *mentir* 'to lie', must have an alternating stem. There are two types of alternations, illustrated in (2) by the present and preterite indicative forms:

(2) Present Indicative

a.	pído	pedímos	b.	miénto	mentímos
	pídes	pedís		miéntes	mentís
	píde	píden		miénte	miénten

Preterite Indicative

a.	pedí	pedímos	b.	mentí	mentímos
	pedíste	pedísteis		mentíste	mentísteis
	pidió	pidiéron		mintió	mintiéron

Pedir has an alternation of *e* with *i*, while *mentir* has an alternation of *e* with *i* and *ie*. Notice that the diphthong in *mentir* occurs just where the diphthong occurs in first and second conjugation verbs, that is, in stressed syllables.

There are approximately thirty verbs conjugated like *mentir*. Harris (1969:113–114) observes that the stems of these verbs all end in *r* or the clusters *nt* or *rt*. There are approximately thirty-five verbs conjugated like *pedir*. Their stems end in single consonants, for example, *repetir*, *elegir*, *concebir*, or in the clusters *rv*, *st*, *nd*, *nch*. There are only two verbs with mid back vowels in third conjugation, and they are conjugated like *mentir*, only with the vowels *o*, *u*, and *ue* alternating. These verbs are *dormir* 'to sleep' and *morir* 'to die'. Most treatments of these alternations postulate a rule for the diphthong/monophthong alternation that also governs this alternation in first and second conjugation verbs, and a rule for the high/mid-vowel alternation. *All* treatments generalize across front and back vowels (Harris 1969 is representative).

Our nonce-probe task tested for the productivity of these patterns. We wanted to know if speakers had internalized the generalization that all third conjugation verbs with a mid-vowel in the infinitive must undergo an alternation. We presented the subjects with the infinitives *pertir*, *rebir*, and *norir* (items 7, 18, and 25) and had the subject complete a sentence using the verb in the 3s or 3p present indicative. We also presented the verbs *rentir* and *sornir* and had the subjects complete a sentence using the 3s preterite form. In the results we will look for general productivity, a possible difference between front and back vowel verbs (since the front vowel pattern is more likely to be productive, given that it has more members), and a difference between *pertir* and *rebir*. Since *pertir* has the phonological shape of the *mentir* class, we expect a 3s present *pierte*, and since *rebir* fits the *pedir* class, we expect responses of *riben*. (In this case we asked for a 3p present.)

In table 4.4 the forms presented first are those that would be produced if the pattern were productive. Since it is not clear how to put an absolute value on productivity, we will discuss the results in terms of degrees of productivity. The following is a rank list with the most productive pattern listed first (using only the twenty-two responses of non-Chicanos):

(3)	pertir	: pierte	18
	rentir	: rintió	12
	norir	: nuere	9
	rebir	: riben	6
	sornir	: surnió	1

First observe that, as expected, the front vowel alternations are more productive than the back vowel alternations. In the case of the diphthong in present indicative, there are more responses *pierte* than *nuere*. In the case of the high vowel in the 3s preterite, there are more responses in *rintió* than *surnió*. This is what we expected, since there are only two verbs with the alternation among back vowels. It shows, however, that even though the front and back vowel alternations can be described in a single generalization, the productivity of the front vowel alternation is not generalizable to the back vowels. This means that for morphologically conditioned rules, the simplest and most general rule may not be the psychologically valid rule. It seems that

TABLE 4.4. Responses on Third Conjugation Verbs

Non-Chicanos		*Chicanos*	
pertir		***pertir***	
pierte	18	pierte	6
perte	1	perte	2
pértie	1	pierta	2
pertirse	1	puerte	1
pierda	1	puertió	1
rebir		***rebir***	
riben	6	riben	3
rieben	5	rieben	0
reben	4	reben	3
rebiben	3	rebiben	1
rebien	3		
rebiran	1	rebiran	4
		reban	1
		rebiden	1
		ruben	1
norir		***norir***	
nuere	9	nuere	5
nore	8	nore	4
nuera	3	nuera	2
nuega	1	noriega	1
noriza	1	norle	1
		noriba	1
rentir		***rentir***	
rintió	12	rintió	2
rentió	6	rentió	9
rentó	3	rentó	3
rentuó	1		
sornir		***sornir***	
surnió	1	surnió	3
sornió	20	sornió	10
sornó	1	senió	1

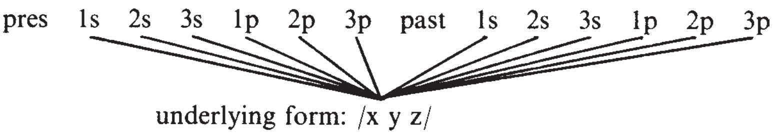
the verbs *dormir* and *morir* are learned by rote and treated as exceptional forms, while the verbs with the same alternation among front vowels are governed to some extent by rules or generalizations. (We would like to leave the question of the nature of these rules open to further investigation.)

The alternation of the mid-vowel with a diphthong in the present tense is more productive than the alternation with a high vowel in the 3s preterite. Thus there are more "productive" responses for *pertir*:*pierte* than for *rentir*:*rintió* and more for *norir*:*nuere* than for *sornir*:*surnió*. The reason for this is that the diphthong/mid-vowel alternation is found in verbs of all conjugation classes and is thus a pervasive phenomenon of Spanish verb morphophonemics. The results suggest that this part of the third conjugation alternation pattern is associated with the general diphthong/mid-vowel alternation and is not considered a separate phenomenon.

Finally, consider the difference between the responses for *pertir*:*pierte* and *rebir*:*riben*. The identification of *pertir* with the *mentir* type of alternation is very strong, much stronger than the identification of *rebir* with the *pedir* type of alternation. This is partly because of the pervasiveness of diphthongization. Indeed, there were five responses with a diphthong (*rieben*) for *rebir*. But the other factor is the more constrained phonological shape of the stems of the *mentir* class. The stems of this class end in *-r*, *-rt*, *-nt* (and one in *-rv*), while the stems in the *pedir* class end in almost everything else: a variety of single consonants (*t*, *d*, *b*, *m*, *g*, *x*) and some clusters (*-st*, *-md*, *-rv*). *Pertir* fits unambiguously into the *mentir* class, just as *rebir* fits unambiguously into the *pedir* class, but the assignment of *pertir* to this class is stronger (note that there were no responses *pirte*), because the *mentir* class is a coherently defined class while the other is an elsewhere, miscellaneous class. (On the productivity of well-defined classes and their internal structure, see Bybee and Slobin 1982.)

3. Testing for the hierarchical organization of paradigms

In generative phonology and similar theories that set up one underlying form for a paradigm, the actual surface forms of the paradigm do not have any official relation to one another; rather, they are each separately related to the abstracted underlying form, as diagrammed in (4). There is evidence, however, from child language and historical change, that there are indeed hierarchical relations among surface forms in a paradigm that suggest quite a different type of grammar from that diagrammed in (4):

(4)


In this section we will discuss this evidence and describe the way we tested for relatedness among surface forms in our nonce-probe task.

The hypotheses underlying this part of the research are that (1) there are indeed relations among surface forms of paradigms; (2) some forms are more closely related to one another than other forms are; (3) the degree of relatedness of two forms depends upon the semantic categories expressed by the forms. A very specific prediction is the following: that the person/number forms within one tense or aspect are more closely related to one another than they are to forms outside that tense or aspect. The traditional representation of verbal paradigms, as in (5), gives a fairly adequate diagram of these relations, as opposed, for instance, to the arrangement in (6), which represents the same person/number forms in different tenses as more closely related to one another than they are to different person/number forms in the same tense or aspect:

(5)

Present		Past	
s 1	p 1	s 1	p 1
2	2	2	2
3	3	3	3

(6)

1st s	2nd s	3rd s	1st p	etc.
Present	Present	Present	Present	
Past	Past	Past	Past	
Future	Future	Future	Future	

In earlier articles (Hooper 1979a; Bybee and Brewer 1980) it has been suggested that paradigms have this hierarchical organization because tense and aspect categories have a greater effect on the meaning expressed by the verb than subject agreement categories do. The reason for this is that tense and aspect deal directly with the event described by the verb and either place it in time relative to the moment of speech or modify its internal temporal constituency, while person/number markers refer to the participants in the event, not to the event itself. Thus tense and aspect are more relevant to the verb than person/number marking, as evidenced by the finding of Greenberg (1963)—that languages that have person/number marking on verbs also have tense, aspect, or mode marking on verbs. The greater effect that tense and aspect have on the meaning of the verb, then, makes person/number forms within one tense or aspect more closely related semantically to one another than they are to person/number forms in another tense or aspect. The hypothesis is that this closer semantic relatedness is reflected in a closer relation in overt form or phonological shape.

This hypothesis can be tested in several ways. One way is to survey a large number of languages in order to determine whether verb forms tend to be more alike within tense/aspect lines than across them. In Hooper (1979a) it was hypothesized that alternations in verb stems would correspond more often to tense/aspect oppositions than to person/number oppositions across tense/aspect lines. That is, it was predicted that there might be a special stem form for a present, past, or future, but not a special stem form for 1st singular that occurs in the present, past, and future. This hypothesis was tested on forty-four languages from a stratified probability sample of languages of the world, and the prediction was borne out overwhelmingly. Of the languages that inflect verbs for both person and number, and tense or aspect, almost all had some stem alternations. These corresponded to tense, aspect, mood, and occasionally number, but never to person oppositions across tense, aspect, or mood lines.

The small amount of data available on the acquisition of verb stem alternations in languages that inflect for person and number also gives evidence for this hierarchical organization of paradigms. It appears that children acquire an initial tense/aspect distinction in verbs before they acquire a person/number distinction in verbs. Furthermore, they seem to elaborate person/number forms *within* tense/aspect lines, not across them, as shown in (7). That is, they will substitute a 3s present form for a 1s present, or a 3s preterite for a 1s preterite, but they do not substitute a 1s present for a 1s preterite (Hooper 1979a):

(7) Direction of substitution in child language

Present ——— Past
s 1 ↑ ——— s 1 ↑
3 | - - - - - - → 3 | 3

If these hierarchical principles of organizing paradigms are real, then one should be able to find evidence for them experimentally. Fortunately, the diphthong/mid-vowel alternation gives us the perfect opportunity for testing this hypothesis. As we reported above, when presented with a nonce verb with a diphthong/mid-vowel alternation subjects varied in their responses—about 25% of the time the response was a diphthong, and about 75% of the time, it was a mid-vowel. Since we expected this, we also hypothesized that the percentage of responses with a mid-vowel would vary according to the morphological category of the form with a mid-vowel, and its degree of relatedness to the form required in the response. Thus consider two situations, as illustrated in (8), I and II. In I the subject hears a 3s present indicative form *puénza* and the infinitive *ponzár* and is asked to form the 3s preterite, either *ponzó* or *puenzó*; in II the subject hears the 3s present indicative *puénza*, but this time the form with the mid-vowel is the 3s preterite *ponzó*, and the subject is asked to form the 1s preterite, either *ponzé* or *puenzé*. Given the morphological relations within the paradigm, we expected more responses with a mid-vowel in the preterite form when the allomorph with the mid-vowel was presented in a preterite form as in B than when it was presented in the infinitive form as in A. The reason for this, of course, is that the preterite forms are more closely related to one another than to the infinitive form. Although the infinitive is used as a citation form in grammars, it is somewhat outside the paradigm, being non-finite and having so many nominal uses.

(8)

I. infinitive *ponzár*					II. infinitive			
Pres. Ind.	1s		1p		Pres. Ind.	1s		1p
	2s		2p			2s		2p
	3s	*puénza*	3p			3s	*puénza*	3p
Pret. Ind.	1s		1p		Pret. Ind.	1s	???	1p
	2s		2p			2s		2p
	3s	???	3p			3s	*ponzó*	3p
??? = *ponzó* or *puenzó*					??? = *ponzé* or *puenzé*			

As we mentioned earlier, there were four nonce verbs used: two with front vowels, *megar* and *nelar*, and two with back vowels, *ponzar* and *monar*. Each subject was asked to use each of these four verbs, but there were two tests, so that the verbs were matched with different pictures in the two tests. Also, the verbs appeared in the two

TABLE 4.5. Percentage of Diphthong and Mid-Vowel Responses of Non-Chicano Subjects according to the Form in Which the Mid-Vowel Was Presented

I. inf, 3s pres → 3s pret			*II. 3s pres, 3s pret → 1s pret*		
	e/ie	*o/ue*		*e/ie*	*o/ue*
Diphthong	32%	23%	Diphthong	18%	14%
Mid-Vowel	64%	73%	Mid-Vowel	82%	82%
Overall	Diphthong	27%	Overall	Diphthong	16%
	Mid-Vowel	68%		Mid-Vowel	82%

$X^2 = 3.07$ (significant at the 0.05 level)

different frames, as in (8) I and II, and the order of presentation of the two allomorphs was varied.

The results for the non-Chicano subjects are presented in table 4.5. It can be seen that the form in which the mid-vowel was presented did significantly influence the percentage of mid-vowels in the responses. If the subjects heard the mid-vowel in a preterite form they were more likely to use it in forming another preterite than if they had heard it in an infinitive form.[4]

These results support the hypothesis that there are relations among surface forms of paradigms, because if there were no such relations, then the subjects' task would be simply to ascertain whether the nonce verb had an alternation or not and the infinitive should be just as good an indicator of alternation as the preterite form. In this case there would be no difference in the subjects' performance in conditions I and II.

The finding of a significant difference in this test, along with the evidence from child language and the cross-linguistic survey mentioned earlier, converge to demonstrate a hierarchical organization of surface forms of verbal paradigms that cannot be accounted for in a grammar that treats each surface form as related to only the one underlying form. The details of this hierarchical organization have not been ascertained, but the results obtained so far give hopeful indication that these details can be worked out by using this method of testing, along with historical (as in Bybee and Brewer 1980), cross-linguistic, and child language data.

3.1. Results for Chicano speakers

Interestingly enough, we did not obtain comparable results with the Chicano speakers. Although, as we reported in table 4.2, the Chicanos used about the same percentage of diphthongs and mid-vowels as the non-Chicanos, the Chicano subjects did not show variation that correlated with the two conditions described in (8); rather, the variation in their responses correlated with the front vowel/back vowel opposition, as shown in table 4.6.

Table 4.6 shows that the percentage of diphthongs remains the same and the percentage of mid-vowels varies across the two conditions only because in II there were two responses for *monar* with a high vowel, *munió*, which we counted as irrele-

TABLE 4.6. Percentage of Diphthong and Mid-Vowel Responses by Chicano Subjects according to the Form in Which the Mid-Vowel Was Presented

I. inf, 3s pres → 3s pret			*II. 3s pres, 3s pret → 1s pret*		
	e/ie	*o/ue*		*e/ie*	*o/ue*
Diphthong	(5) 36%	(2) 14%	Diphthong	(7) 50%	(0) 0%
Mid-Vowel	(9) 64%	(12) 86%	Mid-Vowel	(7) 50%	(12) 86%
Overall	Diphthong 25%		Overall	Diphthong 25%	
	Mid-Vowel 75%			Mid-Vowel 68%	

vant. The reason that there are no significant differences between tasks I and II for Chicanos is that for the back vowels they avoided unstressed *ue* almost categorically and for the front vowels the responses seem to be random. Given the randomness of the responses on items with front vowels, our hypothesis predicts that there should be a tendency for more diphthongs under conditions I than II: this is not the case for front vowels, but we have too little data to conclude that Chicanos have a totally different system from non-Chicanos.

4. Velar-insert verbs

4.1. The alternations

Spanish has a class of second and third conjugation verbs that add a velar stop after the stem in the 1s of the present indicative and throughout the present subjunctive:

(9)

Infinitive	*crecer*	'to grow'	*salir*	'to leave'	*caer*	'to fall'
Pres. Ind.	crezco	crecemos	salgo	salimos	caigo	caemos
	creces	crecéis	sales	salís	caes	caéis
	crece	crecen	sale	salen	cae	caen
Pres. Sub.	crezca	crezcamos	salga	salgamos	caiga	caigamos
	crezcas	crezcáis	salgas	salgáis	caigas	caigáis
	crezca	crezcan	salga	salgan	caiga	caigan

(The letter *c* is pronounced as /s/ in Latin America and as /θ/ in Castilian when it occurs before a front vowel and as /k/ before a back vowel. *z* is pronounced as /s/ or /θ/, depending on the dialect: *crece* [krese] or [kreθe]; *crezco* [kresko] or [kreθko].)

A number of common verbs have this "velar insert," and these verbs form a phonologically defined class since their stems end in a vowel or a vowel followed by *c* (/θ/ or /s/), *n*, or *l*.

The largest group of these have stems ending in *c*. The number of such verbs is greatly increased by the fact that a common verbalizing suffix, *-ecer*, carries with it the velar insert. There are probably more than three hundred verbs ending in *-cer*, and all but a very small number (e.g., *mecer* 'to stir') take the velar insert. Here are a few examples:

(10)	satisfacer 'to satisfy'	conocer 'to know'
	yacer 'to lie, recline'	aborrecer 'to grow dark'
	obedecer 'to obey'	enrojecer 'to redden'
	permanecer 'to remain'	envejecer 'to age'
	aparecer 'to appear'	amanecer 'to dawn'

In addition, there are a dozen or so words built on *-ducir*, for example, *introducir*, *producir*, that also take the velar insert and a half dozen constructed with *lucir* 'to illuminate'.

There are only three verbs whose stems end in *n* that take the velar insert; two are second conjugation (*tener* 'to have' and *poner* 'to put or place') and one is third conjugation (*venir* 'to come'). These are, of course, three of the most frequent verbs of the language, and each one appears in a number of derived verbs, for example, *obtener* 'to obtain', *proponer* 'to propose', *provenir* 'to originate'. Thus the entire class of verb stems ending in *n* that take the velar insert has more than fifty members, although they are all based on these three verbs. There are apparently no second conjugation verbs in *-ner* that do not take the velar insert and only a few third conjugation verbs, all derivatives of *finir*, for example, *definir* and *unir*, that do not.

Verb stems ending in *l* that take the velar insert are *salir* 'to learn' and *valer* 'to be worth' and one or two derivatives of each. There are second conjugation verbs ending in *-ler* that do not belong to the velar-insert class. They are *oler* 'to smell', *doler* 'to pain', *moler* 'to grind', *soler* 'to be used to'. Among third conjugation verbs there is regular *pulir* 'to polish'.

The velar-insert verb stems that end in vowels actually take a palatal glide plus a velar (see [9]). They are:

(11)	caer 'to fall'	traer 'to bring'
	roer 'to gnaw'	corroer 'to corrode'
	raer 'to scrape'	oír 'to hear'

In addition, there are a few derivatives of these verbs. There are no verbs in second or third conjugation that end in *a* or *o* that do not belong to the velar-insert class.

To summarize the situation, most verbs of second and third conjugation that end in *c*, *n*, *l*, or a vowel *a* or *o* belong to the velar-insert class. These verbs are rather few in number except for those ending in *-cer*.

There are two sets of questions we can ask about this class of verbs using the type of nonce-probe task we have described above. The first set of questions concerns the way speakers treat a verb class such as this one. Do they make a generalization concerning the phonological shape of verbs in this class? What is the nature of this generalization and to what extent is it extendable to nonce verbs? The results described in section 4.2 bear on these questions. The second set of questions concerns the organization of the paradigms of these verbs. In particular, do speakers grasp the connection between the 1s present indicative and the present subjunctive forms? The results bearing on this question are described in section 4.3.

4.2. Testing the boundaries of the verb classes

To test whether or not Spanish speakers internalize a phonological generalization concerning the verbs that belong to the velar insert class, nonce verbs corresponding to each of the phonological shapes were constructed. They were *lece*, *lezco*; *rone*, *rongo*; *pale*, *palgo*; and *fae*, *faigo*. These verbs were presented under two conditions. In condition I, the 3s and 1s forms of the present indicative, for example, *lece* and *lezco*, were presented and the 3s of the present subjunctive was elicited. If the subjects assigned *lecer* to the same class as *crecer*, then the response would be *lezca*. Since it seems possible that 1s indicative might be chosen as the base for *any* other form, not just the present subjunctive, another condition was set up. In condition II, the subjects were presented with the infinitive and the 1s present indicative, *lecer* and *lezco*, and were asked to form a 3s present indicative, which would be *lece*, if this verb is assigned to the same class as *crecer*. Two of the four verbs appeared in condition I on test A and in condition II on test B, and the converse held for the other two verbs.

To make a more definitive test of the boundaries of this class, two nonce-verb types that might be considered marginal to the class were included on the test. One was modeled after *decir* 'to say' and the other after *hacer* 'to make, to do'. These verbs have an alternation of /s/ or /θ/ with /g/, and the velar appears in the same verb forms as the inserted velar in the *crecer* class. We actually used *two* verbs with this alternation pattern, one on test A and one on test B. One verb was exactly like *decir* except for the initial consonant: *nice*, *nigo*. The other differed from the two existing verbs of this type by both the initial consonant and the vowel: *proce*, *progo*. We wanted to know if these verbs would be treated like those of the *crecer* class. We suspected that with only two verbs to serve as models, these nonce verbs would be treated in unpredictable ways, so we attempted to tighten the association with the real verbs by making one of the experimental items very close phonologically to a real verb.

The second type of marginal item is not the counterpart of any real verb. Rather, this verb was constructed to see if speakers would consider it a possible verb. As we mentioned earlier, the verbs that add a velar consonant in 1s present indicative and present subjunctive all have stems that end in *c*, *n*, *l*, or a vowel. This is not a natural phonological class based on the inherent phonetic properties of the sounds, but it is a distributional class, since these are the only consonants of Spanish that can appear at the end of a syllable before a consonant, with one exception. The exception is *r*, which does occur at the end of a syllable. But there are no verb stems ending in *r* that take the velar insert, although there are some older dialectal reports of such verbs (Baxter 1974:182ff). If speakers of Spanish form a phonological generalization about the members of the velar-insert class (as speakers of English do about the *sting/stung* class [Bybee and Slobin 1982]), then we might expect them to be able to extend this class to verbs whose stems end in *r*. To test this possibility, the verb *nare*, *nargo* was created.

Table 4.7 shows the responses recorded when subjects were asked to use nonce verbs that resemble the velar-insert class, in the ways we have discussed earlier. The percentages given in the right column of table 4.7 are the percentages of responses

TABLE 4.7. Responses by Chicano and Non-Chicano Adults on Verbs Resembling the Velar-Insert Class

					Percentage of Expected Responses
Input forms	lezco, lecer		lezco, lece		
Expected response	lece	9	lezca	14	64
Other responses	lace	2	lece	3	
	lezcua	1	lesa	1	
	lezca	4	lezque	1	
	lecendo	1			
Input forms	faigo, faer		faigo, fae		
Expected response	fae	8	faiga	12	56
Other responses	faerle	2	fague	1	
	fale	1	faa	1	
	faele	1	fae	2	
	faé	1	faya	1	
	faise	1			
	faiya	1			
	faye	1			
	faigo	1			
	faya	1			
	fuega	1			
Input forms	palir, palgo		palo, palgo		
Expected response	pale	7	palga	3	28
Other responses	palea	2	palgue	7	
	palee	1	pale	8	
	pala	3	pala	1	
	palga	3			
	palió	1			
Input forms	roner, rongo		rone, rongo		
Expected response	rones	1	ronga	6	19
Other responses:	rongas	7	ruenga	1	
	ruengas	6	rone	5	
	ruenas	4	ruene	2	
	roneas	1	rongue	3	
Input forms	nargo, nare				
Expected response	narga	5			14
Other responses	nare	11			
	nargue	18			
	niera	1			
	nargara	1			
Input forms	progo, proce		nice, nigo		
Expected response	proga	5	niga	5	28
Other responses	progue	6	nigue	3	
	pruegue	3	niegue	1	
	proce	4	nice	3	
	prozca	1	nizca	2	
	prosa	1	nisa	1	
			necile	1	

that were expected if the subject assigned the verb to the velar-insert class. The nonce verb most often assigned to this class was *lecer*. This is just what we would expect given the fact that there are comparatively such a large number of verbs ending in *-cer* that take the velar insert. Note that it is the type frequency (the number of verbs in the class) and not the token frequency (the frequency of the individual verbs) that determines productivity. The other classes, that is, of verbs ending in *n* or *1*, contain some of the most frequent verbs in the language.

The nonce verb *faer* was also assigned to the velar-insert class quite often. The reason for this is also clear. Although there are only a small number of verbs of this phonological shape, there is simply no other way to treat them. The difficulty is illustrated well by some of the responses given for this verb. First of all, there was a large number of different miscellaneous responses, showing that if the subjects did not assign this verb to the velar-insert class, they did not quite know what to do with it. Second, the problem with a verb ending in /a/ is that if it is assigned to the second conjugation, its subjunctive suffix must be /a/ also, giving a "regular" form *faa*, which was an actual response by one subject. A form such as this, however, is phonotactically not Spanish. Sequences of two identical vowels are rare (occurring, for example, in the verbs *creer* and *leer*), and cases of two /a/s in a row simply do not occur. Thus *faer* is assigned to the velar-insert class because phonologically there are few other real alternatives.

The other verbs in table 4.7 were assigned to the velar-insert class with considerably less frequency. In fact, there is a strong tendency *not* to assign them to the velar-insert class. We must conclude from this that the set of verbs ending in *c*, *n*, *l*, or V does not form a unified class for speakers of Spanish. Verbs ending in *-cer* do and apparently verbs ending in *a* do also, but verbs ending in *n* or *l* do not. It seems likely that this small number of verbs, *tener*, *venir*, *poner*, *valer*, *salir*, are learned by rote and stored with their irregular forms listed in the lexicon rather than derived by rule. Given that there is no general class, it is not surprising that a verb with stem-final *r* was not included in the class.

Another way to look at the same facts is to consider the number of times a nonce verb was assigned to the productive first conjugation class. Since velar-insert verbs are all second or third conjugation verbs, a sure sign that the subject has not assigned the verb to this class is the use of first conjugation suffixes, even if the subject uses the allomorph with the velar consonant in it. In table 4.8 we see the percentages of first conjugation responses. Again *lecer* and *faer* distinguish themselves from the rest of the nonce verbs by being assigned to the first conjugation with significantly less frequency than the other verbs.

4.3. Testing the conditioning factors

The other set of questions we will consider about velar-insert verbs concerns the generalizations that speakers make to account for the distribution of the velar element. One question we might ask is whether the velar belongs to the stem, to the suffix, or to neither. Another question is, what conditions the alternation? There is a possible phonological conditioning—the velar occurs before back vowels (the *o* of 1s present indicative and the *a* of the present subjunctive) and is absent before front

TABLE 4.8. Percentage of Time a Nonce Verb Was Assigned to the First Conjugation

	Percentage
faer	17
lecer	25
nicer	50
palir	64
procer	65
roner	78
narer	81

vowels. Do speakers use this phonological generalization or a morphological one, or do they learn the forms by rote? If they use a morphologically based generalization, is it extendable to other verbs?

Some analysts (Foley 1965) have analyzed the velar as belonging to the stem and occurring in underlying form, only to be deleted before a front vowel. Under this analysis it is difficult to explain the historical spread of the velar to different stems, so an insertion analysis seems preferable (Baxter 1974; Hooper 1979b). Harris (1972) chooses an insertion analysis for synchronic reasons. In this analysis the velar is inserted between the stem and the suffix. Another possible segmentation would consider the velar as part of the suffix, part of the marker of 1s and present subjunctive. It is difficult to distinguish this analysis from the insertion analysis if the insertion is morphologically conditioned (as we believe it is; see below), for in either case the velar will be part of the marker for 1s and present subjunctive.

Some of the responses given by our subjects suggest an association of the velar element with present subjunctive. Two verbs we used in this part of the experiment did not contain velars in either allomorph. They were *lastro*, *lase* and *seto*, *sade*, purposely constructed with non-Spanish alternations (again, see below). When subjects had to make present subjunctive forms for these verbs, some responded by putting a velar in the form: for the first verb there were two responses *lazca* and one *laga*; for the second verb there was one *sezca* and two *saga*. These were minority responses (out of a total of thirty-six responses for each verb), but there were no other responses in which some consonant that did not occur in one of the stimulus forms was inserted. This suggests an association of the velar element with the subjunctive that transcends individual verb stems and thus suggests a rule of insertion.

Concerning the phonological or morphological conditioning of the appearance of the velar, one would expect that if there were a rule associating the velar with a back vowel, in all the responses for verbs with a velar in one allomorph, there would be a tendency to use the velar before a back vowel. The relation between the backness of the vowel and the morphological category is independent in these data, since many subjects put the verbs in the first conjugation where the back vowel occurs in the indicative and the front vowel in the subjunctive. However, no tendency to associate the velar with back vowels was found: out of all the cases where the velar was used, it occurred before a front vowel in 52% of the cases and before a back vowel in 48% of the cases.

As we have mentioned earlier, the morphological categories in which the velar appears are the 1s of the present indicative and all of the present subjunctive. That it should occur just in these categories is historically the result of the occurrence of back vowels in these categories and so is morphologically arbitrary. However, this distribution has been historically stable in Spanish. There are no changes in which the velar is lost from the 1s present indicative but retained in the subjunctive. The question is, then, does the velar occur in some sort of morphological "natural class"?

We will not claim that the 1s present indicative and the present subjunctive have any semantic properties in common (besides "present"), but we will suggest on the basis of the theory presented in Bybee and Brewer (1980) some reasons that the occurrence of the velar in just these categories should be stable. First, it is normal for all person/number forms of a given tense and mood category to have phonological properties in common (Hooper 1979a). Thus the velar throughout the present subjunctive should be an optimal feature if there is going to be a velar anywhere in the subjunctive. Second, it is also common for the marked members of a category to be formally derivable from the unmarked members (Vennemann 1972). So the present subjunctive should be based on the present indicative. But the whole of the present subjunctive cannot be based on the whole of the present indicative. Rather, it is based on a single form of the present indicative, the most autonomous form (Bybee and Brewer 1980). Evidence from historical change indicates that the most autonomous forms among person/number forms are the 3s and the 1s. It is also possible for the 1s form, because of its relative autonomy, to be distinct from other person/number forms in the same tense and mood. Thus, while there are other distributions of the velar that might also be motivated or even better motivated, the historically inherited distribution in Spanish is not totally at odds with the morphological system.

Pedagogical grammars of Spanish teach that the present subjunctive forms can be made from the 1s of the present indicative. This holds not only for verbs with a velar insert but also occasionally for other verbs, such as *caber* 'to fit' with 1s present indicative *quepo* and subjunctive *quepa*, *quepas*, and so on. The question we addressed was whether or not speakers abstract a generalization such as this one that goes beyond the velar-insert class of verbs. To test this hypothesis we created two nonce verbs with alternations that do not occur in Spanish verbs and presented one allomorph in the 1s present indicative and the other in the 3s. The subjects then supplied a present subjunctive form. If the subjects have a generalization concerning the relation between 1s present indicative and present subjunctive, then we would expect more of them to choose the allomorph presented in the 1s form. This was in fact the case, as we will see later.

The alternations in the two verbs were at different degrees of variance from real Spanish alternations. One was patterned on *caber*, *quepo*, but the alternating consonants were *d* and *t* rather than *b* and *p*. This verb was *sade*, *seto*. The other verb was

TABLE 4.9. Responses by Chicanos and Non-Chicanos to Two Nonce Verbs with Non-Spanish Alternations

			Percentage of Expected Responses
Input forms	lase, lastro		
Expected response	lastra	4	11
Other responses	lastre	19	
	lase	6	
	lasa	1	
	lazca	2	
	laga	1	
	laste	1	
	lastrara	1	
	lastrasteis	1	
Input forms	sade, seto		
Expected response	seta	11	31
Other responses	sete	7	
	sade	6	
	sate	4	
	seda	2	
	sede	1	
	saga	2	
	sezca	1	
	sieta	1	
	siete	1	

created so as to have phonologically possible forms, but an alternation unlike anything occurring in Spanish. This verb was *lase*, *lastro*. The only resemblance this bears to a Spanish alternation is that the 1s form is a fuller form; it has more consonants in it than the 3s form.

Table 4.9 presents the responses for these two verbs. The "expected responses" are second or third conjugation forms based on the allomorph that was presented in the 1s present indicative. The percentage of expected responses is quite small. However, the number of forms using the 1s present indicative stem is much larger, if we consider in this the forms put in the first conjugation (ending in *-e*) along with the ones in the second and third conjugations. Then we find that of the responses that used one stem or the other (rather than something different entirely or a combination, as in *sate*), there were twenty-five responses with *lastr-* and seven with *las-*, a ratio of 78% to 22%, and eighteen with *set-* and six with *sad-*, a ratio of 75% to 25%. It seems, then, that there is a tendency to use the stem of the 1s over the 3s present indicative when forming the present subjunctive.

4.4. Summary

The use of nonce verbs with varying degrees of similarity to real Spanish verbs gave us a chance to see if speakers have a notion of possible morphophonemic alternation in Spanish. In particular we found subjects had an adverse reaction to *lase*, *lastro* and *sado*, *seto*, saying these were not like Spanish or, after several repetitions, throwing their hands up. There is no indication that subjects made an association between *caber*, *quepo* and *sade*, *seto* despite their similarity on the feature level.

5. Conclusions

The nonce-probe tasks we have described here asked a number of questions about morphophonemic alternations. The work was exploratory in the sense that we wanted to learn what kinds of questions about morphophonemics could be answered using this method. Of course, many other tests of this type need to be done on Spanish morphophonemics and on other languages, but we feel that the methodology has proven to be successful, for it has led us to some tentative conclusions, which we will discuss briefly in closing.

The results of these tests confirmed in several different ways the profound dissimilarity between processes with phonetic conditioning and those with morphological conditioning. For instance, the results show that speakers do not always capture all the phonological generalizations available (even on the surface) in morphophonemic alternations. The Chicano subjects treated the *back* diphthong/mid-vowel alternation quite differently from the *front* alternation of the same type. Further, all the subjects treated the front vowel alternations of *i/e* and *i/e/ie* as more productive than the *u/o/ue* alternation, even though these could easily be handled by a single rule that generalizes over front and back vowels. This suggests that speakers treat *morir* and *dormir* as having exceptional alternations, while verbs with the front vowel alternations are treated as a class to which rules apply. From the point of view of descriptive grammar, this analysis is inelegant and uneconomical, missing a generalization as it does. However, we cannot assume that a generalization that might be appropriate for a phonetically conditioned process will also be appropriate for a morphologically conditioned rule. Rather, to gain understanding about morphophonemic alternations and to make predictions about how they will change, they must be investigated as a phenomenon distinct from purely phonological phenomena.

Hooper (1976) claims that speakers prefer to treat an alternation as having morphological conditioning rather than phonological, because this allows for the formulation of the most direct association between sound and meaning. However, from observing the surface facts of Spanish it is very clear that the diphthong/mid-vowel alternation can be predicted much better by stress than by any other factor, particularly any morphological factor. Yet, in nonce-probe tasks, Spanish speakers seem perfectly willing to produce diphthongs in unstressed syllables in forms such as *duentó* and *biercó*. Even when there are two allomorphs to choose from, as in *monár*, *muéna*, there were still many responses that go counter to the rule (e.g., *muenó*). This may not necessarily mean that the diphthong/mid-vowel alternation is conditioned by morphology rather than by stress placement, but it certainly does bring into question the status of the "rule" that governs this alternation, for it appears that this "rule" is not independent of the lexical items to which it applies. We return next to a discussion of lexically bound rules.

In the velar-insert class there is a clear choice of phonological versus morphological conditioning: the appearance of the velar could be linked with a back vowel or with the morphological categories 1s present indicative and present subjunctive. Our results indicate that to the extent that speakers have a generalization about this alternation, it is not a phonologically conditioned one but rather a morphologically conditioned one.

Another property of morphophonemic rules that emerges clearly from our study is the extent to which such rules are lexically bound. We found that overall speakers did not apply these rules to nonce items to the extent predicted by their statement in a generative grammar, even when two allomorphs of the nonce verb were given, so that all clues necessary for correctly applying the rule were available. This means that most morphophonemic alternations are closely tied to the particular lexical items that manifest them, and do not have an independent existence in the grammar. This is the claim made implicitly by the descriptive model proposed in Vennemann (1974) and in Hudson (1974 and 1975), in which all alternations are listed lexically. We hesitate, however, to accept the proposition that *all* morphophonemic rules are lexically bound. MacWhinney (1985) points out that there are some free patterns in Hungarian, rules that apply in a morphological context, but without lexical restriction. Spanish verb stress may be a comparable case. In all of our responses, we found very few errors in stress placement in verbs. Stress in verbs is morphologically conditioned (Hooper 1976) but has only one type of lexical restriction, the "strong preterites," which are irregular in other ways as well.

There is a continuum, then, from alternations that are considered exceptional and listed in the lexicon without any associated rules, to alternations governed by rules that are bound to particular lexical items, to alternations governed by free rules. The question is, what are the factors that determine to which of these three types an alternation belongs? Hudson claims that such distinctions are impossible to make empirically, and therefore treats all alternations as suppletion. The data presented earlier, however, show that there are some differences in alternations—some are much more readily extendable to nonce items than others. There are at least two factors playing a role in determining productivity.

One is, of course, the number of lexical items to which a rule applies. It seems unlikely that there is some absolute number of lexical items below which a rule is unproductive, but it might be possible to specify a range. For instance, on the one hand, in our tests of the velar-insert class we learned that three lexical items (those ending in *n*) were not enough, even if these three entered into nearly fifty derivative verbs. On the other hand, as we saw with verbs ending in *c*, three hundred is quite clearly enough. Delimiting the range better is the *mentir* class, with approximately thirty members, which showed productivity on our test. The cutoff point lies somewhere between three and thirty.

But the number of lexical items to which a rule applies cannot alone predict productivity. Another important factor is how well the phonological material lends itself to generalization. We found that speakers did not always generalize across parallel series of front and back vowels. We found they did not generalize across syllable-final, preconsonantal segments, /s/ or /θ/, /n/, /l/, and V. However, there is evidence of a generalization for the vowel-final stems of the velar-insert class, although they are few in number. (There are six.) Further, the subjects correctly assigned *pertir* to the *mentir* class, which has a restricted phonological criterion that the stem end in *r*, *rt*, or *nt*, but correctly assigned *rebir* to the larger, more miscellaneous *pedir* class much less often. A certain narrowness of phonological defining properties seems necessary for making correct class assignments and, by inference, for the formation of coherent, productive classes of lexical items, as opposed to isolated suppletive lexical items.

Finally, another different sort of issue we approached with this test was the issue of relatedness among surface forms of a paradigm. We were able to show that the theory that surface forms of a paradigm are not related to one another but each independently related to an abstracted underlying form cannot predict certain significant differences in speakers' performance on the task. In particular we showed that a preterite form is a more useful guide to the formation of another preterite form than the infinitive is. This suggests a closer relation of forms of the preterite among one another than to the infinitive. Further testing of this nature might lead to a better understanding of the organization of verbal paradigms.

Appendix 1

A modified version of the same test was given to Spanish-speaking children (see below). The test was modified in that more time was taken to teach the child the nonce verb and the child was asked to respond to a question rather than fill in a blank. Even with these modifications the test was extremely difficult for the young children we had access to in the schools of San Jose, California. The youngest children tended to repeat the nonce verb uninflected or to switch to a real verb. As a result we tried older and older children until we found some that could do the task. These children were between the ages of 8 and 11 years. There were fifteen subjects: eight boys and seven girls, taken from bilingual classrooms, all of whom had Spanish as their first language and used it at home.

One part of the test given to the children corresponded to the adult test items discussed in section 2, the alternations in third conjugation verbs. Recall that the verbs *pertir*, *rebir*, *norir*, *rentir*, and *sornir* were all presented in the infinitive and the subjects were asked to make present indicative forms for the first three and preterite indicative forms for the last two. Here are some of the problems we encountered. Unfortunately we presented *pertir* as a reflexive verb, and in the infinitive the reflexive pronoun *se* follows the verb: *pertirse*. Fourteen of the fifteen subjects gave us the unchanged form *pertirse* as an inflected form, even though the experimenter provided the *se* in the correct preverbal position for an inflected word. Only one subject produced the inflected form, and it was the expected one, *pierte*. With *rebir*, the children also tended to avoid stressing the first syllable and produced forms such as *rebín*, *rebían*, *rebiran*, *rebiera*, *rebinsen*. This indicates that they were not quite sure how to go from an infinitive to a present indicative form in a nonce verb. (Note, however, that for the real verb *servir* they all replied with the correct *sirve*.) There were only three relevant responses, two with a diphthong, for example, *rieben*, and one with a high vowel *riben*. They were considerably more successful at conjugating *norir*, although some still avoided stem stress. There were eight responses with *o*, for example, *nore*, and two with the diphthong, *nuere*. We obtained many more

TABLE 4.10. Children's Responses to Nonce Verbs with Diphthong/ Mid-Vowel Alternations under Two Conditions

I. inf, 3s pres → 3s pret			*II. 3s pres 3s pret → 1s pret*		
	e/ie	o/ue		e/ie	o/ue
Diphthong	14%	38%	Diphthong	36%	20%
Mid-Vowel	86%	62%	Mid-Vowel	64%	80%
Overall	Diphthong	26%	Overall	Diphthong	27%
	Mid-Vowel	74%		Mid-Vowel	73%

TABLE 4.11. Children's Responses to Nonce Verbs with Diphthong/Mid-Vowel Alternation in Two Different Orders of Presentation

Diphthong, Mid-Vowel			*Mid-Vowel, Diphthong*	
	e/ie	o/ue	e/ie	o/ue
Diphthong	15%	7%	33%	54%
Mid-Vowel	85%	93%	67%	46%
Overall	Diphthong	11%	Diphthong	44%
	Mid-Vowel	89%	Mid-Vowel	56%

relevant responses when we asked them to go from the infinitive to the 3s preterite form. This is probably because the stress remains the same. But the expected vowel changes were not made for the most part: there were eight responses *rentió* and two *rintió*; ten responses *sornió* and no *surnió*.

We also gave the children nonce verbs with the diphthong/mid-vowel alternation under the two conditions discussed in section 3. Tables 4.10 and 4.11 show that the children's responses did not vary according to the two conditions of presentation of allomorphs as the non-Chicano adults' responses did, but they were rather affected strongly by the order of presentation. The children tended to use the form they heard last as the basis for the form they were asked to create. This seems to be the major difference between the children's and adults' responses. Note that overall percentages of diphthongs and mid-vowels in the responses were much the same for the children as for the adults (approximately 26% diphthongs, 74% mid-vowels).

Appendix 2

Osito y sus amigos

1. La mamá de Osito siempre *fana* por la ventana.
   Osito quiere *fanar* por la ventana también.
   Le dice a su mama, 'Mamá quiero ________________ por la ventana.'
2. En las noches la mamá de Osito le *pila* cuentos.
   Al Osito le gusta que su mama le *pile* cuentos.
   Anoche su mamá le __________ un cuento.

3. La mamá de Osito lo *bierca* mucho.
   Cada vez que se enferma lo *bierca*.
   Ayer Osito se enfermó y su mamá lo __________.
4. La mamá de Osito *duenta* cada día.
   Ella *duenta* mucha ropa.
   Ayer Osito rompió una camisa y su mamá se la ____________.
5. Cuando Osito prepara una sopa El Pato siempre la *pale*.
   El Pato le dice a Osito, 'Cada vez que *palgo* tu sopa me siento muy contento.' Ahoro El Osito y El Pato quieren que la gallina también la ___________ .
6. Osito *muena* sopa todos los días. Le gusta mucho *monar* sopa.
   Ayer en la tarde se __________ un plato grande.
7. Al Osito le gustaría *pertirse* en las nubes.
   En sus sueños siempre se ___________ en las nubes.
8. Al Osito le encanta *roner* en el suelo.
   Le dice a su mamá, 'Mira Mamá, mira como *rongo* en el suelo.'
   Pero su mamá le contesta, 'Ven acá, Osito, te ensucias mucho cuando _________ en el suelo.'
9. La mamá de Osito siempre *dierra* el patio por las manañas.
   Pero ayer lo *derró* en la tarde también.
   Mientras Osito jugaba, ella lo __________.
10. Mientras Pepito juega con los palillos, Osito *rone* en el túnel.
    Ayer Osito *ruso* en el túnel hasta muy tarde.
    Su mamá no quería que se cansara; por eso le dijo que no_________ tanto.
11. Osito le cuenta a un amigo: 'Yo *lezco* a mi mamá, y mi mamá me *lece* a mí. Me encanta que mi mamá me __________.'
12. Cada día Osito *puenza* del mismo árbol. Ayer *ponzó* del árbol.
    Le dice a su maestra: 'Ayer yo __________ del árbol.'
13. En les tardes Osito *lase* afuera de su casa. Le pregunta a su mamá, 'Mamá, ¿*lastro* adentro o afuera de la casa?'
    Su mamá le contesta que __________ afuera de la casa.
14. Todos los sábados, al levantarse, Osito tiene que *rentir* el piso de la cocina. Pero ayer sus papás lo llevaron a 'Great America' y no lo _________.
15. La mamá de Osito le dice, 'Ahorita te *faigo* un cuento.'
    Le gusta mucho *faerle* cuentos.
    Cada noche le _________ un cuento.
16. Ayer Osito *ruvo* la mesa para la cena. Cuando *rade* la mesa lo hace con mucho entusiasmo. Antes de empezar, su mamá le mandó que la __________ con mucho cuidado.
17. En las noches, la gallina se *sade* en la cabecera de la mesa.
    La gallina le explica a Osito, 'Yo siempre me *seto* en la cabecera de la mesa.' Pero el Osito se queja de que la gallina siempre se_________ en la cabecera de la mesa.
18. Al Osito le gusta *rebir* bolas de lodo. Todos sus amigos también lo hacen.¿Qué hacen los amigos de osito? Ellos __________.

19. Anoche Osito *megó* a su mamá. Antes de acostarse siempre *miega* a su mamá. En la manaña Osito le dice, 'Mamá anoche yo te ________.'
20. Al Osito le enseñaron a *sornir* el carro. Ayer, por primera vez, lo __________.
21. Osito sabe *nelar* en las nubes. Por eso, cada día *niela* en las nubes. Ayer, en la tarde, __________ en las nubes.
22. En las mañanas la gallina *nare* la miel. Osito le dice, 'Yo también *nargo* la miel en las manañas.' Osito quiere que los dos __________ la miel juntos.
23. El sábado pasado Osito *noló* en el bosque. Todos los sábados *nuela* en el bosque. Ayer, mientras su mamá cocinaba, Osito ________ en el bosque.
24. La mamá de Osito le dice. 'Aquí te *progo* un sombrero para tu cumpleaños.' La mamá de Osito siempre le *proce* sombreros en sus cumpleaños. Al Osito le encanta que su mamá le __________ sombreros.
25. Osito ve el barco *norir* en el agua. Qué hace el barco? El barco _________ en el agua.
26. Osito tiene que *servir* la sopa. Él siempre ___________ la sopa.
27. En este dibujo Osito *vuela* en el cielo. Le gusta mucho __________ en el cielo.
28. A los animales les gusta mucho *divertirse*. El año pasado, en una fiesta de cumpleaños, todos se ___________. En las fiestas, siempre se ___________.
29. Osito *sale* de su casa al parque y antes de *salir* pregunta a su mamá, 'Mamá, ¿ __________ al parque?' Su mamá le contesta, 'No hijo, no quiero que _________ ahorita.'

Notes

While this research was being conducted, the first author was partially supported by a Research Training Fellowship from the Social Science Research Council. We would like to thank Carol Lynn Moder for help with the statistics used in table 4.5.

1. As shown in appendix 2, which is the complete text of one of the tests used, several different types of morphophonemic alternations were tested together. Thus they each served as filler for the others. The verbs *bierca* and *duenta*, which were given only in the diphthongal allomorph, were presented first, before any diphthong/monophthong models were presented, in order that the test itself not influence the response. The real verbs *servir* 'to serve', *volar* 'to fly', *divertirse* 'to enjoy oneself', and *salir* 'to leave' were presented last, again to avoid a priming effect. All of the adult subjects gave responses for the real verbs that were correct according to the standard language, except for two non-Chicano subjects who gave *sierve* for *sirve* and one Chicano who gave *vueler* for *volar* and *salgues* for *salgas*.

2. Chicano subjects are perhaps more accustomed to adapting non-Spanish verbs by this method because of the large number of English words that make their way into Chicano Spanish.

3. MacWhinney (1978) uses a technique in which stems are presented with two allomorphs and the subject in effect makes a choice of allomorphs for the response form.

4. Table 4.5 shows the results for both orders of presentation. There was no significant difference in results that correlated with the order of presentation of the forms.

References

Baxter, A. R. W. 1974. Some aspects of naturalness in phonological theory. B. Litt. diss., Oxford University.

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14, 150–177.

Bybee, J. 1980. Morphophonemic change from inside and outside the paradigm. *Lingua* 50, 45–59.

Bybee, J. L., and Mary Alexander Brewer. 1980. Explanation in morphophonemics: Changes in Provençal and Spanish preterite forms. *Lingua* 52, 201–242. (Reprinted in this volume as chapter 3.)

Bybee, J. L., and D. I. Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58, 265–289. (Reprinted in this volume as chapter 5.)

Foley, J. 1965. Spanish morphology. Unpublished doct. diss., MIT.

Greenberg, J. H. (ed.). 1963. *Universals of language*. Cambridge, Mass.: MIT Press.

Harris, J. W. 1969. *Spanish phonology*. Cambridge, Mass.: MIT Press.

———. 1972. Five classes of irregular verbs in Spanish. In *Generative studies in Romance languages*, ed. Jean Casagrande and Bohaan Saciuk, 247–271. Rowley, Mass.: Newbury House.

Hooper, Joan Bybee. 1976. *Introduction to natural generative phonology*. New York: Academic Press.

———. 1979a. Child morphology and morphophonemic change. *Linguistics* 17-1/2, 21–50. Also in *Historical morphology*, ed. J. Fisiak, 157–187. The Hague: Mouton.

———. 1979b. Substantive principles in natural generative phonology. In *Current phonological theories*, ed. D. Dinnsen, 106–125. Bloomington: Indiana University Press.

Hudson, G. 1974. The representation of non-productive alternations. In *Proceedings of the First International Conference on Historical Linguistics*, vol. 2, ed. J. Anderson and C. Jones, 203–229. Amsterdam: North-Holland.

———. 1975. Suppletion in the representation of alternations. Doct. diss., UCLA.

Kernan, K. T., and B. G. Blount. 1966. The acquisition of Spanish grammar by Mexican children. *Anthropological Linguistics* 8, 1–14.

Kuryłowicz, J. 1949. La nature des procès dits "analogiques." *Acta Linguistica* 5, 15–37.

MacWhinney, B. 1978. The acquisition of morphophonology. Monographs of the Society for Research in Child Development, no. 174, vol. 43.

———. 1985. The acquisition of Hungarian. In *The cross-linguistic study of language acquisition*, ed. D. Slobin, vol. 2, 1069–1156. Hillsdale, N.J.: Lawrence Erlbaum.

Mańczak, W. 1958. Tendances générales des changements analogiques. *Lingua* 7, 298–325, 387–420.

———. 1963. Tendances générales du développement morphologique. *Lingua* 12, 19–38.

———. 1980. Laws of analogy. In *Historical morphology*, ed. J. Fisiak, 283–288. The Hague: Mouton.

Ohala, J. 1974. Experimental historical phonology. In *Proceedings of the First International Conference on Historical Linguistics*, ed. J. Anderson and C. Jones, 353–390. Amsterdam: North Holland.

Skousen, R. 1975. *Substantive evidence in phonology*. The Hague: Mouton.

Stampe, David. 1973. A dissertation on natural phonology. Doct. diss., University of Chicago.

Vennemann, Theo. 1971. Natural generative phonology. Paper read at annual meeting of the LSA, St. Louis, Mo.

———. 1972. Rule inversion. *Lingua* 29, 209–242.

———. 1974. Words and syllables in natural generative phonology. In *Papers from the Parasession on Natural Phonology*, ed. A. Bruck, R. A. Fox, and M. W. LaGaly, 346–374. Chicago: Chicago Linguistic Society.

Chapter 5

WITH DAN I. SLOBIN

Rules and Schemas in the Development and Use of the English Past Tense

1. English irregular verbs

English verbal morphology is rather restricted, compared to that of a full-fledged inflectional language, since it offers only four inflectional morphemes: the 3s present, the past tense, the past participle, and the progressive. Thus it provides no opportunity to study the complex interaction between intersecting inflectional categories within a paradigm, such as person, number, mood, and tense. It does, however, provide the opportunity for a study of a different sort: although English has a demonstrably productive process of suffixation for past-tense formation, in the form of *-ed*, it also has many irregular verbs whose past tense is formed in some cases without suffixation and in others with changes of vowel (or, less commonly, consonant) in the stem.

These irregular verbs are relatively few in number. Bloch (1947) identifies about two hundred (several of which are archaic), but thousands of verbs form their past tense by adding an allomorph of *-ed*. Although irregular verbs are relatively insignificant as to type frequency, the picture is quite different when token frequency is considered: these are among the most frequent verbs of the language. Of the thirty most frequent past-tense forms (Kučera and Francis 1967), twenty-two are irregular. The situation changes radically in the second thirty most frequent past-tense forms, where eight are irregular. From the child learner's point of view, irregular verbs are also prominent: Slobin (1971) finds that, in forty-nine hours of adult speech to Roger Brown's subject Eve, between the ages of 18 and 26 months, irregular past-tense forms account for 292 of the past-tense tokens, while regular verbs comprise only 99. Thus irregular past-tense forms constitute an important core of English verbal morphology.

The difficulty presented by these irregular verbs, both for the learner and for the linguist, is that so many different irregularities occur. Though most irregulars are characterized by a vowel change, a number of different changes occur, for example, *stick/stuck*, *sing/sang*, *keep/kept*, *grow/grew*. In addition, some verbs add a *t* or *d*; and some have consonant changes as well, for example, *make/made*. Jespersen (1942) lists ten classes of irregular past-tense formations and dozens of sub-classes, while Bloch comes up with twenty conjugation types.

Such variety increases the difficulty of the learning task but at the same time provides us with an opportunity to investigate the interaction of rule and rote learning in morphology. It is clear that some rote learning is necessary in acquiring irregular past-tense forms: the pair *go/went* could not be learned in any other way. What is not so clear is whether speakers must learn all irregulars by rote or whether they can organize some of the verbs into classes as Jespersen and Bloch do—and even use some minor rules to generate irregular past-tense forms from a base form, as generative phonologists have suggested (cf. Hoard and Sloat 1973).

The two hypotheses—one, that irregular past-tense forms are learned by rote and stored in the lexicon, and the other, that such forms do not need to be stored in the lexicon but are generated by rule from base forms—have been investigated in the psycholinguistic literature. Thus Kuczaj (1977), examining spontaneous speech data from preschool-age children, takes the position that irregular past-tense forms are learned one by one, by memorization. He shows that chronological age is a better predictor of success on irregular past forms than mean length of utterance (MLU), while the reverse is true for the regular past-tense rule. He argues that chronological age will correlate naturally with amount of exposure to irregular forms—which suggests rote learning, which depends on exposure to particular forms. By contrast, MLU reflects the child's ability to formulate rules and is therefore not an expected predictor of advancement in rote learning. Kuczaj notes further that his data contain no instances of incorrect vowel changes—which would be predicted if the child were formulating vowel-change rules, rather than learning rote forms.

MacKay (1976), however, has examined response times for adults forming the past tense from a base-form stimulus, along with the errors that adults make; on this basis he argues that irregular past-tense forms are not stored in the lexicon but are derived from a base.

The two positions have counterparts in recent debates on phonological theory. In the tradition started by Chomsky and Halle (1968), it is considered desirable to handle as much irregularity as possible by rules, in order to "capture generalizations." Thus Hoard and Sloat formulate a small number of extremely abstract rules that can generate a large percentage of the vowel alterations in irregular verbs. However, a more recent trend in theoretical phonology claims that speakers do not formulate such abstract rules (Vennemann 1971). Evidence from historical change suggests rather that speakers function with surface true generalizations and must learn irregularities by rote.[1]

The data examined in the present study come from three different age groups: preschool children, 8½- to 10-year-old children, and adults. These data suggest to us that neither of the extreme positions is correct. Rather, we find evidence for an alternate view: that irregular past-tense forms are rote-learned and stored in the lexicon, but this does not prevent speakers from formulating generalizations about these forms. These generalizations are not in the form of rules that derive one thing from another by changing features. Thus we will not call them rules but will rather refer to them as *schemas*. A *schema* is a statement that describes the phonological properties of a morphological class (in this case, past tense). It does not relate a base form to a derived one, as a rule does, but describes only one class of forms (the product class, in terms used by Zager 1980). It is not a constraint that rigidly specifies what can and cannot occur but is rather a much looser type of correlation, used in organizing and accessing the lexicon. A schema for which we have found evidence may be stated as either 1a or 1b:

(1) a. A past-tense form may have the vowel [uw].
b. . . . uw]$_{\text{verb}}$
$_{\text{past}}$

Examples are the past forms *drew*, *blew*, *flew*. Note that the base forms of these verbs (*draw*, *blow*, *fly*) do not all have the same vowel, so the class cannot be defined in terms of the input, nor can a rule be stated that changes a single base vowel into the vowel [uw]. Rather, what these verbs share is the single vowel in their past forms. The existence of this schema does not, of course, exclude other shapes for past-tense forms. In the following sections, we outline our methods of data collection and discuss our interpretation of the results.

We mentioned earlier that we have collected data from three different age groups. While we are interested in the acquisition process, it is also necessary to be aware of what the outcome of this process will be, that is, the adult system. Some aspects of the adult system reveal themselves through the errors that adults make when they are forced to perform when tired or under pressure. A comparison of such errors with those made by children reveals both similarities and differences. In particular, both children and adults use schemas, though the particular schemas they formulate are somewhat different.

2. Data collection

The data from the preschoolers are spontaneous speech records from four independent sets, collected by Susan Ervin-Tripp and Wick Miller, by Dan Slobin, and by Zell Greenberg. Thirty-one children, between the ages of 1½ and 5 years, were interviewed; they used a total of sixty-nine irregular verbs.

These spontaneous data are supplemented by an elicitation task designed by Slobin and administered to twenty children, 3 to 4 years old, by Greenberg (1975). This was a sentence-completion task involving thirty-three real English verbs. The experimenter presented the child with a picture of a person engaged in a particular activity, and said, "This is a girl who knows how to______. She is ______ ing. She did the same thing yesterday. What did she do yesterday? Yesterday she ______."
The last blank is the one to be filled in orally by the child.[2]

The data on third-graders come from children between the ages of 8;6 and 10;1—eight girls and seven boys—all of whom were students at Cornell School, Albany, California. The children were given a sentence-completion task utilizing ninety irregular verbs. The verb was read first in its base form, and then a sentence was read using the verb in its base or present-tense form. Each child was then asked to complete a sentence using the same verb. This sentence began with a word that placed the action in the past, for example, *yesterday*. The test sessions were tape-recorded, and the past-tense forms supplied by the child were transcribed later. The test contained only irregular verbs, but a regular verb was used at the beginning as an example of what the child should do. A typical item would be *Blow: When I get a balloon, I always blow it up. Yesterday I . . .*

The task given to adults was designed to produce as many regularizations of irregular verbs as possible. A list was compiled of ninety irregular verbs, mixed in with three times that many regular verbs; thus two, three, or four regular verbs separated the irregular verbs. The subjects were asked to listen to the experimenter read the base form of the verb and then to produce the past-tense form of that verb as quickly as possible. The subjects were told that the object of the experiment was to see how fast they could go through the list, and the experimenters were instructed to put as much time pressure on the subjects as possible, since the purpose was to induce errors.

The list of 360 verbs was divided in half, and each half was administered to twenty subjects in two orders of presentation. All forty subjects were adult native speakers of American English, students at the University of California at Berkeley. The subjects' responses were tape-recorded and transcribed later.

The irregular verbs used in all three sets of data can be divided into eight classes as follows:

I. Verbs that do not change at all to form the past tense, e.g., *beat*, *cut*, *hit.*
II. Verbs that change a final *d* to *t* to form the past tense, e.g., *send/sent*, *build/built.*
III. Verbs that undergo an internal vowel change and also add a final *t* or *d*, e.g., *feel/felt*, *lose/lost*, *say/said*, *tell/told.*
IV. Verbs that undergo vowel change, delete a final consonant, and add a final *t*, e.g. *bring/brought*, *catch/caught.*[3]
V. Verbs that undergo an internal vowel change and whose stems end in a dental, e.g., *bite/bit*, *find/found*, *ride/rode.*

The rest of the irregular verbs are ones that undergo an internal vowel change and which neither end in *t/d* nor add *t/d*. Several different vowel changes occur among these verbs. As Class VI we have extracted two related groups that seem to be the most productive of the vowel-change verbs, while Classes VII and VIII are distinguished according to whether the verb base ends in a consonant or a vowel:

VI. Verbs that undergo a vowel change of /ɪ/ to /æ/ or to /ʌ/ e.g., *sing/sang*, *sting/stung.*

VII. All other verbs that undergo an internal vowel change, e.g., *give/gave*, *break/broke.*

VIII. All verbs that undergo a vowel change and that end in a diphthongal sequence, e.g., *blow/blew*, *fly/flew.*

The two primary criteria used in making these class divisions were (1) the presence or absence of a final *t/d* and (2) the presence or absence of an internal vowel change. Slobin (1971) finds that, in spontaneous data from children 1;6 to 4;0 years, strikingly fewer regularizations occurred with verbs that add a *t/d* in addition to making a vowel change (our Classes III–IV) than in verbs that undergo only a vowel change (Classes VI–VIII). Kuczaj (1977) finds only partial confirmation of this distinction in his data. We were interested in discovering if this distinction would emerge in our preschool, third-grade, and adult data.

It is also clear from Slobin's data that children are not as likely to regularize verbs whose past-tense forms are identical to the base forms (Class I). Similarly, Kuczaj (1978) finds that 4-year-olds are more likely to accept Class I verbs as past tenses than other irregular past forms and more likely to reject regularized verbs of this class (e.g., *hitted*) than regularized verbs of other classes (e.g., *failed*). It is not an accident that all the verbs of Class I have *t/d* as their final consonant. (Examples with *d* are *spread* and *shed.*) As has been pointed out by Stemberger (1981) and by Menn and MacWhinney (1981), it is common among the inflectional languages of the world to avoid adding an affix to a word or stem that already appears to contain that affix. Of course, many English verbs end in *t/d* and *do* add the /ɨd/ suffix, for example, *seated*, *blasted*, *shredded*: but the class of verbs that do not has long been significant and partially productive (Jespersen 1942:34–38). Our hypothesis is that this class exists because speakers apply a schema statable as 2a or 2b:

(2) a. A past-tense verb ends in *t* or *d*.
 b. $\ldots\, t/d\,]_{\substack{\text{verb}\\ \text{past}}}$

Verbs such as *cut*, *set*, and *spread*, then, fit the schema to their base forms and require no change. If the strategy of forming schemas is available to young children, then we might expect fewer regularizations among Class I verbs than among the other classes. Further evidence that encourages this expectation is reported by Anisfeld and Gordon (1968), who find that fifth-grade children (with a mean age of 10;4) accept nonce forms as past-tense forms more readily if the final consonant shares some features with *t/d*—for example, the affricate /č/, which contains an alveolar stop. We were interested to determine the extent to which evidence for schemas could be found among the three different age-groups.

Class V verbs have base forms ending in *t/d*, like verbs of Class I, but they also undergo a vowel change. A possible hypothesis about these verbs is that they might also resist regularization, because of their final consonant.

With regard to Class VI, we expected fewer regularizations because of the partial productivity of this class—as evidenced, for example, by dialectal *brang* or *brung* and by responses obtained by Berko (1958:165) when adult subjects formed the past tense of *bing* as *bang* or *bung* (50%) and the past tense of *gling* as *glang* or *glung* (75%).

Classes VII and VIII were separated to test the hypothesis that one of the important components of the process of acquiring irregular past-tense forms is the matching of base and past forms (Kuczaj 1977). It was hypothesized that this task would be easier for verbs with both initial and final consonants (or clusters) than for verbs that lack final consonants—since, in the CVC type, the base and past form will share more phonological features than in the CV type.

3. Results

3.1. Frequency

Since frequency of exposure is an important determinant of rote learning, we expect a significant correlation of frequency with correct use of irregular forms or an inverse correlation of frequency and percentage of regularizations, in children from both the preschool-age group and the third-grade group. We might even expect such a correlation in the adult errors—since, even though the adults have presumably mastered the irregular forms, some might be "stronger" than others, if (as MacWhinney 1978 suggests) lexical items gain strength with use and if this process continues into adulthood.

These predictions are borne out to some degree in all three age groups. For the preschool spontaneous speech data, it was possible to use past-tense form frequencies within the sample to test the hypothesis. A significant negative rank-order correlation (–.67) was found between the number of times that adult caretakers used a verb in the speech sample and the number of times that the verb was regularized by the children. That is, the more often a verb was used by caretakers, the less likely it was to be regularized.

For the regularizations made by the third-grade and adult subjects, rank-order correlations with frequency were calculated by verb class. The frequencies used were those of the past-tense form as given by Kučera and Francis. For the *hit* class of no-change verbs, the frequency of the unchanging base was used.[4] Significant negative correlations of frequency with percentage of regularizations were found for both sets of subjects in Class I, the *hit* class, and Class III, the *feel* class (I: third-graders –.603, n = 12; adults –.708, n = 16; III: third-graders –.663, n = 12; adults –.668. n = 12). Correlations that approach significance were found for both adults and third-graders in Class VII, the *break* class (third-graders –.546, n = 9; adults –.611, n = 11), and for the third-graders in Class VIII, the *blow* class (–.657, n = 6).

Correlations were not attempted for Classes II (*send*) and 1V (*catch*) because they are too small. No significant correlation was found in Class V (*bite*). We suspect this is because this class contains such a variety of vowel-change types. No significant correlations were found in Class VI (*sing*), even though only two vowel alternations are involved, because these verbs are all quite low in frequency and do not present a proper range upon which to test a correlation.

We conclude from these facts that frequency of usage is an important variable in the learning of irregular past-tense forms, as evidenced in the preschoolers' and the third-graders' usage, and that it is important to the maintenance of irregular forms, especially under pressure, as evidenced by the adult errors. However, frequency of usage alone cannot predict the error patterns of our data.

3.2. Class I: No-change verbs

When we consider the individual verb classes and compare them across age groups, we find, as expected, that the percentage of regularizations goes down between the preschool and third-grade groups for most verb classes (see table 5.1). The most striking exception to this is found in Class I, where the percentage of regularizations across all three groups of subjects is very much the same, between 27% and 29%, even though the over-all percentage of regularizations changes considerably between the preschoolers and the other age groups. Thus, among preschoolers, the percentage of regularizations in Class I is slightly below their over-all average; but for the other two groups, this percentage is considerably above average. In relation to their general performance, then, the third-graders and adults do worse on Class I verbs than preschoolers do.

As mentioned earlier, it is no accident that all the no-change verbs of English end in *t/d*—which may make them appear, at least to the new learner, as though they already contain a marker of past tense. Our hypothesis is that the preschoolers, rather than functioning with a suffixation rule that requires that *t/d* be *added* to the verb, are functioning with a schema like 2a or 2b, which says that if a verb ends in *t/d*, then it is an acceptable past tense. They are not concerned with the *addition* of a marker but rather with how the general shape of the word fits the pattern. Of course, they also show evidence of a suffixation process, but it is not strong enough to apply consistently to words that fit the schema. However, for third-graders and adults, the suffixation process (the requirement that something be *added* to form the past tense) seems much stronger.

The elicited past-tense forms of preschoolers show even stronger evidence for this schema. There were thirty-three verbs in this test. Six were regular: *walk*, *smoke*, *melt*, *pat*, *smile*, and *climb*. Eight had a vowel change: *drink*, *break*, *run*, *swim*, *throw*, *meet*, *shoot*, and *ride*. Eight had a vowel change and a final *t/d* added: *do*, *buy*, *lose*, *sell*, *sleep*, *help*, *teach*, and *catch*. Seven had no change: *hit*, *hurt*, *set*, *shut*, *cut*, *put*, and *beat*. In addition, the list included *go*, *make*, *build*, and *lend*. The most significant

TABLE 5.1. Percentage of Regularizations in the Eight Verb Classes and Three Age Groups

Verb Classes	*I* (*hit*)	*II* (*send*)	*III* (*feel*)	*IV* (*catch*)	*V* (*bite*)	*VI* (*sing*)	*VII* (*break*)	*VIII* (*blow*)
Preschoolers	29	57	13	10	34	55	32	80
Third-graders	27	55	20	7	9	8	9	2
Adults	27	20	10	31	6	20	11	8

Preschool percentages are based on combined data from thirty-one children; third-grade and adult percentages are averages of individual percentages of regularization. Since our analyses are concerned with inter-column comparisons for each age group separately, the differences in computing percentages should not alter the overall patterns discussed.

response patterns for this test were obtained by grouping the verbs into those ending in *t/d* (regular or irregular) versus those that did not and computing the number of responses that produced no change in the verb. The results are summarized in table 5.2. It can be seen that the children have a regular suffixation rule, but systematically fail to apply it to verbs already ending in *t/d*, even if they are regular (*melt*, *put*).

Berko encounters the same phenomenon in both plural formation and past-tense formation in her nonce-probe task. She finds that preschool and first-grade children show a highly significant tendency not to add the plural suffix to nonce nouns ending in sibilants, that is, those nouns that would take the /ɨz/ allomorph, such as *tass*, and not to add the past allomorph /ɨd/ to verbs ending in *t/d*. She interprets this as the late acquisition of the /ɨz/ and /ɨd/ allomorphs. We propose that these results are not related to the acquisition of a particular allomorph but are caused by the phonological shape of the nonce form—which, in verbs, conforms to the past-tense schema and, in nouns, conforms to a similar schema for the plural.

Berko also finds a difference between children and adults in the past tense supplied for nonce verbs ending in *t/d*: the adults consistently add the /ɨd/ suffix to these forms, but the children (as we have just mentioned) add the suffix much less often—only 31% and 33% of the times, respectively. Thus the difference we found between preschoolers and adults is the same found by Berko: the preschoolers seemed to apply a schema and were satisfied that verbs ending in *t/d* were sufficiently marked for past, but the adults applied a suffixation rule.

Since the third-graders treated the no-change verbs much as the adults did, the transition from a schematic analysis to a suffixation process must take place before age 8;6. Berko's subjects ranged between ages 4 and 7, so the change must take place around age 7. Data from nonce-probe tasks conducted by Derwing and Baker (1980) on children aged 3 through 9 give approximate confirmation; see table 5.3. Here it can be seen that the addition of a suffix to bases ending in *t/d* begins around age 6, and is established for some clusters by age 8 for nonce words.

Kuczaj (1978) also finds a dramatic change in the treatment of no-change verbs. He elicits grammaticality judgments from children by asking them to say which sentences produced by puppets were "silly." In the relevant portion of the experiment, the puppets produce pairs of sentences containing verbs such as *hit* and *bitted*. Table 5.4 shows the percentage of time that a form is treated as acceptable by the child. Note first that grammatical no-change verbs (*hit*) are never judged as unacceptable, whereas grammatical vowel-change verbs (*ate*) are so judged at times by groups 1 and 2. Note further that regularized no-change verbs (e.g. *hitted*) are treated as acceptable in many fewer instances than the regularized vowel-change verbs, especially in groups 1 and 3. These two points support the hypothesis of a schema that

TABLE 5.2. Number of Regularizations and No-Change Responses by Preschoolers

	Regular Suffix	*No Change*
Verb does not end in *t/d*	203	34
Verb ends in *t/d*	42	157

$\chi^2 = 180.1$ (significant at .001 level).

TABLE 5.3. Past Tense Responses to Nonce Verbs from Derwing and Baker 1974

				Age				*Total %*
Item	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*Correct*
nt	(Ø)	(ɨd)	ɨd	ɨd	ɨd	ɨd	ɨd	63
t	Ø	Ø	Ø	ɨd	ɨd	ɨd	ɨd	59
l/rt	Ø	(Ø)	Ø	Ø	ɨd	ɨd	ɨd	53
d	Ø	Ø	Ø	ɨd	Ø/ɨd	ɨd	ɨd	46
nd	Ø	Ø	Ø	ɨd	Ø/ɨd	ɨd	ɨd	41
st	Ø	Ø	Ø	Ø	Ø/ɨd	ɨd	ɨd	39
pt	Ø	Ø	Ø	Ø	Ø	Ø	Ø	18
kt	Ø	Ø	Ø	Ø	Ø	Ø	Ø	17
l/rd	Ø	Ø	Ø	Ø	Ø	Ø	Ø	13
ft	Ø	Ø	Ø	Ø	Ø	Ø	Ø	12
mpt	Ø	Ø	Ø	Ø	Ø	Ø	(Ø)	9
kst	–	–/Ø	Ø	Ø	Ø	Ø	Ø	4

says that verbs ending in *t/d* amount to acceptable past-tense forms. Note, however, the increase in the acceptability of regularized forms in the middle age-group; this is the point at which the suffixation process gains strength and partially replaces the schema. The percentage of regularizations judged acceptable in this verb group drops dramatically as the children sort out those verbs that do not take a suffix from those verbs that do. The suffixation process gains strength at a slightly earlier age in these data than in the data we have discussed from Berko and from Derwing and Baker, but this could result from the nature of the task, since Kuczaj asks for acceptability judgments of forms of real words, while the others are nonce-probe tasks.

Further evidence for a change in strategy is evident in Class V verbs (*bite*, etc.), which end in *t/d* but undergo a vowel change. The preschool children treat them much like the *hit* class, but the third-graders and adults show a difference between these verbs and the no-change verbs. The older two groups, under the influence of the suffixation rule, feel that they must somehow modify the base form to produce a past form. Verbs like *bite* can be modified by vowel change, but verbs like *hit* cannot. Therefore, verbs like *hit* are supplied with a suffix more often than verbs like *bite*.

Even though the suffixation process gains prominence around age 7 or 8, the schema remains available—but with a much restricted domain of applicability. Indeed, the schema must remain until adulthood, since the substantial class of no-change

TABLE 5.4. Percentage of Verbs Considered "Silly" by Children in Three Age Groups from Kuczaj 1978

Age-Group	*Grammatical No Change* (*e.g.*, hit)	*Regularized No Change* (*e.g.*, hitted)	*Grammatical Vowel Change* (*e.g.*, ate)	*Ungrammatical Vowel Change* (*base* + ed, eated)
1. 3;4–4;11	100	28.3	84.7	89.3
2. 5;1–6;7	100	55.0	94.7	60.0
3. 7;2–9;0	100	5.0	100.0	26.7
Overall	100	29.4	93.1	58.7

TABLE 5.5. Percentage of Verbs Ending in *t/d* and Percentage of No-Change Errors Among Adult Subjects

	Regular		*Irregular*	
	t/d	*Other*	*t/d*	*Other*
Percentage of verbs	27	73	29	71
Percentage of no-change errors	77.5	22.5	27	73

verbs remains to be handled. The problem is to attach the proper lexical restrictions to the schema, so that it applies only to the proper set of verbs. Positive evidence for the existence of this schema in the third-graders and adults comes from a study of the instances in which they erroneously gave a past-tense form as the unchanged base form of the verb. Recall that the list of verbs given to the third-graders contained only irregular verbs. Of these, fourteen belonged to Class I, where a response with no change in the base form would be correct.[5] Of the remaining seventy-six, final *t/d* occur in twenty-three (30%). The third-graders made thirty-seven errors in which an unchanged base form was given for a past tense, when some change would have been appropriate. Of these no-change errors, twenty-three (62%) occurred with verbs ending in *t/d*. These children were more than twice as likely to produce a no-change past tense with a verb ending in *t/d* than with a verb ending in another segment.

The adults erred in not producing a change where one was required on both regular and irregular verbs. Table 5.5 shows that verbs ending in *t/d* made up 27% of the entire corpus of regular verbs. Among errors made by producing no change in the base, 77.5% were on verbs ending in *t/d*. These subjects were almost three times more likely to produce a no-change error with a verb ending in *t/d* than with a verb ending in something else—which strongly suggests the presence of a *t/d* schema among the adults. Note, however, that the no-change errors made on irregular (vowel-change) verbs were not affected by the final consonant, since the proportion of errors on verbs ending in *t/d* was approximately the same as the proportion of verbs ending in *t/d* in the entire corpus. We cannot explain this difference between regular and irregular verbs.

The evidence suggests that two very different kinds of strategies are used in the analysis of the English past tense: a strict segmentation strategy, which considers verbs to be discretely segmentable into base and suffix and forms the past by suffixing *t/d*, and a strategy that does not require strict segmentation but requires only that a word in its general shape fit certain criteria. The former strategy results in a suffixation rule and the latter in a schema. Both strategies are natural, in that they are both evidenced in the child's original attempts to grapple with English morphology. What must be learned is the place of the two in the English system. The child originally overuses the schema strategy and must learn to restrict it.

3.3. Class II

Class II consists of a small set of verbs that form the past tense by changing final *d* to *t*. The preschool data contain the three examples *bend*, *build*, and *send*; in addition, the third-grade and adult lists contain *lend*. This small class is anomalous in the

TABLE 5.6. Past Tense Forms for Original Class II Verbs

t	*ed* OR *t*	*ed*
bend	*blend*	*wend*
lend	*geld*	*t* OR *ed*:
send	*gird*	*rend*
spend		NOT LISTED:
build		*shend* 'to shame'

sense that it employs neither of the basic methods of past-tense formation—neither addition of *t/d* nor internal vowel change. The regularization percentages in table 5.1 show all three age-groups responding to the anomalous nature of this class. The preschoolers' regularizations in this class were their second highest; and for the third-graders, this class had the highest percentage of regularizations. The adults' regularizations for this class were also above average.

An examination of the history of this verb class shows that the results reported here are in line with the continuing development of the class, which once had many more members. Table 5.6 is taken from Jespersen (1942) and is grouped according to past-tense forms listed in the *American Heritage Dictionary* (*AHD*).

It will be noted that the five verbs that still have a final *t* (plus *blend*, which has regularized) are all very common, while the remainder are very rare. Indeed, *shend* does not even appear in the *AHD*. This particular verb class is shrinking rapidly—not so much because the verbs are being regularized, but rather because they are falling into disuse. It is even possible that the anomalous past-tense formation may be partly responsible for the attrition in this class of verbs. Among the adults we tested, four gave *loaned* as the past form of *lend*, suggesting an avoidance of these anomalous past tenses by lexical substitution.

Related, perhaps, to this attrition of Class II verbs is the imminent loss of another phonologically similar class of irregulars, namely, verbs ending in *n* or *l* that formerly took a -*t* suffix for the past. These are shown in table 5.7.

TABLE 5.7. Past Tense Forms in American English of Original Class II Verbs

	AHD
burn/burnt	-*ed* or -*t*
learn/learnt	-*ed* or -*t*
pen/pent	-*ed* or -*t*
earn/earnt	-*ed*
dwell/dwelt	-*t* or -*ed*
smell/smelt	-*ed* or -*t*
spell/spelt	-*ed* or -*t*
spill/spilt	-*ed* or -*t*
spoil/spoilt	-*ed* or -*t*

For all these forms except *dwelt*, the regular suffix *d* is given first as the *American* form. The tendency to regularize these verbs in American English is observed by Jespersen. With these verbs regularized, the only verbs ending in *n* or *l* that have a *t* past tense are the few remaining in Class II. This means that a good deal of support for Class II type past-tense formation has been lost in American English.

3.4. Vowel-change verbs

The remaining verb classes, Classes IlI–VIII, all require an internal vowel change to form the past tense. The percentage of regularization for these classes in the preschoolers' corpus covers a large range, from 10% to 80%. The third-graders' corpus, however, shows no differences among these verb classes: they all have a very low percentage of regularization, from 4% to 9%. (The 20% for Class III is caused entirely by the three verbs *weep*, *kneel*, and *deal*, which have relatively low frequency.) This means that difficulties in handling vowel-change past tense for children aged 2 to 5 years are basically overcome by age 8;6. The adults' regularizations are below average in Classes III, V, VII, and VIII; they are higher only in Classes IV and VI. We will return to a discussion of the adult errors after a discussion of what the preschool data reveal about the acquisition of vowel-change past-tense forms in English.

In order for children to learn an irregular past-tense form, they must be exposed to the word (e.g., *heard*, *broke*) and must learn the following facts about it (Kuczaj 1977):

1. The form expresses past tense.
2. The past-tense form must be matched with the appropriate base verb.
3, The irregular form is used to the exclusion of a regularized form.

Consider now verb Classes III–VIII, listed in table 5.8, in descending order according to percentage of regularization by preschoolers.

Note that the two classes with the lowest percentage of regularizations are just those in which a *t/d* is added along with the vowel change. This confirms an earlier finding of Slobin (1971), on a somewhat smaller sample of children's speech, that verbs whose past forms end in *t/d* are regularized less than verbs of other classes. It

TABLE 5.8. Percentage of Regularization in Classes IV–VIII

	%	*Examples*
VIII	80	*blew*
VI	55	*sang*
V	34	*bit*
VII	32	*broke*
III	13	*felt*
IV	10	*caught*

is not immediately obvious, however, why the addition of *t/d* should aid the child in acquiring the vowel change—unless we consider the steps in the acquisition process, as outlined earlier. The final *t/d* added to the verb (*felt*, *heard*, *left*, etc.) can serve as an important clue to the child that the form expresses past tense. By this time, the child has learned the regular rule that adds a *t/d* suffix to a verb, so complete reliance on context is not necessary in order to discover that these forms are past tense. However, with past forms such as *broke*, *bit*, *sang*, and *blew*, the child must rely entirely upon the context. We suggest, then, that the head start provided by the final *t/d* in Classes III–IV accounts for the earlier mastery of the past-tense vowel—since the sooner children succeed in analyzing the form as past, the sooner they can approach the problem of matching the past form with a base form.

This task must depend upon both semantic content and phonological shape. That is, the child must ascertain that *break* and *broke* mean the same thing, except that *broke* is past while *break* is not, and also that *broke* is actually the past form of *break* and must be used rather than *breaked*. This process is apparently aided by shared phonological features of the base and past forms. The phonological clue that the child can use in matching base with past is the consonantal structure of the verb. (This is not unlike the task that faces the Hebrew-speaking child, who must identify consonantal roots for all verbs; see Berman 1981.) Some verbs offer more consonantal structure than others and would therefore be easier to master. Among the verbs considered in this study, we have found a striking difference between Classes V, VI, and VII, which have both initial and final consonants or clusters (e.g., *break*, *sing*, *bite*), and Class VIII, which has only initial consonants or clusters but no final consonants (*blow*, *know*, *see*, etc.) In the former groups, regularization ranges between 32% and 55%; in the latter, regularization is an overwhelming 80%. The problem is especially clear with verbs like *know* and *see*, where the only shared phonological material between base and past is a single initial consonant *n*- or *s*-.

While the matching problem is completely overcome by the time children reach the third-grade level, our data contain hints of various sorts that phonological distance remains a factor in storage and accessing. Some of these hints are anecdotal. Two of the third-grade subjects reacted to the pair *go/went*, which has the greatest phonological distance of any English base–past pair. One boy (age 8;6) gave the form *goed* initially and then corrected himself immediately, saying, "No. Well, you have to use the word *went*, then, because you can't say *goed.*" Note the implication that *went* is a separate word. Another boy (age 9;2) paused for one minute and forty-five seconds without responding to *go*. Finally the experimenter suggested *went*; the subject then acquiesced, somewhat reluctantly, saying *went* was not enough like *go*. One other instance in which a third-grader had difficulty supplying the past form of a common verb was the case of a girl (age 9;7) who could not think of the past form of *make*. She kept saying *maked* and then rejecting it, but she wouldn't accept *made*, either, when it was suggested by the experimenter. These instances suggest that a greater phonological distance makes the accessing task more difficult.

Phonological distance can also be offered as an explanation for the high rate of regularizations by adults in Class IV. This class contains the verbs *catch*, *bring*, *think*, and *buy*, which all have high frequency and should present no problems for adults. But these verbs have past forms that differ from the base in both vowel and consonant.

The high percentage of regularizations in this class is caused primarily by *catch* and *think*, and *make* also has a high number of regularizations. This could well be because the phonological distance makes the past forms more difficult to access under pressure. This suggestion is consistent with the findings of Stanners et al. (1979).

To summarize, shared phonological material is important in the preschoolers' task of matching base with past forms. This can be seen clearly in the high regularizations of Class VIII as compared to V–VII. Class IV is apparently mastered earlier by preschoolers, despite the phonological distance between base and past, because the final *t* facilitates the identification of the past form as such and thus speeds up the matching of past form to base verb. (No conclusion about Class IV should be drawn for the third-graders, because in the week when about half of them were tested they had all the past-tense forms from Class IV on their spelling list.) The factor of phonological distance remains important for the continuing connection between base and past forms, as evidenced anecdotally for third-graders, and in the high regularizations in Class IV among adults.

Finally, we must consider the relatively high percentage of regularizations made by adults in Class VI. These verbs were isolated from other vowel-change verbs because the change of this class shows some productivity. Indeed, common among the incorrect vowel changes made by adults were forms such as *streak/struck* and *clink/clank* (or *clunk*). Because of this productivity, we expected that there would be fewer regularizations for these verbs and more correct responses. This turns out not to be true. In fact, just the reverse is true: there were significantly more regularizations in this class than in the other vowel-change classes (with the exception of Class IV), and we do not know exactly why. It is worth noting, however, that the verbs in this group tend to be of relatively low frequency: ten of the twelve are less frequent than other irregular verbs. To find out if frequency is the relevant factor, we averaged the percentage of regularizations for those verbs having a frequency of 23 or less. The average of regularization was 23.3%. Then we averaged the percentage of regularizations for all the other verbs with a frequency of 23 or less, excluding Class I verbs. This average was 17.85%, which suggests that verbs of Class VI are not behaving much differently from other low-frequency verbs.

4. Novel vowel changes

4.1. Class VI

The novel vowel changes made by the adult and third-grade subjects provide us with information on the storage and accessing of irregular past-tense forms.[6] Approximately half of these innovations occurred among Class VI verbs. All these innovations were very systematic, and they will be discussed first.

Class VI contains two types of verbs: those with three forms, listed as Class VIa, and those with only two, Class VIb. (For complete lists, see appendix.) In the first sub-class, the verbs all contain a nasal consonant following the vowel. This is a velar nasal (either /ŋk/ or /ŋ/) in all the verbs except *swim*, *begin*, and *run*. The verbs of the second sub-class may also contain a velar nasal (*sting/stung*), a velar without nasality (*dig/dug*, *stick/stuck*), or a nasal that is not velar (*win/won*). The psychologi-

TABLE 5.9. Verbs Incorrectly Given an /æ/ Past

	Number of Errors	
	Adults	*Third-Graders*
string/strang	2	1
sting/stang	2	1
slink/slank	1	
clink/clank	1	
clanked	3	

cal reality of these phonological constraints is clear from the verbs that are incorrectly put into these classes by our subjects.

Consider first table 5.9, which lists the four verbs that were incorrectly given an /æ/ past form; only *clink* is not from Class VIb. Note that all these verbs have a velar nasal following the vowel. In no examples did /æ/ appear in a verb of Class VIb that did not have a nasal; for example, we got no forms like *dig/* dag* or *stick/*stack*. The /æ/ is clearly connected to the velar nasal; in fact, we would suggest that it is the whole complex /æŋ/ that creates the past tense. We suggest a schema of the following form:

(3) . . . æŋ(k)$]_{\substack{\text{verb}\\ \text{past}}}$

Note that the schema defines a prototype of the category (in the sense of Rosch and Mervis 1975), in that *sing* and *drink* are the best exemplars—but *swim* and *begin* may also belong to the category, because they end in nasals, although not velar nasals. All the novel uses of /æ/ involved verbs that would be central to the category, those ending in /ŋ/.

A much larger set of innovations involve past-tense forms with the vowel /ʌ/. Two types of verbs show this innovation: those with this vowel in their past participles (the *ring/rang/rung* subgroup) and those of other types. They are listed in table 5.10.

First, it must be observed that many of these responses are not exactly "errors": some are dialectal variants, and the whole set reflects ongoing change in the system. *Shrunk* is, for some speakers, the acceptable past form for *shrink*; it clearly seems to be so for the third-graders. However, the adult subjects were all speakers of standard American English and would not ordinarily use *rung*, *swum*, or *brung* as past-tense forms. For them, these forms are innovative. But the third-graders were not being pushed to make errors but were using these forms in sentences—which should reflect, to some degree, their normal usage. (Some of these same children have been observed using these same forms in their spontaneous speech.)

It is interesting, then, that the adult innovations should show such systematic correspondence with the children's usage. These forms constitute the bulk of novel vowel changes made by the third-graders—only six did not involve Class VI.[7] Errors

TABLE 5.10. Verbs Given with a Non-Standard /ʌ/ Past

		Number of Errors	
		Adults	*Third-Graders*
(a)	*shrink/shrunk*	5	9
	ring/rung	5	6
	swim/swum	5	1
	begin/begun	2	2
	drink/drunk	2	6
	sink/sunk	7	[a]
(b)	*bring/brung*	1	2
	think/thunk	0	1
	clink/clunk	1	[a]
	streak/struck	4	[a]
	eat/ut	0	1
	shake/shuck	0	2

[a] Not given on third-graders' test.

in Class VI constituted almost half, forty-one out of eighty-five, of all incorrect vowel changes made by adults. This correspondence between adult errors and child usage is evidence of very similar grammars, even though normal usage might be quite different (a possibility countenanced by the model proposed in Andersen 1973).

In explaining the errors listed in table 5.10 earlier, we must first discount the possibility that these forms resulted merely from the subjects' responding with the past participle rather than the past-tense form. It is true that the forms given in (a) are the past participles of these verbs and this contributes somewhat to the likelihood that they will be used but this is by no means an adequate explanation in itself. First, the similar errors in (b) are not past-participle forms. Second, the third-graders made no other errors that consisted of supplying a past-participle form rather than a past tense. So, for the third-graders, these forms must be explained in terms of this verb class. Finally, though the adults did make other errors of giving a past participle for a past tense, twenty-six such errors occurred with the six verbs listed in (a), while the other nineteen were distributed over fourteen different verbs. So 58% of such uses occurred with these six Class VI verbs—a fact that needs explaining.

The forms in (a) do not consist of simply supplying a past participle rather than a past tense but rather result from the subjects' treating these verbs, which have three principal parts, as though they belonged to the other sub-class of Class VI with only two principal parts. In other words, the errors made by both adults and third-graders give evidence for a breakdown in the distinction between the two subgroups of Class VI. To understand this, we must analyze the members of Class VIb to determine the schema that fits them.

It is difficult to state a single past-tense schema for this class of verbs, because some of its members share no phonological features other than the past-tense vowel /ʌ/; for example, *dug* and *won* share no other attributes. However, this can be considered a coherent class if it is analyzed in terms of family resemblances (cf. Rosch and Mervis). The most common and the best examples of this class end in the velar nasal. All other members share some features with this prototypical example—either the feature of ending in a nasal or that of ending in a velar or both, as in table 5.11. The forms in (b) of table 5.10 attest to a limited productivity for this class—and further to its family resemblance organization, since the innovations involve verbs ending in /ŋ/, /ŋk/, /k/, and even /t/.

A statement of the schema for this class must involve a disjunction; that is, the verb must end in either a velar or a nasal:

$$(4) \quad \ldots \begin{Bmatrix} \text{nasal} \\ \text{velar} \end{Bmatrix}]_{\substack{\text{verb} \\ \text{past}}}$$

The difficulty encountered by our subjects, and by English speakers in general, is that of distinguishing Class VIa from Class VIb, since the prototypes of both classes are verbs ending in a velar nasal.

There are two reasons that this difficulty in distinguishing classes results in more innovations in the direction of past tense in /ʌ/ than in past tense /æ/. The first is that a past tense in /ʌ/ produces an existing word (the past-participle form); and as we shall see later, a large majority of novel vowel changes actually consist of supplying the wrong existing word. Second, the class with /ʌ/ in the past form is larger and more general than the other class and continues the general trend in English not to distinguish the past and past-participle forms.

In fact, the results we have obtained here for Class VI reflect a change that has been in progress for several centuries, by which the /æ/ past form of a Class VIa verb is lost, producing a Class VIb verb. The source of many Class VIb verbs is precisely this: a loss of the distinction between a past-tense form and a past participle (Jespersen 1942:49–53). The question that we cannot answer at this point is why the participle form, the form in /ʌ/, is the one that survives when the distinction is lost.

To conclude this section, three points emerge concerning the organization of Class VI verbs: (1) the responses show a continuation of the historical trend to collapse the past and participle forms into a single form with the vowel /ʌ/; (2) the phonological shapes of the verbs in this class form a series of family resemblances, rather than sharing a discrete set of phonological features; and (3) the class defined by the schema is one of past-tense forms, not one of base forms. The important property of the past form is that it contains the vowels /æ/or /ʌ/ and certain consonants; it does not matter what vowel the base form contains. This is clear for

TABLE 5.11. Chain of Family Resemblances for Class VIb

Nasal			*Velar*	
/n/	/ŋ/	/ŋk/	/k/	/g/
spin	*cling*	*slink*	*strike*	*dig*

the forms *strike/struck* and *sneak/snuck* (both of which are relatively new to the class)—as well as the novel form in our data, *streak/struck*, which occurred four times. Furthermore, if only the past forms are considered, then *run* and *hang* will also belong in Class VI.

4.2. Novel vowel changes: Other verbs

Considering now only the responses from the adult subjects, forty-four of them showed novel vowel changes. Of these, thirty-five (80%) of the responses were real past-tense forms of some other verb. Of these thirty-five, we find that twenty-nine (83%) were past-tense forms of verbs semantically related to the base forms given as input. To meet the criterion of semantic relatedness, the stimulus and response verbs must either be synonymous (e.g., *search/sought*) or be members of transitive/intransitive pairs (*set/sat*). There were four responses with a verb that was not in the past tense; three of these ended in *d*, thus fitting the past-tense schema. One noun was supplied and four non-words. (We counted /gluw/, which the experimenter spelled *glew*, as a non-word, because we are confident that the subjects did not intend the noun *glue*.) In all, forty out of forty-four (91%) of these novel change responses were real English verbs.

These results are clear evidence that the production of vowel-change past-tense forms is a matter of lexical selection rather than rule application. If speakers apply phonological rules to base forms, changing the vowels to generate past forms, and if the innovations that they create under pressure result from the misapplication of these rules, then we would expect novel vowel changes to produce mainly non-words of English, such as *hept* and *snoze*. However, if these irregular past forms are stored in the lexicon, then the production task involves lexical selection, and we would expect errors made under pressure to be caused by the selection of the wrong pre-existing word. Since it is precisely the latter type of error that we find in the overwhelming majority of cases, we conclude that irregular past forms are stored in the lexicon.[8] The few cases that cannot be accounted for by lexical selection will be discussed later.

An analysis of the responses listed in table 5.12 gives some information about the nature of the process of lexical access. The errors indicate that syntactic category (verb), morphological category (past tense), phonological shape, and semantic field all play important roles in lexical selection. The errors of novel vowel change that are real words are all verbs (except for a single adjective, *proud*). Thus the constraint on syntactic category is very strong. As mentioned earlier, thirty-five out of forty-four of the errors were words of the correct morphological category, that is, past tense. All but one of the errors conform in phonological shape to the initial consonant or consonant clusters of the stimulus word, and only four words do not have matching final consonants: *search/sought*, *sight/saw*, *crawl/crept*, and *glide/glow*.[9] The semantic matches are not quite so strong: of all the real-word responses, 75% were semantically related.

The responses show that speakers can access the lexicon from any one of these four parameters—or, perhaps more accurately, from all four at once. This suggests a complex set of inter-relations among lexical forms. The errors consist mainly of selecting items that match phonologically and syntactically but may be semantically

TABLE 5.12. Novel Responses Involving Existing Words

Semantically related	*Numbers of Errors*
(*a*) *Real past tenses* (*35*)	
seat/sat	9
set/sat	5
lend/loaned	4
raise/rose	4
search/sought	4
cite (*sight*)*/saw*	1
/sought	1
crawl/crept	1
Not semantically related	
flow/flew	2
lean/lent	2
ride/rid	1
shun/shone	1
(*b*) *Real verb* (*not past tense*)	
sled/slide	1
Not semantically related	
glide/glow	1
weed/wed	1
wheel/weld	1
(*c*) *Adjective*	
pride/proud	1
(*d*) *Not real words*	
glow/glew	2
heap/hept	1
snooze/snoze	1

aberrant. The strength of the phonological constraints over the semantic may result from the nature of the task, but it at least shows clearly that speakers of English have a strong and productive strategy that allows past tenses to be formed by holding the consonantal structure of a verb constant, while changing the vowel.

Several past-tense schemas also play a role in lexical selection. The speaker brings to the task of lexical selection a set of notions about what a past-tense verb should sound like. These notions are expressed in schemas that describe certain prototypical phonological structures corresponding to past tense. A word is more likely to be selected if it conforms to one of these schemas. It seems also (though it is clearly less usual) that new words may be created on the basis of the schemas.

The schemas for which positive evidence exists in the responses are those for the past-tense vowels /ɛ/, /uw/, and /ow/. These are the only ones in which a case could be made for rule application, since they are the only vowels that occur in the non–semantically related, non-past-tense, and non-word errors. The innovations are restated in table 5.13, grouped according to past-tense vowel.

TABLE 5.13. Vowel Change Responses Resulting in Non-related, Non-past and Non-word Errors

		Number of Vowel Change Responses
(a)	*lean/lent*	2
	wheel/weld	1
	weed/wed	1
	heap/hept	1
(b)	*flow/flew*	2
	glow/glew	2
(c)	*snooze/snoze*	1
	glide/glow	1
	shun/shone	1

The class of real past forms corresponding to the errors in (b) of table 5.13 are listed in table 5.14 and must also be described by a schema. The defining properties of this class are that these verbs have a past-tense form ending in /uw/ with no consonant. The vowels in these base forms differ. It will be recalled that this class of verbs has the highest rate of regularization among preschoolers and the rate is fairly consistent across verbs. But among third-graders and adults this class has a consistently low rate of regularization. These facts, along with the errors listed in (b) of Table 5.13, suggest this schema:

(5) $\ldots \text{uw}]_{\substack{\text{verb} \\ \text{past}}}$

The schemas are used in two ways—both exemplified by the errors in our corpus. In lexical selection, words are favored that fit one of the past-tense schemas. Thus we have four occurrences each of *rose* and *loaned* among the semantically related words; these words fit the same schema as those in (c) of Table 5.13. Class VI errors include four occurrences of *streak/struck*, plus the past participles of Class VIa that fit the schema for Class VIb. In addition, the schemas (or the prototypes that the schemas

TABLE 5.14. Percentage of Regularization in Class VIII

	Percentage of Regularization		
	Preschool	*Third Grade*	*Adult*
blow/blew	100	7	10
draw/drew	100	0	10
fly/flew	82	7	5
grow/grew	100	7	10
know/knew	86	0	10
throw/threw	82	0	0

describe) can be used to create new words on a limited scale; thus we have innovations such as *snoze*, *hept*, and *glew*.

5. Discussion

Here we attempt to draw together what we have learned about the acquisition and processing of English past-tense verb forms. We have found evidence that bears on the issue of the interaction of rule and rote learning. In particular, we find, as expected, substantial evidence that English irregular verbs are rote-learned and stored in the lexicon, rather than generated by rule from base forms. But the evidence indicates that at least some of the rote-learned forms are organized into classes according to the phonological shape of the past-tense form. These classes are described by *schemas*.

Zager (1980) distinguishes two types of morphological modifications by which new words are created: source-oriented and product-oriented. The former type can be described as a specific modification of a basic form to create a derived one, for example, English regular past-tense formation. Product-oriented modifications are less concerned with the shape of the base form (the source word) and more with creating a product that resembles other words of the same morphological category. We are claiming that English irregular past-tense forms make up product-oriented classes, since these forms are partially similar despite differences in base forms; for example, *strike/struck*, *stick/stuck*, *sneak/snuck* have different vowels in the base forms, but the same vowel in the past forms. Schemas describe such product-oriented classes. It is important to note, however, that we are not denying the existence of source-oriented modifications or of basic derived relationships between forms (see Bybee and Brewer 1980).

Besides defining product classes, schemas have the following characteristics:

1. Their defining properties are phonological and can range over more than one segment. In the Class VI verbs, it is not the vowel that defines the class but the vowel and the consonant together; thus the whole sequence /ʌŋ/ signals that the verb is in the past. This is different from saying that the vowel /ʌ/ marks past in verbs ending in /ŋ/.
2. Classes of items covered by schemas are defined in sets of family resemblances, not by sets of strictly shared properties. Thus *won*, *slung*, and *dug* can belong to the same class, not because they all share the same phonological properties but because each has some property in common with at least one other member.
3. Though schemas do not in themselves change features, they are used in lexical selection; and they may serve as the basis of new formations occasionally, either in speech errors, such as those we elicited, or in so-called analogical formations, such as past *snuck* for earlier *sneaked.*

We envision the following as the use of schemas in past-tense formation in English: When speakers need to use a verb in the past tense, they begin a process of lexical selection, deciding on a verb with an appropriate semantic representation. If the verb arrived at is not marked as irregular, it is pulled out and the regular past-tense suffixation rule is applied.[10] But if the verb has an irregular past form, then the lexical entry will contain a mark that prevents the selection of the base form and ensures the selection of the associated past form.[11] The further lexical search will call on semantic and syntactic parameters, just as the original search did. In addition, phonological properties are important, since the past form should share its consonantal shape with the stem but have a different vowel. The particular vowel is not predictable, but certain vowels are more probable than others, and these are the ones described in the schemas. Thus the schemas aid in the identification of a form as a past-tense form, because they describe the phonological shape expected of such forms. Note, however, that since the forms in question are rote forms, correct lexical selection may proceed without schemas.[12] The schemas simply represent one of many ways that lexical information may be organized for more efficient accessing.

In this postulated lexical search procedure, it is not necessary for a speaker first to access the base form of a verb and then determine whether it is regular. Rather, when the search arrives at an appropriate lexical entry, both the base form and its irregular past tense are immediately available. The search cannot totally bypass the base form. Stanners et al. (1979) have found that accessing an irregular past form can also serve as a prime for the base form. However, it seems that the phonological similarity of the base and past forms determines the extent to which they are accessed together.

The lexical selection of a no-change verb presents a slightly different situation, since the verb must be marked as not undergoing the regular suffixation rule; but there is no associated lexical form that can be chosen as the past-tense form. Under conditions of pressure and fatigue, adult subjects often put these verbs through the regular suffixation rule, despite the mark that should prevent the rule's application. The third-grade subjects responded similarly, as noted earlier. The correlation of correct responses with frequency in this verb class (noted in section 3.1 earlier) indicates that the strength of the mark that bars the regular rule grows with correct application.

Comparing the use of schemas across age groups, we find that both children and adults formulate and use schemas for the English past tense, but each group uses particular schemas. Preschool children orient to both the source and product classes of the regular past tense. They have the regular suffixation rule; but when it appears that a verb already belongs to the product class—that is, if it ends in *t/d* and fits their schema for past tense—the children are satisfied not to apply suffixation. Thus preschool children can analyze a past-tense form from two points of view: in terms of its relation to a base form or in terms of other past forms.

Adults and third-grade children can do the same thing, but they have learned that the suffixation process that modifies the source or base word in a certain way takes precedence. Thus they make more errors of adding *-ed* to zero-marked past forms than preschoolers do. in addition, adults show evidence of having developed a set of schemas for the vowel-change verbs. The strongest of these covers the verbs with past forms in /ʌŋ/ and related verbs. Schemas also exist for past-tense /æŋ/, /ow/, /uw/, and /ʌ/. Many irregular verbs may not be covered by any schemas at all.

These differences between children and adults have important implications for the relation between child language and historical change. It has often been suggested that children's innovations in the morphological system are the source of diachronic changes in this system; but this could he true only if children's creations survived with them into adulthood. Our data show that, in some areas of English verb morphology (e.g., the treatment of no-change verbs), children's and adults' strategies are somewhat different and that a real change in approach takes place before age 8. This means that whatever tendency younger children manifest to apply a schema to verbs ending in *t/d* is nullified by the strength of the suffixation rule in speakers over the age of 8. As things currently stand, if any changes are occurring in Class I verbs, we predict that they will be in the direction of regularization—e.g., *wetted*—rather than the opposite direction, that of adding new members to Class I. However, it should also be pointed out that the preschoolers' strategy of forming schemas is a natural and universal tendency in adult morphology; even in English, the competition between the schema and the suffixation rules survives in adults. Thus child morphology continues to be of great importance to the study of both synchronic and diachonic morphology.

The tendency we found shared by third-graders and adults to use an /ʌ/ past form for verbs with a standard /æ/ past form is a reflection of a long-standing series of changes in English. The important point here is that the errors elicited from adults under pressure reveal the same system shown in long-term historical changes and in the utterances of children aged 8½ to 10. This means that elicited errors may be used with some degree of confidence in investigating other aspects of adult morphological systems and that these errors project both the current system and future changes in it.

The more extensive use of product-oriented schemas in child language may be caused by the acquisition strategy which MacWhinney (1978), as well as Menn and MacWhinney (1984), call the affix-checker. The affix-checker determines whether or not rote forms already contain the desired affix—and if they do, whether it has suppressed the affixation process. We suggest. however, that the phenomena associated with affix-checking (such as the lower regularizations of the *hit* class) are part of a larger phenomenon (i.e., the tendency of children to form product-oriented schemas). We suggest that such schemas are precursors to source-oriented rules that change or add features. In fact, product-oriented schemas may provide the missing step between amalgams and productive rules. As MacWhinney observes, regular affixes are first used in amalgams, that is, rote-learned unanalyzed forms. From such amalgams common properties may he extracted to develop a schema, which describes a class of items. It is a further step to relate this class to some other (perhaps more basic) class by rules that specify how to move from the basic to the derived form. In some cases, this final step may not be possible (as in the vowel-change schemas), and the generalization remains product-oriented. It may be useful, then, to look for other cases in child language of product-oriented generalizations serving as a link between amalgams, or rote formulae, and source-oriented rules.

Appendix: Subgroups of Class VI verbs in English

Class VIa:	*m*	swim/swam/swum	Class VIb:	*n*	spin/spun
	n	begin/began/begun			win/won
		run/ran		*ng*	bring/brung*
	ng	ring/rang/rung*			cling/clung
		sing/sang/sung			fling/flung*
		spring/sprang/sprung			hang/hung*
	nk	drink/drank/drunk			sling/slung*
		shrink/shrank/shrunk			sting/stung*
		sink/sank/sunk			string/strung
					swing/swung
					wring/wrung
				nk	slink/slunk
				k	sneak/snuck*
					stick/stuck*
					strike/struck*
					shake/shuck*
				g	dig/dug
					drag/drug

*Verbs marked with an asterisk were not members of this class in Old English (cf. Jespersen 1942).

Notes

A number of people contributed to the research reported here, and we would like to thank them for their help. Susan Ervin-Tripp, Roger Brown, and Zell Greenberg made their spontaneous speech data available to us; Zell Greenberg tested the preschool children and analyzed the results; Amy Strage and Ann Eisenberg tested the adults; Mr. B. Harris and his third-grade class at Cornell Elementary School, Albany, Cal., played the verb game with us; Tanya Renner, Carol Lynn Moder, and BongHee Choi helped with the statistics; and Brody Hooper let us use his verb forms and his friends.

While this research was being conducted, Joan Bybee held a Postdoctoral Research Training Fellowship from the Social Science Research Council. This research was also supported by the University of California Committee on Research and the Institute of Human Learning, University of California at Berkeley.

1. The True Generalization Condition (Hooper 1976) applies to fully specified strings containing morphological, lexical, and syntactic information, and it can be a condition on morphophonemic rules, even if they have exceptions, so long as the exceptions are marked in the string.

2. The entire body of tests and the full list of verbs studied may be obtained by writing to either of the co-authors.

3. *Buy/bought* is included in this class, even though no final consonant is deleted, because its past tense rhymes with the three other members of the class: *thought*, *caught*, *brought*. Other evidence suggests that these verb classes are defined in the past form, not the base form.

4. Unfortunately, some of these verbs (such as *hit* and *cut*) are also used as nouns, and Kučera and Francis do not distinguish words by syntactic category. Thus our frequency correlations must be considered only a rough approximation. This problem is not nearly as common in the other verb classes where the frequency of the past form can be used, since past forms are less commonly homonymous with nouns.

5. *Spit* was included in this class, since none of the children gave *spat* as the past tense and *spit* is one of the accepted past forms according to the *American Heritage Dicutionary*.

6. It is possible that some of the novel vowel changes were caused by perseveration from earlier items. We are unable to check for this effect, since it is not clear how recent an item would have to be to cause perseveration. However, we will be glad to supply the lists of verbs in their orders to any interested readers who would like to check. We suspect that such intrusions would be minor.

Note also that the task was designed to elicit regularization errors. A different kind of task might show more productivity of vowel-change patterns.

7. Outside of Class VI, the third-grade subjects made only six incorrect vowel changes. These are too few to analyze. but we list them here for completeness: *ride/red* (1), *ride/rid* (1), *set/sat* (2), *fight/fit* (1), *weave/waved* (1).

8. MacKay (1976) notes the same phenomenon in his errors of past-tense formation. In fact, all MacKay's results are consistent with our position (that irregular past-tense forms are stored in the lexicon), even though he argues for rules. When he measured response time for subjects supplying past-tense forms, he found that the fastest responses occurred in verbs with "simple" vowel changes, followed by the regular verbs, then by verbs with "complex" vowel changes, and finally by verbs with vowel change and a *t* suffix. The problem with interpreting these results as favoring lexical selection or rule derivation is that no one knows whether a lexical search takes a longer or shorter time than derivation by rule.

9. In only one case was an error semantically related, but without phonological resemblance to the stimulus word: *let/allow*.

10. It is possible that some regular past-tense forms, particularly of high-frequency verbs, are also stored in the lexicon, but we have found no evidence in the current research to support this possibility.

11. The strength of this particular mark is evident in the small number of "overmarking" errors (errors in which a vowel change *and* suffixation apply to a verb) in the adult and third-grade corpora of errors. The third-graders produced only three: *shrunked*, *shooked*, and *stucked*; the adults produced *clanked* (three times) and *clunked* (once), both for the past of *clink*. Note that both *clanked* and *clunked* could be past-tense forms of other verbs.

12. In this way, schemas resemble "via rules" (Vennemann 1972), except that they define product classes rather than relations among forms.

References

Andersen, Henning. 1973. *Abductive and deductive change*. *Language* 49, 765–793.

Anisfeld, M., and M. Gordon. 1968. On the psychophonological structure of English inflectional rules. *Journal of Verbal Learning and Verbal Behavior* 7, 973–979.

Berko, J. 1958. The child's learning of English morphology. *Word* 14, 150–177.

Berman, Ruth A. 1981. Language development and language knowledge: Evidence from the acquisition of Hebrew morphology. *Journal of Child Language* 8, 609–626.

Bloch, Bernard. 1947. English verb inflection. *Language* 23, 399–418. (Reprinted in *Readings in linguistics I*, ed. Martin Joos, 243–251. Chicago: University of Chicago Press, 1966.)

Bybee, Joan L., and Mary Alexandra Brewer. 1980. Explanation in morphophonemics: Changes in Provençal and Spanish preterite forms. *Lingua* 52, 201–242. (Reprinted in this volume as chapter 3.)

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Derwing. Bruce L., and William J. Baker. 1980. Rule learning and the English inflections. In *Experimental Linguistics: Integration of Theories and Applications*, ed. G. D. Prideaux, B. L. Derwing, and W. J. Baker, 248–272. Ghent: E. Story-Scientia.

Greenberg, Zell. 1975. Overgeneralization of past tense inflections. Senior honors thesis, Department of Psychology, University of California at Berkeley.

Hoard, James E., and Clarence Sloat. 1973. English irregular verbs. *Language* 49, 107–120.

Hooper, Joan Bybee. 1976. *Introduction to natural generative phonology*. New York: Academic Press.

Jespersen, Otto. 1942. *A modern English grammar on historical principles*, vol. 6: *Morphology*. Copenhagen: Munksgaard.

Kučera, Henry, and Nelson Francis. 1967. *Computational analysis of present-day American English*. Providence, R.I.: Brown University Press.

Kuczaj, Stan A. 1977. The acquisition of regular and irregular past tense forms. *Journal of Verbal Learning and Verbal Behavior* 16, 589–600.

———. 1978. Children's judgments of grammatical and ungrammatical irregular past tense verbs. *Child Development* 49, 319–326.

MacKay, Donald C. 1976. On the retrieval and lexical structure of verbs. *Journal of Verbal Learning and Verbal Behavior* 15, 169–182.

MacWhinney, Brian. 1978. *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development, no. 174, vol. 43.

Menn, Lise, and Brian MacWhinney. 1984. The repeated morph constraint: Toward an explanation. *Language* 60(3), 519–541.

Rosch, Eleanor, and Carolyn B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7, 573–605.

Slobin, Dan I. 1971. On the learning of morphological rules: A reply to Palermo and Eberhart. In *The ontogenesis of grammar*, ed. Dan I. Slobin, 215–223. New York: Academic Press.

Stanners, R. F., J. J. Neiser, W. P. Hernon, and R. Hall. 1979. Memory representation for morphologically related words. *Journal of Verbal Learning and Verbal Behavior* 18, 399–412.

Stemberger, Joseph P. 1981. *Morphological haplology*. *Language* 57, 791–817.

Vennemann, Theo. 1971. Natural generative phonology. Paper read at LSA Annual Meeting, St. Louis, Mo.

———. 1972. Rule inversion. *Lingua* 29, 209–242.

Zager, David. 1980. A real-time process model of morphological change. Ph.D. diss., SUNY at Buffalo.

Chapter 6

WITH CAROL LYNN MODER

Morphological Classes as Natural Categories

Recent theories of phonology make a distinction between alternations that are phonetically conditioned and those that occur in environments identified in morphological, lexical, or syntactic terms (Vennemann 1971; Stampe 1973; Skousen 1975; Hooper 1976). The latter type of alternation may occur in conjunction with some morphological category, for example, plural or past tense; or it may be restricted to certain classes of lexical items—either large meaningful classes (e.g., noun and verb) or smaller, more arbitrary classes (e.g., English verbs that undergo vowel changes in the past tense and past participle). The two types of alternations or rules have different properties; they undergo different types of change; they are acquired in different ways; and they have different motivations and hence explanations.

Phonetically conditioned processes are often called "natural" because they recur in the languages of the world and because they seem to have a phonetic (usually articulatory) teleology. It is not so obvious that morphologically conditioned rules are "natural" in a comparable sense. They often seem arbitrary from a synchronic point of view, because they are usually the residue of older, phonetically conditioned processes. Furthermore, such alternations are often unproductive and scheduled for leveling, since they disrupt the one-to-one correspondence between sound and meaning. However, if they were entirely arbitrary and without function, it would be difficult to explain why such irregularities are preserved so well and so long in many languages. Furthermore, the changes that occur in these alternations are not random but patterned. It seems possible that a theory of morphophonemics could be developed in which speakers' treatment of morphophonemic alternations in certain dynamic situations—for example, in the acquisition process, under conditions of diachronic change, or under experimental situations—could be explained by reference to psychological principles, which would correspond to the phonetic principles that explain natural phonological processes. Some principles in a theory of morphophonemics, especially concerning how the structure of paradigms affects morphophonemic alternations, have been discussed in Hooper (1979) and in Bybee and Brewer (1980). We propose to study here classes of lexical items that exhibit the same morphophonemic behavior, such as the English strong-verb classes exemplified by *sing/sang/sung* and *string/strung*. In particular, we want to know how such classes are defined by speakers—that is, what criteria they use in deciding that a set of verbs forms a class. We want to know what determines the productivity of such a class—that is, why some classes become productive and attract new members, while others gradually disappear as their members are regularized. Finally, we want to know what determines the direction in which any of these classes expands.

1. Two related classes of English verbs

The English verbs with which we will deal are listed in table 6.1, divided into two classes and organized according to the final consonant of the base. The distinction between the two classes is not entirely discrete, since the historical trend is for members of the class with three forms (*sing/sang/sung*) to lose their separate past-tense forms and thus become members of the *string/strung* class. Many American speakers do not use the forms *sprang*, *shrank*, and *stank*. The forms marked with one asterisk

TABLE 6.1. English Strong Verbs with Past or Participle in /ʌ/

	Class I				*Class II*	
Final				*Final*		
/m/	swim	swam	swum	/n/	spin	spun
	come	came	come		win	won
/n/	begin	began	begun	/ŋ/	cling	clung
	run	ran	run		fling	flung*
/ŋ/	ring	rang	rung*		sling	slung*
	sing	sang	sung		sting	stung*
	spring	sprang	sprung		string	strung*
/ŋk/	drink	drank	drunk		swing	swung
	shrink	shrank	shrunk		wring	wrung
	sink	sank	sunk		hang	hung*
	stink	stank	stunk		bring	brung**
				/ŋk/	slink	slunk
				/k/	stick	stuck*
					strike	struck*
					sneak	snuck**
					shake	shuck**
				/g/	dig	dug*
					drag	drug**

in these lists belong to verbs that were not in the strong class in Old English but have been attracted into it since the OE period.[1] We have also included certain dialectal forms that also were not members of the OE strong class, such as *shuck* and *drug*; such forms are marked with two asterisks.

The OE strong verbs from which these two classes developed had a three-way alternation: /ɪ/ occurred in the infinitive and present tense, /a/ in the 1s and 3s preterit, and /ʊ/ in the 2s and the plural of the preterit and in the past participle. When the person and number distinctions were lost, the stem-vowel alternation in the preterit was resolved in favor of /a/ (which later became /æ/) in some verbs and /ʊ/ (which later became /ʌ/) in others; thus the two classes of table 6.1 were created. Subsequently, these verbs have also participated in the general trend toward the elimination of the distinction between the past and participle forms, In the verbs of Class I, this is accomplished by substituting the past-participle form for the past; the result is a verb following the pattern of Class II: thus *spin/span/spun* becomes *spin/spun.*

Of all the strong-verb classes that survive in Modern English, Class II of table 6.1 is the most productive. Class I is also of interest, because it is closely related to the other class and because it shows some limited productivity. The evidence for the productivity of the *string/strung* class is the large number of new members that have been attracted to it over the centuries (Jespersen 1942): twelve of the eighteen verbs in Class II were not members of the strong class in Old English. This demonstrates relatively higher productivity than the *sing/sang/sung* class, since only one member of the latter class has been added since the OE period.[2]

Evidence for the current productivity of these two classes is found in the experiments reported in Bybee and Slobin (1982), in which innovations in English past-tense forms were elicited from adults and from children aged 8½ to 10 years. The experiment involving adults was designed to get them to produce innovative verb forms—especially regularizations, that is, *-ed* forms for irregular or strong verbs. The subjects, who were all native speakers of English, were asked to listen to a list of verbs read by the experimenter at a fast pace. The subjects were asked to supply the past-tense form of the verb they heard and to do so as quickly as possible. There were forty subjects, and each one heard half of a list consisting of 90 irregular or strong verbs, interspersed with approximately 270 weak verbs. In addition to the regularizations elicited in this way, some innovations in vowel change also occurred. In all, eighty-five innovations involved vowel change; of these, forty-one involved verbs of the two classes under discussion here. The experiment with the fifteen children, who will be referred to as "third-graders," was somewhat different, in that no pressure was put on them to perform quickly; only the 90 irregular verbs were given to them, and these were presented in a sentence-completion frame. This test was designed to elicit forms close to those that the children actually used. Most of the innovations produced were regularizations; but in addition, the children produced thirty-eight innovations involving vowel change, and thirty-two of these involved the verbs of the classes under discussion here.

These innovations suggest that something is going on in these verb classes. An analysis of the particular innovations recorded reveals the productivity of the *string/strung* class.[3] Three types of vowel change were found:

1. The most frequent was the use of a past form in /ʌ/ for a verb in Class I. Such responses are counted as innovations for the purposes of our discussion, even though they may regularly occur in the subjects' spontaneous speech. We take the forms in table 6.1 to represent a conservative norm, and we want to find out what changes are currently in progress among these verbs. The responses of this type are listed in section A of table 6.2. The responses listed are percentages, out of a possible twenty responses to each verb for the adults and a possible fifteen responses to each verb for the third-graders. These six verbs are the only ones of this class that were given on the tests. The experimental results reflect a clear historical trend toward substituting the past-participle form in /ʌ/ for the past form in /æ/.
2. A second type of innovation involved using a past form in /ʌ/ for a verb that ordinarily has some other type of past tense. These responses are listed in section B of table 6.2. Some of these forms (e.g., *brung*, *thunk*, and *shuck*) are established dialectal forms, while others are more innovative. But they all show that the *string/strung* class of verbs is relatively productive, especially considering that the maximum number of innovations on the pattern of any other strong-verb class was only four.
3. Examples of the third type of innovation were somewhat smaller in number: they involved using a past tense in /æ/ for verbs with some other past tense—specifically, a past tense in /ʌ/ for all the examples except *clink* (see section C of table 6.2). Innovations of this type were

TABLE 6.2. Elicited Past-Tense Innovations in Two Age-Groups: Percentage of Total Responses

			Adults	*Third-graders (%)*
A	swim	swum	25	7
	begin	begun	10	13
	ring	rung	25	40
	drink	drunk	10	40
	shrink	shrunk	25	60
	sink	sunk	35	[a]
B	bring	brung	5	13
	think	thunk	0	7
	clink	clunk	5	[a]
	streak	struck	20	[a]
	eat	ut	0	7
	shake	shuck	0	13
C	string	strang	10	7
	sting	stang	10	7
	slink	slank	5	[a]
	clink	clank	5	[a]
		clanked	15	

[a]Not on third-graders' test.

less frequent, especially among the third-graders. Still, some productivity for this type of past tense is indicated.

If the strong-verb class represented as Class II in table 6.1 is more productive than any others, as the evidence indicates, then two questions arise. First, what are the defining properties that allow speakers to decide that certain verbs belong in these groups, while others do not? Second, why is this class productive, while others are not?

2. Product-oriented schemas

It is apparent that the two classes of verbs in table 6.1 are defined on phonological criteria; syntactic or semantic criteria seem to play no role in defining them. The most obvious phonological property shared by all the verbs concerns the final consonant or consonants of the base. The verbs of Class I all end in a nasal or a nasal followed by a velar; those in II all end in a nasal or a velar. It seems reasonable to expect that this definition would include information about the vowel of the base form; however, although the vowel is usually /ɪ/, it is not always, for example, *hang* and *strike* from Class II. It is possible, however, to make a valid generalization about this class of verbs, if this generalization describes the past and participle forms, rather than the base form. Such a generalization would be *product-oriented*.

Zager (1980) distinguishes two types of morphological modifications (where "modifications" are roughly equivalent to innovations). *Source-oriented* modifications take a source word (a base) and modify it in a very specific way to produce a new, derived form. Thus a source word X undergoes some process Y to produce a form Z, and the process Y is well-defined. The regular English past-tense suffixation process is an example: a verb *walk* has the suffix /d/ added to it, and the result is a form *walked*. The relevant relation in this case is between *walk* and *walked*, and it is "basic-derived" (cf. Bybee and Brewer 1980). In a "product-oriented" modification, the process Y by which the new word is formed is not well-defined, nor is the shape of the source word X necessarily well-defined. However, the product, Z, is well-defined. In our example here, the past form for a verb of the *string/strung* type must end in /ʌ/ followed by a nasal or a velar; but the vowel of the base does not necessarily have to be /ɪ/. In this case, the relevant relations are among the past forms of the different verbs (*strung*, *slung*, *swung*, *wrung*, *hung*, etc.), rather than those between base and derived forms. This is not to deny the existence of relations between *string* and *strung*, *hang* and *hung*—but only to assert that in adding new members to the class, the focus is on the properties of the product (or past-tense forms) and the relevant relations are among these product forms.

Note that the postulation of product-oriented modifications allows for morphological innovations that would be impossible if proportional analogy were the only means of creating new forms, because innovations such as *strike/struck* and *sneak/snuck* have no pre-existing model with appropriate vowels in the base form.

Bybee and Slobin (1982) propose that the synchronic counterpart of a product-oriented modification is a "schema," that is, a generalization about the shape of a lexical item of a certain category. Schemas are not used in generating forms by combination but rather are lexical associations among existing forms. Bybee and Slobin report evidence that English irregular verbs have their past-tense forms listed in the lexicon and that any associations among them are handled by schemas. In the experimental task for adults described earlier, forty-four vocalic innovations were found in addition to the ones we have already described. Of these, only *one* response was not a real English word, only *five* were not real English verbs, and thirty-five were real past-tense forms of other verbs. This indicates that producing a past form of an irregular verb is not a matter of applying a feature-changing rule to an underlying form but is rather a matter of lexical selection. If the innovations that the subjects produced were the result of the misapplication of vowel-change rules, then the results would contain many nonsense words, such as the one example *heap/hept*. Since the responses were not nonsense but real past-tense forms of other verbs (for the most part), it appears that producing an irregular or strong past-tense form in English is a lexical selection procedure. Since Bybee and Slobin also found evidence of classhood in a few cases—especially in the verb classes under discussion here—they propose that schemas define these classes, and aid in the lexical selection process.

3. Defining the classes

The schema for Class I says that the past tense has the vowel /æ/ and ends in a nasal, followed by an optional /k/:

(1) $\ldots$ æ nasal (k)$]_{\substack{\text{verb}\\ \text{past}}}$

The schema for the verbs in Class II states that the vowel of the past tense is /ʌ/ and that the final consonant is either a nasal or a velar:

(2) $\ldots$ ʌ nasal or velar$]_{\substack{\text{verb}\\ \text{past}}}$

This latter description is unsatisfactory for several reasons. First, the disjunction "nasal or velar" is merely a list of the possible finals. Second, the list does not provide a clue as to the reason for this particular selection of final consonants. Third, it neither describes nor explains the fact that the largest number of verbs in this class actually have final consonants with *both* properties, that is, velar nasals. Finally, natural classes in phonology are not usually of this type; that is, they are not usually describable as disjunctions of two features where the usual case is for the two features to occur together.

However, there are other categories that work in precisely this way—namely, the "natural categories" that people use when they label natural or cultural objects. Wittgenstein (1953) proposed that the categories that we create in construing our universe do not consist of sets of items sharing a common set of attributes not shared by other items but rather that members of categories form chains of family resemblances. That is, all that is necessary is that each member of a category share one attribute with some other member. Rosch and Mervis (1975) and Rosch (1978) develop this idea further by proposing that each category has a *prototype* or best exemplar, which has a cluster of attributes characteristic of the category. Thus there can be both *prototypical* members of categories, with a number of relevant attributes, and more marginal members, with perhaps only one attribute that puts them in the category. For example, a wooden, straight-backed chair is a good exemplar of the category "chair," since it has an identifiable seat, back, and four legs and is an object designed to be sat upon; but a bean-bag chair is a poor or marginal exemplar, since the structural attributes of chair-hood are lacking. Nonetheless, it is a chair, since its function is to be sat upon, an attribute shared with other members of the category. Zager (1980) suggests that morphological classes may have the structure of natural categories in this sense.

Indeed, the verbs of the *string/strung* class seem to form a class with the structure of a natural category. The prototypical verb in this class has a velar nasal as a final consonant. The other verbs share at least one property with this prototype, so far as the final consonant is concerned; that is, they end either in a velar or in a nasal. Thus even though the final consonants of *won* and *dug* have no attributes in common, they each share one attribute with the prototype, which is enough to determine class membership.

The evidence that the structure of this class is not just a historical accident but is indeed a real, psychologically valid structure is quite strong. In looking at Class II again, it can be seen that all the original members of the class contained a nasal consonant and that all the members with a non-nasal velar are 'new' members of the class. This means that, as the class has expanded, it has created the particular structure we see today. It has expanded outward, allowing verbs that share only the property of velarness to join the class. This is strong evidence that its family-resemblance structure is psychologically valid. Further evidence for the psychological validity of this type of structuring can be obtained experimentally. We describe later the results of an experiment designed to test our hypothesis about the structure of this verb class.

4. The experiment

The purpose of the experiment was twofold: (1) to determine the defining attributes of the most productive strong-verb class (Class II) and (2) to test the hypothesis that this class has the structure of a natural category.

We hypothesized that this class of verbs is organized around a prototype, such that each member of the class shares one or more attributes with the prototype.[4] We arrived at a working hypothesis about the prototype by examining the various members of the class. As mentioned earlier, the final consonant(s) of the prototype would contain both the features nasal and velar and thus would be either /ŋ/ or /ŋk/. Since there are many more verbs with a final /ŋ/, we might furthermore surmise that /ŋ/ would make a verb closer to the prototype than /ŋk/. The vowel of the prototype base-form is hypothesized to be /ɪ/. Recall, however, that we are suggesting that the prototype is product oriented; it defines the past form, not a base form. Thus we are interested in the importance of the vowel /ɪ/ in determining closeness to the prototype.

The list of Class II verbs in table 6.1 also suggests that the initial consonants of the verb may serve as attributes that define the prototype: it is striking that nine of the eighteen verbs begin with /s/ followed by one or two consonants. We hypothesized, then, that the prototype would begin with /s/ plus one or more consonants, so that a form like *string* would be a likely prototype for the past tense of this class of verbs.

In the experiment, we asked subjects to form the past tense of nonce verbs that were constructed phonologically so as to stand in varying relations to the prototype. We expected the subjects to give more past-tense responses with /ʌ/ for nonce verbs that were closer to the prototype than for nonce verbs that were further from it. We turn now to an outline of this experiment, and the results we obtained from it.

4.1. Subjects

The subjects of the experiment were thirty-four undergraduate students enrolled in an introductory linguistics course at SUNY at Buffalo. All subjects were native speakers of English. Neither the hypothesis nor the verb class had been discussed in the course.

4.2. Method

A list of ninety-three nonce verbs interspersed with sixteen real verbs was presented orally. These verbs were all monosyllabic and varied phonologically along three dimensions: the initial consonants, the vowel, and the final consonants. Initial-consonant segments were of four types: the consonant /s/ followed by a consonant cluster (*sCC*); the consonant /*s*/ followed by a single consonant (*sC*); a two-consonant cluster not containing /s/ (*CC*); and a single consonant (*C*). A variety of vowels were represented, including /ɪ æ ʌ ey iy ay uw ow/. Final-consonant segments fell into four categories: velar and nasal (/ŋk ŋ/); velar (/k g/); nasal (/m n/); other single consonants (*C*). A list of verbs appears in the appendix to this chapter.

Subjects were asked to write the past-tense forms of the verbs presented. At the beginning of the experiment, the following sentence-completion frame was given: *Sam likes to . . . Yesterday he . . .* Subjects were asked to refer to this frame as they felt necessary during the course of the experiment.

4.3. Results

The three phonological parameters—the initial, the vowel, and the final segments—were all significant factors affecting a subject's response. The responses can be divided into two types: (1) regular past tense formed by adding *-ed* and (2) past tense formed by vowel change. In a very small number of responses, the subject changed the vowel and added *-ed*. These did not figure in our analysis. In forming the past by vowel change alone, there was an overwhelming tendency for subjects to use /ʌ/ and /æ/; we will therefore give the percentage of responses with *-ed* and with these two vowels.[5] The overall percentages of vowels which occurred in past-tense responses formed by vowel change was as follows:

(3) /ʌ/ 49%
/æ/ 34%
other 17%

To test the effect of the vowel of the base in eliciting an /ʌ/ vowel in nonce verbs, we constructed such verbs with a variety of vowels. However, our input items were limited by the fact that the number of vowels that can occur before a velar nasal is severely restricted in English. Thus, for the nonce verbs ending in a velar nasal, we present the results for base verbs with the vowels /ɪ æ ʌ/ in table 6.3.

In table 6.4, we present the results for nonce verbs ending in a velar, including verbs containing diphthongs. It can be seen that the highest number of /ʌ/ responses was obtained for nonce verbs that contained /ɪ/ but were otherwise close to the prototype. All the diphthongs tested showed a strong tendency to elicit regular past tenses and a weaker tendency to make vowel changes other than /æ/ or /ʌ/ in the past tense. Since their response patterns did not differ significantly, they are grouped together in the analysis.

The tendency (mentioned earlier) for subjects to use either /æ/ or /ʌ/ in forming the past by vowel change had an effect on the response patterns for verbs with /æ/ or /ʌ/ as input vowels. In general, subjects avoided no-change responses that would make the past tense identical to the present. Thus, for input verbs like /stræŋk/, which contains one of the favored past-tense vowels, subjects rarely gave /stræŋk/ as the past tense. In most cases, this avoidance of a same-vowel response led to an increase in responses with the other favored vowel. For verbs with /æ/, the percentage of /ʌ/ often increased relative to other verbs; for verbs with /ʌ/, the percentage of /æ/ responses was often greater, as was the percentage of regularizations. In spite of this elevation, the percentage of /ʌ/ responses remained greatest for verbs with /ɪ/ as input vowel.

TABLE 6.3. Past-Tense Responses for Nonce Verbs of the Shape *sCC*____Velar Nasal

		-ed (%)	/æ/ (%)	/ʌ/ (%)
Input vowel	/t/	18	34	44
	/ʌ/	84	8	4
	/æ/	67	8	25

TABLE 6.4. Past-Tense Responses for Nonce Verbs of the Shape *sCC*___ Velar

		-ed (%)	/æ/ (%)	/ʌ/ (%)
Input vowel	/I/	56	13	25
	/ʌ/	74	3	0
	/æ/	100	0	0
	diphthong	53	3	15

TABLE 6.5. Past-Tense Responses for Nonce Verbs by Initial Segments

		-ed (%)	/æ/ (%)	/ʌ/ (%)
Input initials	*sCC*	42	23	23
	sC	48	21	23
	CC	63	14	12
	C	56	23	14

The vowel effects may be summarized in the following way: /ɪ/ elicits a significantly higher percentage of /ʌ/ responses, overall, than other vowels; it is followed by /æ/, the diphthongs, and /ʌ/.

As regards initial consonants, nonce verbs of the form *#sCC* and *#sC* were significantly better at eliciting /ʌ/ in the past tense than those with *#CC* or *#C*, as shown in table 6.5.

There was no overall significant difference between *#sCC* and *#sC*; but among the verbs close to the prototype, *#sCC* was significantly better, as shown in table 6.6. No difference was found between consonant clusters within a group.

Among the final segments, the ones that had both the features velar and nasal elicited the highest percentage of /ʌ/ responses, as shown in table 6.7.

Within this group, /ŋ/ was significantly better than /ŋk/, as shown in table 6.8.

Velars and nasals elicited approximately equal percentages of /ʌ/ vowels. However, verbs with final nasals had a higher percentage of /æ/ responses than /ʌ/ responses, while the reverse was true for velars. Single consonants were the poorest at eliciting /ʌ/ responses, but like the nasals, they had a high percentage of /æ/ responses. Within this group, final /t/ or /d/ elicited more vowel-change responses than other finals, as shown in table 6.9. No other differences within groups were found.

Of the three parameters, the final segment had the greatest effect on the number of /ʌ/ responses. When no other attributes of the prototype were present, the presence of the best final (a velar nasal) elicited 34%, the best initial (*#sCC*) elicited 17%, and the best vowel (/ɪ/) elicited 7%. Combining the best final and initial features with a non-prototypic vowel (/æ/) results in 50% /ʌ/ vowels. This percentage does not differ significantly from the percentage of /ʌ/ responses at the prototype, which is 44%.

TABLE 6.6. Responses for Verbs of the Form _______ ɪ + Velar Nasal by Initial Segment

		-ed (%)	/æ/ (%)	/ʌ/ (%)
Input initials	*sCC*	18	34	44
	sC	34	26	37
	CC	42	26	27
	C	40	32	22

TABLE 6.7. Overall Responses by Final Segment

		-ed (%)	/æ/ (%)	/ʌ/ (%)
Input finals	/ŋ nk/	33	29	32
	/k g/	67	10	14
	/n m/	50	21	15
	C	61	19	10

(The percentage at the prototype is slightly lower in this case, because of the unique response features of the verbs with /æ/ as an input verb; see earlier.) This is further evidence for the minimal effect of the vowel alone. The absence of the prototypic initial segment causes a significant decrease in /ʌ/ responses: the best final and vowel, with the poorest initial (*C* + ɪ + velar nasal), elicited only 24%. A still greater decrease was caused by the absence of the final: the best initial and vowel (*sCC* + ɪ) received 4% /ʌ/ responses; see tables 6.10 and 6.11. We conclude that the final velar nasal is the strongest attribute, followed by the initial. The vowel makes a significant difference only when other features of the prototype are present.

On the basis of the experiment, the attributes of the prototype of this class of verbs are:

(4) a. a final velar nasal (/ŋ/ is better than /ŋk/)
b. an initial consonant cluster that begins with /s/
c. a vowel /ɪ/, which has an effect only in conjunction with the preceding elements

The strength of the final and initial parameters may be classified on the basis of their similarity to the prototype. The best final has the features of velar and nasal. The verbs with velars and those with nasals, each of which contains one of these features, can both be considered one step away from the prototype. The single-consonant set bears neither of these features and so is further away from the prototype. The prototypic initial is an /s/ followed by a consonant or cluster. If /s/ + cluster is the prototype, we can say that the *#sC* initial and the *#CC* initial are each only one step away from the prototype. The data make it clear, however, that the /s/ is a more important attribute than the consonant cluster. The single consonant lacks two elements, the /s/ and the cluster; therefore, it is furthest from the prototype. Within this

TABLE 6.8. Overall Responses of /ŋ/ vs. /ŋk/

		-ed (%)	/æ/ (%)	/ʌ/ (%)
Input finals	/ŋ/	29	29	37
	/nk/	38	29	27

TABLE 6.9. Overall Percentages for Final /t d/ vs. /p b v/

		-ed (%)	/æ/ (%)	/ʌ/ (%)
Input finals	/t d/	40	31	18
	/p b v/	82	8	2

ranking, we can compare distance from the prototype with the percentage of /ʌ/ responses elicited, as in table 6.11.

Within each parameter, we find a linear decrease in /ʌ/ responses as the verb's phonological properties are removed from the prototype. We find a similar progression across parameters. As the number of prototypic attributes present decreases, so does the percentage of /ʌ/ responses. The more a set of features within or across parameters diverges from the prototype, the fewer /ʌ/ responses it elicits. This progression is evidence for the operation of a morphological prototype in this verb class in English.

Table 6.12 shows the responses for the real English verbs that were included in the experiment. It can be seen that the subjects gave standard past-tense forms for the large majority. The overall percentage of correct forms was 83%: this indicates that the subjects took the task seriously and were attending to input material sufficiently to determine whether or not the test item was a real verb. The subjects performed better on the ten regular verbs tested (87% correct). Of these verbs, seven verbs received 88% or better correct responses (94% on the average) and the remaining three averaged 72% correct. The verbs *streak* and *stroke* elicited a greater percentage of vowel-change responses than other regular verbs, favoring /ʌ/ (17%). These verbs are also those that share the greatest number of attributes with the prototype of the class we have been discussing. Thus, even for regular verbs, the prototype of this productive class makes its influence felt. The irregular verbs with /ʌ/ in the standard past tense elicited an average of 73% correct responses. The remaining responses were divided between *-ed* regularizations and other vowel changes.

TABLE 6.10. Effect of the Three Parameters

	Responses in /ʌ/ (%)
Prototype: *sCC ɪ* velar nasal	44
Two prototype features present	
Final & initial: *sCC æ* velar nasal	50
Final & vowel: *C ɪ* velar nasal	24
Initial & vowel: *sCC ɪ C*	4
One prototype feature present	
Final: *C æ* velar nasal	34
Initial: *sCC æ C*	17
Vowel: *C ɪ C*	7

TABLE 6.11. Progression Away from Prototype of Initial and Final Segments, Other Parameters Held Constant

		Responses with /ʌ/ (%)
______ɪ + velar nasal		
Initials	*sCC*	44
	sC	37
	CC	27
	C	22
sCCI ____		
Finals	*ŋ, nk*	44
	k, g	25
	n, m	21
	C	4

5. Discussion

The experimental data clearly point to a morphological class consisting of verbs bearing a family resemblance to a prototype. On the one hand, the data show that the verbs that can form a past tense by using the vowel /ʌ/ are not a random set, but that the likelihood of a verb forming a past tense in this way correlates with its phonological similarity to a prototype. On the other hand, we have also demonstrated that this class of verbs does not have discrete boundaries, or phonological features whose presence is absolutely required, but rather consists of both central and marginal members. The latter point is important because it helps us understand how such a class can expand and grow. It explains why a class of verbs ending only in nasals can expand to include items such as *strike/struck*, and *dig/dug*. It explains occasional errors such as *eat/ut* and usage such as *slide/slud* (this latter form is attributed to Dizzy Dean, a baseball celebrity; Madeleine Newfield, p.c.). Membership in morphological classes is not a matter of strict presence or absence

TABLE 6.12. Past-Tense Responses on Real Verbs

	Regular Verbs (%)				*Irregular Verbs (%)*		
	-ed	/æ/	/ʌ/		*-ed*	/æ/	/ʌ/
streak	69	7	14	slink	39	16	42
stroke	77	0	19	bid	70	7	7
link	71	18	7	sting	6	31	63
stack	92	0	4	stick	6	13	72
skim	88	6	3	fling	12.5	12.5	75
skid	90	0	3	dig	12	6	82
trick	97	3	0				
trip	94	3	0				
clip	96	4	0				
tip	100	0	0				

of features, but rather of similarity to a prototype, which may be defined on a number of features. Given the findings reported in section 4, we can predict that a particular verb will join the *string/strung* class on the basis of the number of features it shares with the prototype and the ranking of these features. We can also predict that certain verbs are unlikely to join the class (i.e., those with no features of the prototype, such as *rob* or *type*) and that others have a very low probability of joining the class, since they lack a final consonant that shares features with the prototype (e.g., *stop*, *scratch*, or *strip*).

The experiment demonstrates that, at least to some degree, the initial consonants, the vowel, and the final consonants determine the likelihood of membership in this verb class. The quality of the vowel is the weakest determinant, as we expected, since the class is actually defined by a product-oriented generalization that refers to the past form, not to the base. The existence of new members of the class (e.g., *strike*, *hang*, *drag*, *sneak*) led us to that conclusion, which is confirmed by the experiment. The minor importance of the vowel argues against a description of past-tense formation for these verbs as a process by which a particular vowel /ɪ/ is changed into another vowel, as proposed in Chomsky and Halle (1968) and in Hoard and Sloat (1973). It also argues against the view that new verbs entering the class do so on the basis of strict proportional analogy: thus, before *strike* joined the class, no verb with /ay/ in the base form was available to provide forms for the first two members of the proportion.

Our experiment yields further evidence in support of the product-oriented schema. In some cases, subjects substituted final consonants in the test words that were different from the consonants in the input items. These errors most frequently involved the production of a past-tense form with more attributes of the prototype than the original test word. For example, some subjects responded to the nonce form *spriv* by producing a past tense *sprung* or *sprug*. This pattern of producing a past-tense form with the vowels /ʌ/ or /æ/ and a final consonant closer to the prototype occurred as the most frequent error pattern involving consonants for all the nonce verbs presented (see table 6.13). In these cases, subjects are clearly operating with a product-oriented strategy: they are changing the final consonant as well as the vowel, in order to make the input verb conform to the past-tense schema.

The strongest determinant of class membership, as we predicted, is the final consonant of the stem. This was to be expected, because the existing class of verbs shows

TABLE 6.13. Most Frequent Error Patterns in Final Consonants

Input	*Replaced by*	*Errors*	*Example*
/-ŋk/	/ŋ/	58	spink: sprung
/-g/	/ŋ/	65	trig: trung
/-k/	/ŋk/	60	strick: strunk
/-m/	/n/	82	spim: spun
/-n/	/ŋ/	76	vin: vung
/-C/	/k g/	43	smip: smuk stid: slug

more coherence on this point than on any other. Why should the final consonant(s) of the verb define the class more strongly than the initials? The finals with the characteristic /ʌ/ vowel could make up a unit, which in itself is a marker of past tense. Phonologically, the vowel of a syllable forms a more cohesive unit with the consonants of the coda than with the consonants of the onset (cf. MacKay 1972). Furthermore, in English the inflectional morphology is suffixing—suggesting that, in general, inflectional meaning will be expressed at the end of the word. Indeed, Slobin (1973) has suggested a language-acquisition strategy that directs children to the ends of words to find grammatical material; this strategy may be related to the universal preference for suffixing of inflectional material. It would be interesting to determine whether, in prefixing languages, morphological classes are defined on the beginnings of words.

Given these considerations, the influence of initial consonants on class membership is somewhat surprising. However, it is paralleled somewhat by the findings of Zubin and Koepcke (1981) on gender in German. These authors report that initial consonant clusters in monosyllabic nouns show a significant correlation with the assignment of nouns to the masculine gender. Thus, taking native nouns whose gender is not semantically determined, 73% of those with a two-consonant initial have masculine gender and 82% of those with a three-consonant initial have masculine gender.

What our findings regarding the attributes of this verb class indicate, in general, is that English morphology may be a kind of "whole word" morphology, in which both meaningful and classificatory information is distributed over the whole word. The decisions that our subjects made about how to form the past tense of a particular nonce verb may have been based on how the whole resulting word would sound. Evaluations of how the word sounds may be based on comparisons with whole existing words.

We find in general, then, that the experimental results corroborate the hypotheses that we reached after examination of the synchronic and diachronic data. On one issue, however, the experimental data afford a somewhat different view than the diachronic, namely, the relative productivity of the class with regard to verbs ending in /n/ versus verbs ending in /k g/. While each of these consonants shares a feature with the prototypic final, the diachronic evidence shows that the class is expanding in the direction of /k g/ and not toward /n/. That is, all the verbs ending in /k g/ are new members of the class, while the two verbs ending in /n/ (*win* and *spin*) are original members. However, our experimental results show that subjects are just as likely to assign a nonce verb ending in /n/ to this class as they are to assign a verb ending in a velar to it. Thus we must ask why the class has not added new members with final /n/ over the last few centuries. A possible factor is the low availability of verbs ending in /-ɪn/ as possible candidates for inclusion. A search of the reverse word list in Dolby and Resnikoff (1967) turns up only *spin* and *win* as monosyllabic verbs ending in /-ɪn/, plus a few denominative verbs: *skin*, *grin*, and *pin*. Perhaps there simply have not been many opportunities for verbs in /n/ to be added to the class.

Another area in which the experimental results diverge somewhat from the diachronic patterns is in the use of /æ/ as a vowel in the past tense. An examination of Class I, earlier, shows that only one non-original verb (*ring*) exists in this class.

We took this to mean that this class was not productive. In view of this, it is somewhat surprising that our subjects used /æ/ in 34% of the responses in which they made a vowel change. Table 6.14 displays the percentage of /æ/ responses for nonce verbs ending in each of the final consonants tested. These figures indicate some productivity for Class I verbs, since the /æ/ responses occur in nonce items ending in nasals and velar nasals, which are the consonants occurring on the real verbs of this class.

Note also that the largest percentage of /æ/ responses occurs in nonce items ending in a dental consonant. This is an even more surprising result, since English has so few verbs in English that end in /t d/ and whose past tense is formed with /æ/. *Sit*, *sat* is the only fully standard example, since *spit* and *shit* both have unmarked past forms in addition to forms *spat* and *shat*. We do not think that *sit* forms a productive class by itself but rather that another phenomenon accounts for the results. Looking again at table 6.9, we see that the percentage of regularizations (*-ed* responses) for verbs ending in /t d/ is quite low. In fact, if we compare the percentage of *-ed* responses in tables 6.7 and 6.9, we see that only items ending in the velar nasal have a lower percentage of *-ed* responses. This results, we hypothesize, from the avoidance of adding the dental suffix *-ed* to a word that ends in a dental. Such an avoidance has long been noted among children. who regularize irregular verbs ending in /t d/ much less often than verbs ending in other segments (Slobin 1971; Kuczaj 1978). Furthermore, final consonants in English vowel-change verbs show an asymmetrical distribution such that approximately 31% end in /t d/, 28% belong to either Classes I or II, 13% end in diphthongs, and the remaining 28% are distributed over eight different consonants, /r l n č z v k p/. In other words, a striking concentration of vowel-change verbs occurs among those ending in /t d/. We hypothesize that speakers tend to take verbs that end in /t d/ as irregular in some respect; thus our subjects supplied nonce verbs ending in these consonants with vowel changes. If a general prototype exists for a vowel-change verb in English, then it might end in /t d/. We do not know why so many responses involved /æ/, but this could result from the input vowel /ɪ/, which suggests a reference to *sit*, or from the fact that most of the items on the test resembled verbs that could have either /æ/ or /ʌ/ in their past tenses.

Turning again to the historical development of strong verbs in English, we may ask why this particular class of verbs became productive—while the other classes have tended to remain unproductive and to shrink in size. The answer has to do with

TABLE 6.14. Percentage of /æ/ Responses out of Total Responses for Each Final Consonant Type

	Percentage
/ŋ/	29
/ŋk/	29
/n m/	21
/k g/	10
/p b v/	8
/t d/	31

the fact that this class had certain coherent defining phonological features, while the others did not. Verbs of Classes I and II that were strong in Old English belonged to a group whose stem vowel was followed by a nasal or a liquid, plus another consonant (the Class III of Sweet 1882). Many of the verbs in this class that contained a liquid plus another consonant have become weak or obsolete, but many of those with a nasal plus another consonant have survived. The class included *bindan* 'to bind', *findan* 'to find', *grindan* 'to grind', *windan* 'to wind', *drincan* 'to drink', *scrincan* 'to shrink', *swincan* 'to toil', *ge-limpan* 'to happen', *swimman* 'to swim', *winnan* 'to fight', *springan* 'to spring'. No other OE verb class could be so narrowly defined phonologically by features of the stem. Most of them were like the class that included *ride* and *drive*, in that any consonant could follow the stem vowel.

Regular sound changes narrowed the phonological definition of the class even more. OE short vowels were lengthened when followed by voiced homorganic consonant clusters such as /nd/ (Moore 1968). This change affected verbs such as *findan* and *bindan*, while leaving other members of the class untouched. In Middle English, these long vowels were diphthongized, giving /faynd/, /fawnd/, /baynd/, /bawnd/, in effect removing verbs ending in /nd/ from the class of verbs with short /ʊ/. This left an overwhelming predominance of verbs ending in /ŋ/ or /ŋk/. There simply is no other verb class in English with such a coherent set of phonological defining features, and this was a major factor that allowed the class to become productive.

Another factor must also be considered, however—namely, what Rosch terms *cue validity*. A category has high cue validity if the features associated with it frequently occur with members of the category and rarely occur with members of other categories. The sequences /ʌŋ/ and /ʌŋk/ can serve effectively as signals of the past tense so long as these sequences do not, in other items, signal a variety of other meanings. The class of verbs under consideration is favored in this respect also. There are no verbs that, in their present tense or base form, end in /ʌŋ/, and there are very few nouns or adjectives: *tongue*, *lung*, and *young* are the most common examples. Thus /ʌŋ/ is almost uniquely a signal of past tense. The sequence /ʌŋk/ is also highly distinguishable; only two verbs have this sequence in their bases, *flunk* and *dunk*. It seems, then, that this class of verbs has been especially well suited for productivity, having both a coherent phonological definition and a past-tense shape with high cue validity.

Note further that the attributes of the finals found in this class of verbs have higher cue validity than the attributes of initials, since *#sC*(*C*) frequently occurs in words of all types in English. This may account for the fact that the features of the final provide the strongest determinant of class membership.

Since our experiment dealt primarily with nonce verbs, it does not yield data to bear on the issue of how the product-oriented schema, incorporating the prototype, is used in ordinary processing. We can only present a speculative sketch—which, without details, outlines the role we see for schemas in language use.

Schemas should be thought of as associations among lexical items. There are associations on many levels: on the phonological level, words may be associated by initial segment, by rhyme, by stress pattern, or by number of syllables; on the syntactic level, they may be associated by membership in categories such as noun or verb; on the semantic level, they are associated by being similar or opposite in meaning or by belonging in the same semantic field. The morphological schema that we propose associates words on two levels simultaneously: those with a particular phonological shape and those with a particular morphological function. We assume that these lexical associations function in language use to facilitate lexical accessing. We have no evidence that bears on the question of how a schema interacts with the regular suffixation rule for the English past tense, but we assume that because of the identity of function, the schema and the regular past-tense rule are closely associated.

This close association probably accounts for the new words that are formed on the basis of similarity to the prototype expressed in the schema. When a discourse is framed in a past-tense context, the morphological function of past tense is constantly being accessed. We suggest that even if regular verbs are used, the schemas associated with past tense are made available. This availability makes it possible for innovations to be based on the prototype. These innovations, which form new members of the verb class, are like other innovations that result in new lexical items (such as blends), in that they are not the result of combination by rule but are nonetheless made likely by strong lexical associations.

Perhaps the most important finding from this combination of synchronic, diachronic, and experimental data is that speakers of natural language form categorizations of linguistic objects in the same way that they form categorizations of natural and cultural objects. These verbs are not categorized by properties of their meaning or by properties that associate them with external reality; rather, they are categorized by their form, that is, their phonological properties. We do not know how pervasive this type of categorization is in morphological classes. This example is, however, not unique, since Zager (1980) makes a quite convincing case that Old Spanish strong preterites are organized in this same way. These cases suggest that the psychological principles that govern linguistic behavior are the same as those that govern other types of human behavior and that explanations of linguistic structures can be sought outside the linguistic system itself.

Appendix

		Nonce Verbs		*Real Verbs*	
		Class I	*Other*	*Class I*	*Other*
1.	*sCC ɪ ŋ/ŋk*	spling skring sprink strink skrink	strank scronk		
2.	*sCC ɪ k/g*	strick strig	sprayk spreyk sprock sprook streyk strug skreek scrag		streak stroke

		Nonce Verbs		*Real Verbs*	
		Class I	*Other*	*Class I*	*Other*
3.	*sCC ɪ n/m*	skrim	splam		
		sprin	spreen		
4.	*sCC ɪ C*	spriv	sprat		
		skrit			
5.	*sC ɪ ŋ/ŋk*	sming	spang	sting	
		sking		slink	
		smink			
		spink			
6.	*sC ɪ k/g*	smick	steeg	stick	stack
		skig	skeyk		
			stayk		
			spug		
			spock		
			spak		
7.	*sC ɪ n/m*	spim	slan	skim	
		stin	skeen		
			stom		
8.	*sC ɪ C*	stid	stap		
		smip	speeb		
			skeyp		
			skeep		
9.	*CC ɪ ŋ/ŋk*	pling	glank	fling	
		tring			
		krink			
		glink			
10.	*Cc ɪ k/g*	trig	krag	trick	
		glick	treek		
			playk		
			glok		
			trook		
			treyk		
			kreeg		
11.	*CC ɪ n/m*	krin	greem		
		plim	tran		
12.	*CC ɪ C*	trib	treep	trip	
		clid	trad	clip	
13.	C *ɪ ŋ/ŋk*	shink	tang	link	
		tink			
		ping			
		ving			
14.	*C ɪ k/g*	gick	pook	dig	
		sig	deyk		
			gack		
			gok		
			seeg		
15.	*C ɪ n/m*	tim	tam		
		vin	veem		
16.	*C ɪ C*	sid	toop	bid	
		kib	bive	tip	
			peet		
			seyb		
			dop		
			tad		

Notes

We are grateful to Brian MacWhinney and David Zager for discussing earlier versions of this essay. Joan Bybee was supported by a stipend from the National Endowment for the Humanities during the period in which this paper was written.

1. Some of the additions to the class were originally weak, e.g., *dig*, the Old Norse borrowing *fling*, and the denominative *string*. A few others belonged to other OE strong classes, e.g., *strike*, of which the original past form was *stroke* (cf. Jespersen 1942).

2. Of course this productivity is extremely limited compared to the weak-verb pattern of suffixing *-ed*, the productivity of which overshadows that of any strong-verb patterns. However, the fact that so many members of the *string/strung* class are new additions needs to be explained.

3. The productivity of this class is also evident in the work of Berko (1958) and Newfield (1981).

4. In the following, the term *feature* is used in its usual sense of "phonologica1 feature," while *attribute* refers to a particular feature that characterizes the prototype. *Parameter* is used to refer to a dimension of which the attribute is a part. For example, the attribute "velar nasal" belongs to the parameter of final consonants.

5. All statistical differences discussed in the text are significant at the .05 level.

References

Berko, Jean. 1958. The child's learning of English morphology. *Word* 14, 150–177.

Bybee, Joan L., and Mary Alexandria Brewer. 1980. Explanation in morphophonemics: Changes in Provençal and Spanish preterite forms. *Lingua* 52, 201–242. (Reprinted in this volume as chapter 3.)

Bybee, Joan L., and Dan I. Slobin. 1982. Rules and schemas in the development and use of the English past tense. *Language* 58, 265–289. (Reprinted in this volume as chapter 5.)

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Dolby, J. L., and H. L. Resnikoff. 1967. *The English word speculum*. The Hague: Mouton.

Hoard, James, and Clarence Sloat. 1973. English irregular verbs. *Language* 49, 107–120.

Hooper, Joan Bybee. 1976. *Introduction to natural generative phonology*. New York: Academic Press.

———. 1979. Child morphology and morphophonemic change. *Linguistics* 17, 21 –50. [Also in *Historical morphology*, ed. Jacek Fisiak, 157–187. The Hague: Mouton, 1980.]

Jespersen, Otto. 1942. *A modern English grammar on historical principles*, vol. 4: *Morphology*. Copenhagen: Munksgaard.

Kuczaj, Stan A. 1978. Children's judgments of grammatical and ungrammatical irregular past tense verbs. *Child Development* 49, 319–326.

MacKay, Donald. 1972. The structure of words and syllables: Evidence from errors in speech. *Cognitive Psychology* 3, 210–227.

Moore, Samuel. 1968. *Historical outlines of English sounds and inflections*. Revised by Albert H. Marckwardt. Ann Arbor, Mich.: Wahr.

Newfield, Madeleine. 1981. *Options in language: The productive formation of the English verb preterit*. Voorlopige Publikatie 78. Leuven, Belgium: Katholieke Universiteit Leuven.

Rosch, Eleanor. 1978. Principles of categorization. In *Cognition and categorization*, ed. E. Rosch and B. Lloyd, 27–48. Hillsdale, N.J.: Erlbaum.

Rosch, Eleanor, and Carolyn B. Mervis. 1975. Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology* 7, 573–605.

Skousen, Royal. 1975. *Substantive evidence in phonology*. The Hague: Mouton.
Slobin, Dan I. 1971. On the learning of morphological rules: A reply to Palermo and Eberhart. In *The ontogenesis of language*, ed. Dan I. Slobin, 215–223. New York: Academic Press.
———. 1973. Cognitive prerequisites for the development of grammar. In *Studies of child language development*, ed. Charles Ferguson and Dan Slobin, 174–208. New York: Holt, Rinehart and Winston.
Stampe, David. 1973. A dissertation on natural phonology. Doct. diss., University of Chicago.
Sweet, Henry. 1882. *Anglo-Saxon primer*. Oxford: Clarendon Press. (Reprinted, 1965.)
Vennemann, Theo. 1971. Natural generative phonology. Paper read at annual meeting of the LSA, St. Louis, Mo.
Wittgenstein, Ludwig. 1953. *Philosophical investigations*. New York: Macmillan.
Zager, David. 1980. A real-time process model of morphological change. Ph.D. diss., SUNY at Buffalo.
Zubin, David A., and Klaus-Michael Koepcke. 1981. Gender: A less than arbitrary category. *CLS* 17, 439–449.

Chapter 7

WITH JEAN E. NEWMAN

Are Stem Changes as Natural as Affixes?

1. One processing type or two?

Two types of processing have been proposed to account for linguistic behavior associated with morphologically complex words. A traditional generative account of affixation and other morphological processes uses symbolic rules located in a module separate from the lexicon and acts on lexically represented morphemes to create words (Chomsky and Halle 1968). Various recent versions of this model treat morphology as rule derived, analogous to syntax (e.g., Saddock 1985). The other types of processing proposed for morphology are various versions of pattern network associators. Bybee and Slobin (1982) and Bybee and Moder (1983) presented evidence that English irregular past-tense verbs are stored in the lexicon and organized into networks centered around prototypes. Bybee (1985, 1988) proposed that even regular morphological patterns can be accounted for in the same way and that a separate module for morphological derivation was not necessary. The fact that some morphological processes are productive and others are not is a reflex of various factors, but primarily of the type frequency of the particular process, that is, the number of lexical items that participate in the process. A similar claim has been made in the literature on connectionist models that rely exclusively on building up networks of associated patterns based on learned items (Rumelhart and McClelland 1986; MacWhinney et al. 1989). Another recent proposal is that both types of processing are available and utilized and a strict distinction exists between irregular morphological formations, which are stored in the lexicon and organized into networks, and regular morphological processes, which are symbolic rules located in a separate module of the grammar (Pinker 1991; Marcus et al. 1992; Prasada and Pinker 1993). In this proposal, "regular" processes are defined as the productive ones, that is, the morphological processes that are most easily extended to new items.

It is perhaps not accidental that in English and many other languages (including other Germanic languages and Romance languages) the regular processes are affixal (and often agglutinative in their structure) while the irregular ones often involve changes in the stem or a high degree of fusion between stem and affix. It is possible that the substance of the process—stem change versus affixation—is also associated with the processing type. Since affixation bears a greater resemblance to syntactic concatenation than stem change does, it might be more appropriately handled by symbolic rules and stem changes might be more appropriately handled by lexical connections.

A prominent claim made in the theory of natural morphology, as conceived of by Dressler and others, is that among morphological processes, affixation is more "natural" than internal stem changes. In this theory the criteria for "more natural" include the following: (1) more common in the world's languages; (2) more common within a single language; (3) easier to acquire, and thus acquired earlier; (4) later lost in aphasia; and (5) easier to process in general (Dressler 1985). If affixation generally correlates with regularity, its greater "naturalness" could be explained by a difference in processing type: the conclusion would be that processing by symbolic rules is to be preferred over processing by lexical pattern networks. Marcus et al. (1993) describe symbolic rules as concatenating, that is, affixing, rules, which suggests an association of affixation with regularity and stem change with irregularity.

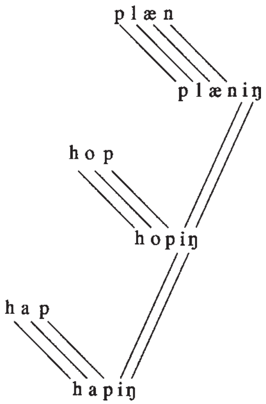
In contrast, a model in which all morphology is handled in the lexicon via associative pattern networks does not distinguish affixation from stem change. Affixes are found in the lexicon attached to words and are analyzed via lexical connections, just as stem changes are. For instance, in the model proposed by Bybee (1985, 1988) a regular suffix is represented by the fragment in (la) and the semiproductive pattern whose prototypical member is *string*, *strung* is partially represented as in (lb) (adapted from Bybee 1988).

If this model is correct, there should be no preference for affixation over stem change from the point of view of acquisition or processing.

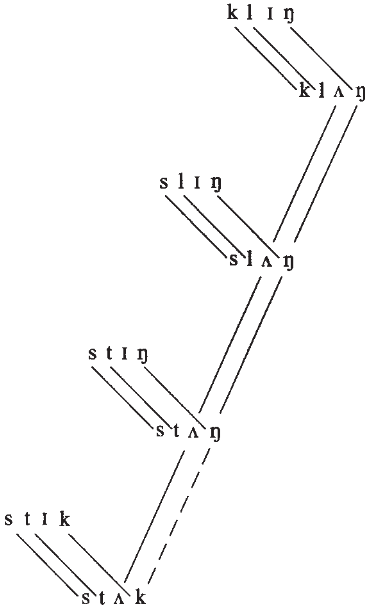
However, such differences *are* found in the acquisition process. Clark and Berman (1987) report that children acquiring Hebrew master morphology requiring suffixation before that requiring stem changes. They point to two reasons for this: stem changes obscure the shape of the lexeme more than affixes do, and stem changes tend to be restricted to lexically arbitrary classes of stems. In the research reported on in this article we attempted to manipulate these two factors in order to determine whether stem changes per se are more difficult to acquire than affixes. Our hypothesis is that stem changes themselves are no more difficult than affixes to store and process, but it is the fact that stem changes, in all reported cases, are lexically arbitrary and of low type frequency that makes them more difficult to master.

In order to find out whether it is stem change itself or other factors that delay acquisition, it was necessary to experiment on artificial languages in which stem change could be made independent of lexical arbitrariness and type frequency. Before describing this experiment and its results, another matter needs to be examined.

(1) a.



b.



If we are claiming that affixation is no easier to acquire and process than stem changes, we need to explain why it is that affixation is so much more common both across the languages of the world and within individual languages. The answer to this question lies in the diachronic phenomenon of grammaticization.

2. The diachronic perspective

Grammaticization is the process by which lexical morphemes or phrases develop into grammatical morphemes and continue their development, often resulting in the development of new affixes. Now that recent research has examined this process in a wide variety of languages (Heine and Reh 1984; Bybee 1985; Heine, Claudi, and Hünnemeyer 1991; Traugott and Heine 1991; Bybee, Perkins, and Pagliuca 1994), it is clear that the primary sources of grammatical affixes are free words or phrases. (Although there is no direct evidence available on the source of Semitic stem changes, they could have originally been conditioned by the presence of affixes that have long since disappeared. See table 7.1 below for an example of how this occurs.) It is also clear that the grammaticization process is ongoing at all times in all languages and that new grammatical morphemes are constantly arising, some of which are destined to become affixes. Some accessible examples are the development of the English regular past-tense suffix *-ed* (OE *-ede*) from the preterite form of the verb *don* 'to do', *dyde*, which in Proto-Germanic followed the main verb, leading to its development as a suffix. In derivational morphology, the productive affix *-ly*, which appears both on nouns to form adjectives and on adjectives to form adverbs, developed from the OE word *liç* meaning 'body'. Its frequent occurrence in compounds led to its development into an affix.

The development of internal stem changes is a related phenomenon and also occurs during the grammaticization process. However, it is much less common, and more important, the development of stem changes depends entirely upon the prior existence of affixes. Stem changes tend to develop very late in the grammaticization process. The longer an affix has been attached to a stem and the more the combination is used, the greater will be the phonological fusion between the affix and the stem. Sometimes the stem conditions changes in the affix (as when the *-d* of the English past tense is devoiced following a voiceless consonant). At other times, the affix will condition a change in the stem. In Middle English, long vowels were shortened before certain stop clusters so that the past-tense suffix in words such as *slept*, *kept*, *left* caused shortening of the stem vowel, giving them a different vowel from their base forms *sleep*, *keep*, and *leave*. If an affix conditions a change in the stem and then is itself deleted, the stem change becomes the signal for the morphological category. For instance, an old Germanic noun plural suffix was *-i*, and it appeared on nouns such as *mūs* 'mouse', *gōs* 'goose', and *fōt* 'foot'. This suffix, which consisted of a high front vowel, caused the vowel of the stem also to become a front vowel, yielding the long front rounded vowels [ü] and [ö]. At about the same time, this suffix vowel reduced to schwa. Later the schwa deleted entirely and the front rounded vowels became unrounded (to [i] and [u]). Then these long vowels of the singular and the plural were subject to the Great Vowel Shift of early Modem English, giving us the pairs *mouse*, *mice*; *goose*, *geese*; and *foot*, *feet* (see table 7.1). The entire process by which a stem change arose under the influence of an affix that was later lost

TABLE 7.1. The Development of an English Stem Change for Plural

	Singular	*Plural*	
Proto-Germanic	* mūs	* mūsi	
Proto-Germanic	* mūs	* mȳsə	
Old English	mūs	mȳs	[müs]
Middle English	mūs	miese	[mi:s]
Modem English	mouse [maus]	mice	[mais]

took many centuries. Since grammaticization is a continuous process and new grammatical morphemes are constantly arising and supplanting older ones, most affixes do not last long enough to produce stem changes, and even if they do, they are usually being replaced by some other affix, leaving the stem change only as residue in highly frequent items. Thus grammaticization explains why affixes are common, but stem changes are rare.

Another historical dimension of our argument involves grammaticization as well. As we have mentioned, during grammaticization a once free grammatical morpheme can become an affix. Becoming an affix involves a number of changes:

1. The morpheme in question must become fixed in its position. In particular, it must have a grammatically determined position with regard to a lexical stem. (If it is fixed with respect to a phrase, it will be a clitic and it will not become an affix.)
2. No open-class items can come between the grammatical morpheme and the lexical morpheme.
3. When these two criteria are met, the grammatical and lexical morpheme combination may be reanalyzed as a single unit, that is, a word. This reanalysis depends upon the combination constituting a coherent conceptual unit (Bybee 1985).
4. Phonological reduction in the grammaticizing morpheme is ongoing at all times (loss of stress or high tone, reduction and loss of consonants and vowels), forcing it to become more dependent on surrounding material.
5. Phonological fusion between stem and affix takes place.

A well-documented instance of this type of change is the development of the future tense in Romance languages such as Spanish and French. A periphrastic construction in Latin consisting of an inflected auxiliary *habere* 'to have' and an infinitive yielded a meaning of obligation or predestination:

(2) amare habeo
love + inf aux + 1s
'I have to love, I am to love'

The auxiliary reduces phonologically and comes consistently to appear after the infinitive (where previously it could occur in various places in the clause). In Old Spanish we find the construction indicating future:

(3) amar he
love + inf aux + 1s
'I will/shall love'

The auxiliary is written separately from the infinitive because at this stage other morphemes could come between the two, for instance, the object pronoun:

(4) amar lo he
love + inf him aux + ls
'I will love him'

Later this possibility disappears and the auxiliary becomes an actual suffix to the verb:

(5) lo amaré
'I will love him'

In this process the grammaticizing morpheme undergoes phonological reduction (e.g., from *habeo*, to *he*, to *é*), its position becomes fixed, it fuses with the verb, and the whole construction takes on a more abstract, grammatical meaning.

If all regular morphology were handled in a module with symbolic rules that resemble syntactic rules there would be no motivation for the process of affixation ever to take place. Free grammatical morphemes inserted by rule could simply continue to be inserted by rule. However, if affixation is the lexical representation of two units (morphemes) as one unit (a word), then we can understand the affixation process as being the result of the two units occurring together frequently enough and constituting a coherent enough conceptual unit that speakers remember and store them together as a single unit in their mental lexicon. As we said earlier, if affixes are represented lexically, then there is no qualitative difference between affixes and stem changes.

3. The experiment

Because in all languages affixes develop historically more quickly and more easily than stem-internal changes, it is normally the case that affixes are used in the productive patterns and have high type frequency, while stem changes, if they are used at all, occur in lexically arbitrary classes and have low type frequency but high token frequency. It is also often the case (as in English past tense) that there are several types of stem changes, each one restricted to an arbitrary class of verbs. In order to determine whether or not there is an acquisition or processing advantage to affixes, it is necessary to create artificial languages in which the relevant variables can be manipulated independently of one another. Artificial languages have been used in psycholinguistic studies to explore other linguistic variables such as the acquisition of syntax (e.g., Green 1979; Mori and Moeser 1983). Such artificially constructed languages have the advantage that they can be used to study the linguistic strategies that speakers bring to their task without examining the properties of a specific, existing language.

3.1. Background

A study by Bybee, Fox, and Klinck (1987) serves as a point of departure for the current experiment. In the earlier study Bybee and co-authors asked whether affixes were more "natural" than stem changes. Their operational definition of *natural* was the frequency with which subjects generalized either stem changes or suffixes to novel words in an artificial language–learning task. Twenty subjects were presented with sixteen singular and plural noun pairs in an artificial language. The novel words were accompanied by pictures of their referents in order to facilitate the learning of these forms. All of the singular forms were single syllables with a CVC form. Half of the stimuli (eight pairs) used a suffix schwa in the plural, whereas the other eight pairs had an internal vowel change to schwa to indicate the plural form. Both the suffix and the stem had equal type frequency, and both were equally regular, in the sense that there was no allomorphy to learn.

Subjects underwent three learning trials in which they were presented the singular form and were asked to recall the plural; by the last trial, performance was at 90% accuracy. In the final, critical phase, subjects were subsequently tested on a generalization task. They were presented with six novel words in singular form and were asked to supply what they thought would be an appropriate plural form.

The results indicated that there was no preference for affixation over the stem change. Of the 120 possible responses in the experiment (20 subjects × 6 novel nouns), 63 were affixes (52.5%). When the results were examined by subjects, there was also no clear tendency to choose affixes more frequently than stem changes. The results showed that eight subjects had more affix than stem-change responses, six subjects had an equal number of affixes and stem changes, and six subjects had more stem-change than affix responses.

These data, then, show that there is no obvious advantage for affixation over stem change to indicate the plural and therefore suggest that the former is not a more "natural" process. Accepting these results, however, means that we are accepting the null hypothesis, which is definitely problematic. Furthermore, it may be that Bybee and co-authors did not detect a difference between the two types of plural markers for a different reason that was masked by the design of the study. The experiment did not systematically address the possibility that type frequency, not the type of morphological process, determines the generalization of a suffix or a stem change. In this experiment, the stem change and the suffix were regular, that is, there was only one form for each, and both stem changes and suffixes had equal type frequency. Because the researchers held regularity constant, we cannot determine whether the equivalent generalization of the two types of plural markers was the result of their being equally "natural" or of the fact that they had equal type frequency, which is not usually the case with natural languages where suffixes typically have high type frequency and stem changes typically have low type frequency. The following experiment addresses this issue by using the same basic paradigm but manipulating both type frequency and type of morphological process while keeping token frequency constant in all conditions.

The following variables were controlled:

1. Affixation versus stem change. In each condition, half of the nouns form their plural by the addition of a suffix and half by a change in the stem vowel.

2. Phonetic simplicity (Clark and Berman 1987) is the degree of change between two morphological forms. We recognize that a derivation involving multiple differences between two forms, such as *bring*, *brought*, will be more difficult to acquire than derivations involving only one, such as *string*, *strung*. Therefore, the stem changes used in the experiment involved changing only the vowel of a monosyllabic word.
3. Lexical arbitrariness. In each condition, half of the nouns used required a stem change to form the plural and half used a suffix. The nouns were constructed so that there were no phonological or semantic reasons for the choice. In two of the conditions stem changes took on four different patterns, again lexically arbitrary, and in two conditions the suffix had four arbitrary allomorphs.
4. Type frequency. Type frequency is related in this experiment to lexical arbitrariness. The more variants of the stem change or suffix there are, the lower the type frequency. The conditions in which there are four variants with low type frequency will be called the *irregular* conditions, and the conditions in which there is only one variant with consequent higher type frequency will be called the *regular* condition.
5. Token frequency. Paradigms with irregularities that also have low type frequency tend to have high token frequency. In our experiment we hold token frequency constant by presenting each noun and its plural the same number of times.

There were two independent variables in this experiment: *suffix type* (regular or irregular) and *stem-change type* (regular or irregular), which were crossed with each other in a 2 × 2 design. Each subject received one of four possible combinations of conditions: (1) suffix regular, stem change regular; (2) suffix irregular, stem change irregular; (3) suffix irregular, stem change regular; and (4) suffix regular, stem change irregular. Therefore, all subjects received both types of plurals, suffixes and stem changes, but what varied between conditions was the regularity of the plural for the two morphological processes. This design allows us to assess the effects of regularity on the ease of acquisition and the generalization of suffixes and stem changes while at the same time determining whether there is an overall preference for one morphological process over the other.

Subjects were randomly assigned to one of the four combinations of conditions, resulting in ten subjects in each group. The dependent variables in the generalization phase were percent choices of plural forms (suffix or stem change). Analyses on the learning data used percent correct.

The predictions of the "naturalness" hypothesis are clear: subjects should produce suffixes more frequently than stem changes in the generalization phase, regardless of the type frequencies of the two forms. If type frequency is an important constraint, however, then we expect to see the rate of generalization vary: high type frequency in the input condition could result in more generalization of the form as a plural marker, with the reverse true for plural forms with low type frequency. Of course the picture could be more complex, such that one morphological process might be favored under conditions of high type frequency and the other under conditions of low type frequency. Finally, given our argument that stem changes are as "natural" as suffixes in marking the plural, we expect the overall rate of stem changes in the experiment to be equivalent to that of suffixes.

4. Method

4.1. Subjects

Forty students at the University of New Mexico served as subjects in this experiment. They were members of introductory linguistics and psycholinguistics courses who were given extra course credit for a written description of their participation in the experiment. In order for their data to be included in the experiment, subjects had to be native speakers of English (defined as their primary language, acquired before the age of 6 years).

The use of English speakers is justified by the fact that English has precisely the pattern that we are hypothesizing will *not* show up in our results. Thus the use of English speakers biases the experiment against the result we expect to obtain.

4.2. Materials

Two sets of materials were constructed for this experiment, corresponding to the two phases, learning and generalization. Sixteen novel words were generated for the learning phase (see table 7.2). These novel words were all of the form CVC and were constructed with a variety of consonants and ensured that vowels were used with equal frequency (see table 7.3).

The novel words were presented to subjects as the "singular" form of a word in a novel language, with the referent depicted in a line drawing. Subjects were also presented with a "plural" form of the same word, accompanied by an appropriate line drawing. The subject's task was to learn the singular and plural forms, such that when presented with the singular forms and pictures followed by the pictures for the plural they could provide the plural forms. They were also given a generalization task at the end of the experiment in which they were presented with a new set of ten novel words in the singular and asked to supply, for each, what they thought would be an appropriate plural form.

The stimuli were constructed in the following manner. The set of words was divided in half, with eight words being assigned to the affix condition and the other

TABLE 7.2. Example Stimuli for Stem-Change and Affixation Conditions

Singular		*Plural*	*Type*	*Picture*
faut	→	fət	stem change	girl(s)
dit	→	ditə	suffix	flower(s)

TABLE 7.3. Stimuli Used in the Learning Phase of the Experiment

Plural form	*Regular*	*Irregular*
Stem change		
waiš	wəš	wəš
zib	zəb	zəb
raik	rək	ruk
sauf	səf	suf
vaum	vəm	voum
faut	fət	fout
ǰik	ǰək	ǰɪk
naid	nəd	nɪd
Affix		
hauk	haukə	haukə
mir	mirə	mirə
sail	sailə	sailəm
nauz	nauzə	nauzəm
ǰaib	ǰaibə	ǰaibni
dit	ditə	ditni
vaud	vaudə	vaudət
faiš	faišə	faušət

eight to the stem-change condition. The other variable was regularity: in the case of regular suffixes and stem changes, there was only one form of the plural; thus each of the eight words (in either the stem-change or affix condition, as appropriate) took the same form. For the *regular* conditions this meant either the addition of [ə] as a suffix or the replacement of the vowel with [ə] for the stem change, for all eight instances of each condition (see table 7.2). By using schwa for both the stem-change and suffix condition, we controlled the vowel such that any differences in the two conditions could not be attributed to the specific vowel used. While we think it unlikely that the specific vowel that is used for the regular conditions would make much difference, future studies could replicate the study using a different vowel or vowels.

In the *irregular* suffix and stem-change conditions, four different plural forms were used, resulting in two novel words in each set of eight whose plural was signaled in the same manner. This contrasts with the regular conditions in which, within each of the eight-item sets, the plural form was marked by the same form. For the irregular suffix condition the four forms used for the affix were [ə] and three other suffixes: [əm], [ət], and [ni]. The irregular plural stem changed to [ə], [u], [ɪ], and [ou]. Care was taken to ensure that the resulting forms did not already exist in English, and that the choice of a particular stem change or suffix in the irregular condition did not correlate with any phonological or semantic properties of the stem (see table 7.3).

Each novel word and its plural form were recorded by a female native English speaker using a Bell and Howell Language Master. The Language Master is an audiotape recorder that reads from, and records onto, thin cardboard cards that have a strip of audiotape attached to the bottom. Pictures of the referents (both single and multiple forms) were copied onto the card so that in the initial phase subjects could both hear the singular and plural forms of each word and see their referents. Because each item was recorded on a separate card, it was possible to shuffle the stimulus set and present each subject with a unique order of stimuli.

Ten additional novel words were generated for use in the generalization phase of the experiment:

(6) Novel items used in the generalization phase of the experiment
biš
waub
zaud
šaim
lauǰ
kaiǰ
tis
zaif
vin
raus

These were constructed in the same manner as the stimuli for the learning phase. In all cases, for learning and generalization trials both, the materials were carefully constructed so that there were no consistent cues from the structure of the CVC that indicated whether it would take a stem change or suffix in the plural.

4.3. Procedure

4.3.1. Learning phase

Subjects were told that they would be presented with singular and plural forms of words from a language that they did not know. They were instructed to learn the novel words so that they could, when given the singular form, recall the correct plural form.

The set of sixteen words was randomized by shuffling the audio cards and was then divided into four separate sets of four items each. On the initial trial, subjects heard the four items in both the singular and plural form, with their appropriate pictures. They were then presented with the singular form only and asked to generate the plural form. The responses were recorded (both by the experimenter and on audiocassette) and corrected, if wrong. The same four singular items were presented again for two more recall trials of four items, and on each trial the subject was again asked to provide the plural form. Corrections were provided. Following the third iteration of the first four-item set, the subject was presented with the second set of four items (in both singular and plural form), and the same procedure, with three recall trials, was followed. The initial phase of the learning task was repeated for the remaining two sets of four items in the list. For purposes of analysis, each of the three recall trials is referred to as a practice list (numbered from 1 to 3), and the results will be reported across all sixteen words for each list.

Following the initial phase, subjects were presented with the full list of sixteen singular items and instructed to generate the correct plural for each. They were given feedback and corrected where appropriate. The final learning trial consisted of the same sixteen singular items, with instructions to generate the plural form. In this last trial of the learning phase no feedback was given.

4.3.2. Generalization phase

In the last phase of the experiment, subjects were given, one at a time, the new list of ten novel words, accompanied by pictures of their referents. They were told that these were in singular form and, following the presentation of each, were asked to supply what they considered to be appropriate plural forms. Obviously there were no correct answers since these were new items, and subjects were not given feedback on their responses. Following this phase subjects were thanked and debriefed.

4.4. Results

4.4.1. Learning phase

Analysis of the learning data was conducted on the mean percent correct recall across the conditions. In order to observe the pattern of acquisition, the recall data were examined separately for the three practice lists and the two final, full lists. The results revealed that the four conditions differed in initial learnability (see figure 7.1);

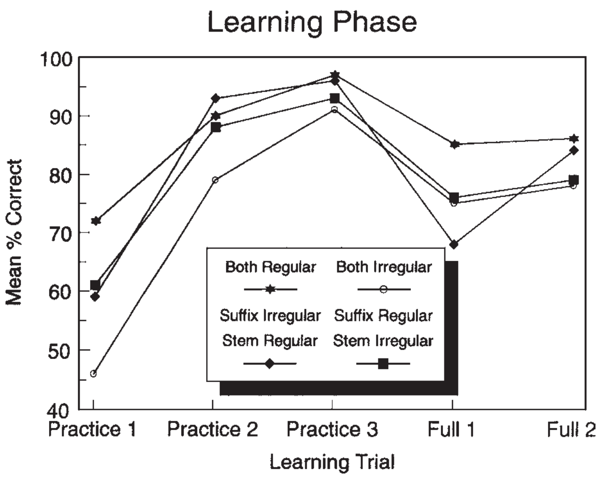


FIGURE 7.1. Correct performance by conditions in the learning phase.

the regular conditions led to higher initial learnability, whereas the irregular conditions led to the lowest initial performance. The effect of regularity was present for both suffixes and stem changes as shown by an analysis of the percent correct data for practice list 1. There was a significant main effect of suffix regularity, $F(1,36) = 5.941$, $p < .02$, MSe = 2065; and a marginally significant main effect of stem-change regularity, $F(1,36) = 4.054$, $p < .052$, MSe = 1409. The interaction of suffix regularity × stem-change regularity was not significant, $F < 1$.

Pairwise comparisons revealed that the source of the initial difference was a significant difference between condition 1 (suffix and stem change both regular) and condition 2 (suffix and stem change both irregular), such that condition 2 led to considerably fewer correct responses (condition 2: 46%, condition 1: 72%), two-tailed $t(18) = 3.31$, $p < .004$. Condition 1 was also marginally different from condition 3, suffix irregular, stem regular (59%): two-tailed $t(18) = 2.03$, $p < .058$. None of the other pairwise comparisons between conditions for practice list 1 approached significance.

Performance on the remaining practice lists quickly reached near ceiling levels. There was a marginally significant effect of stem-change regularity on practice list 2, $F(1,36) = 4.020$, $p < .053$, MSe = 612. No other effects approached significance for either practice list 2 or 3. Mean percent correct on practice list 3, collapsed across conditions, was 94%.

Analyses were also conducted on the percent correct scores for the two full lists. The analyses on the data for full list 1 revealed that there was a marginally significant interaction between suffix and stem-change regularity, $F(1,36) = 3.239$, $p < .08$, MSe = 712, and a marginally significant effect of suffix regularity, $F(1,36) = 3.732$, $p < .061$, MSe = 821. Although neither effect achieved the conventional level for rejecting the null hypothesis, both effects are strong and approaching significance, so that they should not be dismissed. The locus of the interaction can be seen in figure 7.1, where irregular suffixes in the context of a regular stem change (i.e., condition 3: suffix irregular, stem regular) lowered relative performance considerably more than did irregular stem changes in the context of regular suffixes (i.e., condition 4: suffix regular, stem irregular). The suffix-regularity effect is shown in the graph by the fact that the mean of the two suffix-regular conditions (both regular, and suffix regular, stem irregular) is higher than the mean of the two suffix irregular conditions (both irregular, and suffix irregular, stem regular).

Performance on full list 2 was quite high (82%) though not quite at the levels achieved by the end of practice list 3, probably because subjects had to successfully perform a paired-associate task for the entire sixteen items in the list, rather than recalling each set of four items separately. Importantly for our purposes, there were no significant differences in the rate of correct responses across the four conditions: suffix regularity × stem-change regularity, $F < 1$; suffix regularity, $F < 1$; stem-change regularity, $F(1,36) = 1.656$, $p < .206$, MSe = 389.

Thus the learning data indicate that there is no inherent difference in difficulty in terms of acquiring stem changes rather than suffixes. Although there were initial differences in percent correct as a function of regularity for both suffixes and stem changes, these quickly disappeared.

4.4.2. Generalization data

The rate of suffix generalization was 48.75%, whereas the rate of stem-change generalization across subjects in all four conditions was actually 1.5% higher, at 50.25%. Two other types of responses were noted in the generalization task—the use of an English suffix (.25%) and no change to the singular form (.75%). Combined, these two unusual response types totaled only 1% of the generalization data and were not included in the statistical analysis. Thus the overall results of this experiment look very much like those of the Bybee, Fox, and Klinck study, showing that stem changes and suffixes appear to be equally learnable and thus equally "natural." This study, however, differs in that we manipulated regularity as well as morphological form and are thus able to explore whether the apparently equivalent rates of use of the two types of morphological processes hold when regularity varies.

Figure 7.2 presents the percentages for the use of suffixes and stem changes in the generalization phase as a function of stem-change and suffix regularity. The pattern of results shows that the regularity of the stem-change condition during the learning phase had no effect on the use of either suffixes or stem changes in the generalization phase, but that suffix regularity did. Specifically, the rate of suffix use was considerably higher than stem-change use when the suffix input in the learning phase was regular, but lower than stem-change use when the suffix input was irregular. The asymmetry of this pattern is clearly illustrated in figure 7.3, which plots difference scores formed by subtracting the percentage of stem-change choices

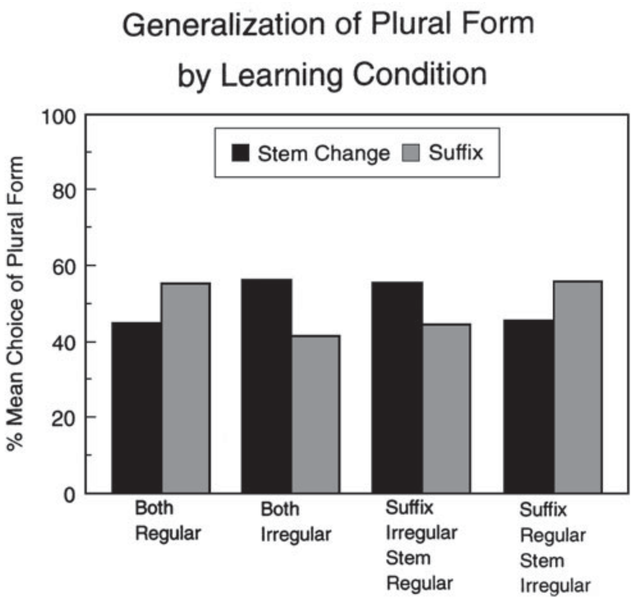


FIGURE 7.2. Generalization of suffixes and stem changes by learning conditions.

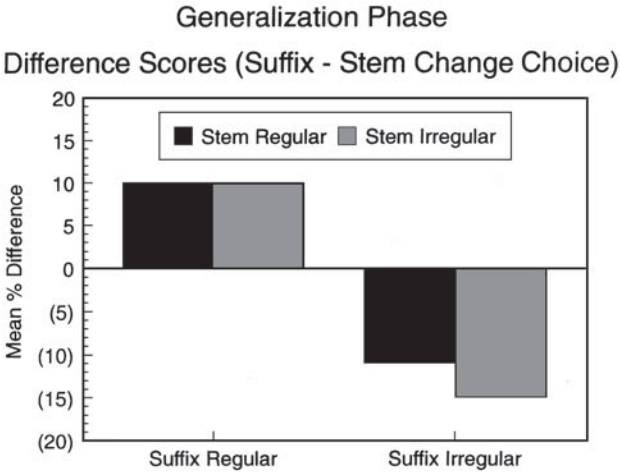


FIGURE 7.3. Generalization difference scores (suffix–stem change) as a function of regularity.

from the percentage of suffix choices. If suffix and stem-change generalization are equivalent, then the difference scores will be zero, regardless of input condition. Figure 7.3 clearly demonstrates that there were more suffix choices than stem-change choices when the suffix was regular (positive difference scores) and more stem-change than suffix choices when the suffix was irregular (negative difference scores).

Because the two dependent measures, percent suffix choice and percent stem-change choice, sum to 100%, thus ensuring that knowledge of one score unambiguously determines knowledge of the other, it would not be appropriate to conduct analyses of variance on both dependent measures. Thus a 2 (suffix regularity) × 2 (stem-change regularity) analysis of variance was conducted on the mean percent suffix choices. There was a main effect of suffix regularity, $F(1,36) = 5.009$, $p < .031$, MSe = 1562. Neither the main effect of stem-change regularity nor the interaction between suffix and stem-change regularity was significant, both Fs < 1. Thus the initial observation that there was no overall difference between the rates of suffix and stem-change generalization must be qualified by the finding that suffix regularity, but not stem-change regularity, determines the rate of suffix and stem-change generalization.

The pattern of results clearly shows that subjects' acquisition of the plural is affected by suffix regularity, but it also demonstrates that acquisition of a stem change to mark the plural is not inherently more difficult than learning to add a suffix as the plural marker. In the two conditions that contained irregular suffixes, subjects chose the stem change more frequently as a plural marker during the generalization phase (condition 2, suffix and stem irregular: 56%; and condition 3, suffix irregular, stem regular: 55%). This level is equal to that of suffix choice in the two regular suffix conditions (condition 1, suffix and stem regular: 55%; and condition 4, suffix regular, stem irregular: 55%), again showing that it is regularity of the suffix, not difficulty of the stem change, that determines the use of affixation versus stem change. Were affixation a more "natural" process overall, then we would expect to see a much greater advantage for suffixes under the input conditions that most favor their use, namely, the two regular suffix conditions. Indeed, we would expect the greatest use of suffixes in condition 4, in which the stem changes are irregular, but the suffixes are regular; but as we noted, that condition did not result in any higher suffix use than did condition 1, where both the suffix and the stem change were regular.

5. Discussion

Based on a theory in which all morphology is treated in basically the same manner, we hypothesized that stem changes as morphological markers could be learned as easily as affixes. The reason that stem changes seem difficult to acquire in languages such as English, German, and Hebrew is that most stem-change patterns are lexically arbitrary and have very low type frequency. By creating an artificial language and asking subjects to learn it we could distinguish the effects of stem change from the effects of type frequency. The results showed no overall advantage of suffixes over stem changes, but an effect of irregularity or type frequency of the suffixes was revealed. In the learning phase, irregularity in both affix and stem-change conditions led to initial difficulties. In the generalization phase, the overall rate of stem changes was as high as, in fact slightly exceeded, the rate of suffix usage. In this phase also, we see the effects of regularity: regularity of suffixes led to greater use of suffixes, while irregularity of suffixes led to greater use of stem changes.

Thus claims that suffixes (and affixes in general) are somehow psychologically easier to encode may be misguided. Our data suggest instead that regularity, which for diachronic reasons happens to be correlated with affixation rather than stem change, affects the ease of generalization of the suffix. It is also interesting to note that under the right circumstances, such as the two cases of irregular suffixes set up by our artificial language, stem changes can be acquired as easily as, or even more easily than, suffixes.

Our results also show the effects of type frequency on determining generalizability. In the two mixed conditions, one in which suffixes were regular and stem changes irregular and the other in which suffixes were irregular and stem changes regular, the regular morphological process was generalized more frequently than the irregular one. Our interpretation of these results is that the subjects were basically involved in a rote-learning task, setting up lexical representations for this fragment of an artificial language. Where half of the singular/plural pairs they learned had the same morphological process, this process was strengthened. In generalizing, the subjects were accessing the stored units and relations and could form new plurals on the basis of any stored pattern, but they were more likely to use the pattern reinforced by more separate entries.

For the conditions in which both processes were regular or both were irregular, we expected that responses would be approximately evenly divided among suffixes and stem changes. It is logically possible that subjects might have chosen a single suffix or stem change and generalized it throughout, but it appears that in producing new forms the subjects were in general replicating the distribution of forms in the input. Thus in the all-regular and all-irregular conditions, subjects gave responses of both types. But here we found an interesting effect of regularity. When both processes were regular, subjects used significantly more suffixes, but when they were both irregular, subjects chose significantly more stem changes. Taken together, then, there is no overall suffixation preference, but we did find a difference that needs to be explained.

This result may be due to the subjects' prior experience with natural language(s), in which, as we have said, there is a strong tendency for stem change to be associated with irregularity and for affixes to be associated with regularity. Thus in the very difficult condition in which all plurals were irregular, the subjects might have been more comfortable with the stem-change solution, since stem changes tend to be irregular in English (and in all languages). Conversely, in the all-regular condition, the suffix has an edge since regularity tends to be associated with suffixes.

Because of their experience, English speakers should have been predisposed to use suffixes. The data clearly show that subjects did not, overall, prefer suffixes to stem changes. Since the subjects did generalize suffixes to the novel forms when one of the two input conditions was regular, one possible explanation of this result is that they were speakers of English. The best way to avoid this inherent bias would be to use as subjects speakers of an isolating language such as Mandarin Chinese. In this case we would predict effects of regularity for both suffixes and stem changes, but once again, no overall advantage for one form of morphological expression over the other. If there were a language whose pattern was opposite to that of English, rather than neutral, we would expect to find the opposite results, that is, an effect of stem-change regularity but not suffix regularity, and again, no overall advantage for either form. It is clear that the persistent refrain in our predictions, and in fact our findings, is that there should be no overall advantage for suffixes over stem changes as, in our view, neither form is more "natural" than the other.

In a model in which all morphological relations are treated as lexical connections, there is no particular advantage to affixation over internal changes. Even in such a model, it would still be reasonable to suppose that morphemes that consist of continuous phonological material, rather than interrupted material (as in *sing*, *sang*, *sung*, where the root is *s-ng*), have an acquisitional or processing advantage (Dressler 1985), yet our experiments turned up no evidence of such a tendency. The suggested association of concatenation with regularity and stem change with irregularity (Marcus et al. 1993) was not supported in our experiment, where we demonstrated that the English speakers behave as though these are separate variables.

On a more general theoretical level, we set out to demonstrate that some aspects of the structure of natural language are explainable diachronically but not synchronically. All human languages are the result of long, gradual processes of development. Only certain avenues of development are available, and these lead to certain structures. The fact that some structures are more common than others in the languages of the world does not necessarily mean that the more common structures are optimal from a psycholinguistic point of view. It may only mean that these structures arise more easily and more frequently than others. Since the primary sources of affixes are words and stem changes can arise under conditioning from affixes, the fact that affixes are more widespread and common than stem changes has a straightforward diachronic explanation.

Note

We would like to thank Laurent Thomin, Patricia Escarraz, Sally Weller, and Valerie Daniel for their assistance in running subjects in this study. Correspondence address: Department of Linguistics, University of New Mexico, Albuquerque, New Mexico 87131.

References

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Philadelphia, Penn.: John Benjamins.

———. 1988. Morphology as lexical organization. In *Theoretical morphology*, ed. Michael Hammond and Mickey Noonan, 119–141. San Diego, Calif.: Academic Press.

Bybee, Joan L., Patsy Fox, and Patrick Klinck. 1987. Are affixes more natural than stem changes? Paper given at the "Naturalness and Iconicity: Explanation in Historical Morphology" workshop at the VIIIth International Conference on Historical Linguistics, Lille, France, September.

Bybee, Joan L., and Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59, 251–270. (Reprinted in this volume as chapter 6.)

Bybee, Joan L., Revere D. Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, Joan L., and Dan I. Slobin. 1982. Rules and schemas in the development and use of English past tense. *Language* 58, 265–289. (Reprinted in this volume as chapter 5.)

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Clark, Eve V., and Ruth A.Berman. 1987. Types of linguistic knowledge: Interpreting and producing compound nouns. *Journal of Child Language* 14, 547–567.

Dressier, Wolfgang U. 1985. *Morphonology: The dynamics of derivation*. Ann Arbor, Mich.: Karoma.

Green, T. R. G. 1979. The necessity of syntax markers: Two experiments with artificial languages. *Journal of Verbal Learning and Verbal Behavior* 18, 481–496.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Heine, Bernd, and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

MacWhinney, Brian, Jared Leinbach, Roman Taraban, and Janet McDonald. 1989. Language learning: Cues or rules? *Journal of Memory and Language* 28, 255–277.

Marcus, Gary F., Ursula Brinkmann, Harald Clahsen, Richard Wiese, Andreas Woest, and Steven Pinker. 1993. *German inflection: The exception that proves the rule*. Center for Cognitive Science, Occasional Paper 47. Cambridge, Mass.: MIT.

Marcus, Gary F., Steven Pinker, Michael Ullman, Michelle Hollander, T. John Rosen, and Fe Xu. 1992. *Overregularization in language acquisition*. Monographs of the Society for Research in Child Development 57, 4, Serial No. 228. Chicago: University of Chicago Press.

Mori, Kazuo, and Shannon D. Moeser. 1983. The role of syntax markers and semantic referents in learning an artificial language. *Journal of Verbal Learning and Verbal Behavior* 22, 701–718.

Pinker, Steven. 1991. Rules of language. *Science* 253, 530–535.

Prasada, Sandeep, and Steven Pinker. 1993. Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes* 8, 1–51.

Rumelhart, David E., and James L. McClelland. 1986. On learning the past tenses of English verbs. Implicit rules or parallel distributed processing? In *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 2: *Psychological and biological models*, ed. James L. McClelland, David E. Rumelhart, and the PDP Research Group, 216–271. Cambridge, Mass.: MIT Press.

Saddock, Jerrold. 1985. Autolexical syntax: A theory of parallel grammatical representations. *Natural Language and Linguistic Theory* 3, 379–439.

Traugott, Elizabeth C., and Bernd Heine (eds.). 1991. *Approaches to grammaticalization*, vols. 1 and 2. Amsterdam: John Benjamins.

Chapter 8

Regular Morphology and the Lexicon

1. The controversy

A long-standing debate in the linguistic and psychological literature centers around the representation of morphologically complex words in the grammar and lexicon. It seems as if every conceivable position on this issue has been argued for seriously and debated vigorously at some time in the last thirty years.[1] The view that is now emerging from a consideration of child language development, diachronic change of morphology in the history of languages, and psycholinguistic experimentation shows the representation of morphology to be both complex and very interesting, reflecting basic cognitive organizational principles.

Studies of the inflectional system for English verbs show a highly productive regular pattern by which *-ed* is suffixed to verb bases, as well as a variety of vowel changes that affect relatively frequent verbs to form their past tense. Even though the latter set is deserving of the label "irregular," since there are several different vowel-change patterns, certain classes of verbs within the irregulars show some degree of productivity—that is, the ability to be extended to new items. An important fact about the distribution of regular and irregular verbs in English is that the irregulars constitute a very small proportion of the entire set and they are mostly of very high token frequency. The regular class has by far the highest type frequency.

Several different proposals concerning how regular and irregular morphological formations are stored and processed have emerged in the recent literature. In this article, I will describe and defend my own proposed model (Bybee, 1985, 1988) and compare it with connectionist proposals and with the dual-processing model of Pinker, Marcus, and their colleagues. The main issue to be discussed is the role of type frequency in the determination of regularity or productivity. My model and the connectionist model claim that type frequency is an important determinant of productivity, whereas the dual-processing model denies any role of type frequency in productivity.

In the next section, I outline briefly the properties of the dual-processing and connectionist models and in somewhat more detail outline my own model, showing which properties it shares with these models and how it differs from each of them. I then go on to discuss the motivation for considering type frequency as pivotal in the determination of productivity and make certain predictions concerning the relationship between type and token frequency.

Next, I examine the evidence against type frequency as a major determinant of productivity, first taking the claims of Clahsen and Rothweiler concerning German participles and showing that, in fact, under the most reasonable method of counting type frequency the productive process does have the highest type frequency. I then discuss the case of the German *s*-plurals and argue that to a large extent the productivity patterns of German plurals also reflect their type frequency. I then go on to address the claim that the processing of regulars is of a qualitatively different type from the processing of irregulars. Here data from Hausa come into play, for the Hausa system shows clearly that even the productive patterns have the same properties as lexical patterns: they are organized in a prototype structure and they are product-oriented generalizations (Lobben 1991). Next, I discuss the diachronic origins of productive morphological patterns that, I argue, explain why productive patterns tend to be more transparent and agglutinative in structure than non-productive ones. I then consider whether the study of the acquisition of English past-tense verbs as presented in Marcus et al. (1992) demonstrates a difference between regular and irregular inflection. I will argue that it demonstrates just the opposite. Finally, I discuss evidence presented in Losiewicz (1992) that shows that regularly inflected forms have differential lexical strength, just as irregularly inflected forms do.

2. The models

Several recent studies have reached the conclusion that English irregular past-tense forms are listed in the lexicon and that their listing is structured in such a way that generalizations or schemas emerge from the similarities among verb forms (Bybee & Moder 1983; Bybee & Slobin 1982; Marcus et al. 1992; Prasada & Pinker 1993). One group of researchers, whose model I will refer to as the dual-processing model (Clahsen & Rothweiler 1992; Marcus et al. 1992, 1993; Pinker 1991; Pinker & Prince 1994; Prasada & Pinker 1993), argue that irregular inflections are lexically represented, but regular inflections are derived by a symbolic rule applying to an underlying form to produce the complex surface word. Thus they claim that two highly distinct types of processing are involved in dealing with regular versus irregular inflection. Unlike the other two approaches to be discussed here, these researchers' argument is that the distinction between regular and irregular morphology is a structural difference independent of the type frequency of the morphological patterns.

Connectionist models do not formulate autonomous symbolic rules for regular inflection but rather treat all inflection in the same way: a network of mappings from base form to past-tense form is constructed. The network memorizes individual patterns and their transformations and generalizes on the basis of regularities found in the network. Rumelhart and McClelland (1986) show that a network so constructed can produce output similar to that of a child acquiring English past forms without formulating an explicit rule. In this simulation, type and token frequency are manipulated to reflect, at least approximately, the distribution of irregular and regular verbs in the English lexicon and the way children might encounter them: in the first phase, a small set of mostly irregular verbs is presented and learned, followed by a much larger set of mostly regular verbs.

Subsequent connectionist models have responded to criticisms of this first simulation and have brought connectionist modeling closer to representing the actual situation encountered by children acquiring their language. The models of Plunkett and Marchman (1991) and MacWhinney and Leinbach (1991) each used input that was closer to the actual English input to children. Plunkett and Marchman (1991) also showed that learning was enhanced when phonological sub-regularities (such as those found in *ring*, *rang* and *sing*, *sang*) could be taken into account by the model. MacWhinney and Leinbach (1991) incorporated semantic features into their model so that it could differentiate homophonous pairs such as *ring*, *rang* versus *wring*, *wrung*. Cottrell and Plunkett (1991) devised a model that does not just learn mappings between base and past forms but also can learn and access past forms directly. Each improvement in modeling makes more plausible the claim that all English past tense could be acquired and produced by an associative network without the formulation of a symbolic rule.

The model of Bybee (1985, 1988) was developed to account for cross-linguistic, diachronic, and acquisition patterns in complex morphological systems. The basic proposal is that morphological properties of words, paradigms, and morphological patterns once described as rules emerge from associations made among related words in lexical representation. A major difference between this model, which I will call the *network* model, and structuralist models that contain rules is that actual usage in terms of both type and token frequency plays an important role in establishing and maintaining representation. Langacker (1987, 1988) has also proposed a model of grammar based on usage with the same properties as the model I will describe here. Connectionist simulations could be thought of as testing some of the properties of the network model and Langacker's cognitive grammar, but the model itself is more complex and accounts for more phenomena than any existing connectionist model. The relevant properties of this model are described here.

Words entered in the lexicon have varying degrees of *lexical strength*, due primarily to their token frequency. Words with high lexical strength are easy to access, serve as the bases of morphological relations, and exhibit an autonomy that makes them resistant to change and prone to semantic independence. Thus lexical strength explains why irregular formations are usually of high token frequency: In a language such as English, with both irregulars and a strong regular pattern, irregulars will tend to regularize unless they are sufficiently available in the input to create a strong lexical representation. Thus if the irregular past has low token frequency and is thus more difficult to access, a regular form might be created. Note that in this model it is to be expected that a few paradigms might have inflectional doublets such as *weeped*, *wept* and *creeped*, *crept*. Within paradigms, words with higher lexical strength serve as the basis for the formation of new words (Bybee 1985).

Words entered in the lexicon are related to other words via sets of lexical connections between identical and similar phonological and semantic features. These connections among items have the effect of yielding an internal morphological analysis of complex words, as shown in figure 8.1. Even though words entered in the lexicon are not broken up into their constituent morphemes, their morphological structure emerges from the connections they make with other words in the lexicon. Parallel sets of phonological and semantic connections, if they are repeated across multiple sets of words, constitute morphological relations (in figure 8.1, these are represented by heavier lines). Note that in figure 8.1 connections between base ([kæt]) and complex form ([kæts]) exist, as well as connections among complex forms ([ræts], [mæts], [kæts], [kæps]).

Lexical connections can vary in strength according to the type and number of features shared. Variations in strength of connections due to semantic features in verbal paradigms are discussed in Bybee (1985). Weaker semantic connections are reflected in phonological form by a greater degree of stem change or even suppletion among related forms (Bybee 1985). For example, stem changes and suppletion are more common among forms that differ in tense or aspect (*break*, *broke*) than among forms that differ only by person/number (*break*, *breaks*).

Strength of lexical connections can also be affected by token frequency, as reflected in lexical strength. Words that have high token frequency have greater lexical autonomy, and one reflection of this is that such words form weaker connections with other items. This relation is based on the commonsense observation that items that are of high frequency in the input can be learned on their own terms, while lower

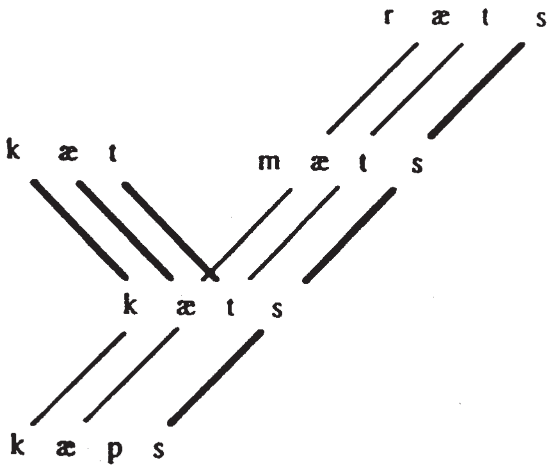


FIGURE 8.1. Sets of lexical connections yielding word-internal morphological structure.

frequency items are better learned in relation to existing items. One of the facts accounted for by the hypothesis that high-frequency words form weaker connections is suppletion. True suppletion (paradigms consisting of stems of different etymologies, such as *go*, *went* or *am*, *was*) occurs only in the most frequent of paradigms. Suppletion requires the splitting of paradigms—*went* had to separate from *wend* in order to become the past tense of *go*. My hypothesis is that the increase in the frequency of *went* and its consequent greater lexical autonomy contributed to its split from *wend*.

Sets of words having similar patterns of semantic and phonological connections reinforce one another and create emergent generalizations describable as schemas. New items or items whose connections are not known or are weak can be fitted into these schemas. The likelihood of the schema being extended to new items is directly dependent upon two factors: (1) the defining properties of the schema and (2) its strength, the latter property being derivable from the number of items that reinforce the schema. If the defining properties of the schema are very specific, the schema will be restricted in its application to new forms and result in lower productivity. If the schema is very open, placing few restrictions on the items to which it can apply, its productivity will be greater. The other determinant of productivity is the strength of the schema, which is based directly on its type frequency—the higher the type frequency of the pattern described in the schema, the greater are its chances of applying to new items.

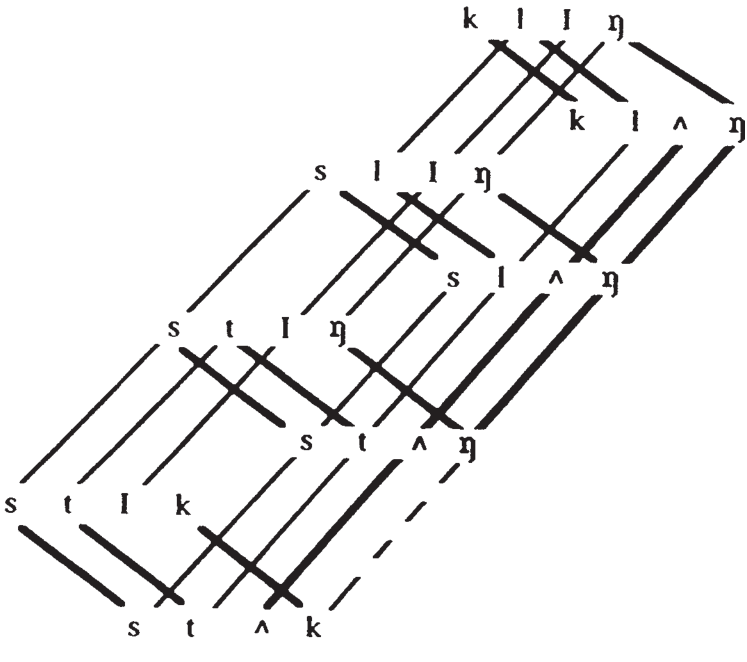
There are two types of schemas, corresponding to the two ways that morphologically complex forms can relate to other forms. The first, source-oriented schemas, are generalizations over pairs of basic and derived forms, such as *wait*, *waited*. These correspond roughly to generative rules, since they can be thought of as instructions for how to modify one form in order to derive another. The second, product-oriented schemas, have no counterpart in generative theory. They are generalizations over sets of complex or "derived" forms, such as *strung*, *stung*, *flung*, *hung*, and so on, which show what features these derived forms have but without stipulating the operations it takes to produce such forms (Zager 1980). Membership in product-oriented classes is based on family resemblances: there are more and less central members of the classes, with the central members, such as *strung*, sharing more phonological features with other members of the class than the more marginal ones, such as *dug*.

Product-oriented generalizations were studied by Bybee and Moder (1983) via the class of verbs that form their past tense like *strung*, and it was found that among nonce verbs the least important feature of the input string was the vowel, even though it appears to be the vowel that undergoes a change. In other words, the defining features of the class are not based on the present form, *string*, but rather on the past, *strung*. The data from diachronic change support this view—new members of this class (in certain dialects) include verbs that do not have [ɪ] as their base vowel: *strike* [ai], *struck*; *sneak* [i], *snuck*; *drag* [æ], *drug*. Wang and Derwing (1994) showed in a nonce-probe task that generalizations governing vowel changes in English past tense, *-ity* nominalizations, and pluralization are based on product-oriented schemas, since certain vowels are favored for each morphological pattern without regard to the quality of the input vowel.

Since in this model lexical connections relate basic and derived forms as well as derived forms of different paradigms and since there is no attempt to avoid redundancy, both source-oriented and product-oriented schemas may exist for the same morphological relation. Lexical connections representing phonological and semantic similarities, as shown in figure 8.2, provide the basis for both types of schemas. This figure also shows the product-oriented schema for the *strung*-verbs of English.

The regular English past-tense suffixation could be described as a source-oriented schema, since there is a standard operation for forming a past from a base stem ("add /t/, /d/, or /ɨd/"). But this pattern also appears to be product oriented in the case of the set of verbs that undergo no change in the past tense (e.g., *put*, *set*, *quit*, *cut*, *spread*, etc.). The verbs of this class all end in /t/ or /d/ and have lax vowels (with the exception of *beat*, which has a tense vowel). They could be described as fitting a product-oriented schema that requires past tense to end in a lax vowel and alveoar stop.

I will now compare the network model with the dual-processing model and connectionist models.[2] The connectionist models and the network model are similar in not postulating a discrete cutoff point between regular and irregular morphology,



$$\left[\mathrm{C}\ (\mathrm{C})\ (\mathrm{C})\ \wedge \begin{Bmatrix} \text{velar} \\ \text{nasal} \end{Bmatrix} \right]_{\text{past}}$$

FIGURE 8.2. Lexical connections and the schema for a part of the *strung* class.

as the dual-processing model does. Connectionist and network models attribute differences between regular and irregular morphology to quantitative differences, in particular to the relatively greater type frequency of regular patterns.

In the dual-processing model, irregular past-tense forms are listed lexically, but regular ones are not (except in special circumstances; see Pinker & Prince 1994). The regular/irregular distinction, then, corresponds to an important structural distinction that is part of the innate architecture of the language acquisition device: this model has a rules component that is separate from the lexicon. In my network model (and in connectionist models), regular and irregular forms and patterns are treated in the same way and there is no separate component for morphological rules. Differences in degrees of productivity are not attributed to different processing types but to differences in type frequency and the openness of defining schemas.

Another difference between the dual-processing model and the network model concerns the treatment of irregulars. The dual-processing model is said to handle these in associative networks, but the nature of the generalizations made about them is not specified. In particular, I find no mention of product-oriented generalizations in the literature on dual-processing.

For the purposes of this article, connectionist models and the network model make essentially the same claim: type frequency is a major determinant of productivity. However, there are differences between my model and existing connectionist ones. In particular, the treatment of token frequency in connectionist models is taken to be the frequency of the mapping between base and derived form (Plunkett & Marchman 1991; Hare & Elman 1995). In the network model, it is the frequency of the derived word itself that is significant and is represented as lexical strength, not a mapping between two forms. In fact, the higher the frequency of the derived form, the weaker the mapping between it and the basic form. High-frequency irregulars are resistant to regularization not because their connections with their base are better established but because they are themselves lexically stronger. As I explained earlier, postulating that high-frequency irregulars have high lexical strength explains not only their resistance to change but also their propensity to undergo the paradigm splits that lead to suppletion.

Another difference between existing connectionist models and the network model is that connectionist models at present form generalizations only over relations between base and derived forms and do not form generalizations over sets of derived forms. That is, connectionist models do not form product-oriented generalizations. Moreover, I see no sense in which connectionist models abstract schemas from relations among words, as proposed in the network and cognitive grammar models.

3. The role of type and token frequency in determining productivity

The evidence for a pivotal role for type frequency in determining productivity comes from child language and from examining the distribution of morphological patterns in many different languages. In Spanish, French, and other Romance languages, verbs are distributed across several conjugation classes, with the largest number of types belonging to what is termed the "first conjugation" (Spanish *-ar* verbs and French verbs such as *chanter* 'to sing'). The other conjugations, though not necessarily irregular, tend to contain fewer verbs and these verbs are typically of high token frequency. Guillaume (1927/1973) observed that the innovations in verb forms created by children acquiring French most commonly involved the use of the first conjugation inflections on verbs belonging to other conjugations in the adult language. Counting the verbs used by nursery school children during play revealed that they were generalizing not the verb class that was used most often but the class that had the highest type frequency, as shown in table 8.1. All of the children studied overgeneralized the first conjugation. Furthermore, in the adult language the first conjugation is the one most often used for inflecting new verbs.

Forms with high individual token frequency can be learned by rote, and they can be autonomous, even from other members of their own paradigms. With sufficient availability in the input, they can be acquired without forming relations with other items and without undergoing internal analysis (many high-frequency forms are more fused than low-frequency ones) (Guillaume 1927/1973; Bybee 1985).

Forms with lower token frequency will be learned more easily if they can be related to other stored forms. If the input contains a large number of distinct items that share an affix, these words will be related to one another in the lexicon and the existence of the affix will emerge, as shown in figure 8.1 (Bybee 1985). The more forms that bear the affix, the stronger the representation of that affix. The stronger the representation of the affix, the easier it will be to access when a new word needs to be inflected, and the greater likelihood that that affix will be productive.

If these hypotheses are correct, then forms of high token frequency will be more autonomous and more likely to be unanalyzed and less likely to participate in schemas; high token frequency forms will thus not contribute to the productivity of a pattern (Guillaume 1927/1973; Bybee 1985; Moder 1992). For example, two English irregular classes of approximately the same type frequency exhibit differences in productivity because one has more high-frequency members than the other. The *strung* class with thirteen members, whose total token frequency according to Francis and Kučera (1982) is 199, is much more productive than the *swept* class, which has fourteen members with a total token frequency of 656 (Moder 1992). Baayen and Lieber (1991) also argue, based on intuitive judgments of productivity of derivational morphology, that productivity is governed by type frequency and that high token frequency can actually detract from productivity. Wang and Derwing (1994) on English and Lobben

TABLE 8.1. Count of Verbs Used by French Nursery School Children during Play

Conjugation Class	*Number of Uses*	*Number of Verbs*
First (*chanter*)	1060 (36.2%)	124 (76.0%)
Second (*finir*)	173 (6.0%)	10 (6.1%)
Third (*vendre*)	1706 (57.8%)	29 (17.9%)

From Guillaume (1927/1973).

(1991) on Hausa (to be discussed below) find the likelihood of the use of a pattern in nonce-probe tasks corresponds to the type frequency of that pattern in existing forms.

Given this interaction of type and token frequency, it follows that a high-frequency member of a (semi-)productive class does not contribute to that productivity. Thus *begin*, the highest frequency member of the *ring*, *rang*, *rung* class, may not actually be associated with this class in the mental lexicon of real speakers. (We can observe that no new members of this class have two syllables or end in the alveolar nasal.) Moder (1992) tested this hypothesis by priming subjects with high- and medium-frequency members of English irregular verb classes before presenting them with lists of nonce verbs resembling members of these classes. When primed with medium-frequency irregulars, the subjects produced more nonce irregular past-tense forms than when primed with high-frequency irregulars. While it might seem intuitively that individual high-frequency irregular forms would better serve as models for new formations, the fact that high-frequency items tend to be stored unanalyzed makes them less usable than items that enter into relations with other lexical items.

Having now explained how high type frequency contributes to productivity, it is necessary to add that an additional factor needs to be taken into account. A pattern cannot attain full productivity if there are restrictions—phonological, semantic, or morphological—on its applicability. The *strung* pattern favored for verbs beginning in consonant clusters and ending in a velar and/or nasal will never be fully productive, since most verbs do not meet these conditions. Thus productivity also depends upon the openness of the defining schema. Only a schema that is totally open, such as the English past in *-ed* or plural in *-s*, can attain full productivity. This second factor explains how new patterns can attain high type frequency: if they have an open schema, they will gradually increase in type frequency. Two cases where openness of the schema and high type frequency do not correspond are discussed on pages 178–182. We turn now to the arguments that have been offered against the role of type frequency in determining productivity.

4. The evidence from German participles

Clahsen and Rothweiler (1992) discuss the German system of past-participle formation in which there are two possible suffixes, *-en* and *-t*. The former suffix occurs on verbs that also undergo some stem changes, and the latter suffix usually entails no stem changes, although a small set of thirteen verbs has both the *-t* suffix and stem changes. Clahsen and Rothweiler present considerable evidence that the productive or regular affix for past participles in German is the *-t* affix. They claim that this fact is theoretically interesting because the *-t* affix does not have a higher type frequency than the *-en* affix. Thus they claim that the determinant of productivity or regularity is not type frequency.

The figures for type and token frequency that they present from Ruoff (1990) are from the first 1,000 verbs he lists, which constitute 96% of all tokens, as shown in table 8.2. If mixed and weak verbs (all those with the *-t* suffix for the past participle) are taken together, their type and token frequencies are very similar to those

TABLE 8.2. Type and Token Frequency of the Most Frequent 1,000 German Verbs

	Tokens	*Types*
Strong verbs	47%	50% (502)
Mixed verbs	32%	5% (50)
Weak verbs	17%	45% (448)

Mixed verbs have the *-t* suffix and some internal stem changes. *Examples*: Strong: *schlafen*, *schlief*, *geschlafen* 'sleep'; Mixed: *rennen*, *rannte*, *gerannt* 'run'; Weak: *kaufen*, *kaufte*, *gekauft* 'buy'.

Based on Ruoff (1990).

of the strong verbs. Thus, Clahsen and Rothweiler argue, the type frequency of the weak and mixed verbs cannot be responsible for the productivity of the *-t* suffix.

This argument depends to a great extent on the way type frequency is counted for German verbs, and I argue that the method chosen by Clahsen and Rothweiler (1992) artificially inflates the type frequency of the irregular verbs by counting the same verbs with productive prefixation as instances of different types. When they first introduce the strong verbs, they note that there are approximately 160 of them. However, by counting verbs with productive prefixes as different types, they substantially increase the number of strong verb types. In addition, by cutting off the count after the first 1,000 verbs, they count only one-quarter of the total number of verbs in the Ruoff database, which substantially lowers the number of weak verb types because they tend to be of lower frequency than the strong verbs. Let us consider these two problems.

The productive prefixes on German verbs are known as "separable prefixes" because they occur in two positions: prefixed to the infinitive or participial forms of the verb, or at the end of the clause when the verb is in a finite form of the present, past, or imperative, as shown in the following examples:

(1) Seine Frau wird morgen *abreisen.*
his wife FUT tomorrow away-travel
'His wife will set out tomorrow'.

(2) Er *reist* heute *ab.*
he travels today away
'He sets out today'.

Even when prefixed, the separable prefix comes before the past-participle prefix *ge-*, as in the form *abgereist.* This German formation is productive, even though not all combinations still have predictable meaning. Thus the prefix *aus*, meaning 'out', occurs in combinations such as: *ausgehen* literally 'out-to go'), with the meaning 'go out' or 'end, finish'; *ausdenken* ('out-to think'), with the meaning "think out' or 'elaborate'; and *ausschreiben* ('out-to write'), meaning 'write out', 'announce', or 'advertise'. This German construction is fully analogous to the English verb-particle combinations that are distinguishable by the fact that the particle may be separated from the verb by an object pronoun or noun phrase. The following are some examples:

(3) I *wrote* the number *down.*

(4) We *thought* it *out* carefully.

Even though the combination of the verb and particle constitutes a lexical item by virtue of its semantic coherence, there is no question that the verbs *write* and *think* in these examples are identified as instances of the simplex verbs *write* and *think* and not different lexical types. Part of the evidence for this is that they have the same morphological properties as the verbs that occur without the particles, while verbs homophonous with strong verbs but constituting different types are regular: English *ring*, *ringed* (as in "to form a ring around") and *grandstanded.* The same is true for German verbs with and without their separable prefixes: strong verbs stay strong with and without the prefixes.

In their count of the type frequency of German verbs, Clahsen and Rothweiler (1992) counted the same verbs with different prefixes as different types. The verb *schreiben* ('to write'), for instance, was counted nineteen times because it occurred as a simplex verb and in eighteen combinations with separable prefixes. This would be comparable to counting each instance of English *write* in *write up*, *write down*, *write over*, and so on, as distinct verb types rather than as the same verb.

The frequency count used by Clahsen and Rothweiler, that of Ruoff (1990), also gives a count of *Grundverben*—that is, the base verbs, in which all instances of a verb, with and without separable prefixes, are counted as one type. In this list, there are 1,258 base verbs, of which 150 are strong verbs and have the *-en* suffix. The remaining 1,108 are weak or mixed and have the *-t* suffix. Thus when only base verbs are counted, weak and mixed verbs constitute 88% of the total and strong verbs only 12%. Contrary to the claim made by Clahsen and Rothweiler, the relative distribution of strong to weak verbs in German is very similar to that of English. It is no surprise, then, that the weak verb participle suffix *-t* is the productive suffix in German.[3]

As for their token frequency, if we exclude the auxiliaries *haben* and *sein*, which account for about one-half of the verb tokens, there are 56,331 verb tokens in Ruoff's database of 500,000 words. Of these, 23,608 are strong verbs, leaving 32,723 weak verbs (and a few mixed ones). Given only 150 strong verb types, it is clear that, as in English, the strong verbs have relatively high token frequency. As argued earlier, high token frequency actually detracts from productivity, because lexical items that are extremely frequent can be learned more easily and stored without forming lexical connections with other related items. Thus forms with high token frequency do not necessarily contribute to the strength of schemas even if by their form they appear to fit these schemas. It is, then, those schemas that are reinforced by medium- and low-frequency items that tend to be the strongest and have the greater productivity. For this reason, it is not legitimate to cut off the count of German verb types after the most frequent 1,000. Ruoff (1990) lists 4,414 verbs. Among the 3,414 that were not considered by Clahsen and Rothweiler, there would be many more weak verb types than strong verb types, since the former tend to be less frequent. Thus by counting only one-quarter of the database they have not given an accurate representation of the relative number of weak verb types.

The conclusion here, then, is that the German past participles do not provide evidence against the proposal that productivity of a morphological pattern is directly related to the number of lexical items that display the pattern. In fact, this case constitutes another instance in which productivity corresponds to high type frequency.

5. German and Arabic plurals

As mentioned earlier, there appear to be at least two main factors that determine the productivity of a morphological pattern: the type frequency and the extent to which the defining schema of the pattern is open to a wide variety of phonological types (see Janda 1990). In the case of English past-tense *-ed* and the German past participle in *-t*, both of these factors are operative—both patterns have high type frequency and a completely open schema. In the cases to which I now turn, only the latter factor is operative. The German plural in *-s* and the Arabic "sound" plural have open schemas but relatively low type frequency. These cases allow us to assess the role of type frequency in determining productivity.

Marcus et al. (1993) argue that these two morphological formations are governed by "symbolic rules," since they demonstrate productivity, and that their productivity cannot be the result of type frequency. However, their experimental results on German, those of Köpcke (1988), and German child language data show that the German plural *-s* has a much lower degree of productivity than English past-tense *-ed.* I will argue that the difference in productivity is a direct result of the difference in type frequency.

For the following discussion, it is useful to distinguish between the terms *productive* and *default*. I have been using a gradient notion of productivity coinciding with the ability of a pattern to apply to novel forms. The term *default* as used by Marcus et al. (1993) and Prasada and Pinker (1993) is different: it is not intended to be gradient but is meant to single out the methods of inflection that are used in various "emergency" circumstances when a plural or other inflected form is non-existent or unknown. The default pattern must be capable of occurring with words of any phonological makeup, so it must be completely open in the way it defines the items to which it may apply. It is this property that leads Marcus and co-authors (1993) and Prasada and Pinker (1993) to regard default patterns as processed by symbolic rules. Marcus and co-authors (1993:7) say: "The hallmark of a symbol is that it can uniformly represent an entire class of individuals, suppressing the distinctions among them." Thus, they would regard the *-s* suffix of German pluralization as applying only to words marked as nouns, with no further specification. Symbolic rules are qualitatively different from schemas that emerge from lexical patterns. Symbolic rules are independent of any of the forms to which they apply and are unaffected by lexical patterning.

In the following, we will see that the German *-s* pluralization is used in various emergency situations as a default. However, this does not mean that this pattern is processed by symbolic rule. Rather, evidence indicates even this default pattern functions like a lexical schema—its application is highly affected by existing lexical items.

Table 8.3 shows the five German suffixes used for plural formation and indicates their type frequency according to a study by Janda (1990) based on a corpus of 600,000 words in tape-recorded interviews (Pfeffer 1964). The percentages given are from a count of the 200 most frequent nouns in that corpus. The table shows that there are two fairly robust patterns of pluralization for German nouns. What it does not show is that there is also a strong association of pluralization in -(*e*)*n* with feminine nouns, of *-e* with masculine nouns, and of *-er* with neuter nouns (Köpcke 1988). In general, pluralization in German is heavily influenced by gender assignment. The conditioning for the zero allomorph is phonological: it most commonly occurs with nouns ending in *-en*, *-el*, and *-er*. However, again gender comes into play: the zero occurs with masculine or neuter nouns with these endings but not with feminines.

The *-s* alternative has extremely low type frequency in this sample of 200 nouns. It should be noted that 200 nouns is an extremely small sample and may not be relevant at all for determining productivity. (German and English both have nearly 200 strong verbs each, but these do not include very many productive patterns.) In a larger corpus of 6 million words, where there are 4,571 noun stems, the type frequency of *-s* increases to 7.2% of the types (Marcus et al. 1993). Such an increase is expected of a pluralization type that tends to occur with infrequent words. The type frequency is still very low, but as noted earlier, high token frequency detracts from productivity, so that it is the type frequency among infrequent words that is the most important determinant of productivity. However, for the sake of the argument, we will assume

TABLE 8.3. German Noun Plural Formation in the 200 Most Frequent Nouns

	Singular	*Plural*	*Gloss*
-(e)n 42%	die Strasse	die Strassen	'the street'
	die Frau	die Frauen	'the woman'
	das Bett	die Betten	'the bed'
-e 35%	der Hund	die Hunde	'the dog'
(+ umlaut)	die Kuh	die Kühe	'the cow'
zero 12%	der Daumen	die Daumen	'the thumb'
(+ umlaut)[a]	die Mütter	die Mütter	'the mother'
	das Leben	die Leben	'the life'
-er 10%	das Kind	die Kinder	'the child'
(+ umlaut)	der Wald	die Wälder	'the forest'
-s 1%	das Auto	die Autos	'the car'
	der Park	die Parks	'the park'

[a]Umlaut (the fronting of a back vowel) is a marker of plural sometimes used in conjunction with a suffix.

Based on Janda (1990).

that the type frequency of *-s* is too low to account for its productivity, and we now turn to the question of just how productive *-s* is.

The evidence for productivity shows some competition among the pluralization patterns. Children generalize the *-(e)n* plural most frequently (Park 1978; Mills 1985; but see also Clahsen, Rothweiler, & Woest 1992). Recent loan words favor *-(e)n* for feminines and *-s* with masculines and neuters. The plural *-e* is also strongly represented in masculine gender borrowings (Köpcke 1988:325). Marcus et al. (1993) summarize Köpcke's analysis by saying that about half of the 182 recent loan words he analyzed took *-s* as the plural suffix. In Köpcke's nonce-probe task, *-n* was strongly favored for nouns ending in schwa in all genders, and for monosyllabic nouns *-en* was favored for feminines and *-e* for masculines; neuters were split between *-en* and *-e*. Fewer than 15% of monosyllables had responses with *-s*. The *-s* suffix was favored only in nouns ending in a full vowel, an environment that almost always takes *-s* in the German lexicon (Janda 1990).

Compared with English past-tense *-ed* and German past participle in *-t*, which strongly outweigh other alternatives in borrowed words and nonce forms, the German plural *-s* shows only weak productivity and in fact is used less often than other alternatives. It does seem to be favored under certain conditions, however. In a task in which subjects evaluated the naturalness of the plurals of nonce nouns, a distinction was made among ordinary nonce nouns, proper names, and borrowings, by introducing the nonce words in appropriate contexts that identified their status (Marcus et al. 1993). The results showed that for nonce words that were intended to be German common nouns, a suffix other than *-s* was preferred; for proper names, the *-s* suffix was preferred; and for borrowings, a suffix other than *-s* was preferred if the nonce root rhymed with an existing root and *-s* was preferred if the nonce root did not rhyme with an existing root. These results can be summarized by saying that if the nonce form could be treated as an ordinary common noun, then the preferred method of pluralization was one of the suffixes other than *-s*, but if it could not be regarded as a German common noun, the suffix *-s* was preferred.[4]

From these facts, Marcus et al. (1993) appropriately conclude that *-s* is the "default" marker of plural in German. They go on to argue that since *-s* does not have a high type frequency, its status as the default allomorph cannot be due to its having formed a strong schema in an associative network. They argue that the pluralization process that uses *-s* is a symbolic rule that is qualitatively different from the lexical patterns that account for the productivity of the other pluralization patterns. As such, it is on a par with the English Past Tense *-ed* and the German past participle *-t*, even though it does not have a comparable type frequency. Their point is that type frequency is not a determinant of productivity or "default" status.

While Marcus et al. have made an adequate case for regarding the *-s* of German plural as an "emergency" or default pluralization process, the data do not support the conclusion that morphological formations with *-s* are unaffected by type frequency nor that *-s* pluralization is qualitatively different from the other pluralization patterns, in that it is an independent, symbolic rule. The data on German *-s* pluralization confirm the prediction that the openness of the phonological definition of a schema is independent of the type frequency of the schema and that it is type frequency that is important in determining productivity. Even though the *-s* pluralization pattern has totally open defining conditions, it shows reduced productivity compared with the English past-tense *-ed* or the German past participle in *-t*, because of its low type frequency. In particular, since it is restricted to loan words and proper names, it is productive only within these lexical domains and for certain special formations that have no established plural. Since type frequency affects lexical schemas but not symbolic rules, the evidence suggests that the *-s* pluralization pattern is a lexical pattern, represented by a schema rather than a symbolic rule.

If *-s* pluralization is a lexical schema rather than a symbolic rule, then existing lexical items will have an effect on its application to nonce items. This prediction is also borne out in Köpcke's (1988) experiment: when he asked subjects to supply plural forms for nonce words that were intended to be normal German nouns, the percentage of *-s* plurals was very low (ranging from 1 to 20%, depending upon the phonological shape and gender of the stem), except for nouns that ended in a full vowel, where *-s* plurals were used 69% of the time. This result reflects the fact that almost all existing German nouns that end in a full vowel form their plural with *-s*.[5] Prasada and Pinker (1993) and Marcus et al. (1993) state that a primary difference between associative lexical patterns and symbolic rules is that the former are affected by the distribution of items in the lexicon while the latter are not. Since *-s* pluralization is favored for nonce nouns matching the phonological shape of a cluster of real nouns that take *-s*, the conclusion must be that *-s* pluralization is not a symbolic rule but rather a lexical pattern. Even though it can be used as the default plural, it is not qualitatively different from other pluralization mechanisms.

Another case that has been invoked to argue in favor of a qualitative difference between schemas and rules that is not based on type frequency is the pluralization process in Arabic. The productive pattern in Arabic, called the "iambic broken plural," requires changes internal to the noun root and affects nouns of different phonological shape in different ways (McCarthy & Prince 1990). The following are some examples from McCarthy and Prince (1990:217):

5.	*singular*	*plural*	*gloss*
	nafs	nufuus	'soul'
	rajul	rijaal	'man'
	jundub	janaadib	'locust'

McCarthy and Prince (1990:214) cite this pattern as "truly productive" because it will apply to any loan word that has a stem that meets its canonical shape criteria. However, there is another pattern, involving suffixation (called the "sound plural"), that can be applied in what might be regarded as "default" circumstances, such as with proper names, transparently derived nouns or adjectives, diminutives, non-canonical or unassimilated loans, and the names of the letters of the alphabet, which are mostly non-canonical (McCarthy & Prince 1990:212). The sound plural has a very low type frequency, since it is only used with non-canonical stems. As might be expected, however, children overgeneralize both patterns (Omar 1973).

This situation appears to be very similar to the situation with German pluralization. There is no argument about the sound plural being the default pattern, but it does appear that its productivity is extremely limited, due to its low type frequency.

Again, it appears that type frequency does affect the degree of productivity of a default pattern.

In this section, we have seen that the two cases put forth as evidence that type frequency does not determine productivity are both cases where the morphological pattern used in emergency situations has very low type frequency. In both cases, however, the patterns in question have very limited productivity compared with patterns with high type frequency in other languages (e.g., English plural *-s*, past-tense *-ed*, and German weak participles). I have argued that this limited productivity is due to the low type frequency. Furthermore, in the case of the German *-s* plural, there is experimental evidence that shows that existing lexical patterns (the fact that nouns ending in full vowels pluralize with *-s*) affect the application of the default pattern to new items. Thus even though the *-s* plural is the default pattern, it is represented in a lexical schema rather than as a symbolic rule.

6. Product-orientation and Hausa plurals

As explained earlier, lexical schemas involve both product-oriented and source-oriented generalizations. Product-oriented schemas are generalizations over non-basic forms rather than generalizations about the relation of a non-basic form to some underlying stem or base form. The product-oriented schema specifies only the phonological shape of the member of the category but does not specify how that shape is to be achieved. In the nonce experiment of Köpcke (1988), some subjects pluralized nouns ending in full vowels, such as *der Treika* and *die Kafti*, by replacing the final vowel with the suffix *-en*, yielding *die Treiken* and *die Kaften.* The resulting plurals matched those in the lexicon, and yet the procedure required more than just adding a suffix to the input noun. The subjects were aiming at a product that matches other lexical items, and they adjusted the nonce words to achieve this result.

A generative "symbolic" rule corresponds to a source-oriented schema in that it specifies the input shape and details the procedure for changing it. An affixation rule that makes no changes to the stem could be described by a source-oriented rule: for example, for German plurals, suffix *-s* to the end of the stem. A symbolic rule cannot describe a product-oriented generalization. It appears that the dual-processing model does not allow for product-oriented generalizations. as they are not describable by symbolic rules and the lexical network used for irregulars in this model generalizes over base–past pairs such as *swing–swung* and *string–strung* (Marcus et al. 1992:122–123).

In the cases I have examined so far, the default inflection has been more agglutinative or concatenative than the other lexical patterns. Thus the English irregular verbs are formed with vowel and sometimes consonant changes to the stem, while the regular formation involves a suffix and no stem change; the German past participle in *-en* often involves a vowel change, while the pattern with *-t* rarely does; some of the German plural suffixes are accompanied by umlaut, but the suffix *-s* never is; and the Arabic iambic plural involves stem changes while the sound plural is suffixation. Marcus et al. (1993:7) imply that they always expect the default pattern to involve agglutination by calling the processing type for the default a "symbol-concatenating rule." If this is the correct interpretation, then another proposed qualitative difference between lexical patterns and symbolic rules has emerged: lexical patterns may involve affixation or internal changes to the stem, but symbolic rules are strictly affixing. Since stem changes typically involve product-oriented schemas and symbol-concatenation is necessarily source oriented, then we can conclude that default rules in their sense must be source oriented. In this section and the next, I will argue that default status or productivity is not necessarily associated with source-oriented rules. First, I will present data from Hausa that show that the productive morphological patterns for plural are all product oriented. Then, in the next section, I will argue that the usual association of source-oriented patterns with default status is due to the way new affixes tend to develop diachronically.

As just mentioned, in Hausa even the most productive plural patterns are product oriented. One analysis of Hausa nouns proposes ten lexical classes, using a combination of criteria that include properties of the singular as well as the plural (Kraft & Kirk-Greene 1973). More recent analyses by Haspelmath (1989) and Lobben (1991) point out that there are a number of arguments for analyzing Hausa plurals in terms of product-oriented schemas that reference the plural forms only. First, there is little predictability between singular and plural form. Just to cite one problem, Lobben (1991:112–113) notes that of ninety-five nouns that end in *-aa* and have low–high–low tone patterns, forty-eight take a plural in *-uu* with a low . . . high tone pattern and forty-seven take a plural in *-ooCii* with a high tone.

Second, the plural suffix replaces part of the end of the singular noun, yet the part replaced cannot be considered a singular suffix. For instance, in the singular/plural pairs *tàmbáayàa*, *támbáyóoyíi* (‘question’) and *táagàa*, *táagóogíi* (‘window’), the final vowels *àa* are replaced by *óo* + a copy of the last consonant + *íi* (examples from Lobben 1991:3). In these examples, the acute accent (*á*) represents a high tone and the grave accent (*à*) represents a low tone.

But the strongest evidence for product-oriented schemas over plural forms is the fact that for both real and nonce nouns a single plural pattern can be arrived at in a variety of ways. For example, Haspelmath (1989) gives the following examples of the plural pattern HL *-úCàa*:

(6)	a.	rìigáa	ríigúnàa	‘gown’
		kàntíi	kántúnàa	‘store’
	b.	záurèe	záurúkàa	‘porch’
		hánnúu	hánnúwàa	‘hand’
	c.	árzìkíi	árzúkàa	‘prosperity’
		tákòobii	tákúbàa	‘sword’

As can be seen in (6), the plurals all have a consistent tone pattern of high–high–low and they all end in *-úCàa*, where *C* is /n/, /k/, or /w/ or a consonant that is part of the root. In (6a) and (6b), the final vowels of the singular appear to have been removed and a suffix added. But in (6c), the plural morpheme is fitted like a template over the last two syllables of the singular form. Thus it is possible to formulate a single product-oriented schema that describes the shape of the plural form, but it would take numerous source-oriented rules to describe how to derive a plural from a singular.

Since a similar situation holds for all the plural patterns of Hausa and since Lobben found strong evidence for product-oriented schemas in her nonce experiments with Hausa speakers, she argues that Hausa pluralization is processed entirely by product-oriented schemas. A further argument, put forward by Haspelmath, is that the individual plural schemas have properties in common, so that it is possible to formulate a meta-schema that unites the various methods of plural formation. Generalization at this level is not possible using source-oriented rules.

Lobben's nonce word task showed that there is not just one productive pattern but rather three patterns that show some degree of productivity, each one accounting for more than 20% of the nonce responses. Lobben (1991:223) states that the degree of productivity of a pattern is in direct correspondence with the type frequency of that pattern in Hausa nouns, which is precisely what we would expect from lexical patterning. It is not clear whether or not there is just one "emergency" plural formation strategy in Hausa. The most productive schema in the nonce experiment (high *-óoCíi*), garnering 29% of responses, is also the one used with recent loan words, especially from English. For example, *ásìbítìi* ('hospital') has the plural *ásíbítóocíi*, and *téebùr* ('table') has the plural *téebúróoríi.* According to Lobben (1991:223), this schema also has the highest type frequency, and it apparently also has the most open phonological definition for singulars (Lobben 1991:113), occurring with the highest number of combinations of finals and tone patterns (at least for trisyllabic nouns). It is also the only schema that applies to words (which are all loans) that end in consonants (see the word for "table" above).

The evidence, then, points to this one schema as being (or becoming) comparable to the "default" pattern in other languages. Yet this pattern cannot be captured in a "symbol-concatenating" or source-oriented rule; it has the same properties as other patterns that are represented in lexical associations, including a prototype structure. In the nonce experiment, Lobben found extensive evidence for a product-oriented treatment of this pattern. In forming plurals with this pattern, subjects added or omitted phonemes in the singular or made use of the singular consonants and inserted the appropriate vowels but always arrived at a plural that precisely fitted the schema.

It appears, then, that there is at least one language in which an obligatory inflectional category does not have a symbolic rule to fall back on for creating new plurals but rather uses only the patterns that emerge from lexical associations. This is not meant to be a devastating argument against the dual-processing model as proposed by Pinker (1991), Marcus et al. (1992, 1993), and Prasada and Pinker (1993). Dual-processing could exist but simply not be used for this inflectional category of Hausa. This case does show, however, that there is no necessary correspondence between the most productive or default patterns and symbolic rules.

7. Diachronic excursus

Despite the existence of languages such as Hausa, it is often the case that the productive pattern for a morphological category is a concatenative process, describable in source-oriented terms, while the irregular patterns tend to have more of an effect on the stem and are better describable in product-oriented terms. This cross-linguistic pattern in itself might be taken as evidence for the dual-processing proposal were it not for the fact that this pattern has a straightforward diachronic explanation. In many cases, the most productive pattern is the newer one, derived more recently from a periphrastic construction. In such cases, the diachronic source accounts for both the concatenative structure of the formation and its ability to apply to a wide variety of phonological types.

For instance, the English *-ed* derived from the past form of the ancestor of the modern verb *do*, in a periphrasis in which the past form of *do* stood after the main verb form. Such periphrases usually arise as strategies for forming verbs from nouns and adjectives and are extended to use with loan words, as a way of inflecting them. A strategy such as this is particularly welcome when the existing inflectional patterns requiring stem changes are not easily extendable to new lexical items and at times when the language is experiencing an influx of borrowed words. These newer, more transparent, inflections can increase in type frequency and hence productivity to the extent that they replace the older means of inflection. Since they are originally periphrastic—that is, expressed as separate words—they impose no phonological shape requirements on the main verb, which also helps them increase their productivity.

A verb meaning ‘do’ or ‘be’ is usually the auxiliary verb used in such cases (e.g., in Japanese or Basque), but other verbs are possible. A particularly interesting example occurs in the African language Kanuri (Lukas 1937; Hutchison 1981), where there are two inflectional classes for verbs. One class consists of approximately 150 verbs, entails complex phonological and morphological irregularities, and is unproductive. The other more regular, productive class adds affixes to verb stems that are forms of the verb *ngin*, meaning ‘to say’. Hutchison (1981:102–111) argues that the original members of this class must have been verbs derived from onomatopoeic words (hence the appropriateness of the verb “to say”) and that the formation spread from this specific use to use with derived verbs from other lexical items and loan words.

Since the major source of new morphology is grammaticization of separate words, transparent affixes are more easily formed diachronically than patterns that affect a lexical stem (Heine & Reh 1984; Heine, Claudi, & Hunnemeyer 1991; Bybee, Perkins, & Pagliuca 1994). In fact, stem changes are almost always derived from phonological changes in the stem conditioned by an affix (e.g., Germanic umlaut). This means that it typically takes a very long time for stem changes to develop. Since grammaticization of new affixes is occurring at all times, older affixes may be replaced before they have a chance to fuse with the stem to an extent sufficient to create stem changes. Thus the greater frequency of affixes over stem changes is due to the way the two patterns arise diachronically and may have no bearing on the type of processing involved for the two patterns. While some (e.g., Dressler 1985:316–319) would argue that affixes present an acquisition or processing advantage over stem changes, which would coincide with the use of symbolic rules, Bybee and Newman (1995) have shown experimentally that subjects can internalize and generalize stem changes as readily as affixes, suggesting that both are processed in the same manner.[6]

8. English-speaking children's overregularization of past tense

For decades, the most convincing evidence for the existence of productive morphological rules was to be found in children's overgeneralizations, such as English-speaking children's regularizations of irregular past-tense forms. As traditionally understood, children's performance on English verbs appears to yield a U-shaped learning curve. In the first stage, children produce irregular past forms correctly, but after acquiring a number of regular past forms, they begin to produce regularized past forms, such as *goed* and *breaked* (the second stage). Only later do they sort out the irregular from the regular and correctly produce *went* and *broke* again. It is the second stage—the stage in which regularized pasts are produced—that provides the strongest evidence available for rule formation and application in children, because it is at this stage that children produce forms that they do not hear from adults.

Two recent developments now cast doubt on this phenomenon as evidence for symbolic rule formation in children. The connectionist modelers discussed on page 169 have shown that a parallel distributed processing model, which does not formulate symbolic rules, can nonetheless simulate the three acquisition stages described earlier, given input that reflects the type and token frequency of English regular and irregular verbs. In a second development, Marcus et al. (1992) examined regularizations in over 10,000 tokens of children's speech and found that the percentage of regularizations is unexpectedly low—only 2.5% of English irregular tokens are regularized. Even though individual children regularize at varying rates (from less than 1% to as much as 24%), Marcus et al. (1992:35) conclude: "the global data suggest that overregularization is a relatively rare phenomenon."

This result is surprising because regularizations are so salient in children's speech; apparently, correct irregulars go largely unnoticed by adult observers. Since a presumed high rate of regularization has always been taken by linguists as evidence that children formulate rules, the actual rarity of regularizations casts doubt on the proposition that children internalize a symbolic rule for regular past tense. Low rates of regularization demonstrate that the difference between regular and irregular pasts in English is much smaller than originally supposed and more consistent with a uniform treatment of regular and irregular morphology with lexical representation and lexical schemas for all past forms.

Marcus et al. (1992) do not interpret their data this way. Even in the face of this evidence, they maintain the notion that children formulate a symbolic rule for regular past tense. They explain the low rate of application of this rule to irregulars as the result of a very strong blocking device that suppresses regularizations. Thus their model contains (1) lexical representations and schemas (for irregular forms), (2) symbolic rules, *and* (3) a powerful blocking device to keep the symbolic rules under control. This can be compared to the more parsimonious model proposed here that requires only (1)—lexical representations and schemas of varying degrees of strength and generality. The network model does not require a blocking device to prevent the regular pattern from overapplying, since it is not a categorical rule anyway but simply derives its productivity from its type frequency.

In Rumelhart and McClelland's (1986) simulation, the onset of regularization coincided with a sharp increase in the input of regular past-tense formations. The data examined by Marcus et al. did not reveal any chronological relationship between the number of regular verbs in the input (or the child's output) and the onset of regularization. Instead, Marcus et al. argue, convincingly in my opinion, that the onset of regularization is a result of the child's learning that the marking of tense is obligatory in English. Having established that tense-marking must occur, the child is forced into situations in which a past-tense form must be supplied, even if the appropriate form is not represented in the child's lexicon or is perhaps only very weakly represented. In the dual-processing model, the symbolic rule takes over in these cases. In the network model, other lexical patterns are accessed, and the strongest of these is the one regarded as regular in English. The important point here is that it is not the formulation of the symbolic rule that triggers regularization but rather the child's formulation of the (correct) hypothesis that tense-marking is obligatory in English. The existence of regularizations is independent of the issue of whether or not children formulate a symbolic rule.

In the dual-processing model, which includes both lexical schemas and symbolic rules, the question arises as to how the symbolic rule is acquired. In order for the child to discover the rules, forms with and without the suffix (*walk* and *walked*) must be stored, compared, and the pattern extracted from them. In fact, quite a number of such forms must be stored before the child can conclude that the pattern is not in fact appropriate for a lexical schema but must rather be represented in a separate component as a symbolic rule. Thus the child must switch from the lexical schema strategy for representing the regular pattern to the symbolic rule strategy.

Is there any evidence for such a switch in processing strategies? When and how does this switch take place? Since Marcus et al. (1992) did not find any dramatic spurt of regularizations in the data they examined, the point of restructuring from lexical schema to symbolic rule cannot be identified in this way. In fact, for most children the rate of regularization, once begun, remains fairly stable. The emergence of regularization itself cannot be used to identify the point of change in processing type, since this is argued to be due to the emergence of tense-marking as obligatory. Implicit in Marcus and co-workers' argument is the supposition that the symbolic rule is already formed when regularizations begin. While Marcus et al. list a number of possible ways in which the child could determine which past-tense formation process is the productive or default one, they do not address the question of exactly how the rule is extracted and deposited in a separate module. Thus despite the fact that the relevant section in their paper is named "How Might a Regular Rule Be Learned?" the question of how the restructuring from lexical schema to symbolic rule takes place is actually not addressed.[7]

9. Evidence for the representation of high-frequency regulars

The most innovative aspect of the model proposed in Bybee (1985, 1988) is the hypothesis that high-frequency regularly inflected words are stored in the lexicon, while low-frequency regulars are derived in the lexicon by applying the strongest schema to base forms. This follows from the lexical strength proposal—for regular forms just as for irregular forms, lexical strength varies according to frequency of use. Because of the extremely open nature of this schema, experiments have turned up no evidence for effects of lexical distribution such as that found in the case of more restricted schemas (Stemberger & MacWhinney 1988; Prasada & Pinker 1993). An approach taken by Losiewicz (1992), however, has resulted in evidence in support of differential representation for high- and low-frequency regular forms.

In a series of experiments, Walsh and Parker (1983) found that English /s/ in word-final position is longer in acoustic duration if it is the plural morpheme (*laps*) than if it is part of a monomorphemic word (*lapse*). Losiewicz (1992) found this same distinction applies to morphemic /d/ or /t/ (the past tense) as in *rapped* versus non-morphemic final /d/ or /t/ as in *rapt.* Losiewicz further reasoned that if the difference in length is due to the non-morphemic segment being part of a lexical representation, while the morphemic one is added to the stem in processing, then the same difference in length should appear in low- versus high-frequency words with morphemic /d/, since the low-frequency words would be formed by using a schema and the high-frequency words would be accessed directly from the lexicon.

Losiewicz asked subjects to read sentences containing English past-tense forms that constituted rhyming pairs of high- and low-frequency verbs (*covered*, *hovered*; *needed*, *kneaded*). For all subjects and all pairs of verbs, the final past-tense morpheme was longer in the low-frequency verb of the pair. The average difference in duration was 7 msec, a difference that was highly significant and not due to overall differences in word length. It is not proposed that such a length difference is either perceptible or learnable from input but rather that it reflects a difference in processing type. These results, then, can be taken to support the hypothesis that high-frequency inflected verbs are stored in the lexicon, whereas low-frequency inflected forms are produced by applying the strongest schema to base forms. It is conceivable that further investigation of an experimental nature will yield more evidence of differences between high- and low-frequency regular forms.[8]

Faced with the idea that regular morphological forms can be stored in the lexicon just as irregulars are, one could ask why Prasada and Pinker (1993) found no effect of existing lexical verbs in their nonce-probe tasks that used the English regular past tense. The answer is that the regular past tense has a totally open schema together with extremely high type frequency, which means it can apply to virtually any string of sounds and thus shows no effect of lexical distribution.[9]

10. Conclusion

To summarize the arguments given, we have first seen that contrary to the claims of Clahsen and Rothweiler, the case of German past participles is precisely parallel to the case of English past tense in the sense that the productive affix is the one with the highest type frequency. In the case of German plural formation, there is no one pattern with overwhelmingly high type frequency (over 80%, as in the case of English past tense and German participles) and, as predicted, there is no one productive pattern. One pattern of very low type frequency (the *-s* plural) has some productivity, and it is the pattern with the most open phonological definition. The application of this pattern does not exhibit full productivity and does not seem rulelike, since its application to nonce forms is affected by its distribution in the lexicon.

A second property that distinguishes symbolic rules from lexical schemas is source- versus product-orientation. Symbolic rules are necessarily source oriented and schemas may be either source or product oriented. As Hausa plural formation shows, there is no requirement that a language should have a source-oriented means for forming obligatory inflectional categories. All of the means of forming plurals in Hausa are product oriented, and some of these means show productivity. Thus Hausa does not make use of two processing avenues but uses only one—lexical schemas. Even the evidence for a symbolic rule in the case of English past-tense formation is considerably weakened by the finding that the actual rate of regularization of irregular past-tense forms in English-speaking children is quite low. Finally, as to the controversial claim that high-frequency regular past-tense forms are listed in the lexicon, whereas low-frequency forms may not be, Losiewicz (1992) has found a consistent measurable difference in the phonetic shape of the suffix in low- versus high-frequency inflected forms.

The issues discussed here suggest that some systematic distinctions should be made among the terms *productive*, *default*, and *regular*. I propose the following, which are based on the way these terms appear to be used in the relevant literature: *Productive* means that a pattern may be applied to new items. Degrees of productivity are highly correlated with the type frequency of a pattern within a language. *Default* refers to the pattern that applies when all else fails; default patterns have the most open phonological and lexical definitions (where *lexical* refers to features such as loan word and proper noun). The *regular* pattern is the one with the least allomorphy in affix and stem. This term seems also to be used when type frequency and the most open phonological schema converge on the same pattern, as they do in English. It is important to note, however, that productivity, default application, and regularity are independent and may be properties of distinct patterns in some languages (see Bybee 1996).

The data cited here and the arguments given support a functionalist conception of the lexicon, as a store of words and phrases that is highly affected by actual language use (cf. Bates & MacWhinney 1982, 1987). The strength of lexical representations of individual items is in part a reflection of token frequency, while the strength (and one determinant of productivity) of lexical associations or schemas is built up by type frequency. The distinction between storage of complex forms and formation by schema application is based on the availability of stored items, which is determined by frequency of use, not by structural distinctions, such as the classification into regular and irregular patterns. The maintenance of irregularity in high-frequency forms is just one piece of evidence that use overrides structure in determining representation.

The alternative proposal—that structure determines representation—is not supported by the data. The one structural criterion—the openness of the phonological definition of the schema—only gives rise to full productivity in combination with high type frequency. Other aspects of structure, such as agglutinative structure or apparent source-oriented application, are epiphenomenal and due to diachronic patterns. The synchronic reality of varying degrees of productivity is derivable from the way language is used.

Only one "module" is necessary for morphology: a highly structured lexicon. Morphological structure and organization emerge from the connections made among related stored items. Even forms produced by combination are produced in the lexicon by accessing a stem and utilizing a schema that is available in the lexicon. Given that schemas have the same properties as other types of human categorizations, it is distinctly possible that storage and access of morphologically complex words may require very little cognitive ability that is specific to language.

Notes

I am grateful to Martin Haspelmath, Klaus-Michael Köpcke, Beth Losiewicz, and two anonymous reviewers for their comments on an earlier draft of the manuscript.

1. In early generative phonology (Chomsky & Halle 1968), the position was taken that only morphemes were stored in the lexicon and all words with more than one morpheme were formed by rules that concatenated morphemes and by phonological rules that changed feature values. Not long after the presentation of this view, a position that is the polar opposite of it was also defended: in this view, all words of a language, no matter how complex or how regularly inflected, are listed in the lexicon (Vennemann 1974). Naturally, there are a number of possible positions to take between these two poles. I will consider arguments for one of these alternatives in this essay.

2. The network model is compared with level-ordered phonology in Bybee (1996). In the final section of this article, I explain how the network model handles the regular inflection of verbs derived from nouns.

3. The implication of the token frequency distribution of base verbs versus verbs with separable prefixes is that strong verbs occur with separable prefixes more often than weak verbs do. Use with prefixes is one of the factors contributing to their high frequency, which in turn contributes to the preservation of their irregularities.

4. In fact, Köpcke (p.c.) reports that recent loans, which are not yet integrated into the phonological, morphological, and graphemic system, take *-s*. But as soon as they are integrated, the same nouns change from using *-s*, to *-en* (*Rivals* > *Rivalen* 'rival'), *-e* (*Barons* > *Barone* "baron"), or -o (*Computers* > *Computer* 'computer'), depending upon their gender and phonological shape.

5. Köpcke (1993:128–133) shows further that the choice of *-s* versus *-en* pluralization depends upon the quality of the final vowel and the gender assignment.

6. Dressler's discussion and examples also confound morphological transparency with regularity. In Bybee and Newman (1995) we try to disentangle these two variables.

7. Clahsen et al. (1992) argue that German children with "specific language impairment" distinguish regular from irregular inflection. Goad (1996), however, suggests that the network model makes the correct prediction for subjects with specific language impairment. It seems that such subjects are able to memorize highly frequent inflected forms but are not able to form the relations among sets of forms that allow the formation of schemas.

8. Bybee (2000) studied *t/d* deletion in a corpus of American English and found that regular past and participial /t/ and /d/ are more likely to be deleted in high-frequency verbs than in low-frequency verbs. This finding suggests that high-frequency regulars are listed in the lexicon and low-frequency regulars are produced from a schema, in which in the /t/ and /d/ are less reduced because the schema is influenced by its use with low-frequency verbs.

9. Daugherty and Seidenberg (1994) show for a connectionist model that a sufficiently high type frequency and lack of phonological coherence in a pattern leads the model to behave eventually as though the pattern were independent of its lexical distribution.

References

Baayen, R., and R. Lieber. 1991. Productivity and English derivation: A corpus-based study. *Linguistics* 29, 801–843.

Bates, E., and B. MacWhinney. 1982. Functionalist approaches to grammar. In *Language acquisition*: *The state of the art*, ed. E. Wanner & L. Gleitman, 173–218. New York: Cambridge University Press.

Bates, E., and B. MacWhinney. 1987. Competition, variation and language learning. In *Mechanisms of language acquisition*, ed. B. MacWhinney, 157–193. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form.* Philadelphia, Penn.: John Benjamins.

———. 1988. Morphology as lexical organization. In *Theoretical morphology*, ed. M. Hammond and M. Noonan, 119–141. San Diego, Calif.: Academic Press.

———. 1996. Productivity, regularity and fusion: How language use affects the lexicon. In *Trubetzkoy's Orphan*, ed. R. Singh, 247–269. Philadelphia, Penn.

———. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language*, ed. Michael Barlow and Suzanne Kemmer, 65–85. Stanford, Calif.: CSLI. (Reprinted in this volume as chapter 9.)

Bybee, J. L., and Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59, 251–270. ((Reprinted in this volume as chapter 6.)

Bybee, J., and J. Newman. 1995. Are stem changes as natural as affixes? *Linguistics* 32, 33–34. (Reprinted in this volume as chapter 7.)

Bybee, Joan L., Revere D. Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense*, *aspect and modality in the languages of the world.* Chicago: University of Chicago Press.

Bybee, J. L., and D. I. Slobin. 1982. Rules and schemas in the development and use of English past tense. *Language* 59, 251–270.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English.* New York: Harper and Row.

Clahsen, H., and M. Rothweiler. 1992. Inflectional rules in children's grammars: Evidence from the development of participles in German. *Yearbook of Morphology 1992*, 1–34.

Clahsen. H., M. Rothweiler, and A. Woest. 1992. Regular and irregular inflection in the acquisition of German noun plurals. *Cognition* 45, 225–255.

Cottrell, G., and K. Plunkett. 1991. Using a recurrent net to learn the past tense. In *Proceedings of the 13th Annual Meeting of the Cognitive Science Society.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Daugherty, K. G., and M. S. Seidenberg. 1994. Beyond rules and exceptions: A connectionist modeling approach to inflectional morphology. In *The reality of linguistic rules*, ed. S. D. Lima, R. L. Corrigan, and G.K. Iverson, 353–388. Amsterdam: John Benjamins.

Dressler, W. U. 1985. *Morphophonology: The dynamics of derivation.* Ann Arbor, Mich.: Karoma Press.

Francis, W. N., and Henry Kučera. 1982. *Frequency analysis of English usage.* Boston, Mass.: Houghton Miflin.

Goad, H. 1996. Remarks on the network model: Productivity, derivational morphology and atypical populations. In *Trubetzkoy's Orphan*, ed. R. Singh, 270–283. Phildelphia, Benjamins.

Guillaume, P. 1927/1973. The development of formal elements in the child's speech. In *Studies of child language development*, ed. C. A. Ferguson and D. I. Slobin, 240–251. New York: Holt, Rinehart and Winston.

Hare, M., and J. Elman. 1995. Learning and morphological change. *Cognition* 56, 61–98.

Haspelmath, M. 1989. Schemas in Hausa plural formation: Product-orientation and motivation *vs.* source-orientation and generation. *Buffalo Working Papers in Linguistics* 89–101, 32–74.

Heine, Bernd, and Mechtild Reh. 1984. *Grammaticalization and reanalysis in African languages.* Hamburg: Buske.

Heine, Bernd, Urike Claudi, and Frederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework.* Chicago: University of Chicago Press.

Hutchison, J. 1981. *The Kanuri language: A reference grammar.* Madison: University of Wisconsin, African Studies Program.

Janda, R. D. 1990. Frequency, markedness, and morphological change: On predicting the spread of noun-plural *-s* in Modern High German and West Germanic. In *ESCOL '90*, 136–153.

Köpcke, K.-M. 1988. Schemas in German plural formation. *Lingua* 74, 303–335.

Köpcke, K.-M. 1993. *Schemata bei der Pluralbildung im Deutschen: Versuch einer kognitiven Morphologie.* Tübingen: Narr.

Kraft, C. H., and A. H. M. Kirk-Greene. 1973. *Teach yourself Hausa.* London: Hodder and Stoughton.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical prerequisites*. Stanford, Calif.: Stanford University Press.

———. 1988. A usage-based model. In *Topics in cognitive linguistics*, ed. B. Rudzka-Ostyn, 127–161. Amsterdam: Benjamins.

Lobben, M. 1991. Pluralization of Hausa nouns, viewed from psycholinguistic experiments and child language data. M. Phil. University of Oslo.

Losiewicz, Beth L. 1992. The effect of frequency on linguistic morphology. Ph.D. diss., University of Texas, Austin.

Lukas, J. 1937. *A study of the Kanuri language: Grammar and vocabulary.* Oxford: Oxford University Press.

MacWhinney, B., and J. Leinbach. 1991. Implementations are not conceptualizations: Revising the verb learning model. *Cognition* 29, 121–157.

Marcus, G. F., S. Pinker, M. Ullman, M. Hollander, T. J. Rosen, and F. Xu, 1992. Overregularization in language acquisition. *Monographs of the Society for Research in Child Development* 57(4).

Marcus, G. F., U. Brinkmann, H. Clahsen, R. Wiese, A. Woest, and S. Pinker. 1993. *German inflection: The exception that proves the rule.* Occasional Paper No. 47. Center for Cognitive Science, MIT, Cambridge, Mass.

McCarthy, J., and A. Prince. 1990. Foot and word in prosodic morphology: The Arabic broken plural. *Natural Language and Linguistic Theory* 8, 209–283.

Mills, A. E. 1985. The acquisition of German. In *The cross-linguistic study of language acquisition*, vol. 1: *The data*, ed. D. I. Slobin, 141–254. Hillsdale. N.J.: Lawrence Erlbaum Associates.

Moder, C. L. 1992. Productivity and categorization in morphological classes. Ph.D. diss., SUNY at Buffalo.

Omar, M. 1973. *The acquisition of Egyptian Arabic as a native language*. The Hague: Mouton.

Park. T.-Z. 1978. Plurals in child speech. *Journal of Child Language* 5, 237–250.

Pfeffer. J. 1964. *Grunddeutsch: Basic* (*spoken*) *German word list: Grundstufe.* Engelwood Cliffs, N.J.: Prentice-Hall.

Pinker. S. 1991. Rules of language. *Science* 253, 530–535.

Pinker, S., and A. Prince. 1994. Regular and irregular morphology and the psychological status of rules of grammar. In *The reality of linguistic rules*, ed. S. D. Lima, R. L. Corrigan, and G. K. Iverson, 353–388. Amsterdam: John Benjamins.

Plunkett. K., and V. Marchman. 1991. U-shaped learning and frequency effects in a multi-layered perceptron: Implications for child language acquisition. *Cognition* 38, 43–102.

Prasada, S., and S. Pinker. 1993. Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes* 8, 1–56.

Rumelhart, D., and J. McClelland. 1986. On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In *Parallel distributed processing: Explorations in the microstructure of cognition*, ed. J. McClelland, D. Rumelhart, and the PDP Research Group. Cambridge, Mass: MIT Press.

Ruoff, A. 1990. *Haufigkeitswörterbuch gesprochener Sprache.* Tubingen: Niemeyer.

Stemberger, J. P., and B. MacWhinney. 1988. Are inflected forms stored in the lexicon? In *Theoretical morphology*, ed. J. M. Hammond and M. Noonan, 101–116. San Diego, Calif.; Academic Press.

Vennemann, T. 1947. Words and syllables in natural generative phonology. In *Natural phonology parasession*, 346–374. Chicago: Chicago Linguistics Society.

Walsh. T., and F. Parker. 1983. The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics* 11, 201–206.

Wang. H. S., and B. L. Derwing. 1994. Some vowel schemas in three English morphological classes: Experimental evidence. In *In honor of Professor William S.-Y. Wang: Interdisciplinary studies on language and language change*, ed. M. Y. Chen and O. C. L. Tzeng, 561–575. Taipei: Pyramid Press.

Zager, David. 1980. A real-time process model of morphological change. Ph.D. diss., SUNY at Buffalo.

This page intentionally left blank

PART III

PHONETIC CHANGE: FREQUENCY IN CONTEXT

This page intentionally left blank

Introduction to Part III

The chapters of part III address frequency effects in the spread of sound change, the other topic that was part of the early work represented in chapter 2. Chapter 9 looks for and finds evidence that high-frequency words undergo a well-studied sound change in American English—final [t]/[d] deletion—before low-frequency words. Similar results are also found in Gregory, Raymond, Bell, Fosler-Lussier, and Jurafsky (1999), a study of final [t]/[d] deletion in a different corpus. The data also show that the higher rate of deletion found in past-tense verbs that also have a stem change (e.g., *tell/told*) is largely a frequency effect, since these verbs are generally of high frequency. In addition, a frequency effect within this class of verbs is uncovered. A follow-up on the issue of the lower deletion rate in regular past-tense forms is found in chapter 11.

Chapters 10 and 11 examine sound changes that occur at the boundaries between two words, setting up more than one variant for a word. What happens in such cases is interesting, because it is common for bound morphemes to have multiple variants but not so common for words to have variant forms. In the cases discussed in these two chapters, it is shown that the more frequent variant has an impact on the outcome of the change for a word. Chapter 10 takes up final [s] reduction in Spanish dialects, arguing that the spread of the reduced variant from preconsonantal position (the original phonetic environment) to prevocalic position is the generalization of the more frequent variant—that which occurs before consonants. In this chapter, word-internal changes in alternating environments are also studied, showing again an effect of the frequency of the different environments. This chapter and the next make use of an exemplar model to record the range of phonetic variation for individual words. This model is further developed in Bybee (2001) and Pierrehumbert (2001, 2002). This chapter shows the power of the usage-based approach in explaining linguistic phenomena: the effect of frequency can explain why a word boundary so often behaves like a consonant. The reason is that word boundaries are most often followed by consonants and the more frequent variant eventually takes over. A second point is that the data show that individual words tend to resolve a large range of variation toward having a single variant. This phenomenon explains "word-level phonology": the fact that morphemes tend to have variants in different words, but words (unless they are of extreme high frequency) tend to have only a single form. Finally, in this chapter and the next, it is argued that the data on gradual lexical diffusion show that lexical representations cannot be reduced to phonemic representation.

Chapter 11 examines in more detail how phonetic change interacts with words and affixes, using the case of [t]/[d] deletion again and focusing on the reasons that the regular past-tense affix shows a lower rate of deletion than [t] and [d] that belong to lexical morphemes. The data show that past-tense [t] and [d] tend to occur before vowels more than monomorphemic [t] and [d]. Since the context before vowels does not favor deletion, the affix is undergoing the reduction and deletion process at a slower rate. Subsequent work that also shows that it is the context in which the word occurs that determines how readily it undergoes a change can be found in Brown (2004), Alba (2005), and chapter 14 of this volume. Also in this chapter it is shown that the reduction of a consonant in an affix in a uniform environment provides evidence that the affix and the words that contain it have a memory representation that can be affected by frequency of use.

References

Alba, Matt. 2005. Hiatus resolution between words in New Mexican Spanish: A usage-based account. Ph.D. diss., University of New Mexico.

Brown, Esther. 2004. Reduction of syllable-initial /s/ in the Spanish of New Mexico and Southern Colorado: A usage-based approach. Ph.D. diss., University of New Mexico.

Bybee, Joan L. 2001. *Phonology and language us*e. Cambridge: Cambridge University Press.

Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *CLS* 99, 151–166.

Pierrehumbert, Janet. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency and the emergence of linguistic structure*, ed. Joan. L. Bybee and Paul Hopper, 137–157. Amsterdam: John Benjamins.

———. 2002. Word-specific phonetics. In *Laboratory phonology*, vol. 7, ed. Carlos Gussenhoven and Natasha Warner, 101–139. Berlin: Mouton de Gruyter.

Chapter 9

The Phonology of the Lexicon

Evidence from Lexical Diffusion

In many areas of linguistics, examining the nature of diachronic changes leads to a more accurate modeling of synchronic systems.[1] Lexical diffusion—the way a sound change spreads through the lexicon—has, as yet, not been exploited as a potential source of evidence about the phonological shape of lexical representations. This study contains evidence that sound change is both phonetically gradual and lexically gradual and that the rate at which words undergo sound change is positively correlated with their text frequency. This correlation is found in monomorphemic words, regularly inflected words, and irregularly inflected words. I argue that frequency effects in sound change may be explained by assuming that cognitive representations are affected by every token of use.

The sound change about which new evidence will be provided here—t/d deletion in English—has been vigorously studied by sociolinguists seeking to understand the factors controlling the variation found in this process. One of their findings is that the morphological status of the final /t/ or /d/ affects the rate of deletion. Special properties of phonological segments serving as morphemes have also been studied experimentally by psycholinguists. This chapter proposes to bring together evidence from variation, experimental findings and data on text frequency to refine a usage-based model of lexical representation and morphological organization.

The first part of this chapter examines lexical diffusion in general and the claim that frequency affects the diffusion of sound change. The evidence that t/d deletion is more advanced in high-frequency words is presented, and a view of phonetic variation compatible with the data is examined.

The second part of the chapter addresses the morphological effects on the variation in t/d deletion. My morphological model (Bybee 1985, 1988) makes certain predictions about the effect of t/d deletion on regular verbs, and these are tested on the data. It is found that the frequency effect demonstrated for the corpus overall is also found among regularly inflected verbs, suggesting that at least high-frequency regulars are listed lexically. A frequency effect is also found in the double-marked pasts, such as *told* and *slept*. It is argued that the apparent special properties of this class are due to the high frequency of many of its members.

1. Frequency effects in the lexical diffusion of sound change

1.1. Lexical diffusion

The study of lexical diffusion has primarily been directed at accounting for irregular results of sound changes, in particular, lexical items that appear not to have undergone a change (Wang 1969; Krishnamurti 1978). Changes that have occurred over very long periods of time have also been the subject of lexical diffusion studies (Krishnamurti 1978). Such changes can be said to be taking place at the phonemic level and have led Wang and Cheng (1977) to postulate that most sound change is lexically gradual and phonetically abrupt.

A lexically abrupt sound change would affect all lexical items at the same time, while one that is lexically gradual moves through the lexicon affecting individual items or groups of items in sequence. A phonetically gradual sound change takes place in very small increments, moving gradually through phonetic space, while a phonetically abrupt change takes place in a single large, discrete step. If the phonological units of the lexicon are phonemes or generative underlying representations, a sound change that affects words differentially must be phonetically abrupt, since small phonetic increments do not exist in the structuralist or generativist lexicon. Thus the changes reported on in Wang (1977) are changes affecting phonemes, such as Chinese tones, final nasals, and English stress as used in diatone pairs (*récord*, *recórd*).

Similarly, Labov (1981, 1994) proposes that some sound changes exhibit lexical diffusion while others are "Neogrammarian" changes; that is, they are regular in the sense that they affect all lexical items with appropriate phonetic conditioning at the same time. The changes that exhibit lexical diffusion, according to Labov (1994:542–543), are phonetically discrete, bringing about phonemic changes in particular words. Labov further proposes a typology of Neogrammarian versus lexical diffusion changes: the latter, Labov proposes, are found in shortening and lengthening of segments, diphthongization of mid- and low vowels, changes of place of articulation of consonants, metathesis of liquids and stops, and deletion of obstruents. Regular sound changes occur in subsystems of vowels, including vowel shifts and the diphthongization of high vowels; also regular are changes in the manner of articulation of consonants, vocalization of liquids, and deletion of glides and schwa. Such changes are most often phonetically gradual and lexically regular.

Phillips (1984) does not accept Labov's claims, arguing that his dichotomy is too simplistic, as many changes affecting subsystems of vowels and the manner of articulation of consonants show evidence of lexical diffusion. Phillips implies that lexical diffusion is much more widespread than previously thought and, based on Hooper (1976), provides a distinction between two types of lexical diffusion. In one the direction of diffusion is from the most frequent to the least frequent words, while in the other the opposite direction of change is found. High-frequency words are affected earlier by vowel and consonant reduction or assimilation processes (Leslau 1969; Fidelholtz 1975; Hooper 1976). Phillips characterizes this type of change as "motivated by physiological factors, acting on surface phonetic forms" (320). Changes that affect the least frequent words first are more likely motivated by conceptual factors. These include grammatical conditioning, as in analogical leveling and diatone formation, as well as changes affecting the sequential constraints of the language.[2]

This chapter concerns only the first type of change—phonetically motivated change. Change of this type usually appears to be lexically regular once the change is complete, but evidence presented here shows that such changes affect lexical items at different rates while the change is in progress. Using *t/d* deletion as the example, it will be argued that sound change can be both lexically gradual and phonetically gradual. The consequences of this finding for lexical representation is that lexical entries for words undergoing gradual change must include detail about the range of phonetic variation associated with each word.

1.2. Word frequency in lexical diffusion

There exists a considerable body of evidence showing that phonetically conditioned sound change proceeds through the lexicon, gradually working from high-frequency words to low-frequency words. The effects of frequency have been shown for vowel reduction and deletion in English (Fidelholz 1975; Hooper 1976), for the raising of /a/ to /o/ before nasals in Old English (Phillips 1980), for various changes in Ethiopian languages (Leslau 1969), for the weakening of stops in American English and vowel change in the Cologne dialect of German (Johnson 1983), for ongoing vowel changes in San Francisco English (Moonwomon 1992), and for tensing of short *a* in Philadelphia (Labov 1994:506–507).

Some of these changes are both lexically gradual and phonetically gradual. For instance, the schwa-deletion process discussed in Hooper (1976) gradually reduces a schwa following the stressed vowel and preceding a sonorant, preferably /r/ or /l/. This reduction is more advanced in higher frequency words. Thus high-frequency *memory*, *salary*, *summary*, and *nursery* have a more reduced penultimate syllable than low-frequency *mammary*, *artillery*, *summery*, and *cursory* (table 9.1). Of course all of these words have multiple variants in actual use, but their ranges of variation differ systematically.

Changes that are both lexically and phonetically gradual create a problem for traditional theories of phonemic or lexical representation: individual words with distinct ranges of variation cannot be represented by a simple set of contrastive phonological units. In the schwa-deletion example, the number of distinct lexical sets is unknown, but there must be at least three: (1) those with the schwa totally deleted, *every*, *evening* (noun); (2) those in which the /r/ or /l/ can be syllabic or non-syllabic, *memory*, *salary*, *summary*, *nursery*; and (3) those in which the schwa or syllabic

TABLE 9.1. Words Undergoing Reduction at Differential Rates due to Word Frequency

No Schwa	*Syllabic* [*r*]	*Schwa* + [*r*]
every (492)	memory (91)	mammary (0)
	salary (51)	artillery (11)
	summary (21)	summery (0)
	nursery.(14)	cursory (4)
evening (149)		evening (0)
(noun)		(verb + *ing*)

Frequency figures from Francis and Kučera 1982.

resonant is always present, *mammary*, *artillery*, *cursory*, *evening* (verb + *ing*). These three lexical categories range over only two phonemes, since /r/ and syllabic /r/ are not distinguished phonemically, nor are syllabic /r/ and /ər/. This example shows that lexical items can have subphonemic detail associated with them. The fact that all three classes of words have variable pronunciations suggests that an appropriate model of the lexicon must allow for the representation of ranges of phonetic variation, and these ranges do not necessarily coincide with traditional or generative phonemes. The problem is even more severe if, as is extremely likely, there are not just three classes, but rather there is a continuum of degree of reduction, with each word exhibiting its own range of variation.[3] Moreover, the deletion of schwa and the reduction of the syllabicity of [r] are not separate processes; rather the reduction in all of these words is one continuous process.

1.3. Word frequency and *t*/*d* deletion

The variable deletion of final /t/ and /d/ in English has been well studied in the last two decades (Labov 1972; Guy 1980; Neu 1980). The factors influencing the deletion of final /t*1* or /d/ are phonetic, grammatical, and social:

- Phonetic: final /t/ and /d/ are deleted more often if a consonant follows in the next word than if a vowel follows.
- Grammatical: final /t/ and /d/ are deleted less often if they function as the regular past-tense marker; they are deleted more often if they constitute the past-tense suffix in words that also have a vowel change (*told*, *kept*).
- Social: final /t/ and /d/ are deleted more often by younger people, males, and members of lower social classes (Labov 1972; Neu 1980).

The data analyzed for this study were generously supplied by Otto Santa Ana from his study of phonological variation in Chicano English speakers of Los Angeles (Santa Ana 1991). The speakers used in this study were all native speakers of English. The values for final /t/ and /d/ were transcribed by Santa Ana from recordings made in interviews. For this research, 2,000 tokens of final /t/ and /d/ from forty-one speakers were used. The speakers ranged in age from 13 to 62 years; approximately half were male and half female.

No evidence of lexical diffusion or a general frequency effect has been reported for *t/d* deletion, despite the fact that it is a well-studied example of phonological variation. However, in studies of *t/d* deletion the words *and*, *just*, and *went* are often excluded because of their high rates of deletion. These words are clear examples of lexical diffusion conditioned by high token frequency, leading us to suspect a more general effect of word frequency on the diffusion of this change.

In order to test for the effects of word frequency on deletion, for each word that potentially ends in a /t/ or /d/ following a consonant we recorded the text frequency as listed in Francis and Kučera (1982). *Just*, *went*, and *and* were excluded because of their very high frequency and high rate of deletion. All other words were divided into two groups: a high-frequency and a low-frequency group. The cutoff point of 35 per million was chosen, because a number in this range is used in the psycholinguistic literature when frequency effects are measured. Particularly relevant to our analysis of past-tense forms is the fact that Stemberger and MacWhinney (1988) report that a frequency of 35 per million divides English inflected forms in half: half the tokens of inflected forms in Francis and Kučera (1982) have a frequency greater than 35 and half have a frequency of less than 35.

Dividing the current corpus in this way, 20% of the tokens fall in the low-frequency group and 80% in the high-frequency group. Table 9.2. shows the deletion figures for the entire corpus by frequency grouping.[4] As the table shows, word frequency is a significant factor in the variation in *t/d* deletion. It will be shown later that frequency has a general effect even when verbs are separated from nouns and adjectives. Before discussing the data on verbs and their inflections, however, let us examine the implications of phonetically and lexically gradual sound change for phonological representation.

1.4. Discussion: phonetic variation in lexical representation

Genuine lexically determined variation cannot be represented in a variable rule; it must be represented lexically, especially if it affects all or most of the words of a particular phonological shape. Trying to build such variation into a variable rule would require repeating the contents of the lexicon in the rule.

If memory for words resembles memory for other types of sensory input, then we would not expect that words are stored with details removed, but rather that even

TABLE 9.2. Rate of *t/d* Deletion for Entire Corpus by Word Frequency

	Deletion	*Non-Deletion*	*% Deletion*
High frequency	898	752	54.4%
Low frequency	137	262	34.3%

Chi-squared: 41.67, $p < .001$, $df = 1$.

redundant features of words contribute to the formation of a prototype structure for the phonological representation. The fact that speakers consistently produce *every*, *memory*, and *mammary* with three distinct ranges of variation means that more distinctions than those that are strictly considered phonemic exist in the lexicon.

One way of accounting for the continuous nature of such sound changes and the frequency effect they show in lexical diffusion is to assume that sound changes affect words opportunistically each time they are used. The sound change "rules" apply in real time, so that a frequent word has more exposure to the sound change than an infrequent one. If the effects of the sound change are cycled back into the lexicon, as the speaker monitors his/her own speech and the speech of others, the lexical representations for the words gradually adjust to the new productions (Moonwomon 1992). One reason that frequent words are affected at a higher rate than infrequent ones is because they have more exposure to the pronunciation pattern. However, there are two other factors that enhance change in high-frequency words when that change is reductive:[5] First, low-frequency words occur less often in casual speech and thus have even less exposure to the reductive processes. Second, words repeated in the same discourse are shorter than in their first mention (Fowler and Housum 1987). Thus, given the model of sound change outlined here, words that are repeated often within a discourse, which would also be words that have a high overall frequency, undergo sound changes at a faster rate.

In the case of *t/d* deletion, I am not suggesting that each word has an index of deletion probability related to its frequency. Rather, I propose that deletion takes place very gradually as the shortening of the coronal gesture and this shortening is represented lexically. (For measurements of this shortening, see Losiewicz 1992, discussed later.) I also propose, following Moonwomon, that phonological processes, such as the weakening of final /t/ and /d/, apply in real time and have an effect on the words to which they apply. Thus when a word with a final coronal stop is used, that articulatory gesture is usually compressed and shortened. The output of that process is registered in memory through self-monitoring and decoding the speech of others. Thus small changes may be periodically made in the prototype representation for each word. The more a word is used, especially in contexts that are appropriate for reduction, the more reduced the word becomes. In this way, the effect of word frequency on the propagation of a change through the lexicon is explained.

Proposing that variable rules affect the lexicon is not as radical an idea as it may sound. Harris (1989) discusses the increasing "lexical depth" of variable rules, which he regards as a natural consequence of the development of a rule. In his view, as rules develop they progress both phonetically and lexically. His examples involve phonemic change. Guy (1991a, 1991b) proposes that *t/d* deletion, usually considered postlexical, applies at the deep levels of Lexical Phonology, again an indication that the variable process has penetrated the lexicon. (Later I will give evidence for representing *t/d* weakening in the lexical entries rather than in the lexical rules.)

The proposal that words have their own ranges of phonetic variation rather than being composed of units that belong to a (systematic) phonemic set, specifiable for each language, would seem to allow an infinite range of phonetic representations for the words of a language. Why is it that the phonetic properties of words are strictly controlled by language-specific constraints rather being infinitely variable? The reason is that what goes into the lexicon is regulated by the phonetic processes or pronunciation patterns whose function it is to automate the production of phonetic strings. The efficiency of such patterns depends upon their producing a regular and highly constrained inventory of phonetic strings. Eventually these phonological processes carry sound changes through to completion and regularize lexical representations.

Phonemes, then, do not exist in the representations of words; they are not units of lexical representation. Instead, phonemes are abstract patterns that emerge in the phonological organization of the lexicon (see Langacker 2000). To the extent that distinct phonetic units are grouped together into more abstract units, this is done on the basis of the phonetic implementation schemata and is not a strict matter of complementary distribution, as can be seen from the examples presented so far.

1.5. Phonetic gestures

The phonetic representations and mode of phonetic implementation that best accommodate the workings of sound change that I am trying to characterize here are gestural phonology. I am considering phonological representations to be sets of commands to the articulators and phonetic output to be a series of partially overlapping articulatory gestures (Pagliuca and Mowrey 1987; Browman and Goldstein 1990). This conception provides an explanation for the phonetic conditions that favor *t/d* deletion, that is, the presence of a following consonant.

Browman and Goldstein (1990) traced the articulatory gestures in the phrase *perfect memory* to see what happens to the coronal gesture in the /t/.[6] They found that the lingual gesture was present, but the preceding velar and following labial gesture overlapped the coronal one, obscuring it entirely. It would thus be perceived as deleted, even though articulatorily it is still present. This does not mean that deletion is a myth; it does mean that there is no variable rule of *t/d* deletion. Rather there is a gradual process of shortening or reducing the lingual gesture (for which we will see evidence later). As the gesture shortens, the likelihood that it will be overlapped and thus acoustically obscured by surrounding gestures increases. Thus in words of higher frequency, where the coronal stop is shorter, perceived deletion is more likely.

Perceived deletion leads to actual deletion, that is, loss of the coronal gesture. This loss is ongoing, but at a much slower rate than the perceived deletion. Perceived deletion leads to the restructuring of underlying forms and proceeds item by item. Actual deletion occurs because preconsonantal and prepausal environments are three times as common as prevocalic ones. If lexical representations are prototype representations of all input variations, the final coronal stop in a lexical entry will shorten and gradually move toward deletion.

This particular sound change has articulatory, perceptual, and lexical dimensions. Articulatorily a gestural reduction is taking place. At the same time, the perception of the reduced consonant is masked by surrounding consonants. Lexical entries containing a final coronal stop are gradually accommodating to the changing input and will gradually restructure, losing the stop entirely. Thus there are three sources of the surface variation: the articulatory change is gradually reducing the gesture involved, the phonetic environment conditions whether or not the gesture can be perceived, and the lexical items themselves have different degrees of reduction.

2. Morphological effects on sound change and lexical diffusion

2.1. Past tense of verbs and *t/d* deletion

Studies of *t/d* deletion consistently find that the final coronal stop is less likely to delete if it is the past-tense morpheme than if it is part of a monomorphemic word (Labov 1972; Guy 1980). Moreover, a consistent difference is found between the regular past, as in *walked* and *learned*, and the past suffix on verbs that also have a stem change, such as *kept* and *told*. This difference is consistent with an account that attributes the lower rate of deletion of past-tense *t/d* to the need to retain the information encoded in this morpheme. In the verbs with vowel change, that information is not lost. In the discussion that follows, we will see that this "information preservation" assumption explains only part of the data.

In the data examined for the current study, it was also found that regular past tense and past-participle forms (which were grouped together) had a lower rate of deletion than nouns and adjectives in the same frequency range. Regular *-ed* verbs in our corpus had Francis and Kučera frequencies ranging from 0 to 403 per million. The rate of deletion for all words with frequencies of 403 or less is 45.8%; the rate of deletion on *-ed* verbs is 22.6%, as shown in table 9.3.

Guy (1991a, 1991b) offers an interpretation of *t/d* deletion in the context of Lexical Phonology (Kiparsky 1982). Guy proposes that the variable rule of *t/d* deletion applies at all levels of a Lexical Phonology: it applies at the base level to non-derived forms; it applies after Level 1 affixation and again after Level 2 affixation, assuming only two levels for the purposes of his analysis. Thus monomorphemic words have three chances to undergo *t/d* deletion; irregular past forms, such as *kept* and *told*, have two chances—one at Level 1 after affixation occurs and again after Level 2 affixation. Regular past tense forms are only exposed to the rule once, after Level 2 affixation. Guy proposes that the rule applies at a constant rate and that the different rates of deletion in the three categories of words is an exponent of the number of exposures the word has had to the rule.

Guy's analysis demonstrates a very interesting point: certain words are behaving as though *t/d* deletion has applied to them more than once. In his analysis, this variable rule has penetrated the lexicon and is treating words differentially according to their morphological structure; words that have higher rates of deletion can be thought of as undergoing the rule repeatedly. Note, however, that Guy's hypothesis refers only to morphological structure and cannot explain the overall effect of word frequency on deletion rates.

TABLE 9.3. Rate of Deletion for Regular Past Tense Compared to All Other Words of Comparable Frequency (403 or Less)

	% Deletion
All words	45.8%
-ed verbs	22.6%

The proposal being made here, that *t/d* reduction applies in real time and has a permanent effect on the phonological shape of words in the lexicon, achieves approximately the same result. If a word already had a weakened [t] or [d] and then underwent further reduction when that word was used again, the chances that it would be perceived as deleted (preceding a consonant) would be increased. Furthermore, Guy (1991a) shows that the effect of the environment preceding the [t] or [d], which is internal to the word and thus always present, increases the rate of application of deletion in words subject to multiple applications, while the following environment does not increase the rate of application in words subject to multiple applications, because the following conditioning environment is only present in the last application of the rule, the postlexical application. The same result is predicted for the model in which rule application has a permanent effect on the lexicon, but for different reasons. The preceding environment has a stronger effect on the lexical representation because it is present each time the word is used. The following environment—a following consonant—is an "alternating environment" and affects the lexical representation more slowly since it is not always present when the word is used (see Timberlake 1978 and Bybee 2000 for other examples of variable processes that apply less often or to a lesser extent in alternating environments).

In many ways, then, the two models make similar predictions. The difference is that the Lexical Phonology interpretation attributes differential rates of rule application to structural differences among words, while the usage-based interpretation attributes much of the differential reduction to word frequency. In the following I will show that word frequency is a significant factor in *t/d* reduction, even inside of morphological classes. These facts support the usage-based account over the Lexical Phonology account, since the latter has no way of describing differential application based on word frequency. The following discussion will be situated in the model proposed in Bybee (1985, 1988), which makes certain predictions for the application of *t/d* reduction to morphologically complex words.

2.2. Morphology as lexical organization

In Bybee (1985, 1988) it is proposed that the lexicon consists primarily of words and even words that are morphologically complex may have lexical storage. It is also argued that token frequency has an effect on the way words are stored and processed, as words have varying lexical strength according to their frequency of use. Even at resting levels, high-frequency words have stronger representations in memory, making them, among other things, easier to access. Morphologically complex words of high frequency are more likely to maintain their irregularities because they have a stronger representation in memory, and lower frequency words are more likely to be replaced by those formed on a regular pattern. Irregular forms are stored in the lexicon, and to the extent that they form classes with other forms, patterns of similarities based on prototype structure create lexical schemas (Bybee and Moder 1983; Prasada and Pinker 1993).

Figure 9.1 illustrates the relevant portion of the proposed lexical network. Similarities among words are shown as lexical connections, which can be phonological or semantic. Parallel phonological and semantic connections, such as those shown

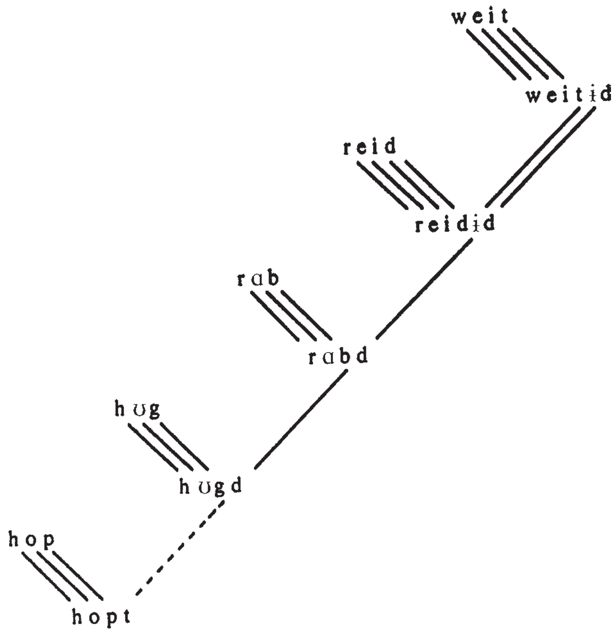


FIGURE 9.1. Emergent Morphological Structure and the Regular Past Tense Schema

in figure 9.1, constitute morphological relations if they are repeated across items. A dotted line indicates similarity but not identity.

It follows from the premise that frequency of use determines representation in the lexicon that even regular morphological formations of sufficient frequency are represented lexically. The internal structure of morphologically complex words derives from the connections made with other words. Thus the past-tense morpheme emerges from the patterns of lexical connections as a schema. Because of its high type frequency, this schema is very powerful, leading to its high productivity (Bybee 1995):

. . . (ɨ) alveolar stop]$_{\text{past tense}}$

Very low-frequency verbs may not have their past-tense forms represented in memory, or they may not be represented with sufficient strength for easy retrieval. In that case, the base verb is supplied with a past-tense inflection by the general schema. This means that there are two ways of processing regular past-tense verbs—selecting them whole from the lexicon and deriving them by applying the schema to a base.

Losiewicz (1992) tested the hypothesis that high- and low-frequency regular verbs are processed in these two ways. She started from the finding that consonants that are parts of monomorphemic words are produced with shorter acoustic duration than those that constitute morphemes; that is, the /s/ in *lapse* is shorter than the one in *laps* (Walsh and Parker 1983). She found that the same situation holds with respect to past tense [t] and [d]; that is, in word pairs such as *swayed* and *suede* the *-ed* suffix is on average 5 msec. longer than the non-morphemic [t] or [d]. The interpretation of this durational effect given by Walsh and Parker and by Losiewicz is that segments that are represented in the underlying form as part of the word tend to be shorter than those added in morphological processing. Losiewicz reasoned that if lexically represented consonants are shorter than added ones and if high-frequency past-tense verbs are stored as units, but low-frequency verbs are constructed by applying a schema to the base, then the past-tense [t] or [d] on high-frequency verbs should be shorter in acoustic duration than the same consonants on low-frequency verbs.

Losiewicz constructed eight pairs of rhyming regular past tense verbs in which one member of the pair had a frequency of 100 per million or greater and the other had a frequency of 10 or fewer per million. Some examples are *called*, *mauled*; *covered*, *hovered*; *needed*, *kneaded*; *expected*, *inspected*. The stimuli verb pairs were presented in sentences that differed only in the verb. (For example, *The workers expected/inspected all the mail at noon every day.*) Each subject read only one of each pair of stimuli sentences. Subjects read the sentences into a recorder and the length of the consonants in question was measured from the waveform. The results show that on average the *-ed* suffix on low-frequency verbs is 7 msec. longer than on high-frequency verbs. Losiewicz concludes that high-frequency multimorphemic verbs are represented as unitary wholes in the lexicon, while low-frequency multimorphemic verbs are processed as stem + affix.

A difference in the length of the final [-t] or [-d] suffix varying with the frequency of the verb is predicted from the interpretation of the lexical diffusion process for *t/d* reduction given earlier. Among verbs of sufficient frequency to have their past tense represented lexically, we would expect that higher frequency verbs will have undergone more reduction than lower frequency verbs. Thus a frequency difference even within lexically listed verbs is predicted, in addition to the length difference found by Losiewicz. The results of the test of this hypothesis are given in the next section.

2.3. *t/d* Reduction in regular past tense

For the test of the effect of frequency on *t/d* deletion among regular *-ed* verbs, we control the phonological environment, taking only those that favor deletion—non-prevocalic conditions. Dividing the regular verbs into two groups—those with frequencies of their *-ed* forms of 35 or fewer per million and those with frequencies over 35 per million—we find a significant effect of frequency, as shown in table 9.4.

The results in table 9.4 present strong confirmation of the hypothesis that regular verbs of higher frequency have lexical listing. If the *-ed* suffix were always added to the verb by morphological rules, there would be no reason for the length of the suffix to be affected by the token frequency of the whole unit, that is, the past-tense verb form. The fact that the frequency of the whole unit affects the deletion rate of the suffix [t] or [d] can best be accounted for by listing the units in the lexicon and allowing gradual phoneti change to affect the lexical entries.

TABLE 9.4. The Effects of Word Frequency on *t/d* Deletion in Regular Past Tense Verbs (Non-prevocalic Only)

	Deletion	*Non-Deletion*	*% Deletion*
High frequency	44	67	39.6%
Low frequency	11	47	18.9%

Chi-squared: 5.00313, $p < .05$, $df = 1$

These data also argue against the dual-processing model proposed by Pinker (1991), in which irregular forms have lexical listing, but regular ones are produced by a rule that is independent of the lexicon. In Pinker's model all regular verbs are formed in precisely the same way and there is no way for the token frequency of the whole unit to affect the phonetic shape of the suffix.

2.4. Double-marked past-tense verbs

Studies of the effects of morphology on *t/d* deletion consistently show that double-marked past-tense forms such as *kept*, *found*, *told*, and so forth, have a higher rate of deletion than regular past tense. In our data, the rate of deletion for these verbs is 42.8%. This high rate of deletion is consistent with the fact that most of these verbs have relatively high token frequency. All but two of them that occur in our corpus have frequencies over 35 per million in Francis and Kučera.

In Guy's treatment of *t/d* deletion in Lexical Phonology, the past tense of these verbs is formed at Level 1. After they are formed, *t/d* deletion is available to apply to them (variably) at Level 1 and then again (variably) at Level 2. Because the variable rule has two chances to apply to these verbs, they have a higher rate of deletion than the past-tense forms derived at Level 2, which only have one chance to undergo the rule. Thus Guy's treatment claims that the different rate of deletion is due to a structural difference in the derivation of irregular and regular past tense.

However, it is entirely possible that the different rate of deletion is due to the high frequency of these verbs and not to a structural difference between them and other verbs. If a structural difference governs the rate of deletion of these verbs, then we would expect the rate to be similar across the different lexical verbs. However, if it is frequency that conditions the rate of deletion, we would expect to find differences even among the double-marked verbs conditioned by frequency. The latter is what the data show: the higher frequency verbs of this class undergo deletion at a higher rate than the low-frequency verbs, as shown in table 9.5.

Table 9.5 shows a strong effect of word frequency among the words of this class, with high-frequency verbs generally showing deletion at a higher rate than low-frequency verbs. The data also suggest conditioning by phonological factors: a labial stop preceding the /t/ favors deletion since the labial gesture can overlap and obscure the perception of the /t/. This explains why *kept* and *slept* have a higher rate of deletion than some of the other more frequent verbs.

TABLE 9.5. Double-marked Pasts Ordered by Frequency in Our Data, with Ties Ordered by Francis and Kučera Frequency

Total Tokens	*Verb*	*# of Deletions*	*% Deletions*
32	told	22	68%
9	felt	5	55%
8	left	2	25%
6	kept	4	66%
4	sent	1	25%
4	built	0	0
3	held	0	0
3	heard	0	0
2	slept	1	50%
2	lent	0	0
1	found	0	0
1	lost	0	0
1	meant	0	0

Spearman rank order correlation: $p = .696$ (significant at the .01 level [two-tailed or one-tailed]).

The frequency effect among the words of this class cannot be accounted for in the structural account that proposes that these words are derived at Level 1 and have a higher rate of deletion because the rule has two chances to apply to them. If this account were correct, the rate of deletion for each verb would be approximately the same. I conclude, therefore, that the primary factor that accounts for the higher deletion in this class is word frequency.

2.5. Modeling

We have seen that the facts of *t/d* deletion are consistent with the hypothesis made by the usage-based model that high-frequency words are represented in the lexicon even if they are morphologically complex. The facts also support the hypothesis that sound change occurs in real time, with its effects being registered in the lexicon as small incremental changes, such that words that are used more often will undergo change at a faster rate.

With these two points now established it is possible to turn to a consideration of the question of why the past-tense suffixes /t/ and /d/ reduce more slowly than non-suffixal /t/ and /d/. The factor traditionally invoked to explain the conservatism of morphemic /t/ and /d/ is that the information value of the past-tense /t/ and /d/ is greater than that of non-morphemic /t/ and /d/. Past-tense marking is often redundant; however, in cases where it is carrying some information crucial to the discourse, speakers are quite capable of slowing down their speech and adjusting junctures in such a way as to make the /t/ or /d/ perceptually salient. Suppression of the reduction and coarticulatory overlap will also serve to retard reduction of the phonetic substance of the lexical schema for past-tense /t/ and /d/.

The data also reveal another factor that may help explain why the past-tense /t/ and /d/ reduce more slowly than non-morphemic /t/ and /d/. For some reason, regular past tense /t/ and /d/ occur more often before a vowel than final /t/ and /d/ in non-verbs. For all words in the corpus ending in /t/ or /d/, 22.7% occur before a vowel; of verbs ending in past-tense /t/ or /d/, 40.1% occur before a vowel. Since prevocalic position is the most favorable position for /t/ and /d/ perceptually, their more frequent occurrence in this position would retard the reduction that is conditioned perceptually.

2.6. Theoretical implications

Two important points for usage-based models are established by the data presented here. It has been argued that sound change, or phonetic change, affects the lexicon gradually by spreading out across lexical items from the most frequent to the least frequent and by producing small gradual changes in the phonetic representation of lexical items. One important consequence of this view is that sound change permanently affects the lexicon and that lexical items are not represented in an abstract phonemic notation. A second consequence is the more general point that the phonetic properties of lexical items, like their morphosyntactic and semantic properties, are affected by language use.

"Phonemes" do not exist as units; the phenomena that phonemes are intended to describe are relations of similarity among parts of the phonetic string. These relations of similarity can be captured by lexical connections and schemas just as other relations of similarity are. Complementary distribution, rather than a criterion for deciding on lexical status of a phone, is just a consequence of the fact that articulatory adjustments are conditioned by the surrounding environment. Registering such phonetic variation in the lexicon even if it is predictable provides a means for accounting for the establishment of new phonemes from distinctions that were formerly predictable, such as the [x] – [ç] distinction in German (Bybee 1994).

For most of this century, linguists have assumed the existence of units representing the phonemic status of certain segments. They have also assumed that both predictable and unpredictable phonetic variation is caused by rules acting on static underlying representations. Unaccounted-for variation in phonetic strings has either been ignored (as low-level phonetic phenomena) or relegated to unexplained "inherent variation." But just as the study of details of actual language use is yielding surprising and important results in the study of morphosyntax and lexicon, the study of the phonetic details of actual language use will yield a new view of phonological phenomena.

Notes

1. The data used in this chapter were generously supplied by Otto Santa Ana. I am also grateful to Valerie Daniel for inputting data, to Greg Thomson, who coded and analyzed the data and ran statistical tests on it, and to Jean Newman, who provided input at various points. Thanks are also due to Janet Pierrehumbert, who provided many valuable comments and suggestions on a previous version of the essay.

2. Phillips (1981) discusses the case of the diffusion of the loss of the [j] glide in words such as *tune* and *Duke* in the English of Georgia. Phillips attributes this change, which affects the least frequent words first, to a change in sequence structure conditions. I suspect, however, that it is a dialect borrowing or accommodation to the standard dialect. Words learned at the mother's knee, so to speak, would be the most conservative, while the least frequent words would be affected first.

3. Other types of representation run into the same problems. If the three classes of words were represented with different moraic structure, separate variable rules could delete the schwa and delete the mora associated with the syllabic consonants. However, since morae are discrete units, there is a limit on the number of distinct classes that could be represented.

4. The table presents percentages for convenience. The chi-squared value was not computed on the percentages.

5. Probably all phonetic change is due to reductive mechanisms—reduction of gestures and the compression of gestures causing overlap (assimilation) (Pagliuca and Mowrey 1987); however, the surface results in terms of segments may not always seem reductive.

6. When read from a list by the same speaker, *perfect* had a released final [t^h].

References

Browman, Catherine P., and Louis M. Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in laboratory phonology*, vol. 1: *Between the grammar and physics of speech*, ed. John Kingston and Mary E. Beckman, 341–376. Cambridge: Cambridge University Press.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins.

———. 1988. Morphology as lexical organization. In *Theoretical morphology: Approaches in modern linguistics*, ed. Michael Hammond and Michael Noonan, 119–141. San Diego, Calif.: Academic Press.

———. 1984. A view of phonology from a cognitive and functional perspective. *Cognitive Linguistics* 5(4), 285–305.

———. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5), 425–455. (Reprinted in this volume as chapter 8.)

———. 2000. Lexicalization of sound change and alternating environments. In *Laboratory phonology*, vol. 5, ed. Michael Broe and Janet Pierrehumbert, 250–268. Cambridge: Cambridge University Press. (Reprinted in this volume as chapter 10.)

Bybee, Joan L., and Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59, 251–70. (Reprinted in this volume as chapter 6.)

Fidelholtz, James. 1975. Word frequency and vowel reduction in English. In *Papers from the Eleventh annual regional meeting of the Chicago Linguistic Society*, 200–213. Chicago: Chicago Linguistic Society.

Fowler, Carol A., and Jonathan Housum. 1987. Talkers' signalling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26, 489–504.

Francis, W. Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Guy, Gregory. 1980. Variation in the group and the individual: The case of final stop deletion. In *Locating language in time and space*, ed. William Labov, 1–36. New York: Academic Press.

———. 1991a. Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3, 223–239.

———. 1991b. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3, 1–22.

Harris, John. 1989. Towards a lexical analysis of sound change in progress. *Journal of Linguistics* 25, 35–56.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In *Current progress in historical linguistics*, ed. William Christie, 95–105. Amsterdam: North Holland. (Reprinted in this volume as chapter 2.)

Johnson, Theodore. 1983. Phonological free variation, word frequency and lexical diffusion. Ph.D. diss., University of Washington.

Kiparsky, Paul. 1982. Lexical phonology and morphology. In *Linguistics in morning calm*, ed. In-Seok Yang, 3–91. Seoul: Hanshin.

Krishnamurti, Bhadriraju. 1978. Areal and lexical diffusion of sound change: Evidence from Dravidian. *Language* 54, 1–20.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

———. 1981. Resolving the Neogrammarian controversy. *Language* 57, 267–308.

———. 1994. *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.

Langacker, Ronald W. 2000. A dynamic usage-based model, In *Usage-based models of language*, eds. Michael Barlow and Suzanne Kemmer, 1–63. Stanford, Calif. CSLI.

Leslau, Wolf. 1969. Frequency as a determinant of linguistic change in the Ethiopian languages. *Word* 25, 180–189.

Losiewicz, Beth L. 1992. The effect of frequency on linguistic morphology. Ph.D. diss, University of Texas, Austin.

Moonwomon, Birch. 1992. The mechanism of lexical diffusion. Paper presented at the annual meeting of the Linguistic Society of America, Philadelphia, January.

Neu, Helene. 1980. Ranking of constraints on /t,d/ deletion in American English: A statistical analysis. In *Locating language in time and space*, ed. William Labov, 37–54. New York: Academic Press.

Pagliuca, William, and Richard Mowrey. 1987. Articulatory evolution. In *Papers from the Vllth international conference on historical linguistics*, ed. Anna Giacolone Ramat, Onofrio Carruba, and Giuliano Bernini, 459–472. Amsterdam: John Benjamins.

Phillips, Betty S. 1980. Old English *an – on*: A new appraisal. *Journal of English Linguistics* 14, 20–23.

———. 1981. Lexical diffusion and Southern *tune*, *duke*, *news*. *American Speech* 56, 72–78.

———. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–342.

Pinker, Steven. 1991. Rules of language. *Science* 253, 530–535.

Prasada, Sandeep, and Steven Pinker. 1993. Generalisation of regular and irregular morphological patterns. *Language and Cognitive Processes* 8, 1–51.

Santa Ana, Otto. 1991. *Phonetic simplification processes in English of the barrio: A cross-generational sociolinguistic study of the Chicanos of Los Angeles*. Ph.D. diss., University of Pennsylvania.

Stemberger, Joseph P., and Brian MacWhinney. 1988. Are inflected forms stored in the lexicon? In *Theoretical morphology: Approaches in modern linguistics*, ed. Michael Hammond and Michael Noonan, 101–116. San Diego, Calif.: Academic Press.

Timberlake, Alan. 1978. Uniform and alternating environments in phonological change. *Folia Slavica* 2, 312–328.

Walsh, Thomas, and Frank Parker. 1983. The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics* 11, 201–206.

Wang, William S.-Y. 1969. Competing changes as a cause of residue. *Language* 45, 9–25.
———, ed. 1977. *The lexicon in phonological change*. The Hague: Mouton.
Wang, William S.-Y., and Chin-Chuan Cheng. 1977. Implementation of phonological change: The Shaungfeng Chinese case. In *The Lexicon in Phonological Changes*, ed. William S.-Y. Wang, 86–100. The Hague: Mouton.

Chapter 10

Lexicalization of Sound Change and Alternating Environments

1. Usage-based theory

Over the last twenty years a significant functionalist trend has developed in the study of morphosyntax with the aim of explaining the nature of grammar by studying how language is used in context. The basic premise of this work is that frequently used patterns become conventionalized or fossilized as grammatical patterns; that is, grammar is emergent from language use (Givon 1979; Hopper and Thompson 1980, 1984; DuBois 1985; Hopper 1987; and many more). Haiman (1994) has discussed the process by which repeated patterns become part of "grammar" in terms of ritualization, showing that the effects that repeated stimuli or repeated action has on an organism—automatization, habituation, and emancipation—are also operative in the process of grammaticalization or the creation of new grammar (see also Boyland 1997 for a discussion of the psychological mechanisms involved).

Some comparable research in the phonological domain has begun to appear in recent years. For instance, several studies have shown that speakers' judgments of the grammaticality of phonotactic patterns is based on the frequency of consonant and vowel combinations actually occurring in the language (Pierrehumbert 1994, Frisch 1996). In addition, speakers' ability to access the lexicon may involve a complex interplay between the frequency of words and the number and frequency of words with similar phonological shape (Pisoni et al. 1985). Connectionism offers the possibility of formally modeling the effect of use on mental representations of language, and such models have been tested in the phonological and morphological domains (for example in Dell 1989 and Daugherty and Seidenberg 1994).

This chapter is intended to contribute to the general effort to show how language use can be recruited to help explain some well-known properties of phonological patterns. The data examined, which are from phonological variation, also provide some evidence for the size and nature of the phonological memory representations for words and phrases.

First, I present evidence that many, if not all, sound changes progress in lexical items as they are used, with more frequently used words undergoing change at a faster rate than less frequently used words. Then I examine "alternating environments"—cases in which a sound in a particular word or morpheme is sometimes in the environment for the change to take place and sometimes not. In cases in which the targeted sound is at the edge of a word, the change can go through even where the sound is not in the appropriate phonetic environment and thus no alternation is produced. In such cases, we have evidence for the restructuring of the lexical representation of the word. However, when the alternating environment is inside of a word, the change can be retarded even in the appropriate environment, but eventually an alternation can be created, showing, again, restructuring of the lexical representation of the word. I argue that lexical representations are restructured gradually on the basis of actually occurring variants of a word and that postulating words and frequent phrases as the units of representation explains the development of word-level phonology. In addition, it will be argued that reference to the frequency with which words begin in consonants explains why a final word boundary often conditions changes as though it were a consonant.

2. The frequency effect on sound change

One of the aims of this chapter is to explore from a phonological perspective the size and nature of storage and processing units. I will present evidence that words and often longer units, such as frequent phrases, are the units of lexical storage. For the moment, however, I assume that words are the units of lexical storage. It is reasonable to assume that lexically stored words are in many ways like other mental records of a person's experience. First, there is no reason to believe that these memorial records have details and predictable features abstracted away from them (Langacker 1987; Ohala and Ohala 1995), and second, it is reasonable to believe that new experiences are categorized, to the extent possible, in terms of the already-stored record of past experiences (see Klatzky 1980).

Each use of a word requires retrieval by the speaker and a matching of the incoming percept to stored images by the hearer (and the speaker, who is monitoring his or her own speech). My thesis in this chapter is that the act of using a word, in either production or perception, has an effect on the stored representation of the word. We already know this is true in terms of the degree of entrenchment of a word (or the resting level of activation): high-frequency words have stronger representations that make them easier to access, more resistant to change on the basis of other patterns, and more likely to serve as the basis for the creation of new forms (Bybee 1985).

In addition, certain levels of use affect the stored representation of words by actually changing their shapes. That is, along with the entrenchment effect of frequency, there is also an automation effect: words and phrases that are used a lot are reduced and compressed. This effect is very salient in grammaticizing phrases (such as *going to* becoming *gonna* and *want to* becoming *wanna*) and more conventionalized contractions (such as *won't* and *didn't*), but it also occurs in a more subtle form across the lexicon when a sound change is taking place. Sound changes (phonetically motivated changes, which are usually the reduction of the magnitude of gestures or retiming of gestures; Browman and Goldstein 1992) tend to be phonetically gradual and also lexically gradual: high-frequency words undergo change at a faster rate than low-frequency words. The effects of frequency in the diffusion of a sound change through the lexicon have been shown for vowel reduction and deletion in English (Fidelholtz 1975; Hooper 1976b), for the raising of /a/ to /o/ before nasals in Old English (Phillips 1984), for various changes in Ethiopian languages (Leslau 1969), for the weakening of stops in American English and vowel change in the Cologne dialect of German (Johnson 1983), for ongoing vowel changes in San Francisco English (Moonwomon 1992), and for tensing of short *a* in Philadelphia (Labov 1994:506–507). In a recent paper, I have shown that there is also a frequency effect in the application of t/ d deletion in American English (Bybee 2000). Deletion occurs more in high-frequency words, including of course monomorphemic nouns and adjectives, but also regular past tense verbs, a point to which I will return later.

My interpretation of the frequency effect in the diffusion of sound change (following Moonwomon 1992) is that sound change takes place in small increments in real time as words are used. The more a word is used, the more it is exposed to the reductive effect of articulatory automation. The effects that production pressures have on the word are registered in the stored representation, probably as an ever-adjusting range of variation. Thus words of higher frequency undergo more adjustments and register the effects of sound change more rapidly than low-frequency words.

The frequency with which a word is subject to the ravages of articulation is not the only factor that encourages sound change. We also have to take into account the fact that certain speech styles allow more reduction and compression than others. In particular, casual speech among familiars typically shows more reduction. Thus words that are used in casual situations will also undergo change at a faster rate (see this volume, chapter 2, and D'Introno and Sosa 1986). Of course, these words are also likely to be those that are of higher frequency overall.

Another factor affecting reduction is the status of the word within the discourse. Fowler and Housum (1987) found that the first use of a word in a spoken text was longer than in subsequent uses. This means that speakers articulate more clearly in the first use of a word, where identification by the hearer might be more difficult, and then allow the reductive processes to apply later when identification by the listener is aided by the context and the fact that the word has already been activated. In fact, speakers may use reduction to indicate that a referent is *not* new but rather one that has already been accessed in the discourse. Words that are used more often within a text are produced in reduced form more often. If the produced form affects the stored form, then words that are repeated more often in a discourse will reduce at a faster rate than words that are repeated less often.

3. Exemplar-based representations

The account of phonetically gradual lexical diffusion of a sound change given in the preceding section requires a model of memory storage for linguistic units based on actual tokens of use. Each experience of a word is stored in memory with other examples of use of the same word. These memories of specific tokens are organized into clusters with more frequently ocurring exemplars, and tokens that share many properties with high-frequency exemplars are treated as more central, while less common or more deviant tokens are treated as more marginal. Thus linguistic experiences are categorized in the same way as other types of perceptual experiences. Rather than conceiving of stored representations as abstractions from the phonetic tokens, representations are considered to be the result of the categorization of phonetic tokens. This proposal, which will be referred to as "the exemplar model," adapts proposals made by Miller (1994) for phonetic segments and Johnson (1997) for larger units. Similar arguments for phonological representations have been made by Hooper (1981) and Cole and Hualde (1998). Note that this model does not distinguish between phonetic and phonemic features in lexical representation (see Steriade 2000). Further implications of this model for sound change will be discussed in the next sections.

4. Alternating environments

Given that produced tokens affect stored representations, what would happen when a word or morpheme occurs in different environments, such that it is subject to a change in one environment but not in another? In the case of such "alternating environments" (as Timberlake [1978] calls them) two or more different surface forms map onto a stored form. How are such alternate mappings resolved?

Here I approach this question by examining cases of sound change in progress. It is necessary to distinguish the phonetic variation that goes on while a sound change is in progress from the conventionalized alternations that can eventually arise from such sound change. By *alternation* I mean that a word or morpheme has two or more variants that are not phonetically continuous or variable but rather constitute discrete alternants conditioned by specific phonological, grammatical, or lexical contexts. An alternation, then, roughly corresponds to the level of variation generated by a classical phonological rule. By studying the conditions under which such alternations are conventionalized and the conditions under which they are not, we learn something about how variants of words and morphemes are organized in memory.

The study of alternating environments in cases of sound change in progress reveals that the outcome differs according to whether it is a word or a morpheme that is in the alternating environment. When the same morpheme is in an alternating environment in different words, a change is retarded even in the conditioning environment (the Timberlake Effect; see below), but an alternation can eventually arise. When the alternates are in two forms of the same word, alternations arise only under special conditions, but ordinarily only one alternate survives. I will argue that the differential behavior of morphemes and words with respect to sound change in progress provides strong evidence for the stored representation of words and frequent phrases.

4.1 Sound changes at word boundaries

The fact that phonological phenomena occurring around word-boundaries often have a different effect from phenomena inside of words can be attributed to the use of forms in context and the way this usage affects stored representations. As an example consider the well-studied variable phenomenon of syllable-final *s*-aspiration and -deletion in Latin American Spanish. A syllable-final /s/ before a consonant is subject to loss of its lingual articulation, resulting in what has been termed *aspiration* or a period of voiceless frication, which itself is eventually subject to loss. This change can affect a word-internal /s/ as well as a word-final /s/ before a consonant (see examples in [1]). As this change is ongoing in many dialects of Latin America, we can compare two dialects that appear to be at different stages in order to observe differences in the effect that the sound change has on word-final versus word-internal /s/.

The examples in (1) illustrate the variation found in /s/ in the different phonological environments where /s/ deletes or aspirates in Latin American Spanish (Terrell 1978). Non-phonological factors involved in this variation include the sex and age of the speaker, lexical factors (see below), speaking rate, and register.

(1) a. ____C: word-internally before a C

feli*h*mente	'happily'
e*h*tilo	'style'
denti*s*ta	'dentist'

b. ____##C: word-finally before a C

o se traen animale*h* finos	'or fine animals are brought'
haya mucho*s* temas	'there are many themes'
su*Ø* detalleh	'his details'

c. ____##V: word-finally before a V

y mientra*h* esa sonoridad así	'and during this voicing thus'
no va*s* a encontrar	'you are not going to find'

d. ____//: before a pause

en momentos // libreh //	'in moments, free'

Consider now table 10.1, which shows the rate of deletion in different contexts in Argentine and Cuban Spanish. This table generalizes over thousands of tokens produced by dozens of subjects and obscures some important factors, including some lexical factors that will be mentioned later. For the moment, however, let us focus on one difference between the two dialects.

The Argentine dialect can be considered to be in an early stage of the implementation of this change relative to the Cuban dialect. The important point to note about Argentine Spanish is that at this stage the word boundary appears to have little effect. Looking at the first column, we see that the maintenance of /s/ is largely pre-

TABLE 10.1. Gradual Lexical Restructuring: Spanish *s*-Aspiration

	s	*h*	*Ø*	*Tokens*
Argentine Spanish				
____C	12%	80%	8%	4150
____##C	11%	69%	20%	5475
____##V	88%	7%	5%	2649
____//	78%	11%	11%	2407
Cuban Spanish				
____C	3%	97%	0%	1714
____##C	2%	75%	23%	3265
____##V	18%	48%	34%	1300
____//	61%	13%	26%	1776

Terrell 1977, 1978, 1979; Hooper 1981.

dictable from the phonetic environment: before a C the maintenance of /s/ is at 11–12%, with no significant difference between word-internal and word-final /s/. The maintenance rate at the end of a word before a V (____##V) is at 88%, since a following vowel does not constitute the appropriate environment for the change. Now compare Cuban Spanish: the fact that this dialect is in a more advanced stage of the change is indicated by the fact that the rate of maintenance of /s/ is only 2–3% before a consonant. Again, the presence of the word boundary makes little difference. The major difference between Argentine and Cuban Spanish for our purposes is in the use of /s/ at the end of a word before a vowel. In Cuban Spanish aspiration and deletion are common in this context, with maintenance of /s/ down from 88% in Argentine to 18% in Cuban, even though it is not the appropriate phonetic context for the change and no such change is going on with word-internal /s/ before a vowel.

What would cause a phonetic change to occur outside of its phonetic environment? Obviously it is the position of the /s/ at the end of the word that is crucial: in this alternating environment, the /s/ is sometimes in the position for aspiration and sometimes not. In fact, more than half the time word-final /s/ is in the environment for aspiration as shown in table 10.2. If, as I hypothesized earlier, sound change affects stored representations incrementally each time a word is used, the use of a word-final /s/ before a consonant will have some effect on the stored representation,

TABLE 10.2. Percentage of Occurrence of Word-Final /s/ before a Consonant, Vowel, and Pause

Argentine		*Cuban*	
C	52.0%	C	51.5%
V	25.1%	V	20.5%
//	22.9%	//	28.0%

especially if there is only one representation per word. Thus the /s/ will gradually decay in the stored representation and the reduced form of that consonant will eventually appear even before vowels, as in Cuban Spanish.

Note that this case exhibits the classic feature of a final word boundary behaving like a consonant. The figures in table 10.2 provide us with a usage-based explanation of this common phenomenon: if usage affects stored representations and if twice as many tokens begin with a consonant as with a vowel, phonetic changes conditioned by a following consonant will also take place at a word boundary, though at a slower rate than word-internally, where the environment does not alternate.

Ordinarily, alternations do not develop where the conditioning is across a word boundary. This fact gives rise to the notion of "word-level phonology"—that is, the fact that most alternations occur within words. The explanation being investigated here is that ordinarily there is only *one* stored representation for each word. Where variation arises during change in progress, the variation is resolved in terms of one variant or the other. The exceptions to this arise only in the case of frequently used phrases, to which I will return shortly.

First let us consider how variation at the word level is represented and how cases of sound change in an alternating environment would eventually be resolved. In the exemplar model described earlier, the representation of a word is a cluster of actually occurring tokens, with more frequent tokens accumulating greater weight or strength. Thus each word has its own range of variation dependent upon its frequency and the contexts in which it is used. When little or no sound change is affecting a word, the range of variation in the tokens may be small and relatively stable. During change, however, the range of variation increases and the center of the cluster gradually shifts.

When the same word occurs in both an environment that conditions a change and a non-conditioning environment, as in the Spanish *s*-aspiration case, the cluster for a word may divide into two (or more) subclusters, each one with a strong center of high-frequency tokens. In this case, each subcluster is associated with one environment—the word-final [s] tokens with the environment before a vowel and the word-final [h] tokens with the environment before a C. It appears that such a situation is unstable when the environment is not also part of the representation, because it tends to be resolved in favor of one variant for all environments. That is, the most frequent variant, the weakened consonant, [h], wins out and tends to be chosen even in contexts before a vowel.

In contrast, when the environment is part of the stored unit, an alternation can be established, in the sense that the [s] can remain before a vowel. This happens in frequent phrases. For example, Terrell (1986:129) reported that one illiterate speaker of Dominican Spanish used word-final /s/ only twelve times out of 443 words with orthographic *s* in a taped interview. All cases involved grammatical morphemes in set phrases: /s/ is used four times with a plural definite article followed by a stressed vowel (e.g., *las otras* 'the others (fem.)', *las únicas* 'the only ones (fem.)', etc.), three timed in the phrase *más o menos* 'more or less', in the names of two ball teams, *Las Estrellas* 'The Stars' and *Las Águilas* 'The Eagles', in the phrase *es igual* 'it's the same', and in the words *ellos* 'they (masc.)', *tres* 'three'. In the last two cases Terrell does not report what word followed, but it is highly probable that it is a high-frequency vowel-initial verb after *ellos* and a high-frequency vowel-initial noun after *tres*. The conclusion is, then, that in frequent phrases that constitute lexical units the /s/ is preserved before a vowel, even though in other contexts the word may have lost its /s/.

The Dominican case is extreme, as the loss of word-final [s] is quite advanced. In the Cuban case, which is not so advanced, I extracted the examples of orthographic *s* before a vowel (as transcribed by Terrell) from the interview of one speaker. In this interview there were seventy-seven opportunities for the speaker to produce an [s] before a vowel; only twelve of these were transcribed as [s]. (This constitutes 15%—a number close to the 18% reported for all speakers.) These cases of word-final [s] before a vowel were:

(2) a. the phrase *las once* meaning 'eleven o'clock'

b. four cases of a quantifier with the noun *años* 'years'
seis años 'six years', *tres años* 'three years', *muchos año(s)* 'many years'

c. the adjective-noun phrase *determinadas época* 'certain times'

d. five occurrences with the third singular copula *es*, followed by a definite or indefinite article: *es al medio día* 'it's at noon', *es el centro* 'it's the center', *es el sureste* 'it's the southeast', *es en lo general* 'that's in general', *es un* . . . (pause) 'it's a . . .'

The phrase in (a) is a standard phrase used in speaking about the time of day. The phrases in (b) and (c) are all quantified nouns designating temporal units. The [s] on the quantifier is retained before the frequently used vowel-initial noun *años*. The cases in (d) are all examples of the third singular copula *es* followed by a grammatical morpheme—either a preposition or an article. All of these sequences (with the possible exception of [c]) occur commonly enough to constitute processing and storage units. The [s] is retained in these phrases because the phrases are acting like words, and the [s] is in a sense "word-internal" in these cases.

Thus my hypothesis is that words and frequent phrases are storage units and that ordinarily there is only one representation per word, so that variations in the form of a word are normally reconciled to a single form and no alternation is created through sound change. Exceptions to this occur when a word is used in high-frequency phrases and/or phrases involving grammatical morphemes, such as pronouns and articles. This hypothesis makes strong predictions about the conditions under which sandhi phenomena will develop. It predicts that sandhi processes will only occur in phrases of high frequency and most commonly in those involving grammatical morphemes or other high-frequency words. This prediction is borne out by the most famous cases of sandhi, such as French liaison (Tranel 1981).[1] The hypothesis also predicts that cases of reduction restricted to certain "syntactic" environments, such as English auxiliary contraction and the reduction of *don't*, will occur only in the most frequent contexts in which the form appears. This prediction is also borne out where it has been tested (Krug 1999; Bybee and Scheibman 1998; see also Bybee 1998).

4.2. Sound change inside of words

Of course, alternations do develop inside of words. A morpheme inside a word may undergo phonological change, producing a new allomorph and thus an alternation. This fact follows from the hypotheses presented earlier: if sound change permanently affects stored units and words—even morphologically complex ones—are the units of storage, then the same morpheme in different words will take on different phonological shapes.

Further evidence for the hypotheses developed here is the fact that the effect of an alternating environment inside a word is very different from the effect across word boundaries. Inside a word, a variable process never applies outside of its phonetic environment (as, say, the aspiration of /s/ in Spanish occurs even ____##V). Instead the effect is the reverse: there is evidence that a change can be retarded even in its phonetic environment if it occurs in a morpheme that also has alternates that appear outside of the phonetic environment. Timberlake (1978) has made this point by presenting examples of changes that progress faster in uniform environments and are retarded in alternating ones. I thus dub this the "Timberlake Effect." Consider the following examples from the Mazowian dialect of Polish (Timberlake 1978:313–314). Transcriptions are from Timberlake and his sources.

In root-internal position velars are always more or less palatalized before /i/ (the symbol [^{j}] indicates full palatalization and [’] indicates partial palatalization):

(3)	k^{j}il	‘stick’
	sk^{j}iba	‘ridge’
	g^{j}ipsu	‘gypsum’
	k’ilômetr	‘kilometer’
	g’ɪrie	‘bends’

In alternating environments, represented here by position before a suffix, stops are palatalized before /i/ in only about half of the recorded examples, in (4).

Where the velar is stem-final, it will sometimes occur in a palatalizing context and sometimes not. This is the reason, according to Timberlake, that the palatalization process is retarded even when the conditioning is present.

The same effect may be observed with affixes. Affixes, such as English past-tense [t] or [d], are in an alternating environment, because they may occur after vowels, as in *played*, or *tried*, and they may occur after consonants, as in *walked* or *learned*.

(4)	prog’ɪ	‘hearths’ (Gen. Sg.)		
	drog’ɪ	‘roads’ (Nom. Sg.)		
	durak’ɪ	‘beets’		
	jarmak’ɪ	‘fairs’	morg’ɪ	‘acres’
	rog’ɪ	‘horns’	gruskɪ	‘pears’
	kackɪ	‘ducks’	zmarsckɪ	‘wrinkles’
	drugɪ	‘other’	robakɪ	‘worms’

TABLE 10.3. Rate of Deletion for Regular Past Tense Compared to All Other Words of Comparable Frequency (403 or Less)

	Percentage Deletion
All words	45.8
-*ed* verbs	22.6

The suffix is only in a context for deletion when it is preceded and followed by a consonant. A monomorphemic word also has an alternating following environment, but its preceding environment is uniform. It could be for this reason that past-tense [t] and [d] delete less often than [t] or [d] in monomorphemic words, as shown in table 10.3 (data from Bybee 2000).[2]

In contrast, sound changes in affixes in internal and non-alternating environments go through before all others: deletion of Spanish /d/, which is allophonically [ð], is much more advanced in the present-participle *–ado* than in other contexts (excepting high-frequency words, such as *lado* 'side' (D'Introno and Sosa 1986). Similarly, a very early instance of the loss of intervocalic /d/ occurred in the second plural morpheme -*ades*, which became -*ais* in Old Spanish.

These examples show that the implementation of a sound change in particular words (that is, its lexical diffusion) depends heavily on the contexts in which the sound is used. Since I have been arguing that the unit that serves as the context for a sound as it undergoes change is the word, then we must now consider how to account for the fact that the environment of a morpheme in one word affects the rate of change for the same morpheme in a different word. To understand this issue, we must understand the nature of morphological relationships, a matter to which I now turn.

5. A network model

In various works (Bybee 1985, 1988, 1995) I have proposed that the lexicon is organized into a complex set of relations among words and phrases by connections drawn among phonologically and semantically similar items. Parallel phonological and semantic connections constitute morphological relations if they are repeated across multiple pairs of items. As Dell (2000) points out, morphological relatedness is the joint effect of the organization of words into phonological and semantic neighborhoods.

In this model, the relations between base and past-tense forms of English verbs are diagrammed as in figure 10.1,[3] where semantic relations are not explicitly shown and where relations of similarity (rather than identity) between segments are shown with broken lines. Affixes are not explicitly listed in storage but emerge from sets of connections made among stored words and phrases. The very high type frequency of the regular English past tense strengthens its representation in memory and makes it highly productive. It can then apply to verbs whose past-tense forms are not

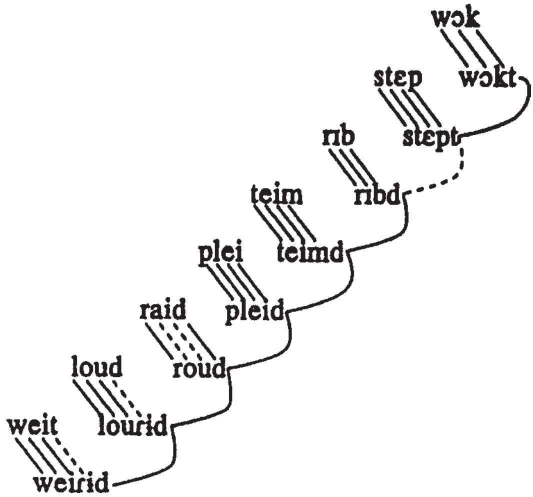


FIGURE 10.1. Model of relations between English Past Tense forms and their bases.

accessible because they have never been encountered or are of such low frequency as to not be easily accessible.

Past tense constitutes a category, but not one that can be accessed independently of a particular verb, because it is a category to which verb forms may belong or not belong. How then do individual tokens of the past-tense suffix relate to one another? This is, of course, an empirical question, and here is the evidence we have so far.

First, instances of the suffix attached to a verb are affected by the token frequency of the whole verb form: the rate of deletion for a final [t] or [d] on a past-tense verb is affected by the frequency of the form, as shown in Bybee (2000). In that study, past-tense forms with a frequency in Francis and Kučera (1982) of 36 or greater are considered high frequency and those with a frequency of less than 36 as low frequency, following a suggestion by Stemberger and MacWhinney (1988), who establish that the mean frequency of inflected verbs in Francis and Kučera is 35. Using this cutoff point, I (Bybee 2000) find that there is a significant

TABLE 10.4. The Effects of Word Frequency on t/ d Deletion in Regular Past Tense Verbs (Non-prevocalic Only)

	Deletion	*Non-deletion*	*% Deletion*
High frequency	44	67	39.6
Low frequency	11	47	18.9

χ^2: 5.00313, $p < .05$, df = 1.

difference between high- and low-frequency verbs in the position for deletion. See table 10.4.

Second, in the data I cited earlier, the overall trend for past-tense [t] and [d] is that they delete *less* often than [t] or [d] in monomorphemic words.

A related third point is that Losiewicz (1992) has shown that not only are monomorphemic [t] and [d] shorter than past-tense [t] and [d], but high-frequency past-tense [t] and [d] are shorter than low-frequency [t] and [d]. Losiewicz proposes to account for her data by a dual-access model in which high-frequency morphologically complex forms are stored and retrieved as wholes, while low-frequency forms are composed by adding the suffix to a base form using a schema. But this model would predict the same rate of deletion in high-frequency regular past tense forms as in monomorphemic forms of comparable frequency, and this prediction is not borne out by the data. Rather, in the data used in Bybee (2000), the rate of deletion for all words with frequencies of 36–403 was 54.4%, while the rate for past-tense forms of the same frequency was 39.6%. Thus we must posit that the past tense in low-frequency verbs, which is longer and phonetically fuller, can have some effect on the past-tense suffix on high-frequency verbs, which is shorter and more prone to deletion but not as short and prone to deletion as the [t] and [d] of monomorphemic forms. Thus the fuller form of the suffix on low-frequency verbs has some impact on the suffix on other verbs.

A proposal that would account for all the facts accumulated so far would be one in which the past-tense suffix forms a category, perhaps in terms of a set of related schemas, not separate from the forms it describes but emergent from them.

$$[\_\_\_\_\ [+\text{voice}]]_{\text{verb}} \qquad \text{d}\ ]_{\text{past}}$$
$$[\_\_\_\_\ [-\text{voice}]]_{\text{verb}} \qquad \text{t}\ ]_{\text{past}}$$
$$[\_\_\_\_\ [\text{t/d}\ ]]_{\text{verb}} \qquad \text{id}\ ]_{\text{past}}$$

Different tokens of past tense contribute to the formation of the category, and the data suggest that the category itself influences the shape of the productions. Since the occurrences of cases in the environment for deletion are many fewer than the cases that are protected from deletion (by surrounding vowels), the deletion is retarded in all members of the category past-tense verb.

Now we return to the Polish examples, in which an alternating environment for a *stem* seems to slow down a change in progress elsewhere. The inflected words affected by the change are embedded in a dense network of related forms, which includes the connections shown in figure 10.2. The closely related forms *progu* and *progi* are separate words and each has a representation; however, the shared stem also forms a category. As the palatalization of [g] before a high front vowel proceeds,

FIGURE 10.2. Lexical connections between two instances of the same stem.

the range of variation for the stem *prog-* begins to widen. The Timberlake Effect is evident now, as the center of the category is affected by the variants in the non-palatalizing environment. Since the two variants occur in different words—their environments are part of the representation—they can eventually diverge into two forms of the stem, creating an alternation.

This account of the Timberlake Effect makes predictions about the circumstances under which the effect will be the strongest. A change will be retarded most noticeably in an alternating environment when the alternates that are not in the environment to undergo the change are the most frequent—either there are more conditions in the paradigm in which the change does not take place or the environments that do not condition a change occur in the unmarked or most frequent categories. Furthermore, it is less likely that the Timberlake Effect will be observed in high-frequency paradigms in which individual forms have a greater lexical strength (accessibility) and weaker connections with related words (see Bybee 1985).

Now compare again these word-internal cases to those where change is occurring at a word boundary. The word may for a time have multiple variants, suggesting either a range of variation in the changing segment or even multiple representations for a single word. However, in this case the tendency is to resolve the variation in favor of a single form for each word, except in the case of high-frequency or grammatical words. It appears that the cases in which distinct alternates become established are just those cases in which the conditioning environment is registered in storage with the alternating item. Thus in the case of *progu*, *progi* each variant can be registered because we are dealing with two different (although related) words, one of which consistently has the palatalizing environment and one that does not. Similarly in frequent phrases, such as *muchos año*(*s*), the [s] preceding the vowel may be preserved (as though it were word-internal) because the conditioning vowel occurs with it in storage and processing. Other instances of the same word may occur without the [s], as [mučoh] or [mučo]. Such variation would not necessarily exist indefinitely. Unless one variant is in a highly entrenched phrase, the variation is likely to be eventually leveled out.

Thus by registering words in the lexicon and establishing connections among them we are able to account for the two different effects on sound change of alternating environments inside of words and across word boundaries.

6. Lexical phonology

Some of the effects of variable processes that I have discussed earlier have been addressed by Guy (1991a, 1991b) in the context of Lexical Phonology. This proposal is relevant here, even though I will argue that it does not work for all the cases at hand, because it incorporates the notion argued for here, that some words behave as if variable processes have applied to them more than once. Guy proposes that variable rules may apply cyclically and at all levels of a Lexical Phonology and offers an account for the variation in t/d deletion that is conditioned by the morphological structure of the word. The facts are as follows: on average, the highest rate of deletion of /t/ or /d/ takes place in monomorphemic words, the next highest rate in pasts with vowel changes (such as *slept*, *left*, *told*), and the lowest rate in regular past-tense forms.

(5)	[perfect]	[sleep]	[check]	
				Level 1
	may apply	______	______	t/d deletion
		[slept]	______	irregular inflection
	may apply	may apply		t/d deletion
				Level 2
			[checkd]	regular inflection
	may apply	may apply	may apply	t/d deletion
	3 chances	2 chances	1 chance	

Guy proposes derivations, as shown in (5), in which t/d deletion applies variably and at the same rate all the time, but it has three chances to apply to monomorphemic words, two chances to apply to vowel-change pasts, and one chance to apply to regular pasts.

The Lexical Phonology approach also works for Timberlake's first example, as shown in (6), where palatalization may be thought of as increasing incrementally each time the palatalization rule applies. This derives the effect of having greater degrees of palatalization in uniform than in alternating environments:

(6)	[kij]	[prog]i	
			Level 1
	[k'ij]	______	Palatalization
			Level 2
		[progi]	Regular Inflection
	[kʲij]	[prog'i]	Palatalization

The advantage of the Lexical Phonology approach is that it does recognize that fairly low-level variable phonology is deeply entwined with the lexicon and morphology. It also suggests that the greater progress of a sound change can be attributed to more applications of the "rule." The problem with it is that it makes incorrect predictions in some cases and it cannot deal at all with frequency effects across lexical items.

Consider first another case of an alternating environment that Timberlake describes. Timberlake brings up this case to show that it is the alternating environment itself, and not the morpheme boundary, that causes the retardation of change. In this case, also from Polish, there is palatalization of /k/ quite regularly and over a large geographic area before /e/ in a uniform environment, such as in the word *sausage* [kʲeu̯basa], but palatalization in [u̯okeć] 'elbow' (Nom. Sg.) is less developed and more geographically restricted, due to its alternation with the form [u̯okća] (Gen. Sg.).

The Lexical Phonology analysis of this case predicts palatalization outside the environment for [ṷokća] and advanced palatalization in [ṷokeć}, as shown in (7):

(7)	[ṷokeć]	[ṷokeć]a	
			Level 1
	[ṷok^j'eć]	[ṷok'eć]a	Palatalization
			Level 2
		[ṷok'ća]	Vowel Deletion
	*[ṷok^{j}eć]	*[ṷok'ć]a	Palatalization

The level-ordering approach could be made to work in this case by formulating the *e/zero* alternation as a vowel insertion rather than a deletion. However, most treatments of Polish regard it as a deletion, and such a change would not solve all the problems with this approach.

The Lexical Phonology approach also makes incorrect predictions concerning sound changes affecting affixes, where the context for the change is non-alternating. Earlier I mentioned the case of Spanish /d/, which is disappearing faster in the non-alternating environment of the past-participle suffix, *-ado*, than in stems. Stems would be subject to the reduction or deletion rule both at Level 1 and at Level 2, but *-ado*, since it is a part of regular inflection, would not be available to undergo the rule until Level 2. It would thus have *less* reduction and deletion than a stem, rather than more. This example shows that the Lexical Phonology approach is fundamentally the wrong approach, for it is a fact of usage, not structure, that is accelerating the change: since the /d/ in /ado/ is in the context for reduction and deletion no matter what verb it is added to and since many verbs with this suffix are of very high frequency, there is a frequency effect to accelerate the change and nothing to impede it.

Finally, Lexical Phonology, as a theory of structure and not a theory of usage, cannot account for the frequency effects demonstrable in the lexical diffusion of sound change. In Bybee (2000) I have shown that t/d deletion occurs more often in words of higher frequency. This is true of all of the 2,000 tokens studied; this relation also holds when nouns and adjectives, semi-weak past-tense verbs, and regular past-tense verbs are considered as well. Since the semi-weak past-tense verbs are all of high frequency, frequency of use alone can account for their higher rate of deletion over the regular past verbs. I conclude, then, that variable rates of phonological change are the product of usage, not of structure.

7. Conclusions

The evidence discussed in this essay bears on two issues regarding the nature of stored memory for linguistic forms. First, the minimal unit of independent storage is the word, which is also the minimal unit of production, since smaller units cannot be used in isolation. I hasten to add, however, that that does not mean that other much longer sequences are not stored and processed as wholes. Here we have seen evidence that frequently used phrases behave like single processing units (just as words typically do) in that they preserve segments that might otherwise be lost at word edges. In various papers I have argued for a highly redundant storage mechanism that includes specific instances of phrases and clauses as well as more generalized constructions as storage and processing units (Bybee and Scheibman 1999; Bybee 2000).

The view of sound change as affecting *sounds in words* according to their context of use allows us to understand why most phonological alternations occur at the word level: alternations can only be established in cases in which the conditioning environment is present in the storage and processing unit. Words or other units that occur in alternating environments that are not part of the stored unit will not have variants but rather will resolve any variation in favor of one form or the other. This proposal also allows us to make interesting predictions about the development of liaison or sandhi phenomena. Conventionalized alternations across traditional word boundaries indicate that at least one alternate is part of a larger stored unit. Thus such liaison alternations can be used to study the nature and size of storage units. Finally, the view of sound change as affecting sounds in words provides an account of the different effects of alternating environments inside of words and across word boundaries.

The second major aspect of the model presented in this essay is that sound change has an immediate and permanent effect on stored representations. This view contrasts with the generative and structural view that underlying representations remain fixed and sound change is "rule addition"—nothing more than a change in the phonological component. The evidence that sound change has an immediate effect on the lexicon is that words change gradually and at different rates according to their token frequency, even while a "rule" is still "variable." The evidence that such change is permanent is the fact that old underlying forms *never* resurface, even when the "phonological rule" becomes unproductive (see Cole and Hualde 1998 for more evidence on this point). Instead the progress of a change is inexorably unidirectional in both a phonetic and a morpho-lexical sense. In the phonetic sense, we see the unidirectionality in chains of reduction and assimilation changes, such as those shown in (8), where one change builds on the other and continues its direction:

(8) t → d → ð → Ø
s → h → Ø
k' → k^j → c

If stored items are changed gradually and the motivation for increased automation remains fairly constant, then the continuous nature and strict directionality of such changes is predicted. If sound change were "rule addition," there would be no explanation for why, for example, after "adding the rule" d → ð / V_V, a language would go on to "add the rule" d → Ø / V_V.

Inexorable unidirectionality is also apparent in the morphologization and lexicalization of the results of sound change. While no one disputes that morphologization eventually takes place, I have shown here and elsewhere that involvement with the lexicon and grammar occurs very early (Hooper 1976a, 1981; Bybee 2000). Examples given earlier are the word frequency effect in lexical diffusion, the lower rate of deletion of morphemic /t/ and /d/ in American English, the appearance of aspiration for earlier /s/ before a vowel at the end of a word in Cuban Spanish, and the examples described as the Timberlake Effect. Once involved with the lexicon and morphology, alternations become more and more entrenched and can only be undone by the strong pattern pressures we know as analogical leveling.

The two hypotheses of words as storage units and the immediate and permanent effect of sound change on words explain why most phonological alternations occur at word level, that is, why word boundaries block phonological "rules" and morpheme boundaries do not. I have also shown here that the tendency of final word boundaries to act like consonants follows from these hypotheses and from the fact that the segment most frequently following a final word boundary is a consonant.

The larger theoretical message is that use impacts representation, a point often made in studies of the discourse origins of syntax and a point that is also being made by connectionist modelers of language. As I have argued here, many cases of what was earlier postulated to be structural turn out to be derivable from the way language is used. I also see many instances where a careful look at use brings to light new data that was ignored before. I suspect that a usage-based perspective will be very productive in generating new questions and new answers in phonology.

Notes

1. In fact, the conditions under which French word-final consonants appear as liaison consonants are strikingly similar to the conditions under which [s] is retained in Caribbean dialects of Spanish (Terrell and Tranel 1978).
2. These data include only cases where a consonant preceded the final /t/ or /d/.
3. Figure 10.1 highlights the emergent morphological relationships that exist between English past-tense forms and their bases. Of course, other relationships exist, but they are not shown here for reasons of simplicity.

References

Boyland, Joyce Tang. 1997. Morphosyntactic change in progress: A psycholinguistic approach. Ph.D. diss., University of California at Berkeley.

Browman, Catherine P., and Louis M. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49, 155–180.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Philadelphia, Penn.: John Benjamins.

———. 1988. Morphology as lexical organization. In *Theoretical morphology*, ed. M. Hammond and M. Noonan. San Diego, Calif.: Academic Press.

———. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5), 425–455. (Reprinted in this volume as chapter 8.)

———. 1998. The emergent lexicon. *CLS* 31, no. 1: *The Panels*, 421–435. University of Chicago: Chicago Linguistic Society. (Reprinted in this volume as chapter 13.)

———. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language*, ed. M. Barlow and S. Kemmer, 65–85. Stanford, Calif.: CSLI. (Reprinted in this volume as chapter 9.)

Bybee, Joan L., and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37, 575–596. (Reprinted in this volume as chapter 14.)

Cole, J. S., and I. Hualde. 1998. The object of acquisition: A UR-free model. *CLS* 34.

Daugherty, K. G., and M. S. Seidenberg. 1994. Beyond rules and exceptions: A connectionist modeling approach to inflectional morphology. In *The reality of linguistic rules*, ed. S. D. Lima, R. L. Corrigan, and G. K. Iverson, 353–388. Amsterdam: John Benjamins.

Dell, G. S. 1989. The retrieval of phonological forms in production: Tests of predictions from a connectionist model. In *Lexical representation and process*, ed. W. Marslen-Wilson, 136–165. Cambridge, Mass.: MIT Press.

———. 2000. Commentary: Counting, connectionism, and lexical representation. In *Papers in laboratory phonology*, vol. 5: *Acquisition and the lexicon*, ed. Michael B. Broe and Janet B. Pierrehumbert, 335–348. Cambridge: Cambridge University Press.

D'Introno, Francisco, and Juan Manuel Sosa. 1986. Elisión de la /d/ en el español de Caracas: aspectos sociolingüísticos e implicaciones teóricas. In *Estudios sobre la fonología del español del Caribe*, ed. Rafael A Núñez Cedeño, Iraset Páez Urdaneta, and Jorge Guitart, 135–163. Caracas: Ediciones La Casa de Bello.

DuBois, J. 1985. Competing motivations. In *Iconicity in syntax*, ed. John Haiman, 343–365. Amsterdam: John Benjamins.

Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. In *Papers from the eleventh annual meeting of the Chicago Linguistic Society*, 200–213. Chicago: Chicago Linguistic Society.

Fowler, Carol A., and Jonathan Housum. 1987. Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26, 489–504.

Francis, W. Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Frisch, S. 1996. Similarity and frequency in phonology. Ph.D. diss., Northwestern University.

Givón, T. 1979. *On understanding grammar*. New York: Academic Press.

Guy, Greg. 1991a. Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3, 223–239.

———. 1991b. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3, 1–22.

Haiman, John. 1994. Ritualization and the development of language. In *Perspective on grammaticalization*, ed. William Pagliuca, 3–28. Amsterdam: John Benjamins.

Hooper, Joan Bybee. 1976a. *Introduction to natural generative phonology*. New York: Academic Press.

———. 1976b. Word frequency in lexical diffusion and the source of morphophonological change. In C*urrent progress in historical linguistics*, ed. William Christie, 96–105. Amsterdam: North Holland. (Reprinted in this volume as chapter 2.)

———. 1981. The empirical determination of phonological representations. *In The cognitive representation of speech*, ed. Terry Myers, John Laver, and John Ander, 347–357. Amsterdam: North Holland.

Hopper, P. 1987. Emergent grammar. *BLS* 13, 139–157.

Hopper, P., and S. A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56, 251–299.

———. 1984. The discourse basis of lexical categories in universal grammar. *Language* 60, 703–722.

Johnson, K. 1997. Speech perception without speaker normalization. In *Talker variability in speech processing*, ed. K. Johnson and J. W. Mullenix, 145–166. New York: Academic Press.

Johnson, Theodore. 1983. Phonological free variation, word frequency and lexical diffusion. Doct. diss., University of Washington.

Klatzky, R. L. 1980. *Human memory*. New York: W. H. Freeman.

Krug, Manfred. 1998. String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26, 286–320.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical prerequisites*. Stanford, Calif.: Stanford University Press.

Leslau, Wolf. 1969. Frequency as a determinant of linguistic change in the Ethiopian languages. *Word* 25, 180–189.

Losiewicz, Beth L. 1992. The effect of frequency on linguistic morphology. Ph.D. diss., University of Texas, Austin.

Miller, J. 1994. On the internal structure of phonetic categories: A progress report. *Cognition* 50, 271–285.

Moonwomon, Birch. 1992. The mechanism of lexical diffusion. Paper presented at the Annual Meeting of the Linguistic Society of America, Philadelphia, January.

Ohala, J. J., and M. Ohala. 1995. Speech perception and lexical representation: The role of vowel nasalization in Hindi and English. In *Papers in laboratory phonology*, vol. 4: *Phonology and phonetic evidence*, ed. B Connell and A Arvaniti, 41–60. Cambridge: Cambridge University Press.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–342.

Pierrehumbert, J. 1994. Syllable structure and word structure: A study of triconsonantal clusters in English. In *Papers in laboratory phonology*, vol. 3: *Phonological structure and phonetic form*, ed. P. A. Keating, 168–190. Cambridge: Cambridge University Press.

Pisoni, D. B., H. C. Nusbaum, P. A. Luce, and L. M. Slowiaczek. 1985. Speech perception, word recognition, and the structure of the lexicon. *Speech Communication* 4, 75–95.

Stemberger, Joseph, and Brian MacWhinney. 1988. Are inflected forms stored in the lexicon? In *Theoretical Morphology: Approaches in Modern Linguistics*, ed. Michael Hammond and Michael Noonan, 101–116. San Diego: Academic Press.

Steriade, Donca. 2000. Paradigm uniformity and the phonetics-phonology boundary. In *Papers in laboratory phonology*, vol. 5: *Acquisition and the lexicon*, ed. Michael B. Broe and Janet B. Pierrehumbert, 313–334. Cambridge: Cambridge University Press.

Terrell, Tracy. 1977. Constraints on the aspiration and deletion of final /s/ in Cuban and Puerto Rican Spanish. *Bilingual Review* 4, 35–51.

———. 1978. La aspiración y elisión de /s/ en el español porteño. *Anuario de Letras* 16, 41–66.

———. 1979. Final /s/ in Cuban Spanish. *Hispania* 62, 599–612.

Terrell, Tracy. 1986. La desparición de /s/ posnuclear a nivel léxico en el habla dominicana. In *Estudios sobre la fonología del español del Caribe*, ed. Rafael A. Núñez Cedeño, Iraset Páez Urdaneta, and Jorge Guitart, 117–134. Caracas: Ediciones La Casa de Bello.

Terrell, T., and B. Tranel. 1978. Parallelism between liaison in French and /s/ aspiration and deletion in Caribbean Spanish dialects. *Montreal Working Papers in Linguistics* 10, 31–50.

Timberlake, Alan. 1978. Uniform and alternating environments in phonological change. *Folia Slavica* 2, 312–328.

Tranel, B. 1981. *Concreteness in generative phonology: Evidence from French*. Berkeley: University of California Press.

Chapter 11

Word Frequency and Context of Use in the Lexical Diffusion of Phonetically Conditioned Sound Change

Lexical diffusion refers to the way that a sound change affects the lexicon. If sound change is lexically abrupt, all the words of a language are affected by the sound change at the same rate. If a sound change is lexically gradual, individual words undergo the change at different rates or different times. Whether sound changes exhibit gradual or abrupt lexical diffusion is a topic of some recent concern. One early contribution to this debate by Schuchardt (1885) is the observation that high-frequency words are affected by sound change earlier and to a greater extent than low-frequency words. In this article, after discussing the recent literature on lexical diffusion, I document this tendency in a variety of sound change types, discuss briefly the opposite tendency (i.e., for low-frequency words to be affected by a change first), and propose explanations for both types of change. I argue that phonetically conditioned changes that affect high-frequency words before low-frequency words are best accounted for in an exemplar model of phonological representation that allows for change to be both phonetically and lexically gradual. Next I show that a word's contexts of use also affect the rate of change. Words that occur more often in the context for change, change more rapidly than those that occur less often in that context. The data for this exploration are from English /t, d/ deletion and involve contracted negation, words ending in unstressed *-nt*, and the regular past tense morpheme. The contexts of use for bound morphemes are examined, using the data from the deletion of [ð] in New Mexican Spanish, to show that sound changes can also progress more rapidly in high-frequency morphemes. The role of morphemes in lexical diffusion is examined further by comparing American English /t, d/ deletion, in which a change is retarded in a high-frequency morpheme, to Spanish [ð] deletion. It is proposed that the contexts of use determine the rate at which a word or morpheme undergoes a sound change.

This article only addresses word frequency as a conditioning factor in phonological variation and assumes that the variation remaining after frequency is considered is due to other well-established social and linguistic factors. My goal is to demonstrate that word frequency and context frequency are factors that can affect variation and should be taken into account in future studies of phonological variation and change.

1. Regular sound change or lexical diffusion?

The hypothesis that sound change is lexically regular seems well supported by the facts of change. When we observe that two languages or dialects exhibit a phonological difference, it is very likely that this difference is regular across all the words that have the appropriate phonetic environment. This observation is fundamental to the comparative method; the establishment of genetic relations and the reconstruction of protolanguages are based on the premise that sound change affects all words equally. Schuchardt (1885) was one of the detractors from this position. When he observed sound change in progress, he noted that all words did not change at the same rate and that the differences were not due to "dialect mixture," as was often claimed by the Neogrammarians, who supported the regularity position.

A major challenge to the regularity position in the twentieth century is expressed in the work of Wang and his colleagues (Wang 1969, 1977; Wang and Cheng 1977), which documented changes that seem to occur word by word over a long period of time. While some of these changes result in lexical regularity, Wang and his colleagues also identified changes that appear to be arrested after affecting only part of the lexicon.

Labov (1981, 1994) also dealt with the issue, availing himself of the data from his numerous studies of sound change in progress. He proposed two types of sound change: "regular sound change" is gradual, phonetically motivated, without lexical or grammatical conditioning, and not influenced by social awareness, whereas "lexical diffusion" change, such as the phenomena studied by Wang, is "the result of the abrupt substitution of one phoneme for another in words that contain that phoneme" (Labov 1994:542). According to Labov, this type of change occurs most often "in the late stages of internal change that has been differentiated by lexical and grammatical conditioning" (ibid.). Labov went so far as to propose that certain changes, such as the deletion of glides and schwa, would be regular changes, while the deletion of obstruents would show lexical diffusion.

A number of researchers have challenged this position. Phillips (1984) argued that even low-level sound changes exhibit gradual lexical diffusion. Similarly, Oliveira (1991) argued that it is likely that gradual lexical diffusion occurs even in changes that turn out to be regular. Krishnamurti (1998) demonstrated that the change of *s* > *h* > *Ø* in Gondi exhibits gradual lexical diffusion but still goes through to completion in some dialects. In this review I present evidence to show that even gradual, phonetically conditioned change exhibits gradual lexical diffusion, though it is perhaps of a more subtle nature than the lexical diffusion studied by Wang and Labov.

The lexical diffusion examined here for reductive phonetic change is highly conditioned by word frequency.

Hooper (1976) identified a lexical diffusion paradox. Reductive sound change tends to affect high-frequency words before low-frequency words, but analogical leveling or regularization tends to affect low-frequency words before high-frequency words. Working from this observation, Phillips (1984, 2001) studied a number of changes that move in each direction and attempted to refine a hypothesis that predicts the direction of lexical diffusion. In this article I focus on the first direction of change, the characteristic of reductive sound change, with the intent of demonstrating that word frequency affects even changes that may have seemed regular to the Neogrammarians. The second type of change, which affects low-frequency words earlier than high-frequency words, is discussed briefly.

2. Frequency effects on regular sound change

Sound changes that are complete can be identified as regular or not, depending upon whether they affected all lexical items existing at the time of the change. Ongoing changes cannot be designated as regular or not, since they are not complete. However, one can reference the typical characteristics of a change to project whether it will be regular or not. That is, a phonetically gradual change with clear phonetic conditioning falls into Labov's first type, and thus we can project its regularity. Here I document lexical diffusion from high-frequency to low-frequency words in ongoing changes that can be expected to be regular, as well as in certain reductive changes that may never be complete because of the nature of the lexicon.[1]

2.1. American English t/d deletion

Consider the deletion of final /t/ and /d/ in American English, which occurs most commonly in words ending in a consonant plus /t/ or /d/, such as *just*, *perfect*, *child*, or *grand*. This much-studied variable process has been shown to be affected by the preceding and following consonant, with more deletion in a consonant environment; by grammatical status, with less deletion if the /t/ or /d/ is the regular past tense; and by social and age factors, with more deletion among younger, lower socioeconomic class speakers (Labov 1972; Neu 1980).

My study (Bybee 2000b) of the deletion of /t/ and /d/, using a corpus of phonological variation in Chicano English speakers in Los Angeles (Santa Ana 1991), focused on lexical frequency as a factor.[2] Using approximately two thousand tokens of final /t/ and /d/ after a consonant, as transcribed by Santa Ana, and referencing word frequency from Francis and Kučera (1982), I found that deletion occurred more in high-frequency words. Table 11.1 demonstrates this effect with a cutoff point of 35 per million. This number was chosen because I was also interested in whether a frequency effect could be found among regular past-tense verbs, and this figure was the median for such forms in Francis and Kučera (1982).

Similar results were obtained by Jurafsky, Bell, Gregory, and Raymond (2001) using 2,042 monosyllabic content words ending in /t/ or /d/ from the Switchboard

TABLE 11.1. Rate of /t, d/ Deletion for Entire Corpus by Word Frequency

	Number of Words with High Frequency	*Number of Words with Low Frequency*
Retention	752 (45.6%)	262 (65.7%)
Deletion	898 (54.4%)	137 (34.3%)

$\chi^2 = 41.67$, $p < .001$.

High-frequency words occurred 35 or more times per million in Francis and Kučera (1982); low-frequency words occurred less than 35 times per million.

corpus, a corpus of telephone conversations between monolingual American English speakers.[3] They found a strong effect of word frequency ($p < .0001$). Another study, using function and content words as well as polysyllabic and monosyllabic words from the Switchboard corpus, found an even higher level of significance for the association of word frequency with final /t/ and /d/ deletion ($p < .00005$) (Gregory et al. 1999).

Will final /t/ and /d/ deletion after a consonant turn out to be a regular sound change? There is certainly precedent for such a change being regular in the end, especially in certain phonetic contexts. Final consonants in Latin were deleted completely in Spanish and other Romance languages. Latin final /m/ was deleted in all words of two syllables or more (e.g., Latin *quindecim* > Spanish *quince* 'fifteen', *caballum* > *caballo* 'horse', *novem* > *nueve* 'nine'). Latin /t/ was also deleted (*caput* > *cabo* 'end', *amat* > *ama* '3s loves'), as were other consonants (Menéndez-Pidal 1968 [1904]). In English, erosion has been working on final consonants for some time. The deletion of /b/ and /g/ after a homorganic nasal, as in *bomb* and *gang*, was completely regular and left English speakers virtually unable to produce final [mb] or [ng] clusters. Final /nd/ could certainly follow and delete regularly as well.

Can the deletion of an obstruent be phonetically gradual? Labov (1994) listed obstruent deletion under "lexical diffusion" changes, presumably because he considered obstruent deletion to involve the phonetically abrupt loss of a phoneme. However, there is evidence that the reduction of final /t/ and /d/ can be gradual. A final /t/ or /d/ may vary in length, and this length variation occurs under the same conditions and in the same direction as the deletion variation. In a laboratory experiment, Losiewicz (1992) found that monomorphemic /t/or /d/ is shorter in duration than regular past tense /t/ or /d/. As mentioned before, monomorphemic /t/ or /d/ is also more likely to delete. Figures from Bybee (2000b) confirm this, as shown in table 11.2. In addition, Losiewicz found that a final past-tense /t/ or /d/ is longer in low-frequency verbs than in high-frequency verbs. Her finding for length variation among past-tense verbs paralleled the deletion data presented in Bybee (2000b), as shown in table 11.3. Regular past tense /t/ and /d/ are more likely to delete in high-frequency verbs than in low-frequency verbs.

The data on this obstruent deletion process, then, suggest both lexical and phonetic gradualness. It thus cannot be said that obstruent deletion is the abrupt deletion

TABLE 11.2. Rate of Deletion for Regular Past Tense Compared to All Other Words of Comparable Frequency (403 or Less)

	Past Tease	*All Other Words*
Retention	218 (77.3%)	831 (54.0%)
Deletion	64 (22.6%)	709 (46.0%)

$\chi^2 = 53.2$, $p < .0001$.

of a phoneme. In fact, these data are problematic for any version of phonemic theory. A model that can accommodate these data is presented later in this article.

2.2. Spanish [ð] deletion

Another good candidate for a phonetically gradual change in progress that exhibits lexical diffusion and that could turn out to be regular is the deletion of intervocalic /d/ or [ð] in many dialects of Spanish. D'Introno and Sosa (1979) regarded the variants ranging from [d] to [ð] to Ø as a continuum, affirming that the reduction is gradual. My own study of approximately one thousand tokens of intervocalic /d/ in Spanish spoken by native New Mexican speakers used samples of Spanish-dominant speakers from the corpus of interviews for the Survey of the Spanish of New Mexico and Southern Colorado (Bills 1997). All medial instances of /d/ were transcribed as present or absent. The results revealed that the rate of deletion is higher among high-frequency words. The frequency count used for table 11.4 is taken from the Corpus Oral de Referencia del Español Contemporáneo (COREC), a 1.1 million word corpus of lectures, discussions, and conversations of middle-class Spaniards (Marcos-Marín 1992). Past-participle tokens were removed because the data showed that past participles have a higher rate of deletion than other items. The cutoff point between high and low frequency was arbitrarily chosen to make the number of tokens in the high and low groups approximately equal.

Again we have evidence that obstruent deletion can diffuse gradually through the lexicon, affecting high-frequency words earlier than low-frequency words. We do not know if this sound change will turn out to be completely regular. The earlier deletion of Latin /d/ between vowels left some residue (Menéndez-Pidal 1904

TABLE 11.3. Effects of Word Frequency on /t, d/ Deletion in Regular Past Tense Verbs (Non-prevocalic Only)

	High Frequency	*Low Frequency*
Retention	67 (60.4%)	47 (81.0%)
Deletion	44 (39.6%)	11 (19.0%)

$\chi^2 = 5.00313$, $p < .05$.

TABLE 11.4. Rate of Deletion according to Token Frequency for All Non-Past-Participle Tokens in the New Mexican Corpus, Using the COREC as a Measure of Frequency

	Low-Frequency Words (0–99)	*High-Frequency Words (≥100)*	*Total*
Retention	243 (91.4%)	287 (78.6%)	530 (84.0%)
Deletion	23 (8.6%)	78 (21.4%)	101 (16.0%)
Total	266	365	631

$\chi^2 = 17.3$, $p < .001$.

[1968]. Thus Latin *credit* gives Spanish *cree* '3s believes', *foedu* > *feo* 'ugly', *pedes* > *pies* 'feet', *fide* > *fe* 'faith', *audire* > *oir* 'to hear', *limpidu* > *limpio* 'clean', and so on, but a few words maintain the /d/, as in *sudare* > *sudar* 'to sweat', *vadu* > *vado* 'ford', *crudu* > *crudo* 'raw', *nidu* > *nido* 'nest', *nudu* > (*des*)*nudo* 'naked'. However, all of these words are attested without the *d* in Old Spanish, which led Penny (1991) to suggest that the forms with the *d* are influenced by the Latin spelling. Others have argued that the deletion was constrained by whether the resulting vowel combination obeyed the phonotactics of Spanish at the time (see Pensado 1984). Thus while the earlier change was not completely regular, it was comparable to the present deletion in being phonetically gradual, and it affected a large majority of words with intervocalic Latin *d*. I suspect the Neogrammarians would have counted it as "regular."

2.3. Vowel shifts

Labov (1994) searched his data on vowel shifts in American English for evidence of lexical diffusion. He was looking for cases in which words with the same phonetic environment for the vowel had distinctly different vowel realizations, as opposed to small differences that could be part of a phonetically gradual continuum. The vowel changes reviewed in his chapter 16 showed detailed phonetic conditioning, gradualness, and lexical regularity. Labov even tested to see if homonyms could split given differences in their token frequency (1994:460–465). This test, on words with very similar phonetic environments, such as *two*, *too*, *to*, *do*, *through*, turned up no lexical diffusion by word frequency. The test, however, was not definitive, since all of the words used occurred three or more times in the interview and thus must be considered high-frequency. In addition, taking the very small set of words necessary to get comparable phonetic environments yielded too few words to discover any general word frequency effect in these gradual changes. Indeed, it might be that for vowel shifts the phonetic environment is generally more powerful than any other effects because there are fewer words in each phonetic environment.

However, some lexical diffusion was found in Labov's data on vowel shifts. The most famous and still unexplained case concerns the raising of short [æ] in Philadelphia, which affects the adjectives ending in [d]: *mad*, *glad*, *bad*, but not *sad*. This same shift provides some evidence for lexical diffusion by frequency.

Labov (1994:506–507) noted that when word-initial short [æ] "occurs before a voiceless fricative, only the more common, monosyllabic words are tensed: tense *ass* and *ask*; lax *ascot*, *aspirin*, *astronauts*, *aspect*, *athletic*, *after*, *African*, *Afghan*."

Similarly, in Moonwomon's (1992) study of the centralization of /æ/ in San Francisco English, she found that this vowel is more centralized in the environment before a fricative than before a nonfricative; it is also more centralized after [l]. The most commonly used word with this pair of phonetic environments is *class*. *Class* shows more centralization than the other words with these two environments (e.g., *glass*, *laugh*). Moonwomon also studied the fronting of /ɑ/ in the same speakers. Here a following /t/ or /d/ conditions more fronting than other consonants. Of the words in the corpus ending in final /t/, *got* occurred most frequently. Moonwomon found that the fronting in *got* is significantly more advanced than in other words ending in alveolars, such as *not*, *god*, *body*, *forgot*, *pot*.

It appears, then, that there is some evidence that high-frequency words undergo vowel shifts before low-frequency words. As previously mentioned, it may be more difficult to discern frequency effects in vowel shifts because of the effects of the preceding and following environments, which narrow each phonetic class to a small number of words.

2.4. Vowel reduction and deletion

In addition to consonant reduction, another type of change that shows robust word frequency effects is vowel reduction and deletion. Fidelholtz (1975:200–201) demonstrated that the essential difference between words that reduce a prestress vowel, such as *astronomy*, *mistake*, *abstain*, and phonetically similar words that do not, such as *gastronomy*, *mistook*, *abstemious*, is word frequency. Van Bergem (1995) found that reduction of a prestress vowel in Dutch also is highly conditioned by frequency. The high-frequency words *minuut* 'minute', *vakantie* 'vacation', *patat* 'chips' are more likely to have a schwa in the first syllable than the phonetically similar low-frequency words *miniem* 'marginal', *vakante* 'vacant', *patent* 'patent'.

Deletion of reduced vowels may also be conditioned by word frequency. Hooper (1976) asked native speakers of American English for their judgments on whether a poststress schwa was usually, sometimes, or rarely deleted in words such as *every*, *memory*, and *family*.[4] The results showed that subjects exercised some phonological constraints on schwa deletion, but aside from these, the contexts in which deletion occurred were highly influenced by word frequency. That is, the subjects judged deletion to be more likely in *nursery* than in *cursory*, in *memory* than in *mammary*, in *scenery* than in *chicanery*, and so on.

These deletions appear to be phonetically gradual in that the variants range from those in which a schwa is followed by a resonant, in these cases [r], to those in which the resonant is syllabic, to those in which the resonant is not syllabic (i.e., all syllabicity has been lost). Such changes, then, are both phonetically gradual and lexically gradual. However, they may never fall into the category of regular sound change. The reason is that as new words enter the language or as low-frequency words become more frequent, there will always be new schwas developing in the context for this deletion process. Similarly, the vowel reductions discussed by Fidelholtz and Van Bergem may never be complete. As new full vowels come into unstressed position there will be new material for the reduction to work on.

2.5. Factors that cause more reduction in high-frequency words

The cases documented so far indicate that high-frequency words tend to change before low-frequency words when the change is the deletion of stops (English /t, d/ deletion), the deletion of fricatives (Spanish ð deletion), some vowel shifts (Moonwomon 1992; Labov 1994), the reduction of vowels to schwa (in both Dutch and English), and the deletion of schwa (in American English). One might therefore predict that, in general, reductive changes tend to occur earlier and to a greater extent in words and phrases of high frequency.

If we take linguistic behavior to be highly practiced neuromotor activity (Anderson 1993; Haiman 1994; Boyland 1997), then we can view reductive sound change as the result of the automation of linguistic production. It is well known that repeated neuromotor patterns become more efficient as they are practiced; transitions are smoothed by the anticipatory overlap of gestures, and unnecessary or extreme gestures decrease in magnitude or are omitted. Recent theories of articulatory change point to precisely these two types of changes—increase in overlap of gestures and decrease in magnitude—as describing all changes that occur in casual speech (Browman and Goldstein 1992) or in sound change (Mowrey and Pagliuca 1995).

If casual speech processes and reductive sound change are the natural result of the automation of linguistic productions, then it follows that such change will be more advanced in productions that are more highly practiced (i.e., in high-frequency words and phrases). In fact, reductive sound change may be just the most salient aspect of an overall reduction of high-frequency words. Using the Switchboard corpus, Jurafsky et al. (2001) found that for 1,412 tokens of monosyllabic content words ending in /t/ or /d/, high-frequency words (at the 95th percentile) were 18% shorter than low-frequency words (at the 5th percentile). That is, the entire articulatory span of high-frequency words may be reduced compared to low-frequency words, a phenomenon that could give rise to some noticeable and some not so noticeable articulatory changes.

Other findings by Jurafsky and colleagues indicate that articulatory reduction cannot run rampant but may be constrained by predictability in discourse. Using the same set of tokens, Jurafsky et al. (2001) found that the predictability of the word given the following word affects duration. That is, the first words in frequently occurring word pairs, such as *Grand Canyon*, *grand piano*, or *Burt Reynolds*, were shorter than words used in less predictable contexts. Gregory et al. (1999) found that among 4,695 monosyllables ending in /t/ or /d/ from the Switchboard corpus, semantic relatedness to the discourse topic affected word duration: words related to the discourse topic were shorter than words that were not. In addition, in Gregory et al.'s study, words tended to be shorter if they were repeated in the same discourse (see also Fowler and Housum 1987). In both studies the findings are explained with reference to predictability: predictable words are more reduced. A fuller interpretation of these results might run as follows. The tendency for articulatory reduction due to increased automation is always present and shows itself more prominently in highly practiced, frequent words. Reduction can be inhibited by the speaker's sensitivity to the predictability of words in the context. If the speaker knows that the word will be easily accessed in the context, because it or related words have already been activated, the reductive automating processes will be allowed to advance. If the word is less predictable in discourse, the speaker is likely to suppress the reductive processes and to give the word a more complete articulation.

3. Changes that affect low-frequency words first

As previously mentioned, Hooper (1976) noted a lexical diffusion paradox: sound change seems to affect high-frequency words first, but analogical change affects low-frequency words first. The first tendency has already been documented. The second tendency is evident in the fact that low-frequency verbs, such as *weep/wept*, *leap/leapt*, *creep/crept*, are regularizing, while high-frequency verbs with the same pattern show no such tendency: that is, *keep/kept*, *sleep/slept*, *leave/left* show no evidence of regularizing. Hooper (1976) argued that changes affecting high-frequency words first have their source in the automation of production, whereas changes affecting low-frequency words first are due to imperfect learning. In the latter category are changes that affect words that do not conform to the general patterns of the language. Such exceptional words can be learned and maintained in their exceptional form if they are of high frequency in the input and in general use. However, if their frequency of use is low, they may not be sufficiently available in experience to be acquired and entrenched. Thus they may be subject to changes based on the general patterns of the language.

The example just given has a clear morphological motivation, but Phillips (1984) showed that some sound changes can proceed from low-frequency to high-frequency words. For instance, the Old English diphthong *eo* monophthongized to a mid front rounded vowel /ö/, with both a long and a short version in the eleventh to twelfth centuries. In some dialects these front rounded vowels were maintained into the fourteenth century, but in Lincolnshire they quickly unrounded and merged with /e(:)/. A text written around A.D. 1200, the Ormulum, captures this change in progress. The author was interested in spelling reform, and so, rather than regularizing the spelling, he represented the variation using two spellings for the same word (e.g., *deop*, *dep* 'deep') in many cases. Phillips found that within the class of nouns and verbs, the low-frequency words are more likely to have the spelling that represents the innovation, the unrounded vowel.

If this were a phonetically motivated reduction that facilitates production, we would expect the high-frequency words to change first. Indeed, the frequent adverbs and function words have changed, suggesting that they might be yielding to production pressures, but the fact that nouns and verbs show more change in low-frequency items indicates a different motivation for the change. Phillips proposed that a constraint against front rounded vowels is operating to remove these vowels. Since there were no other front rounded vowels in English at the time, the majority pattern would be for front vowels to be unrounded. The mid front rounded vowels would have to be learned as a special case. Front rounded vowels are difficult to discriminate perceptually, and children acquire them later than unrounded vowels. Gilbert and Wyman (1975) found that French children confused [ö] and [ɛ] more often than any other non-nasal vowels they tested. A possible explanation for the Middle English change, then, is that children correctly acquired the front rounded vowels in high-frequency words that were highly available in the input, but tended toward merger with the unrounded version in words that were less familiar.

Thus there appear to be at least two directions for lexical diffusion with regard to frequency: some changes affect high-frequency words earlier, and some affect low-frequency words earlier. The former changes are phonetically conditioned and gradual, but they may not all turn out to be lexically regular. The evidence cited here suggests that lexical diffusion is much more common than previously supposed. It may be that all sound change diffuses gradually through the lexicon. For this reason we must reject the dichotomy of regular versus lexical diffusion changes and look at the pattern of lexical diffusion for each change. Each pattern of diffusion is associated with a particular source and mechanism for change, which allows us to use the direction of diffusion as a diagnostic for the cause of change (Hooper 1976; Phillips 1984, 2001; Bybee 2001). Changes that affect high-frequency words first are the result of the automation of production, while low-frequency words change first when the change makes the words conform to the stronger patterns of the language. Low-frequency words, with their lesser availability in experience and consequently their weaker representation, are more susceptible to analysis and change on the basis of other forms.

4. Modeling phonetic and lexical gradualness

The view of lexical diffusion espoused by both Wang and Labov assumes that a change that diffuses gradually through the lexicon must be phonetically abrupt. This is a necessary assumption if one is to accept a synchronic phonological theory that has phonemic underlying representations. Words can change one by one only if the change is a substitution of phonemes in such a theory. The discovery that sound change can be both phonetically gradual and lexically gradual forces a different view of the mental representation of the phonology of words (Hooper 1981; Bybee 2000b). If subphonemic detail or ranges of variation can be associated with particular words, an accurate model of phonological representation must allow phonetic detail in the cognitive representation of words.

A recent proposal is that the cognitive representation of a word can be made up of the set of exemplars that have been experienced by the speaker/hearer. Thus all phonetic variants of a word are stored in memory and organized into a cluster: exemplars that are more similar are closer to one another than to ones that are dissimilar, and exemplars that occur frequently are stronger than less frequent ones (Johnson 1997; Bybee 2000a, 2001; Pierrehumbert 2001). These exemplar clusters, which represent autonomous words, change as experience with language changes. Repeated exemplars within the cluster grow stronger, and less frequently used ones may fade over time, as other memories do.

Changes in the phonetic range of the exemplar cluster may also take place as language is used and new tokens of words are experienced. Thus the range of phonetic variation of a word can gradually change over time, allowing a phonetically gradual sound change to affect different words at different rates. Given a tendency for reduction during production, the phonetic representation of a word will gradually accrue more exemplars that are reduced, and these exemplars will become more likely to be chosen for production, where they may undergo further reduction, gradually moving the words of the language in a consistent direction. The more frequent words will have more chances to undergo online reduction and thus will change more rapidly. The more predictable words (which are usually also the more frequent ones) will have a greater chance of having their reduced version chosen, given the context, and thus will advance the reductive change more rapidly.

The exemplar clusters are embedded in a network of associations among words that map relations of similarity at all levels. Distinct words with similar phonetic properties are associated, as are words with shared semantic features. I have shown in (Bybee 1985, 1988) that morphemes and morphological relations in such a network emerge from parallel phonetic and semantic associations and that schemas or abstractions over relations of similarity can be formulated to account for the regularities and patterns evident in language use. Such a network is illustrated in figure 11.1, where the connection lines indicate parallel semantic and phonological relations among instances of the past tense /d/.

Figure 11.2 shows how complex morphological relations can emerge from the connections based on phonological and semantic similarity. These figures, of course, only show fragments of the network within which these words are embedded, and they do not show all the relations that exist. Other questions about this model with respect to phonology and morphology are answered in Bybee (2001) and Bybee (1985), respectively.

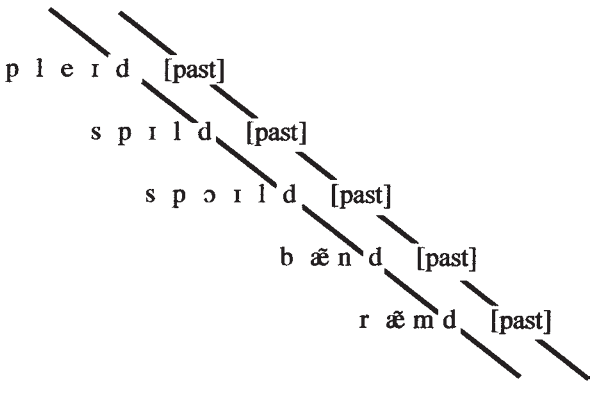


FIGURE 11.1. Network representation for some English past-tense verbs.

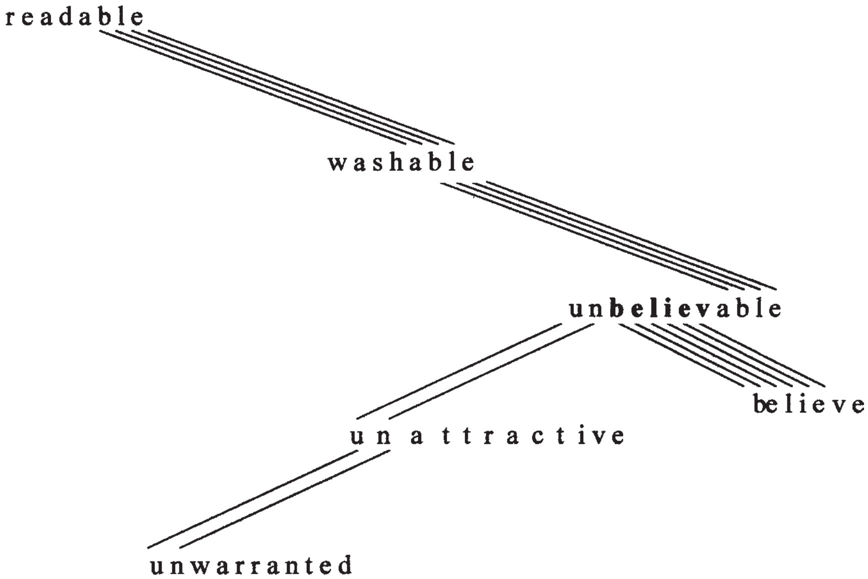


FIGURE 11.2. Lexical connections supply a morphological analysis of words such as *unbelievable.*

An important property of the exemplar model is the emphasis on words as storage units. Various models have proposed that even multi-morphemic words have lexical listing. Vennemann (1974) argued that appropriate constraints on syllable structure can only be applied to whole words, not to morphemes. The common objection to this proposal made in the 1970s was that the human brain does not have sufficient storage capacity for all the words of a language, especially a language with large morphological paradigms. This argument has now been dismissed with the discovery of the huge amount of detail that the brain is capable of recording. Moreover, newer conceptions of the lexicon, not as a list but as a network with tight interconnections, provide the insight that listing two related words, such as *start*, *started*, does not take up as much cognitive space as listing two unrelated words, such as *start*, *flower* (Bybee 1985). Thus connectionist models (Rumelhart and McClelland 1986) and analogical models (Skousen 1989, 1992; Eddington 2000) have storage of whole words with morphological relations emergent from the categorization involved in storage. In addition, the lexical diffusion data provide evidence that multi-morphemic words can have lexical storage. As we saw in table 11.3, high-frequency regular past-tense English verbs are more likely to have their final /t/ or /d/ deleted than are low-frequency regular verbs. In order for a frequency effect to accrue to a word, that word must exist in memory storage. Since multi-morphemic words evince frequency effects, they must be stored in the lexicon. I return to the issue of how morphological relations are expressed in a network model later in this article.

5. The effect of frequency of use in context

Given that the exemplar model tracks tokens of use and the exemplar cluster changes according to the phonetic shape of these tokens, it follows that if the context of use affects the phonetic shape of a word, the exemplar cluster will change accordingly. The effect of context can be best exemplified in changes that take place around word or morpheme boundaries, where the segment affected by the change is sometimes in the context for the change and sometimes not. Timberlake (1978) called this an alternating environment. Since the exemplar model registers phonetic tokens, the probabilities inherent in the alternating environment affect the shape of the exemplar cluster. For example, the aspiration of syllable-final /s/ in many dialects of Spanish is conditioned by a following consonant, but when that /s/ is word-final, it sometimes precedes a consonant and sometimes a vowel.[5] At first, some tokens will have [s] (those that precede vowels) and some will have [h] (those that precede consonants). These variants become part of the exemplar cluster. Except for words of very high frequency, this type of variation in the phonetic shape of a word seems not to be stable. Rather, over time one variant begins to occur in all environments. In the case of Spanish [s] aspiration, the [h] variant begins to occur even prevocalically (Bybee 2000a, 2001). The reason for this is that [h] is the variant for the preconsonantal position, which is the most frequently occurring context in running discourse, as most words in Spanish begin with consonants. The exemplar cluster, then, appears to reorganize itself, with the stronger exemplars being more frequently chosen for use than the less frequent ones despite the context.

Thus along with the general measure of frequency of use, the relative frequency of the immediate linguistic context of use can also affect the lexical diffusion of a sound change. Even holding frequency constant, a word that occurs more often in the right context for a change will undergo the change more rapidly than a word that occurs less often in the conditioning context. The next section examines the effect of specific contexts of use on American English t/d deletion.

5.1. Special contexts for past tense and contracted negation

The deletion of word-final /t/ or /d/ in American English has been studied from a number of different perspectives. Of particular relevance here, Guy (1991a) studied the effects of contextual conditioning on both uniform and alternating environments with respect to /t, d/ deletion. The word-internal context (i.e., the segment preceding the final /t/ or /d/) is uniform or constantly present for each word, whereas the word-external context (i.e., the segment following the word final /t/ or /d/) is an alternating environment because it varies according to what word follows the /t/ or /d/ in question. Guy found that the word-external context has a lesser effect on the deletion of the /t/ or /d/ than does the word-internal context—a finding predicted by the exemplar model. For example, with a word such as *different*, the /t/ always follows /n/, which presents a favorable condition for deletion, but sometimes *different* precedes a vowel, as in *different one*, and other times it precedes a consonant, as in *different story*. Guy found that the preceding environment was the factor with the strongest effect on deletion, conditioning more consistently than the following environment.

As we shall see, this study reveals the same phenomenon. For example, the *t* in words ending in *-nt* actually deletes more often before vowels than before consonants.

The differential effects of uniform and alternating environments find a natural representation in an exemplar model. As tokens of use are registered for a given word, the preceding word-internal context is constant and provides the conditioning environment in each use. The following word-external context varies, sometimes providing a conditioning context and sometimes not. Thus a word with a preceding favoring environment will have more exemplars with deletion than one without such a preceding favoring environment, other factors being equal.

The English regular past-tense /t/ or /d/ has an alternating environment in both the preceding and following contexts. Since past tense occurs on verbs with any final segments, sometimes the past tense follows a vowel (*played*, *tried*), sometimes a resonant (*scored*, *called*), and sometimes an obstruent (*slipped*, *rubbed*). Thus the preceding context alternates, just as the following word-external context alternates. To the extent that speakers generalize over different instances of the same bound morpheme, the instances with a preceding environment that does not favor deletion may retard the deletion across all items, thus providing part of the explanation for the fact that the past-tense /t/ and /d/ delete at a lower rate than other instances of /t/ and /d/. Another important factor for this lower rate of deletion is taken up directly: the environment after the past tense also tends not to favor deletion.

The demonstrated effect that the context has on the deletion rate suggests that a closer look at the contexts in which a word occurs may provide more insight into the rate at which it undergoes a change. For example, in Santa Ana's (1991) corpus, used to study /t, d/ deletion in Bybee (2000b) (henceforth, the LA corpus), 23% of words ending in /t/ or /d/ preceded a vowel. However, in different subsets of words this rate varied from 9 to 40%. Since the position before a vowel is not conducive to deletion, variations in this rate may affect the rate of deletion for certain sets of words.

We coded approximately two thousand tokens of words ending in /t/ or /d/ from the LA corpus. Even though the subjects chosen were dominant in English, there might be some question as to the possible effects of Spanish on their speech. Thus the specific results reported on here were checked in the phonemically transcribed portions of the Switchboard corpus, a corpus of telephone conversations among people living in the Dallas, Texas, area (Greenberg 2000). As we shall see, all results from the LA corpus were replicated in the Switchboard corpus.

I examined three sets of words: (1) auxiliaries with contracted negation, (2) a set of words ending in an unstressed syllable with /nt/ (*government*, *different*, *pregnant*), which was intended as a phonologically matched control for the contracted negation, and (3) the regular past-tense forms in the corpus. The specific hypothesis tested was that words that occur more often before vowels, an environment that phonetically favors retention, would exhibit less deletion overall. This hypothesis assumes that the deletion is no longer purely conditioned by the phonetic environment, but that effects of the change are already represented lexically. The evidence for this assumption is the fact that for two of these groups, the rate of deletion preceding a vowel versus preceding a consonant is not predictable from the phonetics. In fact, table 11.5 shows that, for contracted negation the deletion rates are approximately the same in both contexts and that words ending in *-nt* actually exhibit more deletion

TABLE 11.5. Rates of Deletion for Three Word Groups and All Words before a Pause, a Vowel, and a Consonant

	Prepausal	*Prevocalic*	*Preconsonantal*
n't	16/39 (41%)	31/36 (86%)	254/304 (84%)
-nt	8/21 (38%)	9/13 (69%)	11/25 (44%)
-ed	5/37 (14%)	8/112 (7%)	61/130 (47%)
All words	111/296 (38%)	158/428 (37%)	756/1272 (59%)

before a vowel than before a consonant. Only the past-tense *-ed* shows deletion rates more in line with the original phonetic conditioning, a fact I return to later.

The hypothesis, then, is that, as lexicalization proceeds and the exemplar cluster changes, the overall distribution of the exemplars will reflect their distribution in the experience of the speaker at first, and eventually the more frequent exemplars will take over in more and more contexts. Thus, words that occur more frequently before vowels will show less deletion overall than words that occur more frequently before consonants. Table 11.6 demonstrates that this hypothesis holds for the three sets of words studied in the LA corpus. The overall deletion rates for these word sets are inversely related to their occurrence before vowels in the corpus.

The distributions of these sets of words are strikingly different, as can be observed from table 11.7. Auxiliaries with contracted negation occur overwhelmingly before consonants (approximately 80% of the cases in the LA corpus). Of course, their most common site is before a verb and a large majority of verbs begin with consonants. The contracted negation occurs before a pause in only about 10% of the cases. The higher deletion rate for contracted negation, then, is attributable to its distribution in spoken discourse, where it commonly occurs internal to clauses and closely linked to the following verb, which is highly likely to be consonant-initial.

TABLE 11.6. Percentage Overall Deletion and Occurrence before Vowels for Three Sets of Words in the LA Corpus

	Overall Deletion[a]	*Occurrence before Vowels*[b]
n't	301/379 (79.4%)	36/379 (9.5%)
-nt	28/59 (47.4%)	13/59 (22.0%)
-ed	74/279 (26.5%)	112/279 (40.1%)

[a]For overall deletion, comuarison of *-n't* to *-nt* yields $\chi^2 = 27.90025$, $p < .0001$, and comparison of *-ed* to *-nt* yields $\chi^2 = 10.12929$, $p < .001$.

[b]For occurrence before vowels, comparison of *n't* to *-nt* yields $\chi^2 = 8.073901$, $p < .01$, and comparison of *-ed* to *-nt* yields $\chi^2 = 6.757618$, $p < .01$.

Because exemplars without the final /t/ are highly frequent, they come to dominate the exemplar clusters, giving high rates of deletion overall.

Words that ended in an unstressed syllable ending in *-nt* were examined in an attempt to find a phonological string that matched the contracted negation but without a morphological function. Of course, these words do not provide an exact phonological match to contracted negation in most cases. So although *student* ends in the same consonants as *didn't*, words such as *pregnant*, *government* do not provide exact matches to any contracted negatives. However, these words do constitute a set of nouns and adjectives with which to compare the contractions and the past tense in terms of distribution. As tab1e 11.7 shows, these nouns and adjectives have a very different distribution from the auxiliary and contracted negation complex. They occur prepausally about 36% of the time. Their prevocalic occurrence of 22% is almost the same as the overall average for the corpus. They occur preconsonantally in 42% of the cases. The lower number of preconsonantal cases and the higher number of prevocalic cases amount to fewer deleted cases in the exemplar cluster and thus a lower overall rate of deletion than the contracted negation, as shown in table 11.6. Note again that the lower rate of deletion is not attributable to a higher percentage of prevocalic cases, which would favor retention, because in fact there is not more deletion preconsonantally than prevocalically.

Regular past-tense verbs show a distribution that is even more fascinating. Their occurrence before vowels is extremely high: 40% versus an overall figure for the whole corpus of only 21%. Preconsonantal occurrences are at 47%, and prepausal are at 13%. Thus the exemplar cluster for past-tense verbs contains many more prevocalic occurrences than the other groups and thus would have many more occurrences with the /t/ or /d/ present. In effect, the deletion has not progressed as far for the past-tense word group as for the other groups.

Table 11.6, then, reveals a strong correspondence between the rate of occurrence before vowels and the overall deletion rate for the word group. These facts point clearly to the proposition that sound change occurs in real time as words are used and that its effects are registered in memory, producing a gradual change in lexical items based on the speaker's actual experience with them. It also reinforces what the data on frequency effects demonstrate, as shown earlier in this article. Words that more frequently occur in the context favoring a change undergo the change at a faster rate than those that occur less frequently in the appropriate context. These findings strongly suggest that phonetic change occurs online, in production, as language is used.

TABLE 11.7. Distribution of Word-Final /t/ and /d/ before Pauses, Vowels, and Consonants

	Prepausal (%)	*Prevocalic* (%)	*Preconsonantal* (%)	N
n't	10	10	80	379
-nt	36	22	42	59
-ed	13	40	47	279
All words	15	21	64	1,996

TABLE 11.8. Percentage Overall Deletion and Occurrence before Vowels for Three Sets of Words in the Switchboard Corpus

	Overall Deletion[a]	*Occurrence before Vowels*[b]
n't	423/473 (89%)	64/473 (14%)
-nt	104/144 (72%)	40/144 (28%)
-ed	166/492 (34%)	170/492 (36%)

[a]Differences in rates of deletion are significant at the $p < .0001$ level.
[b]Differences in rates of occurrence before vowels are significant at the $p < .01$ level.

To replicate the findings of the study of the LA corpus, the same word sets were examined in the Switchboard corpus. While the rates of deletion of final /t/ and /d/ were generally somewhat higher in the Switchboard corpus, the correspondence between the overall rate of deletion and the occurrence in prevocalic position held as it did in the LA corpus, as demonstrated in table 11.8. Despite differences in the two corpora in terms of regional and ethnic heritage of the speakers, of the situation, and of potentially different coding criteria, the correspondence between the data from the two corpora is remarkably similar. The differences shown in each column of table 11.8 are statistically significant; the rates of deletion in the three word sets are significant at the $p < .0001$ level and in the occurrences before vowels at the $p < .01$ level.

One additional question concerning these results must be addressed: is it possible that the different rates of deletion are due primarily to the token frequency of the words in each set rather than to the contexts in which they occur? After all, the contracted negation set contains the highly frequent words *don't*, *didn't*, which may be responsible for the higher deletion rates of *-n't* words versus *-nt* words. This is less an issue for the regular past tense, since the earlier test to see if these words deleted less than the general corpus had restricted the general corpus words to those with a frequency of less than 403 per million in the Francis and Kučera count (see table 11.2). To control for frequency for the comparison of *-n't* words with *-nt* words, the frequency ranges of both groups in the corpora were compared. Both groups contained words with only one occurrence in each corpus, while the highest frequency word in the *-nt* set was *different*, with eighteen occurrences in the LA corpus and twenty-three in the Switchboard corpus. This amounted to a total of forty-one occurrences in the two corpora. For this test, the data from the two corpora were pooled and all *-n't* words with a frequency of greater than 41 in the two corpora were removed. The affected words were *don't*, *didn't*, *can't*. The differences between the overall deletion rates and the occurrence before vowels were tested again, and the difference in deletion rate was found to be significant at the $p < .01$ level ($\chi^2 = 7.987$), while the occurrence before vowels was significant at the $p < .05$ level ($\chi^2 = 4.158$); see table 11.9 for details. Thus we see that it is not simply a word's frequency that determines the rate at which it undergoes a sound change but rather that word's frequency in the context that conditions the sound change.

TABLE 11.9. Percentage Overall Deletion and Occurrence before Vowels for *-n't* and *-nt* Words in the LA Corpus and the Switchboard Corpus (Excluding *-n't* Words with a Frequency Greater than 41)

	Overall Deletion[a]	*Occurrence before Vowels*[b]
-n't	181/223 (81.1%)	40/223 (17.9%)
-nt	132/203 (65.0%)	58/203 (28.6%)

[a]For overall deletion, $\chi^2 = 14.20575$, $p < .001$.
[b]For occurrence before vowels, $\chi^2 = 6.784548$, $p < .01$.

The data just examined now raise two other questions. First, why do regular past tense verbs occur so much more frequently before vowels? Second, to what extent does a token of a bound morpheme in one word, such as the contracted negation and the regular past tense, have an effect on other instances of the same morpheme in other words?

5.2. The distribution of regular past-tense verbs

A closer examination of the context of use of regular past tense shows that the high rate of prevocalic instances arises from the frequent combination of verbs with particles or prepositions, many of which begin in vowels. In the LA corpus sixty-two (55%) of the vocalic environments following a regular past tense verb were prepositions or particles. The examples were: *kicked out*, *lived in*, *lived on*, *looked at*, *moved in*, *opened up*, *picked up*, *picked on*, *pulled away*, *pulled out*, *raised in*, *stepped in*, *talked about*, *walked into*, *walked out*, *worked in*, *worked at*. Another twenty-three (20%) were vowel-initial pronouns, such as *it*, *us*, *him*, *her*, *them*, with the first consonant deleted. Another ten (9%) of these cases consisted of *and* in addition to the commonly used phrases *looked alike*, *changed a lot*, *learned a lot*. And seven (6%) of the prevocalic contexts had the indefinite article after the verb. In sum, about 90% of the vowel-initial items that follow the past-tense verb form what Erman and Warren (2000) called a prefabricated unit with the verb. That is, these sequences are conventionalized units that are very likely stored and processed together.[6]

This chunking and the establishment of semi-autonomous representations protect these verbs from the word-level generalization that leads to deletion outside of the appropriate phonetic contexts. It also suggests that the relevant units for the diffusion of the change very likely involve specific verb + particle and verb + pronoun sequences, such as *looked at*, *picked up*, *liked it*, and so on. These sequences present a uniform environment that does not favor the deletion of the /t/ or /d/. This would explain the very low rate of deletion (7%) of /t/ or /d/ when it is a regular past tense occurring before a vowel.[7]

5.3. Other accounts of the past-tense phenomena

The lower rate of deletion in the regular past tense has been observed in all studies of English /t, d/ deletion and various explanations have been proposed to account for it. One idea is that since the past-tense /t/ and /d/ have the specific grammatical function of marking past time reference, the deletion of the consonant would entail loss of that information from the utterance. While it may be that in certain contexts this information is uniquely marked in the suffix and that a speaker might make an effort to suppress reduction in such cases, in most cases the information conveyed by this consonant is redundant. Furthermore, we know from the history of many European languages that final consonants with morphological functions often fall to deletion processes. For instance, singular/plural verb agreement in English was lost with the deletion of final unstressed /n/, and the distinction between Latin first and third singular verb agreement fell with the reduction and loss of final /t/ and /m/. In addition, detailed studies of the interaction of functional load with reduction and deletion processes, such as Poplack (1980a, 1980b) and Labov (1994, chapter 19), have revealed that other factors are more powerful than the preservation of morphological information and that the predictions of the functional account are often not borne out in the data on variation and change.

Guy's (1991a, 1991b) account attributed the deletion rates for /t/ and /d/ to the structure of the lexical and morphological components of the grammar, adopting the theory of Lexical Phonology (Kiparsky 1982). In his account, t/d deletion can apply at a fixed rate at any level of the grammar. Thus some words have a chance to undergo the variable deletion two or three times. Monomorphemic words have a chance to undergo deletion three times, giving them the highest rate of deletion. Semi-weak past-tense verbs (those verbs with both a vowel change and a final /t/ or /d/, like *leave*, *left*; *tell*, *told*) are formed at Level 1 and thus have two chances to undergo deletion. Regular past tense is formed at Level 2 and thus has only one chance to undergo deletion. Since each exposure to the variable deletion rule increases the rate of deletion, Guy's model provides an account for the fact that monomorphemic words have a higher rate of deletion than semi-weak verbs, which in turn have a higher rate of deletion than the regular past tense verbs.

Guy's analysis coincides with the exemplar account in two ways. First, it models the cumulative effect of the variable rule applying multiple times. So does the exemplar model, which predicts that the more often a word is exposed to online reduction, the greater rate of deletion it evinces. Second, Guy's analysis recognizes that even a variable rule can interact in significant ways with the lexicon and the morphology. However, his analysis is unable to account for the major role played by word frequency at every level of the lexicon and morphology, and it misses the fact that the behavior of the semi-weak verbs can be accounted for simply by noting their high frequency with respect to the average regular verb. As mentioned earlier, Bybee (2000b) and Gregory et al. (1999) showed that for all words there is a significant effect of frequency on final /t, d/ deletion, such that high-frequency words undergo more deletion than low-frequency words. In addition, Bybee (2000b) showed that there is an effect of frequency on regular past-tense words occurring in preconsonantal or prepausal position (see table 11.3). Lexical Phonology has no means of tracking the token frequency of words. Even if words in the lexicon could accrue frequency, the regular past tense verbs are not in the lexicon but are formed at Level 2, leaving no means of tracking their token frequency. As for the semi-weak verbs, many of them, such as *told*, *felt*, *left*, *kept*, are of very high frequency, which alone can account for their higher rate of deletion. A striking confirmation of the role of frequency is that even within the semi-weak verbs, the high-frequency verbs undergo deletion much more than the low-frequency verbs, as shown in table 11.10, taken from Bybee (2000b).

A final argument against the Lexical Phonology account of variable rule application refers to data to be presented later in the article, where it is shown that Spanish intervocalic [ð] is deleted more in the past-participle morpheme than in lexical morphemes. This is counter to the predictions made by Lexical Phonology since the past participle in question is perfectly regular and would be formed at Level 2. According to the logic of Guy's model, it should undergo deletion fewer times than instances of lexical morphemes do. As I shall argue, the different rates of change for suffixes are not due to where they appear in the derivation but to how consistently they occur in the environment for change in language use.

5.4. A summary of the usage-based account

The model that can handle the facts just described must allow for gradual changes in the phonetic representations of words, and it must also allow for the expression of certain relations within and among exemplar clusters. As already demonstrated, reductive change occurs in production as already automated sequences of linguistic elements are further reduced. These online reductions feed back into memory repre-

TABLE 11.10. Semi-Weak Past-Tense Forms from the LA Corpus Ordered by Frequency, with Ties Ordered by Francis and Kučera Frequency

Total Tokens	*Verb*	*Tokens with Deletions*
32	told	22 (68%)
9	felt	5 (55%)
8	left	2 (25%)
6	kept	4 (66%)
4	sent	1 (25%)
4	built	0 (0%)
3	held	0 (0%)
3	heard	0 (0%)
2	slept	1 (50%)
2	lent	0 (0%)
1	found	0 (0%)
1	lost	0 (0%)
1	meant	0 (0%)

Spearman rank order correlation: $p = .696$ (significant at the .01 level [two-tailed or one-tailed]).

sentations, since the language user's accumulated experience is represented in memory. Words are represented as clusters of exemplars, and the relative weight of exemplars with different patterns may change over time as reduction proceeds. If the distribution of words in actual discourse contexts differs, the rate at which their exemplar clusters change, and thus the rate at which they undergo a change, may differ.

Exemplar clusters contain information about the contexts in which different exemplars are used. In some cases, this information is lost and exemplars spread from the more frequent context to less frequent ones. This has apparently occurred in the contracted negation and in *-nt* words, where the rate of deletion before vowels and consonants shows no phonetic effect. In other cases, contextual information is explicit and strong and allows the contextual exemplars to be maintained. This appears to be the case in prefabricated, conventionalized units, such as verb + particle and verb + pronoun sequences, where the maintenance of the regular past-tense /t/ or /d/ is high. Note that there is almost no deletion within such units (only 7%), whereas deletion of past-tense /t/ or /d/ before consonants correlates with the frequency of the verb, as shown in table 11.3. This suggests that the /t/ or /d/ in these sequences is behaving more as a word-internal consonant than as a final one.

Thus the lexical representation for a verb such as *live* may also contain the more specific entries *lived at*, *lived in*, *lived for* with information about the phonetic shape of these sequences. Figure 11.3 shows how a portion of the complex network involving verbs and particles might appear. The parallel semantic and phonetic categorization creates exemplar clusters at various levels: all instances of *live* are categorized in a cluster; all instances of *lived* are categorized together, as are all instances of *lived in*. A cross-cutting categorization would put instances of *in* together to the extent that some semantic affinity could be identified among its various uses (e.g., *moved in*, *played in*, and so on). Similarly, it would follow that instances of the regular past tense might be categorized together, with some latitude allowed for the fact that the allomorphs, /t/, /d/, and /ɨd/, differ from one another in voicing and syllabicity.

6. The effect of bound morphemes on changes in words

6.1. Alternating environments

So far I have examined two groups of words that were defined by their suffixes, the regular past tense and contracted negation. An issue I have not yet addressed is the extent to which an affix in one word can have an effect on the same affix in another word. In the model adopted for this study, words can be independent lexical units, but bound morphemes occur only within words and thus are always part of larger units (i.e., they have no independent existence). However, one way that words are categorized in the lexicon is through their shared parts. If these shared parts have both semantic and phonological similarity, they will be identified as morphemes. Thus the regular past-tense occurrences in different words would be represented as in figure 11.4, where a dotted line indicates only partial phonological similarity.

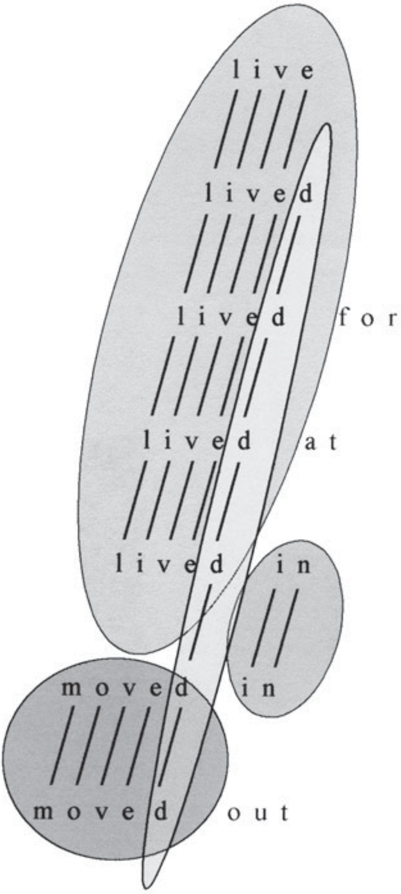


FIGURE 11.3. A complex network of verbs and particles contains parallel semantic and phonetic categorizations.

Another way of conceptualizing the unity of a suffix, such as regular past in English, would be to consider that it also constitutes an exemplar cluster (as shown in the circled portion of fig. 11.4). Because the regular past suffix occurs before a vowel about 35% to 40% of the time, the full /t/ or /d/ exemplars remain strong, especially those in conventionalized sequences such as *looked at*, *opened up*. Thus the overall rate of deletion is slowed in regular past tense, even before a consonant or pause.

The evidence that one instance of a past tense can affect another, however, cannot be taken as evidence for an independent unit, the past-tense suffix. Remember that preconsonantal and prepausal instances of regular past tense show a frequency effect: namely, high-frequency past-tense verbs in these contexts are more likely to undergo deletion than low-frequency ones, as shown in table 11.3.

Thus deletion of the past-tense /t/ or /d/ is never independent of the lexical verb it occurs with.

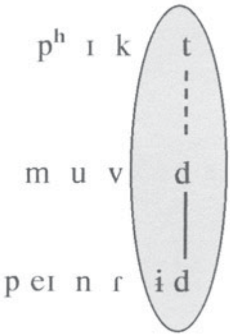


FIGURE 11.4. Regular past-tense morpheme in the words *picked*, *moved*, *painted*.

Similar cases of alternating environments slowing down the progress of a change were presented in Timberlake (1978), and examples of alternating environments blocking a change completely were examined in Bybee (2001).

6.2. Uniform environments

The effect of exemplars of bound morphemes upon one another can also be seen in uniform environments, specifically in ones that always uniformly present the context for change. High-frequency morphemes that contain within them the context for the change can undergo the change more rapidly. A case in point is the deletion of intervocalic [ð] in Spanish dialects, which has been in progress for several centuries. This [ð] derives from Latin *t*, which is voiced in Romance; in Spanish it is variably pronounced as a fricative intervocalically or deleted. As mentioned earlier, a study of the variable deletion of intervocalic [ð] in New Mexican Spanish shows a frequency effect for lexical instances (see table 11.4). This effect was also tested on a much larger corpus of peninsular Spanish (COREC). In this 1.1 million-word corpus of spoken language from a variety of situations, deleted elements were noted in the orthographic transcription.[8] I counted all instances of deleted and retained intervocalic *d*. Table 11.11 shows that the rate of deletion is much lower than for the New Mexican corpus, but that the frequency effect still holds.[9]

In addition to lexical instances of [ð], this sound occurs intervocalically in past participles of all conjugation classes, as in *hablado* 'spoken', *comido* 'eaten', *vivido*

TABLE 11.11. Rate of Deletion in High- and Low-Frequency Non–Past Participles from the COREC

	Low (0–226)	*High (239–2642)*	*Total*
Retention	20,560 (97.6%)	19,171 (91.3%)	39,731 (94.4%)
Deletion	508 (2.4%)	1,830 (8.7%)	2,338 (5.6%)
Total	21,068	21,001	42,069

$\chi^2 = 796.1$, $p < .0001$.

TABLE 11.12. Deletion Rate of [ð] in First Conjugation Past Participles versus All Other Words in the New Mexican Corpus

	First Conjugation Past Participles	*All Other Words*
Retention	38 (42%)	507 (83%)
Deletion	53 (58%)	101 (17%)

$\chi^2 = 79.9$, $p < .0001$.

'lived'. The rate of deletion for the first conjugation suffix (58%) is higher than the rate for all non–past-participle words (17%). This is shown in table 11.12 for the New Mexican corpus and in table 11.13 for the COREC. The figures here include past participles used either as verb forms, such as the various forms of the perfect, or as adjectives.[10] The second and third conjugation forms of the past participle, *-ido*, also delete at a higher rate than other words with *-id-*, with a deletion rate of 29% ($\chi^2 = 11.345$, $p < .001$). In the COREC, with its lower rate of deletion, the only second or third conjugation past participle that shows deletion is *sido*, the past participle of *ser* 'to be'.

The apparent explanation for the higher rate of deletion in the past-participle morpheme is its high frequency. It appears more than any other single morpheme, lexical or grammatical, with etymological /d/ in it. These data provide evidence for the representation of the past-participle suffix (or suffixes) in exemplar clusters. However, again we have evidence that the suffix is not an independent unit, because the rate of deletion in the suffix is higher with high-frequency verbs. The effects of frequency on the deletion rate for the first conjugation past participle are shown in table 11.14 for the COREC data. The New Mexican corpus shows a trend in the same direction, which, because of the small number of words coded, does not achieve statistical significance (see Bybee, 2001).

The frequency effect within past participles demonstrates that the suffix is not an independent unit but exists as a part of a whole word. However, the higher rate of deletion in this suffix suggests some relations among exemplars of the suffix, as indicated in the model with connecting lines for exemplar clusters for parts of words.

TABLE 11.13. Deletion Rate of [ð] in First Conjugation Past Participles versus All Other Words in the COREC

	First Conjugation Past Participles	*All Other Words*
Retention	4,903 (78.3%)	39,731 (94.4%)
Deletion	1,355 (21.7%)	2,338 (5.6%)

$\chi^2 = 1999.5$, $p < .0001$.

TABLE 11.14. Rate of Deletion by Frequency for First Conjugation Past Participles from the COREC Corpus

	Low (0–33)	*High (34–499)*	*Total*
Retention	2,631 (85.1%)	2,275 (71.7%)	4,906 (78.4%)
Deletion	459 (14.9%)	896 (28.3%)	1,355 (21.6%)
Total	3,090	3,171	6,261

$\chi^2 = 165.8$, $p < .0001$.

Another interesting issue that these data address is the segmentation of the past-participle morpheme. Pedagogically, one can teach the past participles as *–ado* for first conjugation and *-ido* for second and third conjugation. However, the first vowel should be associated with the verb stem, as it is predictable from the lexical classification of the verb. This leaves *-do* for the perfect tenses and *-do*, *-da*, *-dos*, *-das* in adjectival uses, depending on the number and gender of the noun. Thus the only constant for the past participle in general is the *-d-*. The difference in deletion rates in *-ado* versus *-ido*, however, suggests that the vowel is segmented with the suffix, and that each two-syllable unit is undergoing change at its own rate.

6.3. Conclusions concerning bound morphemes within words

The data discussed here have shown that bound morphemes can have an effect on the diffusion of a sound change, suggesting some association within the lexicon among instances of the same morpheme in different words. We have seen that the English past tense changes more slowly, while the Spanish past participle and, incidentally, the contracted negation in English change more quickly. The reason for these differences can be directly related to variation in the contexts of use of each of the morphemes. The importance of the morpheme in these cases does not overshadow the importance of the words the morpheme appears in. The frequency of the word containing the morpheme, in both the English past tense and the Spanish past participle, contributes to the rate of deletion, strongly pointing to lexical representations for multimorphemic words, which can accumulate their own frequency values or lexical strength.

7. The effects of extremely high frequency

One final effect of context and high frequency of use should be mentioned. In cases of extremely high frequency of words within certain phrases, a new exemplar cluster can be formed. Bybee and Scheibman (1999) studied the reduction of *don't* in conversation. We found that the most common and the most extreme reductions of *don't* occurred in the contexts in which *don't* occurred most often, namely, when *I* was the subject and the following verb was *know*. In fact, the vowel in *don't* was reduced to schwa only when the subject was *I* and the following verbs were *know*, *think*, *have*, *want*, *like*, *mean*, *feel*. The only other context for reduction to this extent was in the phrase *why don't you*, when used as a suggestion. This special reduction suggests that these phrases are separate from the independent representation of *don't* and that they form exemplar clusters of their own, which allow the reduction of *don't* to progress at a faster rate in these phrases than in other uses. A phrase will form an exemplar cluster of its own when it is used frequently; this frequency of use will also allow it to register phonological, semantic, and pragmatic changes.

Another robust example of the development of variants of words in frequent contexts is liaison in French, where, for example, the definite articles take on special forms depending upon the initial segment of the following noun, as in *le chien* 'the dog' versus *l'ami* 'the friend' and *le*(*s*) *chiens* 'the dogs' versus *le*[*z*] *amis* 'the friends'. While other final consonants were being deleted, liaison consonants remained within highly frequent contexts that were either within a grammatical construction or within a frequent, fixed phrase. As expected, liaison is being lost in a progression from the least frequent to the most frequent contexts (Bybee 2001).

8. Consequences for a usage-based theory

The study of the diffusion of sound change in the lexicon contributes to a better understanding of the nature and causes of sound change. Changes that affect high-frequency words first are a result of the automation of production, the normal overlap and reduction of articulatory gestures that comes with fluency (Browman and Goldstein 1992; Mowrey and Pagliuca 1995). The strong directionality of such changes indicates that they are not the result of random variation, but that they stem from reduction processes resulting from repetition and the normal automation of motor activity. If a sound change does not proceed from the most frequent to the least frequent words, then we should seek its explanation in some other mechanisms of change.

Moreover, I have proposed a model in which variation and change are not external to the lexicon and grammar but inherent to it (Pierrehumbert 1994). Sound change is not rule addition—something that happens at a superficial level without any effect on the deeper reaches of grammar. Rather, lexical representations are affected from the very beginnings of the change. Indeed, they supply an ongoing record of the change since they track the details of the phonetic tokens experienced. Further evidence for sound change having an immediate impact on representation is the fact that sound changes are never reversed or undone (Cole and Hualde 1998; Bybee 2001). The morphological structure of words also plays a role from the initial stages of a change, but less because morphemes have some special status with respect to change and more because of the contexts in which they appear. Alternating contexts retard change, while uniform ones allow change to hurry ahead.

Effects of frequency and context demonstrate that language use has an effect on mental representations. In this view, representations and the grammatical structure that emerges from them are based on experience with language. New linguistic experiences are categorized in terms of already-stored representations, adding to the exemplar clusters already present and, at times, changing them gradually. Various levels of abstraction emerge as exemplars are categorized by phonological and semantic similarity—morphemes, words, phrases, and constructions can all be seen as the result of the categorization of linguistic experiences.

Notes

For help analyzing the data on /t, d/ deletion I would like to thank Greg Thomson and Keri Holley; for coding and analyzing the data on Spanish [ð] deletion I would like to thank Dawn Nordquist, Rena Torres Cacoullos, and Gabriel Waters. Thanks also for the suggestions of the anonymous reviewers for the journal. Parts of this article were published in Bybee (2002).

1. There might be some question as to whether the variation discussed is sound change in progress or stable variation. In either case, word frequency must be taken into account along with other factors in describing the variation.

2. The English recorded in this corpus is a contact variety, as is the Spanish discussed in connection with [ð] deletion. However, the data from these corpora are checked against monolingual varieties in both cases, and the results with respect to the effects of frequency are very much the same, strengthening the evidence for the approach of this work.

3. The datasets were drawn from 38,000 words that were phonetically hand-transcribed (Greenberg, Ellis, and Hollenback 1996; Greenberg 2000).

4. The weakness of this experiment is that it relies on self-report, which must always be viewed with caution.

5. In the remainder of this article, I will use square brackets and slash lines in more or less traditional ways, even though the model invoked here does not have an explicit level of phonemic representation. The slash lines designate a broad range of variants and the square brackets a narrower, contextually determined range of variants.

6. That English verbs often occur in collocations with other words is also discussed in Hopper (1991, 1997).

7. There were only six deletions of final past-tense /d/ before a vowel in the LA corpus, four with a particle or preposition (*changed over*, *involved in* (two), *involved into*) and two with a pronoun (*called him*, *changed her*). In all of these cases except *called him*, the /d/ ends a complex consonant cluster.

8. The reliability of the transcription is not known; however, it does represent an independent source.

9. The cutoff between high and low frequency was chosen so as to produce two groups of tokens of approximately equal size. The same criterion was used for the data in table 11.14.

10. Most of the etymological instances of past participles used as nouns seem to be lexicalized and were not counted as past participles (*marido* 'husband', *estado* 'state', *comida* 'meal').

References

Anderson, John R. 1993. *Rules of the mind*. Hillsdale, N.J.: Lawrence Erlbaum.

Bills, Garland D. 1997. New Mexican Spanish: Demise of the earliest European variety in the United States. *American Speech* 72, 154–171.

Boyland, Joyce Tang. 1997. Morphosyntactic change in progress: A psycholinguistic approach. Ph.D. diss., University of California at Berkeley.

Browman, Catherine P., and Louis M. Goldstein. 1992. Articulatory phonology: An overview. *Phonetica* 49, 155–180.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Philadelphia: John Benjamins.

———. 1988. Morphology as lexical organization. In *Theoretical morphology*, ed. M. Hammond and M. Noonan, 119–141. San Diego, Calif.: Academic Press.

———. 2000a. Lexicalization of sound change and alternating environments. In *Papers in laboratory phonology*, vol. 5: *Acquisition and the lexicon*, ed. Michael Broe and Janet Pierrehumbert, 250–268. Cambridge: Cambridge University Press. (Reprinted in this volume as chapter 10.)

———. 2000b. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language*, ed. M. Barlow and S. Kemmer, 65–85. Stanford, Calif.: CSLI. (Reprinted in this volume as chapter 9.)

———. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

———. 2002. Lexical diffusion in regular sound change. In *Sounds and systems. Studies in structure and change*: *A Festschrift for Theo Vennemann* (Trends in Linguistics), ed. David Restle and Dietmar Zaefferer, 59–74. Berlin: Mouton de Gruyter.

Bybee, Joan, and Joanne Scheibman. 1999. The effect of usage on degree of constituency: The reduction of *don't* in American English. *Linguistics* 37, 575–596. (Reprinted in this volume as chapter 14.)

Cole, Jennifer, and José Ignacio Hualde. 1998. The object of lexical acquisition: A UR-free model. *Chicago Linguistic Society* 34, 447–458.

D'Introno, Francisco, and Juan Manuel Sosa. 1979. Elisión ae la /d/ en el español de Caracas: Aspectos sociolingüísticos e implicaciones teóricas. *Anuario de la Escuela de Letras*. Caracas: Universidad Central de Venezuela, 33–61.

Eddington, David. 2000. Stress assignment in Spanish within the analogical modeling of language. *Language* 76, 92–109.

Erman, Britt, and Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text* 20, 29–62.

Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. In *Papers from the eleventh annual meeting of the Chicago Linguistic Society*, 200–213. Chicago: Chicago Linguistic Society.

Fowler, Carol A., and Jonathan Housum. 1987. Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26, 489–504.

Francis, W. Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage*. Boston: Houghton Mifflin.

Gilbert, John H., and Virginia J. Wyman. 1975. Discrimination learning of nasalized and non-nasalized vowels by five-, six- and seven-year-old children. *Phonetica* 31, 65–80.

Greenberg, S. 2000. Switchboard transcription project. LVCSR Summer Workshop, Johns Hopkins University. http://www.icsi.berkeley.edu/real/stp/.

Greenberg, S., D. Ellis, and J. Hollenback. 1996. Insights into spoken language gleaned from phonetic transcriptions of the Switchboard corpus. *ICSLP*-96, S24-7, 524–527. Philadelphia.

Gregory, Michelle, William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextua1 predictability in lexical production. *Chicago Linguistic Society* 35, 151–166.

Guy, Gregory. 1991a. Contextual conditioning in variable lexical phonology. *Language Variation and Change* 3, 223–239.

———. 1991b. Explanation in variable phonology: An exponential model of morphological constraints. *Language Variation and Change* 3, 1–22.

Haiman, John. 1994. Ritualization and the development of language. In *Perspectives on grammaticalization*, ed. W. Pagliuca, 3–28. Amsterdam: John Benjamins.

Hooper, Joan Bybee. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In *Current progress in historical linguistics*, ed. William Christie, 96–105. Amsterdam: North Holland. (Reprinted in this volume as chapter 2.)

———. 1981. The empirical determination of phonological representations. In *The cognitive representation of speech*, ed. T. Myers, J. Laver, and J. M. Anderson, 347–357. Amsterdam: North Holland.

Hopper, Paul J. 1991. Dispersed verbal predicates in vernacular written narrative. *Berkeley Linguistics Society* 17, 402–413.

———. 1997. Dispersed verbal predicates in vernacular written narrative. In *Directions in functional linguistics*, ed. A. Kamio, 1–18. Amsterdam: John Benjamins.

Johnson, Keith. 1997. Speech perception without speaker normalization. In *Talker variability in speech processing*, ed. K. Johnson and J. W. Mullennix, 145–165. San Diego, Calif.: Academic Press.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the emergence of linguistic structure*, ed. J. Bybee and P. Hopper, 229–254. Amsterdam: John Benjamins.

Kiparsky, Paul. 1982. Lexical phonology and morphology. In *Linguistics in the morning calm*, ed. I. S. Yang, 3–91. Seoul: Hanshin.

Krishnamurti, Bh. 1998. Regularity of sound change through lexical diffusion: A study of *s* > *h* > *Ø* in Gondi dialects. *Language Variation and Change* 10, 193–220.

Labov, William. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

———. 1981. Resolving the neogrammarian controversy. *Language* 57, 267–308.

———. 1994. *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.

Losiewicz, B. L. 1992. The effect of frequency on linguistic morphology. Doct. diss., University of Texas, Austin.

Marcos-Marín, Francisco. 1992. Corpus oral de referencia del español contemporáneo. Universidad Autónoma de Madrid. http://elvira.1llf.uam.es./docs_es/corpus/corpus.html.

Menéndez-Pidal, Ramón. 1968 (1904). *Manual de gramática histórica española*. Madrid: Espasa-Calpe.

Moonwomon, Birch. 1992. The mechanism of lexical diffusion. Paper presented at the annual meeting of the Linguistic Society of America, Philadelphia, January.

Mowrey, Richard, and William Pagliuca. 1995. The reductive character of articulatory evolution. *Rivista di Linguistica* 7, 37–124.

Neu, Helene. 1980. Ranking of constraints on /t, d/ deletion in American English: A statistical analysis. In *Locating language in time and space*, ed. W. Labov, 37–54. New York: Academic Press.

Oliveira, Marco Antonio de. 1991. The neogrammarian controversy revisited. *International Journal of the Sociology of Language* 89, 93–105.

Penny, Ralph. 1991. *A history of the Spanish language*. Cambridge: Cambridge University Press.

Pensado, Carmen. 1984. *Cronología relativa del castellano*. Salamanca: Ediciones Universidad de Salamanca.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–342.

———. 2001. Lexical diffusion, lexical frequency, and lexical analysis. In *Frequency and* *the emergence of linguistic structure*, ed. J. Bybee and P. Hopper, 123–136. Amsterdam: John Benjamins.

Pierrehumbert, Janet. 1994. Knowledge of variation. *Papers from the Parasession on Variation. 30th Meeting of the Chicago Linguistic Society*. Chicago: Chicago Linguistic Society, 232–256.

———. 2001. Exemplar dynamics: Word frequency, lenition and contrast. In *Frequency and the emergence of linguistic structure*, ed. J. Bybee and P. Hopper, 137–157. Amsterdam: John Benjamins.

Poplack, Shana. 1980a. Deletion and disambiguation in Puerto Rican Spanish. *Language* 56, 371–385.

———. 1980b. The notion of the plural in Puerto Rican Spanish: Competing constraints on (s) deletion. In *Locating language in time and space*, ed. W. Labov, 55–67. New York: Academic Press.

Rumelhart, D., and J. McClelland. 1986. On learning the past tenses of English verbs: Implicit rules or parallel distributed processing? In *Parallel distributed processing: Explorations in the microstructure of cognition*, ed. J. McClelland, D. Rumelhart, & PDP Research Group, 216–271. Cambridge, Mass.: MIT Press.

Santa Ana, Otto. 1991. Phonetic simplification processes in English of the barrio: A cross-generational sociolinguistic study of the Chicanos of Los Angeles. Doct. diss., University of Pennsylvania.

Schuchardt, Hugo. 1885. *Über die Lautgesetze: Gegen die Junggrammatiker*. Berlin: Oppenheim. (Reprinted and translated in T. Vennemann and T. Wilbur [1972]. *Schuchardt, the Neogrammarians and the Transformational Theory of Phonological Change*. Frankfurt: Athenaum.)

Skousen, Royal. 1989. *Analogical modeling of language*. Dordrecht: Kluwer Academic.

———. 1992. *Analogy and structure*. Dordrecht: Kluwer Academic.

Timberlake, Alan. 1978. Uniform and alternating environments in phonological change. *Folia Slavica* 2, 312–328.

Van Bergem, Dick. 1995. *Acoustic and lexical vowel reduction* (Studies in Language and Language Use 16). Amsterdam: IFOTT.

Vennemann, Theo. 1974. Words and syllables in natural generative phonology. *Papers from the Parasession on Natural Phonology. 10th Meeting of the Chicago Linguistic Society*, 346–374. Chicago: Chicago Linguistic Society.

Wang, William S.-Y. 1969. Competing changes as a cause of residue. *Language* 45, 9–25.

———. (ed.). 1977. *The lexicon in phonological change*. The Hague: Mouton.

Wang, William S.-Y., and C.-C. Cheng. 1977. Implementation of phonological change: The Shaungfeng Chinese case. In *The lexicon in phonological change*, ed. W. Wang, 86–100. The Hague: Mouton.

PART IV

FREQUENCY EFFECTS IN MORPHOSYNTAX

This page intentionally left blank

Introduction to Part IV

The chapters in this part of the book apply the ideas and hypotheses developed for morphology and phonology to the level of morphosyntax with the goal of fleshing out a more complete usage-based theory. Chapter 12 illustrates how the Reducing Effect and the Conserving Effect (as introduced here in chapters 1 and 2) can be applied to the constructions of morphosyntax. Similarly, the effect of type frequency on productivity is illustrated for a syntactic construction.

Chapter 13's title, "The Emergent Lexicon," is meant as a partner to Paul Hopper's (1987) "Emergent Grammar." Given the high interconnectedness of lexicon and grammar, I argue that even while grammar is emergent, the lexicon must be, too, since it is highly affected by language use. Here the frequency effects discussed elsewhere in this volume are said to provide evidence for lexical storage. The storage of larger and less frequent units is taken up in the chapter, with brief discussions of idioms, prefabs, and grammaticizing constructions. The frequency effects also suggest, however, that in addition to storage units, words, phrases, and constructions must also be processing units. The fact that some priming effects appear with constructions and that phonological effects, including reduction but also retention (as in French liaison; see Bybee 2001), occur in high-frequency contexts both suggest that words, phrases, and constructions are processing units that become more fluent with more use. This notion is also the foundation of chapters 14 and 15.

Chapters 14 and 15 address constituent structure, primarily in English, arguing that the non-semantic phenomena associated with constituent structure can be attributed to frequency of co-occurrence. Chapter 14 examines the phonological phenomenon of the reduction of *don't*, which is highly dependent upon its frequency of use in different contexts. By phonological criteria the reduction of *don't* is more dependent upon the subject than the following verb. This appears to be due to the higher frequency of *I don't* than any combination with a verb, even *don't know*. Chapter 15 explores the interaction of frequency of co-occurrence with the semantic criteria for defining constituents and argues that the grammatical characteristics associated with constituency are due to frequency of co-occurrence. In both of these chapters, it is argued that constituency is gradient and emergent.

Finally, chapter 16 examines in detail the frequency effects associated with grammaticization using the case study of the development of the English auxiliary *can.* I argue here that many of the phenomena associated with grammaticization relate directly to the frequency increase observed in this process. Semantic bleaching is attributed to habituation, phonological reduction to chunking, and the resulting reduction to the automatization of neuromotor routines. The developing autonomy of the new construction, also due to frequency, accounts for the growing semantic opacity of the parts of the construction and the entrenchment of its morphosyntactic characteristics.

References

Bybee, Joan. 2001. Frequency effects on French liaison. In *Frequency and the emergence of linguistic structure*, ed. Joan Bybee and Paul Hopper, 337–359. Amsterdam: John Benjamins.

Hopper, Paul J. 1987. Emergent grammar. *Berkeley Linguistics Society* 13, 139–157.

Chapter 12

WITH SANDRA THOMPSON

Three Frequency Effects in Syntax

The primary hypothesis of functionalist or usage-based linguistics is that language use shapes grammar. The recognized mechanism by which this language-shaping process occurs is frequent repetition (Givón 1979; DuBois 1985; Hopper 1987; Hopper and Thompson 1993). In this chapter we will look at the three types of repetition or frequency effects that have been identified as important in phonology and morphology and demonstrate that these same effects are operative in the shaping of the grammar of larger units, that is, in syntax.

First it is important to distinguish two kinds of frequency: token frequency and type frequency. Token frequency, on the one hand, is the count of the occurrence in texts of particular words, such as *broken* or *have* or of specific phrases, such as *I don't think*. Type frequency, on the other hand, counts how many different lexical items a certain pattern or construction is applicable to. Using a morphological example, regular English past-tense *-ed* has a very high type frequency because it applies to thousands of different verbs. The vowel-change pattern exemplified by *strung* and *stung* has a much lower type frequency since it applies to fewer than twenty verbs. Type frequency in syntactic constructions would count how many distinct items of a particular lexical or grammatical class (e.g., verbs) can be used in the construction.

In this chapter we discuss two effects of token frequency and show how these effects interact with a third effect, which is due to type frequency. The three effects we have chosen to discuss are not the only frequency effects important in shaping grammar, but we have chosen these three for a specific reason: the Reducing Effect and the Conserving Effect of high token frequency appear to condition opposite results; that is, on the one hand, high token frequency promotes change and on the other hand it renders constructions resistant to change. We demonstrate that these two effects of high token frequency are applicable at different stages in the life of a construction and the two types of change involved are very different in nature. The Reducing Effect plays a major role in grammaticization, where high token frequency promotes all types of reductive change, while the type of change that is resisted by old, highly entrenched constructions is change on the basis of productive patterns with high type frequency.

1. The Reducing Effect

First we consider the Reducing Effect of high token frequency, which has a phonetic, a syntactic, and a semantic dimension. Phonetic change often progresses more quickly in items with high token frequency. This effect is particularly noticeable in grammaticizing elements or phrases that undergo drastic reduction as they increase in frequency. Thus *be going to*, which is becoming a future marker in English, is reduced to ['gʌnə] or even further reduced in phrases such as *I'm gonna* to ['aimənə]. Similarly the conventionalized contractions of English are reduced due to their high frequency: *I'm*, *I'll*, *I've*, *can't*, *don't*, *won't*, and so on. But the effect occurs on a more subtle level as well: regular sound change in many cases progresses more quickly in items of high token frequency. For instance, there is a tendency in American English for the loss of syllabicity in poststress schwa + resonant sequences, as in *every*, *camera*, *memory*, *family*. This reduction is more advanced in words of higher frequency (such as those just named) than in words of lower frequency, such as *mammary*, *artillery*, *homily* (Hooper 1976). The loss of final [t] or [d] after a consonant is also more common in words of higher frequency, such as *went*, *just*, and *and*; in fact, a general effect of token frequency on the rate of deletion has been found for 2,000 tokens of final [t] or [d] (Bybee 2000).

If sound changes are the result of phonetic processes that apply in real time as words are used, those words that are used more often will have more opportunity to be affected by phonetic processes. If representations are changed gradually, with each token of use having a potential effect on representation, then words of high frequency will change at a faster rate than words of low frequency. The streamlining of high-frequency words and phrases has the effect of automating production. Any motor activity that is repeated often will become more efficient (Bybee 2000).

The automation of processing has effects well beyond phonetics: the processing of frequently used sequences as single chunks also leads to the loss of internal constituent structure. This effect is observable in grammaticization (Heine and Reh 1984; Heine, Claudi, and Hünnemeyer 1991). Consider, for example, the grammaticizing phrase (*be*) *supposed to*, which is reduced to [spostə]. It comes from a passive construction with an infinitive, as shown in (1), but now the *to* of the infinitive is phonologically fused with the verb, as in the case of other developing auxiliaries, such as *have to* [hæftə], *want to* [wanə], *going to* [gənə], and so on, perhaps indicating a change in the affiliation of *to*. In addition, in (*be*) *supposed to* the syntactic status as a passive has also been lost, as shown by the unacceptability of a passive agent with the reduced form:

(1) He is supposed by most people to be very knowledgeable.

(2) He's s'posed to be very knowledgeable *by most people.

Where formerly this phrase was an instantiation of a construction involving *be + Past Participle*, functioning as a main verb with an embedded clause complement expressed as an infinitive, it now has been fused into a single auxiliary, *s'posta*, which is structurally parallel to *gonna*, *wanna*, and so on.

Accompanying phonological reduction and loss of internal structure of high-frequency phrases is the well-known phenomenon of semantic bleaching. Bleaching is due to habituation whereby organisms cease to respond at the same level to repeated stimuli (Haiman 1994). Words or phrases that are much repeated lose their semantic force, which, in a spiraling effect, allows them to occur more often, which in turn conditions further semantic bleaching.

Some of the clearest examples of bleaching are cases in which formerly emphatic expressions lose their special status as emphatics and become the unmarked way to express certain concepts. For example, it is common for negative markers to be reinforced by the addition of emphatic elements, such as other negators combined with nouns, as in French *ne . . . pas* 'not a step' or English *ne + a + wiht* 'not at all', which gave rise to *not*. Traugott (1972) reports that in Old English *noht* always had its full emphatic force, but in Middle English it gradually lost this force and came to be used as the normal, non-emphatic negation.

Phonetic reduction, loss of internal constituency, and bleaching all occur together in grammaticization because they are all promoted by frequent repetition, which is the mechanism that produces automation and habituation.

2. The Conserving Effect

The second effect of the repetition of tokens is due to the increased lexical strength (Bybee 1985) or entrenchment (Langacker 1987) of a particular word or phrase. Using a morphological example again, high-frequency forms with alternations resist analogical leveling: while English *weep/wept*, *creep/crept*, and *leap/leapt* have a tendency to regularize to *weeped*, *creeped*, and *leaped*, respectively, the high-frequency verbs with the same pattern, *keep/kept*, *sleep/slept*, show no such tendency (Hooper 1976; Bybee 1985). As a result, morphological irregularity is always centered in the high-frequency items of a language.

In Bybee (1985), this conservative behavior of high-frequency forms is related to the faster lexical access of high-frequency forms: the more a form is used, the more its representation is strengthened, making it easier to access the next time. Words that are strong in memory and easy to access are not likely to be replaced by new forms created with the regular pattern. This effect of frequency on the strength of representation in memory we dub the Conserving Effect.

This conservatism of much-used expressions can also be found on the syntactic level: it has often been observed that pronouns show more conservative behavior than full noun phrases (Givón 1979). Pronouns and full noun phrases are related in the sense that pronouns diachronically derive from nouns, and synchronically, in that pronouns and noun phrases often occupy the same positions. However, a major difference between nouns and pronouns is that the latter are much more frequent than the former. This fact can be used as an explanation for why English pronouns maintain distinct forms for nominative versus dative/accusative case, while nouns have lost these case distinctions. It can also be used to explain why the position of pronouns sometimes reflects an earlier word order, as in Spanish, where object clitic pronouns occur before the verb (reflecting, it is argued, an older OV word order) while full noun phrase objects occur after the verb (Givón 1979).

Similarly, verbal auxiliaries, which are highly frequent, often retain old conservative syntactic characteristics. The English auxiliaries, for instance, retain the ability to invert with the subject (which all verbs had previously, e.g. [3]) and they are followed, rather than preceded by, the negative, another property once shared by all verbs, as shown in (4) and (5) from Middle English (Mossé 1952):

(3) Gaf ye the chyld any thyng?
‘Did you give the child anything?’

(4) My wyfe rose nott.

(5) cry not so

This, then, is the Conserving Effect: the idea that high-frequency sequences take on a life of their own and resist change on the basis of newer productive patterns for juxtaposing words and morphemes.

3. The subjunctive in Canadian French

As an example of the conserving role of token frequency in syntax we take the example of the use of the subjunctive mood in Canadian French, as studied in Poplack (1992, 1995). We choose this example because Poplack’s study is based on a corpus of 3.5 million words of naturally occurring spoken language from 120 adult native speakers and because the use of the subjunctive involves long-distance dependencies of main verb on complement verb, a canonical syntactic issue.

In general, subjunctive verb forms tend to be very old forms that have come to be used in subordinate clauses after a long history of main clause use as futures or other indicative forms (Bybee, Perkins, and Pagliuca 1994). So in Canadian French the meaning and form of the subjunctive have been highly eroded. Indeed, we argue, along with Poplack, that the subjunctive has basically been lost, but residue remains in the most frequent contexts, with some indications of minor productivity.[1] The study we cite here shows that subjunctive verb forms now occur only in the most frequent syntactic contexts and with the most frequent verbs.

Poplack's study focuses on noun clauses embedded as complements to certain matrix verbs. In her corpus, Poplack identified 6,000 sentences with a matrix verb governing the subjunctive at least once. It is important to note, however, that one factor leading to the demise of the subjunctive/indicative distinction is the fact that for most verbs there is no phonological distinction between mood forms. So in this set of sentences about half of the embedded verbs were ambiguous between subjunctive and indicative. This left 2,694 instances in which subjunctive and indicative usage could be distinguished. It is particularly important for our point here to note that the verbs that do maintain a formal contrast between indicative and subjunctive are irregular and among the most frequent verbs of the language. That is, the maintenance of the mood distinction in a subset of verbs is at least in part due to the Conserving Effect of high token frequency.

In the sentences to be analyzed, then, the main verb is one that was used with the subjunctive at least once and the embedded verb was one that distinguishes mood formally. In these sentences the subjunctive was used 77% of the time. The goal of Poplack's study was to determine what factors prevent the subjunctive from occurring in these positions 100% of the time.

A statistical analysis of a number of factors led Poplack to conclude that the indicative/subjunctive distinction is not performing any particular functional or semantic work. This conclusion is supported by examples such as those in (6) and (7) in which the same speaker repeats essentially the same message to the same interlocutor but alternates between indicative and subjunctive:

(6) a. Faut que je lui *dis* (*I*) c'est vrai.
'I have to tell him it's true'.

b. Faut je lui *dise* (*S*) c'est la vérité.
'I have to tell him it's the truth'.

(7) Fallait qu'elle répond (*I*) "oui, tu peux faire trois pas de géant." Fallait qu'elle *réponde* (*S*) la phrase complète.
'She had to say "yes, you may take three giant steps." She had to say the whole sentence'.

The critical question here is: if there is no functional difference in mood choice, why are French speakers still using subjunctive verb forms? The evidence suggests that the answer lies in the fact that most of the subjunctive forms occur in certain highly entrenched phrases with particular matrix verbs and particular embedded verbs. Thus, as we would predict from the Conserving Effect, these high-frequency expressions have maintained their traditional form despite general changes that allow the construction of sentences with indicative verb forms in comparable but less frequent contexts.

The most commonly occurring matrix verb is impersonal *falloir* 'have to', which accounts for 62% of the 2,694 matrices and is followed by a subjunctive verb form in 89% of the cases. The main verb *falloir* is always impersonal, meaning that if the subject is expressed, third person masculine singular *il* is always used, but different tense/aspect forms can also be expressed, such as present *il faut* and imperfect *il fallait*.

The embedded verbs that occur most frequently in the subjunctive are high-frequency irregular verbs. In fact, only ten verbs account for two-thirds of the examples with subjunctive; among these are *avoir* 'to have', *être* 'to be', *aller* 'to go', *faire* 'to make, do', and others. See the examples in (8):

(8) a. Même pour un job aujourd'hui, faut tu *sois* (*S*) bilingue.
'Even for a job these days, you have to be bilingual'.

b. Bien certain, faut qu'ils *aient* (*S*) une place eux-autres aussi pour vivre.
'Well, of course, they should have a place to live, too'.

c. Faut j'*aille* (*S*) voir pour de l'ouvrage.
'I have to go look for a job'.

d. Bien ça, fallait tu *fasses* (*S*) ton huit heures par jour.
'Well, there you had to do your eight hours a day'.

These facts suggest that these main verb-complement constructions are not generated from highly generalized syntactic schemas of the form [verb [S]] but rather that very specific constructions ("routines" in Poplack's terms), with some lexical items indicated, are stored and accessed in production, as shown in (9):

(9)

$$(\textit{il}) \left\{\begin{matrix} \textit{faut} \\ \textit{fallait} \\ \vdots \end{matrix}\right\} (\textit{que})\ \text{PRO} \left\{\begin{matrix} \textit{faire} \\ \textit{aller} \\ \textit{avoir} \\ \textit{être} \\ \vdots \end{matrix}\right\} +\text{SUBJ} \quad \ldots$$

Further support for the position that constructions, complete with very specific lexical items, are accessed in these cases comes from the second factor that Poplack found to be significant, the distance factor. That is, if a word or some parenthetical material intervened between the main verb and the subordinate one, it was more likely that the indicative form would be used. Since we are claiming entrenchment of the sequences in storage, much as though they were lexical items rather than hierarchical syntactic structures, it is significant that intervening material favors the productive form, the indicative. That is, if the speaker gets derailed from an automated sequence such as *Il faut que* . . . , then s/he is less likely to resume with the routinized form and more likely to access the more generally used indicative form.

The importance of particular lexical items is also evident in the other 38% of the matrix verbs. Two verbs, *vouloir* 'to want' and *aimer* 'to like', make up 11% of the remaining cases, and they show a high percentage of subjunctive usage (91% and 67%, respectively) (Poplack 1995). With these verbs, too, the irregular embedded verbs favor subjunctive use.

Our point in this section is that distributions that appear to be very arbitrary should be checked for this conserving effect of token frequency. While grammatical analysis should proceed with the working hypothesis that formal distinctions represent functional distinctions, we also have to bear in mind that not all contrasts and distributions are meaningful or functional. Some patterns represent a lexically arbitrary residue of formerly productive patterns.

4. Type frequency

The type of change that is resisted by words or phrases of high token frequency is change on the basis of combinatorial patterns or constructions that are productive. To return to a morphological example, high-frequency English strong verbs resist regularization by the productive suffixation pattern. But frequency also plays a role in the determination of productivity, where productivity is defined as the likelihood that a pattern will apply to new forms. However, in this case it is type frequency: the type frequency of a pattern determines its degree of productivity (MacWhinney 1978; Bybee 1985, 1995).

It is easy to see why type frequency determines productivity: *type frequency* refers to the number of distinct lexical items that can be substituted in a given slot in a construction, whether it is a word-level construction for inflection or syntactic construction specifying the relation among words. The more lexical items that are heard in a certain position in a construction, the less likely it is that the construction will be associated with a particular lexical item and the more likely it is that a general category will be formed over the items that occur in that position. The more items the category must cover, the more general will be its criterial features and the more likely it will be to extend to new items. Furthermore, high type frequency ensures that a construction will be used frequently, which will strengthen its representational schema, making it more accessible for further use, possibly with new items.

As type frequency can range from one to a very large number, so there are varying degrees of productivity associated with ranges of type frequency. The relationship between type frequency and degree of productivity among English strong verbs has been shown by Bybee and Moder (1983) and Moder (1992).[2] The most productive strong verb class, exemplified by *strung*, is the only class that has added new members since the Old English period and also has the highest type frequency of any strong verb class. As a result, we expect to find a relation between the existing type frequency of a syntactic construction and its ability to apply to novel lexical items.

Some of the most difficult problems in syntax concern the competition among functionally similar constructions with varying degrees of productivity due to type frequency. For instance, Goldberg (1995) argues that type frequency affects the productivity of the ditransitive construction (as in [10]). This construction is the continuation of an old construction that is gradually being replaced by the construction in which the dative is marked with the preposition *to* (as in [12]). In the ditransitive construction, which was common in Middle English, the indirect object noun phrase was formerly marked with the dative suffix, which explains why no preposition is present today. The ditransitive construction has some of the features of an archaic construction preserved through high frequency; in particular, it occurs only with specific lexical verbs, and most of these are of Germanic origin and of very high frequency. Thus, for instance, as has often been observed, this construction is not used with a verb such as *whisper*, as in (11). The prepositional construction in (12), however, is perfectly general, occurring with all verbs that can take a patient and recipient argument. Its high type frequency gives it full productivity, allowing it to be used productively, even with nonce forms (Gropen et al. 1989):

(10) He told the woman the news.

(11) *He whispered the woman the news.

(12) He whispered the news to the woman.

However, this case is made complex by the fact that the ditransitive construction does have some limited productivity, as evidenced by its occurrence with certain new verbs such as to *telephone*, to *e-mail*, and *to Fed-Ex*, (see [15]). Goldberg (1995, chapter 5) points out that this limited productivity corresponds to the type frequency of the semantic verb classes that occur in the ditransitive construction. Classes with large membership are able to attract new members to a limited extent while classes with smaller membership do not appear to attract new members. For example, a small class of verbs that can be found in the ditransitive construction is the class of verbs of permission. Only a very few verbs of this class can occur in the ditransitive, as illustrated in (13) and (14):

(13) Sally permitted/allowed Bob one kiss.

(14) *Sally let/enabled Bob one kiss.

Larger classes of verbs that can be used in the ditransitive construction are verbs of sending and communicating. This high type frequency is what allows the construction to be used with new verbs with related meanings, as in (15).[3]

(15) Sally telephoned/e-mailed/Fed-Exed Sam the offer.

The literature on the ditransitive construction shows that some arbitrariness must be recognized in the categorization of verbs according to whether or not they are permitted in this construction. However, certain generalizations have also been shown to be valid: that verbs indicating eventual possession of the object can be used in the ditransitive construction and that monosyllabic verbs are favored over polysyllabic verbs (Gropen et al. 1989). The existence of exceptions, arbitrariness, and non-categorical tendencies in productivity are properties of lexical classes. In fact, these are the properties found with morphological classes, such as the classes of English irregular verbs. It is not surprising, then, that type frequency in these classes affects productivity, just as it does in morphological classes.

5. Conclusion

Recognizing the two effects of token frequency on syntax and their interaction with type frequency over time allows us to explain many situations in language that otherwise appear very arbitrary. Not only can frequency be viewed as an integral part of an explanation for these situations, but also the effects of frequency have important implications for our notions of mental representation. There is not necessarily just one representation per construction; rather, a specific instance of a construction, with specific lexical items in it, can have its own representation in memory if it is of high frequency. As a result, it can undergo phonological reduction and change in meaning or function independently of other examples of the same construction. At a later stage, highly entrenched examples of constructions with particular lexical items can continue to be used even though new productive patterns have become current in the language.

We are thus recognizing that some constructions are relics from earlier stages of a language, but they are not just arbitrary historical residue; rather, such structures are still subject to categorization and generalization. There are not just two choices: fossilization or productivity, but many intermediate possibilities, depending both on categorization in terms of either meaning or form and on the type frequency of the construction.

On a more general level, we hope to have supported and further articulated John Haiman's (1994) claim that grammatical constructions, as conventionalized patterns of language use, are created and maintained through the same mechanisms associated with repetition as are other patterns of human culture, as well as ritualistic behavior found in animals other than humans.

Notes

1. Poplack's interpretations of descriptions of other dialects of French (including the standard) suggest that the same situation holds elsewhere.

2. In these studies, other factors are also shown to interact with type frequency in determining productivity, i.e., phonological properties and high token frequency of individual types, which can detract from productivity. See also Bybee (1995).

3. Other properties of the two constructions, such as restrictions on the use of pronouns, appear to be less related to their relative age and frequency than to notions of topicality (Thompson 1990). In fact, this example shows that any two competing constructions can be affected by issues of semantics and pragmatics in addition to the complex interaction between token and type frequency.

References

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins.

———. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5), 425–455. (Reprinted in this volume as chapter 8.)

———. 2000. The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language*, ed. Michael Barlow and Suzanne Kemmer, 65–85. Stanford, Calif.: CSLI. (Reprinted in this volume as chapter 9.)

Bybee, Joan L., and Carol Lynn Moder. 1983. Morphological classes as natural categories. *Language* 59, 251–270. (Reprinted in this volume as chapter 6.)

Bybee, Joan, Revere D. Perkins, and William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

DuBois, John. 1985. Competing motivations. In *Iconicity in syntax*, ed. John Haiman, 343–365. Amsterdam: John Benjamins.

Givón, Talmy. 1979. *On understanding grammar*. San Diego, Calif.: Academic Press.

Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Gropen, Jess, Steven Pinker, Michelle Hollander, Richard Goldberg, and Ronald Wilson. 1989. The learnability and acquisition of the dative alternation in English. *Language* 65(2): 203–257.

Haiman, John. 1994. Ritualization and the development of language. In *Perspectives on grammaticalization*, ed. William Pagliuca, 3–28. Amsterdam: John Benjamins.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Heine, Bernd, and Mechtild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Helmut Buske.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In *Current progress in historical linguistics*, ed. William Christie, 96–105. Amsterdam: North Holland. (Reprinted in this volume as chapter 2.)

Hopper, Paul J. 1987. Emergent grammar. *Berkeley Linguistics Society* 13, 139–157.

Hopper, Paul J., and Sandra A. Thompson. 1993. Language universals, discourse, pragmatics, and semantics. *Language Sciences* 15(4), 257–376.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical prerequisites.* Stanford, Calif.: Stanford University Press.

MacWhinney, Brian. 1978. *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development, no. 174, vol. 43.

Moder, Carol Lynn. 1992. Productivity and categorization in morphological classes. Ph.D. Diss., SUNY at Buffalo.

Mossé, Fernand. 1952. *A handbook of Middle English*. Baltimore: Johns Hopkins Press.

Poplack, Shana. 1992. The inherent variability of the French subjunctive. In *Theoretical analyses in Romance linguistics*, ed. Christiane Laeufer and Terrell A. Morgan, 235–263. Amsterdam: John Benjamins.

———. 1995. The sociolinguistic dynamics of apparent convergence. In *Towards a social science of language: Papers in honor of William Labov*, ed. G. Guy, J. Baugh, and D. Schiffrin, 285–309. Amsterdam: John Benjamins.

Thompson, Sandra. 1990. Information flow and dative shift in English discourse. In *Development and diversity: Variation across time and space*, ed. Jerold Edmonson, Crawford Feagin, and Peter Mühlhäusler, 239–253. Dallas, Texas: SIL.

Traugott, Elizabeth C. 1972. *A history of English syntax.* New York: Holt Rinehart and Winston.

Chapter 13

The Emergent Lexicon

1. Emergence

In the traditional view, the lexicon is a storage area for all and only the content words or morphemes of a language. The lexicon is relatively static compared to the grammar, which contains all the moving parts of sentence generation: in the metaphor of a dictionary, lexical items are just passive items on a list that wait to be recruited into syntactic structures. In this conception also, memory for linguistic material is thought to be lodged primarily in the lexicon while the grammar is not so much a matter of memory as it is of abstract structure.

Rather than arguing that a lexicon of this type does not exist, I am going to argue that if such a lexicon does exist it is because it is emergent from the storage of linguistic experience, which is of a very different nature than the traditional conception of a lexicon would suggest. The point of this chapter then, is to explore the properties of stored linguistic experience. The data examined provide support for the proposals that much of linguistic knowledge is procedural knowledge, that chunks of linguistic experience much larger than the analytic units of morphemes or even words are the usual units of storage and processing, that there is no real separation of lexicon from grammar, and that phonological alternations whose domain is larger than a word can indicate the size of processing units.

This chapter obviously owes a great deal to the 1987 paper of Paul Hopper's to which the title alludes: "Emergent Grammar." In fact, my basic point is the same as the one Hopper made in that paper: that the "knowledge" underlying the fluent use of language is not grammar in the sense of abstract structure but rather a large store of categorized and sorted previous utterances that form the basis for the production and comprehension of new utterances. The only difference between Hopper's approach and mine is that his goal was the explication of the view of grammar that arises from such a theory and mine will be the view of the lexicon that is entailed by this theory. Since we would both argue that the grammar and lexicon are not separable, our papers are both about the same set of issues.

2. The lexicon reflects linguistic experience

I have argued in various places (Bybee 1985, 1988, 2000a, 2000b), especially in connection with morphology, that lexical storage is highly affected by language use. We know that highly frequent morphological formations, such as irregular nouns and verbs, tend to resist leveling and maintain their irregularities over time. I have called the mechanism behind this tendency "lexical strength." This notion corresponds to the psycholinguistic notion that high-frequency items have a higher level of resting activation; being easier to access, they are less likely to be replaced by regular formations.

In other avenues of research, I have looked into the role of frequency of use in the phonological and semantic reduction of words and phrases. In a number of cases, documentation on sound change in progress shows a significant frequency effect. High-frequency words undergo sound change at a faster rate than low-frequency words, for example, in t/d deletion in American English (Bybee 2000b) and the examples cited there as well as those cited in Phillips (1984). Furthermore, all of these changes are phonetically gradual. My account of the frequency effect in the diffusion of sound change is as follows: sound change has its source in the reduction and retiming of articulatory gestures that affect words or phrases as they are used in context. The memory representation of the phonetic shape of a word is a categorization of the tokens of use that have been experienced and thus represents a range of variation (Johnson 1997; see also Cole and Hualde 1998). Thus each use of a word or phrase has an effect on the stored representation. Words that are used more often in contexts where reduction and change are favored (familiar speech situations, high-frequency phrases) will undergo a gradual shift in phonetic representation at a faster rate than other words.

Both of the frequency effects I have just described imply that it is actual tokens of use that are stored in memory and not smaller units such as bound morphemes, which do not occur as tokens of use. I have argued in Bybee (1985) that the internal structure of words is derivable from sets of connections made between words that have related parts. Affixes and roots or stems have no separate representation but exist only as relations of similarity among words (see fig. 13.1).

Recurrent patterns such as those shown in figure 13.1 are the emergent generalizations or schemas that can be used to produce new combinations. It has been shown for morphology that one important determinant of productivity is the type frequency of a pattern: that is, the greater the number of distinct stems a pattern applies to, the greater is the likelihood that it will apply to new items (MacWhinney 1978; Bybee 1985, 1995; Lobben 1991). Thus some schemas will be very strong and easily accessible for applying to a great many items, and others will be much less so.

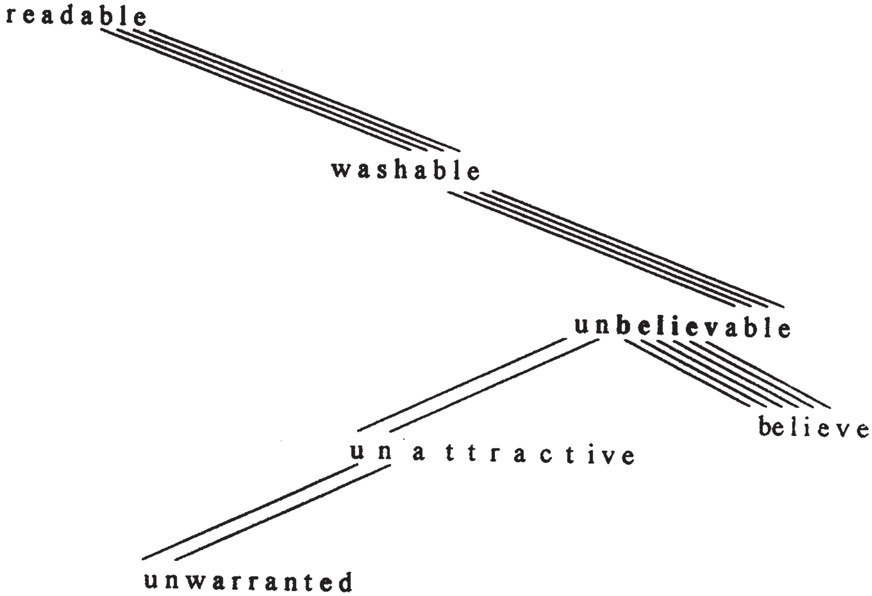


FIGURE 13.1.

The model described here, which I will call the Network Model, is highly redundant, since the same string of features, that is, the same morpheme or word, can occur in many different combinations. This redundancy does not entail that any valid generalizations are being missed: it is, of course, an empirical question what type of generalizations native speakers make, but all of these can be captured in schemas that can be formulated in varying degrees of abstraction. Moreover, redundant representation does not entail that *all* the potential words of a language exist in storage. Not all words have the same status: those that are used frequently have very strong representations, or possibly even multiple representations, but those that are used infrequently, or are potential but have not been used, may have no representation.

3. Storage and processing of larger units

In the following I will extend the model of morphology that I have just reviewed to larger units, arguing that evidence for it is also found in the behavior of phrases and even syntactic constructions. My proposal is that memory for language consists of a large store of units of varying sizes (from word to phrase or even clause) with varying degrees of strength, productivity, and connection with other units. The knowledge represented by these units is not just declarative or propositional knowledge but is also procedural knowledge. According to Anderson (1993), declarative knowledge is factual knowledge that people can report or describe (i.e., "Washington, D.C., is the capital of the USA"), but procedural knowledge is only manifest in people's performance of a skill, such as speaking a language or driving a car. These two types of knowledge entail different memory processes. Procedural knowledge is highly affected by repetition. In fact, with repetition, propositional knowledge can become procedural knowledge. Units of procedure become larger with repetition, as recurring sequences of actions come to be represented as a single action, increasing fluency (Boyland 1997). In addition, frequently used actions become fluent more quickly (Anderson 1993, Boyland 1997). My suggestion is, then, that linguistic units stored in memory are processing units—units of planning, production, and perception.

This proposal for viewing grammar and lexicon as closely tied to performance has far-reaching consequences and, of necessity, is somewhat vague at the moment. To proceed with articulating it more clearly, empirical evidence of many different sorts must be considered. In this chapter I will mention some of the relevant evidence. To begin, I will consider some cases in which evidence exists for memory storage of strings consisting of more than one word, in particular, idioms, high-frequency phrases, and grammaticizing constructions.

3.1. Idioms

It is not controversial to claim that idioms are instances of multi-word sequences that are stored in memory. It is interesting, however, to observe the similarities of idioms to multi-morphemic units that are lexically stored and to further observe that the network model just presented solves some of the problems in the analysis of idioms.

First observe that idioms often contain conservative lexical items and conservative grammatical usage. Examples of words that appear only in idioms are familiar from the literature: *bated* occurs only in *wait with bated breath*; *dint* occurs only in *by dint of*; *hale* occurs only in *hale and hearty*. Each of these phrases preserves a word that has been lost elsewhere from the language. This preservation is evidence that the whole phrase is a storage item in memory.

Such cases are parallel to the morphological cases in which otherwise obsolete affixes are locked into certain formations, as in the old *-er* plural marker that occurs as the *r* in *children*. In inflectional forms, such items occur only in very high-frequency formations; in derivation, extreme high frequency does not seem to be a prerequisite for the preservation of otherwise non-productive or obsolete affixes: in the words *handsome*, *steadfast*, *piecemeal*, for example, the second half of each was formerly a derivational affix. Similarly, idioms do not have to be of especially high frequency to preserve archaic words. A check of Francis and Kučera (1982) revealed no occurrences of *bated*, *dint*, or *hale*, suggesting that the phrases they occur in are not frequently used.

Idioms and set phrases can also preserve obsolete grammar. Observe the SOV word order in the phrase, *with this ring I thee wed* or, as Hopper (1987) points out, the older sense of the indefinite article to mean 'one and the same' in *birds of a feather* or its use to mean 'one' in *a penny saved is a penny earned*. The fact that idioms, like irregular morphological formations, preserve otherwise obsolete lexical items as well as archaic grammar is evidence that they are stored as units.

Just because idioms are remembered as wholes does not mean that their component parts and the semantic contribution they make are not recognized. Nunberg, Sag, and Wasow (1994) have pointed out that even though a phrase such as *pull strings*, as in *John was able to pull strings to get the job*, has a meaning that is different from the literal combination of its parts, speakers recognize the two words in the phrase as the same as those occurring in other combinations and even recognize the semantic contribution of these words. Nunberg and his co-authors point out that *strings* can be thought of metaphorically as connections with certain (influential) people and *pull* can refer to the exploitation or the exertion of pressure on these strings, a verb–object relation. Such an idiom (or "idiomatically combining expression," as Nunberg et al. call it) is not frozen and unanalyzed but connected both lexically and grammatically to other expressions in English. Such connections can be diagrammed as in figure 13.2.

3.2. High-frequency phrases

Like idioms, high-frequency phrases also provide evidence for lexical storage of multi-word items a single units, but units that are nonetheless analyzable into their morphosemantic components. It is well-known that high-frequency phrases change phonetically and semantically in a way that suggests storage as a single unit. In a recent study of the reduction of *don't* in English conversation, Bybee and Scheibman (1999) and Scheibman (2000) found that the reduction of the [o] vowel of *don't* to schwa occurs only in the contexts in which *don't* most frequently occurs, that is, after *I* and before certain verbs. Scheibman also shows that the use of *I don't know* in conversation serves certain pragmatic purposes and does not literally mean *I-don't-know.* Both the phonetic change and the functional change suggest that *I don't know* ([aiɾəɾ̃o]) is a storage and processing unit. Similarly, *I don't think* and *why don't you* are other phrases using *don't* that are phonetically reduced and have a meaning that is not the literal combinations of their parts.

This means that *don't*, which people do still identify as a word, occurs in at least three combinations in storage, the phrases *I don't know*, *I don't think*, and *why don't you* (*verb*). But we can't stop there because *don't* also occurs with its vowel reduced to schwa in other phrases, such as *I don't mean*, *I don't care*, and *I don't have*, sug-

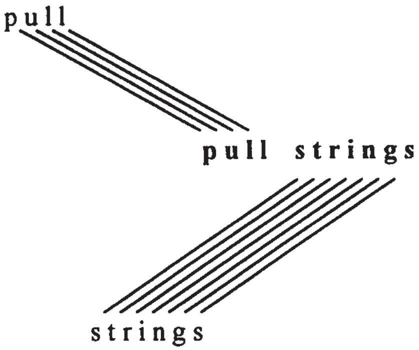


FIGURE 13.2.

gesting that these phrases are also storage and processing units. In fact, since *don't* in our corpus only occurs in the reduced form with schwa after *I*, it might be that *I don't* is also a processing unit as other combinations of pronoun with auxiliary undoubtedly are, for instance, *I'm*, *I'd*, *he's*, *she's*, *we've*, and other contracted forms (Krug 1998). In fact, in a count of 414 tokens of *I* in a conversational corpus, the two most frequent items to follow *I* were *'m* and *don't*, each one accounting for more than 10% of the tokens. The efficiency of processing *I'm* and *I don't* as units is evidence for their representation as units in storage, but it does not rule out the possibility that the sequences *I am* and *I don't* can be compiled from individual units. In fact, such online concatenation probably occurs in the production of these units in their full, unreduced forms. This means that *I*, *don't*, *am*, *I don't*, *I'm*, *I don't know*, and many other phrases are stored in memory and are accessible in production and comprehension. (See figure 13.3.)

3.3. Grammaticizing constructions

In its reduction and change in meaning the phrase *I don't know* resembles a grammaticizing construction—it is undergoing both phonological reduction and functional change that is typical of grammaticization. Boyland (1997) argues that grammaticization is the automatization of a processing unit. The frequency increase characteristic of grammaticizing constructions and the changes conditioned by frequency point to grammaticizing constructions as storage and processing units. Indeed, one fact that was often overlooked earlier in grammaticization studies is that graminaticization takes place in particular constructions. On a synchronic level it is

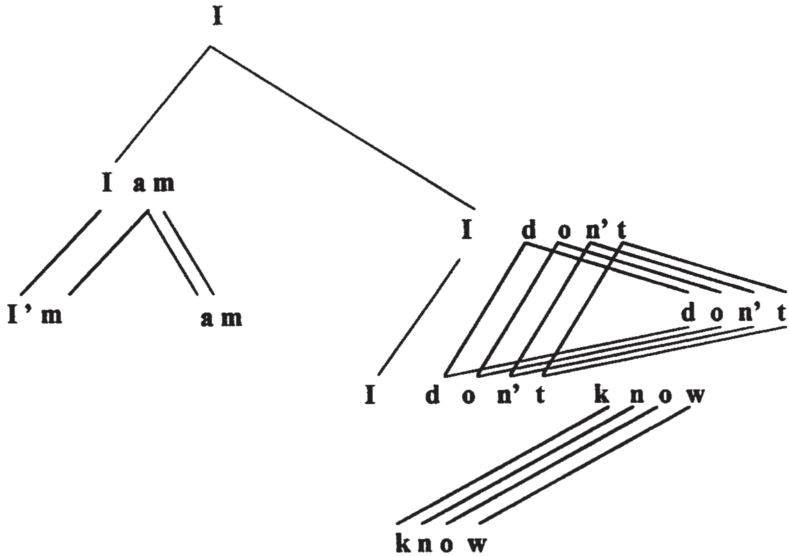


FIGURE 13.3.

a fact that grammatical morphemes are always associated with particular constructions. There is no reason to suppose that grammatical morphemes have any independent existence, since they always occur in and derive their meaning from a specific construction. Thus grammaticization occurs when a new construction or a specific instance of an old construction becomes a processing and storage unit.

For instance, it is often said that English has a *go*-future or, from a diachronic perspective, that a movement verb has become a future marker in English. But in fact, it is not *go* alone that has grammaticized. As is often pointed out, the construction involves the present progressive of *go* plus the goal-marker to and, just as important, a subject and a main verb. Thus *be going to* or *gonna* (with the meaning of future) exists only in the construction seen in (1):

(1) [Subject + BE] *gonna* [Verb]

This sequence is not just a construction but is a processing unit that produces a clause when specific lexical material is supplied for the more open slots in the construction.

4. Constructions as processing units

Croft (1995) analyzes almost two thousand previously coded intonation units from Chafe's pear stories and finds that in 97% of the cases intonation units are grammatical units, the most frequent of these being simple clauses and the next most frequent noun phrases. Croft proposes the Intonation Unit Storage Hypothesis on the basis of this study:

> The constructions that are stored or precompiled are the grammatical units that (normally) occur in a single intonation unit. (1995:872)

He goes on to say that "singly embedded NPs/PPs are almost certainly stored/precompiled constructions. Syntactically, they are relatively simple structures, almost never break across IUs and are fairly frequent" (1995:872). Croft's usage-based approach makes the very plausible suggestion that "the units employed for spoken communication are basically the units stored as constructions in the mind" (1995:872–873). (See also Ono and Thompson 1995 for a similar claim.)

Similarly, Hopper (1987) and Langacker (1987:35–36) note the highly formulaic nature of actual speech and the frequent occurrence of certain phrases such as *the problem is . . .* , *you take . . .* , *a little bit*, *living in a fantasy world*, *one question after another*. Hopper argues that the systematicity of linguistic systems is due to the "lateral associations of real utterances." That is, real pieces of speech are stored, sorted, categorized for both phonological and functional similarity, but "they do not, however, merge into the kind of uniform grammar which would lead one to posit a uniform mental representation to subtend them" (1987:147).

Constructions are not, of course, set phrases but rather abstractions that range over many specific phrases. The slots in a construction are of varying degrees of generality; the different parts of constructions are productive to varying degrees. For example, the *way* construction, as in *we made our way home* (Kemmer 1994; Goldberg 1995), has four parts, each having a different range of possible lexical or grammatical material.

(2) [verb] [possessive pronoun] *way* ([directional adverb])

Each of these positions may contain the following material (examples from Kemmer 1994):

1. The verb is restricted to one that signifies motion, manner, path creation, or means: *went his way*, *swiggled his way*, *made our way*, *cut their way*.
2. The possessive pronoun is a closed grammatical class.
3. *Way* is a unique lexical item that is not replaceable
4. The adverbial is usually a prepositional phrase although it may include other elements: *tooting its way through London*, *singing its way down from the heights.*

In fact, most constructions contain some specific lexical or grammatical material. Here are some of the constructions Ono and Thompson (1995) extracted from seven 5- to 10-minute conversations:

(3)	construction	example
	like to Verbal Expression	*I'd like to have . . .*
	to have NP *V-ed*	*to have my lungs replaced . . .*
	NP *replaced with* NP	*replaced with asbestos . . .*
	NP *or something*	*asbestos or something*
	NP *do a* N *on* NP	*they did a post mortem on her*
	NP *say* CLAUSE	*the doctor said there wasn't any . . .*
	get off ONE's NP (rear body part)	*get off my ass*

The point to note in these examples is that very specific lexical and grammatical material is essential to the construction. In addition note that certain tokens of constructions, with specific lexical material, are clearly stored units due to their frequency. For instance, the prototype *make one's way* could be a stored unit, in addition to *wend one's way*, which must be stored as a unit because *wend* does not occur in other constructions.

Ono and Thompson also propose that the following very generalized constructional schemas exist for producing clauses in English:

(4) NP V NP
NP V NP PP
NP V NP PP PP

Such highly generalized constructional schemas would arise from the representations of much more specific stored units. In my own view, the evidence for the more specific schemas is much stronger than for the very abstract ones proposed in (4).

The fact that lexical and grammatical material is embedded in the construction in which it occurs explains how lexical and grammatical splits occur. For instance, what was originally a single verb *have* is now used in many different constructions and phrases, indicating that it has multiple representations. Phonological evidence suggests that the more lexical uses of *have* in phrases have split from the auxiliary *have* of the perfect construction, since the latter contracts with the subject (*I've done it*) while the former does not in American English (**I've ten dollars*).

These facts all point to the need to investigate the hypothesis that production involves the accessing, concatenation, and overlapping of distinct stretches of speech that have been formed into processing units by repetition. A possible further argument for this hypothesis arises from another observation made by Croft (1995) and others who have studied intonation units. Croft observes that consecutive parallel structures are in different intonation units; that is, conjoined clauses, conjoined verb phrases, and noun phrases in apposition are almost always in different intonation units. A possible interpretation of this finding would follow from the proposal that constructions are processing units; in that case, each use of the same construction would require beginning a new processing unit. Embedded clauses, then, would occur in the same construction as the main clause, since no intonation break necessarily occurs between the main and subordinate clause. Clearly, the relations among intonation units, processing units, and constructions presents a potentially fruitful area of investigation.

5. The intertwining of "lexicon" and "grammar"

Many arguments have been developed over the last few decades for a strong interdependence between lexicon and grammar. Langacker (1987) takes the position that there is no discrete cutoff point between lexicon and grammar, as do many other linguists. The basic observation is that lexical items, in particular verbs, must be stored with a lot of information about the syntactic contexts in which they occur (e.g., Levin 1993; Goldberg 1995). From the other perspective, it is also the case that syntactic constructions often contain very specific lexical material (as shown earlier) or semantically well-defined classes of lexical items (as shown in Goldberg 1995 and other works on construction grammar). Rather than reviewing these arguments here, I would like to add two other pieces of evidence that support the hypothesis that lexicon and grammar are not separate—psycholinguistic and sociolinguistic evidence of syntactic priming and phonological evidence of alternations that apply across word boundaries.

5.1. Priming of constructions

A well-known effect in lexical access is the priming effect. In word recognition tasks, the second instance of the same word is accessed much more quickly by a subject than the first instance. The general explanation for this phenomenon is that words are activated as they are used and their activation remains high for a short period after use, making them easier to access during that period. An important argument for viewing words and constructions as stored in memory and accessed in similar ways is the fact that constructions also show the priming effect (Branigan et al. 1995).

Bock (1986) presented subjects with either one of the sentences in (5), which they were asked to repeat. Then they were shown pictures that could be described using either a prepositional-object or a double-object construction. For example, the picture might show a girl handing a paintbrush to a man. Bock found that subjects tended to produce a sentence with the same structure as the prime.

(5) a. The rock star sold some cocaine to an undercover agent.

b. The rock star sold an undercover agent some cocaine.

Since other experiments have shown that priming occurs between comprehension and production, Branigan et al. conclude that both production and comprehension access the syntactic information associated with the construction, thereby activating it, much as a lexical item is activated in comprehension and production.

Interestingly, this sort of priming effect has been demonstrated to occur in natural discourse as well. Many sociolinguistic studies of variation have turned up the interesting fact that one of the best predictors of the use of a construction is its use in immediately preceding discourse. For instance, Poplack and Tagliamonte (1996) find that in Nigerian Pidgin English, where all tense/aspect forms are optional, the strongest predictor of the occurrence of a tense/aspect marker with a verb is an immediately preceding verb with the same marker.

Weiner and Labov (1983) studied the use of the agentless passive in American English conversation, looking at all the sociolinguistic and pragmatic factors that might influence its use. Their data showed that one of the most powerful factors in the choice of the agentless passive was its use in one of the five preceding clauses and this factor was independent of pragmatic factors such as what was given or new information. These findings are consistent with the idea that access to a construction in comprehension or production activates that construction and makes it available for subsequent use.

Tannen (1989, among others) discusses the pervasive use of repetition in spontaneous conversation. She sees an advantage in the use of repetition for both production and comprehension. She argues that "repetition enables a speaker to produce language in a more efficient, less energy-draining way. It facilitates the production of more language, more fluently" (1989:48). She does not mention the psycholinguistic factor behind this facilitation, but clearly it is the priming factor. It is simply easier to access words, phrases, and constructions that have recently been accessed. Her argument from the point of view of the hearer is similar: repetitions and variations (inserting words into just-used constructions) "facilitate comprehension by providing semantically less dense discourse" (1989:49). Again, this facilitation is due to the priming effect: recently activated material is easier to access again.

5.2. Phonological effects within constructions

It is well known that most phonological or morphophonological alternations occur within words. In particular, most frozen or non-productive alternations (such as Vowel Shift, as in *divine* ~ *divinity*, or in Velar Softening, as in *critic* ~ *criticize*) occur at the word level. That is why the few cases of frozen alternations that appear to be conditioned across word boundaries have attracted so much attention. Much research has been directed toward the syntax/phonology interface with an eye to discovering which syntactic configurations condition phonological alternations. Of course, some alternations of this sort are rather easy to describe, for instance, the alternation in the indefinite article in English, while others, such as the reduction of *don't* or contraction in English, are much more difficult.

Another way of viewing the alternations that appear to take place across word boundaries is to suppose that the units in which such alternations occur are not composed of separate lexical units but rather are themselves unitary with respect to memory storage (Bybee 2000a). That is, as I mentioned earlier, the reduction of *don't* occurs only in fixed high-frequency phrases involving *don't*. The persistence of the alternation in *a/an* suggests that this small grammatical morpheme is not an independent lexical unit but rather a part of a construction: *a/an* (Adj) Noun. Many stored units contain the indefinite article (*an apple*, *an hour*, *an instant*, *a moment*, *a day*, *a friend*), and a more schematic construction may also exist that includes the phonological information about the variants of *a/an*.

Others have also argued that the non-automatic rules of phonology that appear to refer to syntax occur only in precompiled phrases (Zwicky 1987; Hayes 1990). As evidence I note that by far the most common cases of allegedly syntax-sensitive alternations involve particular grammatical morphemes. Since grammatical morphemes exist only in constructions, these alternations exist within constructions, that is, storage units, just as word-level alternations exist within storage units. Probably the most famous case of alternations conditioned across word boundaries is the complex set of alternations in French liaison (Tranel 1981). All of these alternations involve grammatical morphemes in specific constructions: (6) demonstrates liaison codified in the orthography in questions; (7) shows that the final consonant of certain prepositions appears when the following word begins with a vowel; (8) shows the appearance of [z] in articles and the plural marker for nouns and adjectives when the following word is vowel-initial. The liaison contexts that involve classes of lexical items affect only small closed classes of adjectives (those that can occur prenominally) and a few adverbs, as in (9):

(6)	a.	*chante-t-il?*	[šãttil]	'does he sing?'
		cf. *il chante*	[ilšãt]	'he sings'
	b.	*allons-y?*	[alɔ̃zi]	'let's go'
		cf. *allons*	[alɔ̃]	'(we) go'
(7)	prepositions:			
	a.	*dans un mois*	[dãzɛ̃mwa]	'in a month'
		cf. *dans trois mois*	[dãtʁwamwa]	'within three months'
	b.	*pendant un mois*	[pãdãtɛ̃mwa]	'for a month'
		cf. *pendant trois mois*	[pãdãtʁwamwa]	'for three months'
(8)	plural of articles, adjectives, and nouns:			
	a.	*les enfants*	[lɛzãfã]	'the children'
	b.	*les petits animaux*	[lɛptizanimo]	'the little animals'

(9) small closed classes, such as the pre-nominal adjectives (*petit*, *grande*, etc.)
adverbs: *assez* 'enough', *trop* 'too', etc.

It is my view, then, that fixed "lexical" alternations occur only within storage and processing units and could not be maintained if they were indeed applying across boundaries between processing units. Also, the more phonetic, variable changes that occur within phrases, such as the reduction of *don't* discussed earlier, or the palatalization in phrases such as *did you*, occur only in frequent phrases. Both types of phonological alternations that appear to be conditioned across traditional word boundaries actually provide evidence for the size and nature of processing units and point to constructions rather than words as the minimal unit for storage and processing. The traditional studies of such cases, under the rubric of the "phonology-syntax interface," have assumed that the proposed syntactic structure is correct and the phonology is in some way aberrant, not following established syntactic conventions in some cases. However, I would propose that the phonology is providing a more direct indication of the true nature of the processing units, and thus the constituents, involved.

6. Nativelike selection and nativelike fluency

Pawley and Syder (1983) discuss the fact that the rules of grammar generate many more utterances than would be considered idiomatic or nativelike (see also Pawley 1986; Langacker 1987:35–36). For example, they suggest that instead of *I want to marry you*, one could grammatically utter some of the following, though the effect would be decidedly non-idiomatic:

(10) I wish to be wedded to you.
I desire you to become married to me.
Your marrying me is desired by me. . . .

Rather than exploiting the full range of grammatical possibilities, we have certain standardized or conventionalized ways of saying things, not just for specific contexts, such as telling time, where we say *it's twenty to/till six* rather than *it's six less twenty*, but in almost every context and almost every utterance we produce.

We all know that in learning a language learning the vocabulary and rules of grammar is not sufficient preparation for actually using the language for everyday activities. To sound like a native, one must learn a large set of stock phrases. Pawley and Syder argue that actual language production consists of accessing many clauses already, or at least largely, preformed. They estimate that the stock of such phrases in the native speaker may range into the hundreds of thousands.

Being able to access such automated sequences is essential for fluency in both the native and the non-native speaker. Actual language seems to rely on memory much more than on abstract analysis. A significant implication for language acquisition is that along with categorization and generalization as mechanisms for acquisition, there is also an important role for imitation. One must learn, by example, in context, what are the customary ways of formulating one's ideas, requests, questions, and so on. Moreover, the process of automating such sequences requires repetition, the active repetition of production.

None of these points comes as a surprise to anyone who has ever observed a child acquiring a language, nor to anyone who has ever tried to learn or teach a language. However, they are points that have up to now, with the notable exception of the work of Pawley and Syder that I have just cited, been regarded as having practical but not theoretical consequences. Recognizing their general theoretical significance may result in a greater focus on their practical consequences. Clearly the question for both general theory and acquisition theory is how the particular conventionalized phrases and clauses give rise to the abstract generalizations that allow the production of more novel utterances.

7. Conclusions

Clearly, all words, phrases, and utterances of a person's experience are not separately stored in memory! The brain is a powerful categorization device for the efficient sorting and storing of the pieces of our experience, including the units of language use. One type of efficiency is achieved by storing and processing larger chunks rather than smaller ones. It is apparently easier to access, produce, and comprehend a precompiled chunk than to assemble it part by part for production (see Anderson 1993; Boyland 1997). However, this task would be beyond our powerful brains were we not also able to categorize these chunks and generalize over them.

Linguistic knowledge is not just propositional or representational knowledge. A large portion of the stored knowledge that makes language possible is procedural knowledge. Stored chunks are procedural chunks, embedded in context not just cognitively and socially but also physically in the production and comprehension systems along whose paths they run and also physically in the articulatory gestures and the manual gestures that are coproduced with them. Evidence for the procedural nature of linguistic chunks is the fact that they are affected by frequency of use. If linguistic knowledge were abstract, propositional knowledge, frequency of use would not be important.

The model proposed here places a much heavier burden on memory and access of memory than models with simpler lexicons and more complex combinatorial rules. In fact, the task of accessing appropriate units would be impossible were it not for the networks making associations among linguistic units, and between linguistic units and context. The role of the extralinguistic and the linguistic context is extremely important: the context itself activates linguistic units, making relevant words and phrases very easy to access under appropriate conditions. Similarly, the same units in other contexts may be difficult to perceive correctly or to access in production.

Our understanding of both language structure and language use is enhanced by the recognition that memory for language is highly affected by language use. The memory representation of language consists of units that can constitute utterances or intonation units, that is, not just words but also phrases and constructions. The smaller units familiar from structural analysis—stem morphemes, grammatical morphemes—

are not independent units but rather emerge from these larger stored units via a network of connections among them.

References

Anderson, John R. 1993. *Rules of the mind*. Hillsdale, N.J.: Lawrence Erlbaum.

Bock, J. K. 1986. Syntactic persistence in language production. *Cognitive Psychology* 18, 355–387.

Boyland, Joyce Tang. 1997. Morphosyntactic change in progress: A psycholinguistic approach. Ph.D. diss., University of California at Berkeley.

Branigan, Holly P., Martin J. Pickering, Simon P. Liversedge, Andrew J. Stewart, and Thomas P. Urbach. 1995. Syntactic priming: Investigating the mental representation of language. *Journals of Psycholinguistic Research* 24, 489–506.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Philadelphia: John Benjamins.

———. 1988. Morphology as lexical organization. In *Theoretical morphology*, ed. M. Hammond and M. Noonan, 119–141. San Diego, Calif.: Academic Press.

———. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5), 425–455. (Reprinted in this volume as chapter 8.)

———. 2000a. Lexicalization of sound change and alternating environments. *Papers in laboratory phonology*, vol. 5: *Acquisition and the lexicon*, ed. Michael Broe and Janet Pierrehumbert, 250–268. Cambridge: Cambridge University Press. (Reprinted in this volume as chapter 10.)

———. 2000b. The Phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language*, ed. M. Barlow and S. Kemmer, 65–85. Stanford, Calif.: CSLI. (Reprinted in this volume as chapter 9.)

Bybee, Joan L., and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in English. *Linguistics* 37(4), 575–596. (Reprinted in this volume as chapter 14.)

Cole, Jennifer, and José Ignacio Hualde. 1998. The object of lexical acquisition: A UR-free model. Chicago Linguistic Society 34.

Croft, William. 1995. Intonation units and grammatical structure. *Linguistics* 33, 839–882.

Francis, W. Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage*. Boston: Houghton Miflin.

Goldberg, Adele. 1995. *Constructions : A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Hayes, Bruce. 1990. Precompiled phrasal phonology. In *The phonology-syntax connection*, ed. Sharon Inkelas and Draga Zec, 85–108. Chicago: University of Chicago Press.

Hopper, Paul J. 1987. Emergent grammar. *Berkeley Linguistic Society* 13, 139–57.

Johnson, Keith. 1997. Speech perception without speaker normalization. In *Talker variability in speech processing*, ed. Keith Johnson and J. W. Mullinex, 145–166. New York: Academic Press.

Kemmer, Suzanne. 1994. Pattern crystallization in syntactic change. Paper presented at the Symposium on Synchronic and Diachronic Aspects of Grammaticalization, Sophienberg Slot, Rungsted, Denmark, October 9–11.

Krug, Manfred. 1998. String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26, 286–320.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*, vol. 1: *Theoretical prerequisites*. Stanford, Calif.: Stanford University Press.

Levin, Beth. 1993. *English verb classes and alternations*. Chicago: University of Chicago Press.

Lobben, Marrit. 1991. Pluralization of Hausa nouns, viewed from psycholinguistic experiments and child language data. Diss., University of Oslo.

MacWhinney, B. 1978. *The acquisition of morphophonology*. Monographs of the Society for Research in Child Development no. 174, vol. 43.

Nunberg, Geoffrey, Ivan A. Sag, and Thomas Wasow. 1994. Idioms. *Language* 70, 491–538.

Ono, Tsuyoshi, and Sandra A. Thompson. 1995. What can conversation tell us about syntax? In *Alternative linguistics*, ed. Philip W. Davis, 213–271. Amsterdam: John Benjamins.

Pawley, Andrew. 1986. Lexicalization. In *Languages and linguistics: The interdependence of theory, data and application*, ed. Deborah Tannen and James B. Alatis, 98–120. (Georgetown University Round Table on Languages and Linguistics 1985.) Washington, D.C.: Georgetown University Press.

Pawley, Andrew, and Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In *Language and communication*, ed. Jack C. Richards and Richard W. Smith, 191–225. London: Longmans.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–342.

Poplack, Shana, and Sali Tagliamonte. 1996. Nothing in context: Variation, grammaticization and past time marking in Nigerian Pidgin English. In *Changing meanings, changing functions: Papers relating to grammaticalization in contact languages*, ed. P. Baker, 71–94. Westminster: University Press.

Scheibman, Joanne. 2000. *I dunno but*: A usage-based account of the phonological reduction of *don't* in American English conversation. *Journal of Pragmatics* 32(1), 105–124.

Tannen, Deborah. 1989. *Talking voices: Repetition, dialogue, and imagery in conversational discourse*. Cambridge: Cambridge University Press.

Tranel, Bernard. 1981. *Concreteness in generative phonology: Evidence from French*. Berkeley: University of California Press.

Weiner, Judith E., and William Labov. 1983. Constraints on the agentless passive. *Journal of Linguistics* 19, 29–58.

Zwicky, Arnold. 1987. French prepositions: No peeking. *Phonology Yearbook* 4, 211–227.

Chapter 14

WITH JOANNE SCHEIBMAN

The Effect of Usage on Degrees of Constituency

The Reduction of Don't *in English*

In this chapter we take the position that there are many degrees of constituency and that these derive in a direct manner from the frequency with which elements are used together: elements that are frequently found next to each other show a tighter constituent structure than those that collocate less frequently. We use both phonological and functional evidence from conversation to argue that repetition conditions chunking (Haiman 1994), sometimes overriding the syntactic and semantic logic of the organization of utterances. While a generalized constituent structure may be an emergent property arising from many analogous utterances, specific combinations that are frequently used may diverge from the general pattern because frequency conditions autonomy in storage and renders internal analysis unnecessary (Bybee 1985, 1988, 1995). This phenomenon reveals the essential role of repetition in the creation of constituent structure: while semantic and pragmatic factors determine what occurs together in discourse, the actual repetition of stretches of talk triggers the mechanism that binds them into constituents. The general priniciple we propose to predict the degree of cohesion between elements is the following: the greater the probability that one element will follow another, the tighter the grammatical cohesion between them. We further propose that this principle derives from the nature of memory storage of linguistic experiences.

1. Grammaticization

First consider the fact that constituent structure changes in grammaticization. Heine, Claudi, and Hünnemeyer (1991) point out that changes in constituent structure are common in grammaticization; in particular, some instances of grammaticization involve a change in dependency relations within constituents, as, for example, when a head noun in construction with a genitive becomes a preposition, or when a main verb with a verbal complement becomes an auxiliary and the erstwhile embedded verb becomes the main verb. Such changes are accompanied by decategorialization (Hopper 1991)—the loss of morphosyntactic behavior characteristic of the major lexical categories, noun and verb. All of these changes involve a downgrading of constituent boundaries—from those separating higher level constituents to those separating lower level ones. Heine et al. (1991) find the claim that grammaticization downgrades constituent boundaries problematic, but our examination of their examples suggests that in every case, higher level constituents become lower level ones: for example, two clauses become one, V–NP becomes P–NP, and so on. All such changes are gradual and conditioned by repetition.

The gradual nature of diachronic changes in constituent relations manifests itself as a continuum of constituent types in synchronic language. As Heine (1993) points out, the gradual loss of feature is characteristic of verbs and the gradual acquisition of grammatical properties is precisely what makes auxiliaries so difficult to categorize and leads to the controversy among synchronic grammarians as to whether auxiliaries are best considered to be verbs or a category unto themselves.

The role of frequency or repetition is also well documented in the study of grammaticization. Enormous frequency increases are observed during grammaticization. However, frequency increases should not be viewed as simply a consequence of the increase in the appropriate contexts of use of a form undergoing change, but as Haiman (1994:14) has observed, repetition itself can be considered a major force behind the creation of rituals and conventions that in language are manifested as grammar. Among the effects of repetition on grammar (see Bybee and Thompson 1997), two are of special importance to our understanding of the change of constituent structure in grammaticization: chunking (automatization) and lexical autonomy. In chunking (Haiman 1994), a frequently repeated stretch of speech becomes automated as a processing unit. The original internal structure becomes less important and can be obscured by phonological change, making the unit more efficient to process (Boyland 1997). At the same time, high levels of usage also lead to autonomy. High-frequency collocations weaken their association with related items, as when, for example, *be supposed to* [spostə] (showing phonological assimilation) takes on functions increasingly less related to the meaning of *suppose*. In the example to be studied in this chapter, various phrases with *don't* in them, such as *I don't know* and *why don't you*, fuse phonologically and the whole phrase takes on special discourse functions. Both chunking and autonomy lead to the loss of internal structure within constructions (cf. Goldberg 1995). Therefore, the grammaticalization of phrasal patterns based on usage (which includes changes in constituent relations) may be viewed as the creation, modification, or dissolution of constructions.

2. Natural chunking

The more often two elements are used together, the more tightly they will be fused or bonded phonologically and semantically, and thus the tighter their constituency.

Pragmatic and semantic factors determine the frequency with which any two particular items will be contiguous. But in looking at constructions as generalizations over sequences of elements, another factor comes into analytical play: the chances that the same two elements will occur together is a function of the number of elements there are to select from in any two contiguous positions. Within a construction, some positions may be occupied by a single gram, a small set of grams, a semantically defined set of nouns or verbs, or a totally open set of nouns or verbs. For example, consider the construction for normal negation in English:

NOUN PHRASE Aux + *n't* VERB

Here NOUN PHRASE and VERB are fully open classes,[1] AUX designates a set of some twenty members (counting all inflected forms), and the negative position holds only one member (*n't*). Note further that the two positions with the fewest options are the most fused phonologically—the AUX and the negative—and they are the least likely to be separated by other intervening items, such as adverbs or parentheticals.

Constituency follows from our experience with language—the way it is used and the way it is recorded in memory. Consider a storage mechanism for language, proposed for morphology (Bybee 1985) but applicable to larger stretches of language as well, in which representation is highly affected by language use (see also Langacker 1995). Fragments of our linguistic experience are recorded, sorted out, categorized, and associated with other identical or similar records of linguistic experience. Activation of a linguistic element in production or decoding strengthens its representation in memory. Use of elements in sequence strengthens their syntagmatic relations. Elements very frequently used together fuse (e.g., *going to* > *gonna*). Classes of elements used together create constructions; bondedness within and between constructions depends on how frequently two contiguous elements occur together.

3. Reduction of *don't*

Phonological indicators of constituency arise as part of the automatization process. Bybee (2000a) argues that alternations arise only *within* storage units, which are very often word-level units, giving rise to the concept of "word-level phonology." Constructions are storage units as well, and alternations can arise within constructions if they contain specific phonological material, say in the form of grammatical morphemes, as in French liaison (Tranel 1981), or the alternation in *a/an* in English.

It is also well known that ongoing phonological changes often affect high-frequency words to a greater extent or at a faster rate than they do low-frequency words (Hooper 1976; Phillips 1984; Bybee 2000b). Bybee (2000b) proposes that phonological processes of reduction and coarticulation affect lexical items on line in production and that the effects of these phonological processes gradually change the shape of the stored representation. Since there is variability in all phonetic productions and the range of phonetic variants may differ by lexical item, lexical representations must contain a range of variation for each item based on actual productions. This range gradually shifts as change proceeds. Given that high-frequency items have more exposure to the articulatory forces of production, they undergo change at a faster rate.

The case we discuss here, the case of the reduction of *don't* in spoken American English conversation (Scheibman 2000), is a case characterized by variation, which, when examined in context, allows us to make inferences about the usage factors influencing constituent structure. We find that the reduced variants of *don't* are not just favored in but actually restricted to the most frequent environments in which *don't* occurs. Syntactic treatments of the reduction of *don't* attempt to trace the conditions that allow reduction to constituent relations (Kaisse 1985), but even these treatments recognize that reduction depends heavily upon the subject's being a pronoun and the verb following *don't* being one of a small set of verbs frequently used with *don't*. While previous analyses of the reduction of *don't* have been based on speakers' intuitions, Scheibman (2000) studied the variants of *don't* in natural conversation.

3.1. Description of the database

One hundred and thirty-eight tokens of *don't* were taken from approximately 3 hours and 45 minutes of naturally occurring conversation tape-recorded by Scheibman. The conversations took place on three separate occasions and represent the speech of six participants, four females and two males, all residing in Albuquerque, New Mexico. The usages were transcribed using a conventional tape recorder and earphones; therefore, it was often difficult to hear subtle phonetic differences such as the nasalization of a vowel or the presence of a nasal consonant. For this reason, a decision was made to group the tokens into the following four categories.

Group 1

Those tokens with a full-stop consonant and a full vowel, schematically represented as *stop + o*. This group includes [dõt], [dõn], and [dõ] and their oral vowel counterparts. The presence or absence of vowel nasalization does not contribute to the analyses given in this chapter.[2]

e.g., we *don't* see him all ¤winter
[dõ]
most of the time he lives underground. (G1.34S)[3]

Group 2

Those tokens with a reduced consonant, specifically an oral or a nasal flap, and a full vowel. This group is represented as *flap + o* and includes the variants [ɾõt], [ɾõ], and [ɾ̃õ] and their oral vowel variants.

e.g., J: who did it?
L: well they `*don't* ¤know who `did it. (G2.13N)
[ɾo]

Group 3

Those tokens with both a reduced consonant and a reduced vowel. This group is represented as *flap + ə* and includes [ɾəÖ] and [ɾ̃əÖ] and oral vowel variants.

e.g., can you imagine?
I *don't* know if I could ¤do that. (G3.496S)
[ɾ̃ə]

Group 4

Those tokens with just a reduced vowel, represented as ə and including [əÖ] and [ə].
e.g., I *don't* know ¤anything about `guns.
[əÖ]
but god damn it,
I'm getting sick of this shit. (G3.181O)

The majority of *don't* usages appear in declarative constructions in a simple clause or in a main clause of a complement constructions (e.g., *I don't* [ɾ̃ə] *¤know if it would be such a `great job*), 8 of the 138 tokens are found in interrogative structures (e.g., *well why `don't* [ɾõ] *you do a ¤survey?*), and six usages are part of negative imperative forms (e.g., *don't* [dõ] *assume ¤I'm `guilty*).

3.2. Patterns in the data

Our first prediction—that *don't* will be more reduced in the immediate linguistic contexts in which it is most used—is borne out by the data in a very robust way. First consider the requirement formulated by Kaisse (1985), that the subject be a pronoun for reduction to occur. In the data examined here and presented in table 14.1, it can be seen that when flapping occurs, the subject is a pronoun and never a lexical noun phrase. The further reduction, of the [o] to [ə], however, occurs only with *I*, the most frequent of the pronouns, and once in the phrase *why don't you*. Thus it is not enough to single out the class of pronouns as conditioning reduction: this is neither entirely accurate nor explanatory. Instead we see a frequency effect: not only do pronouns occur with *don't* more often in the data than do other subject

TABLE 14.1. *Don't* Variants by Type of Item Preceding or by Type of Construction (n = 138)

Preceding/Type	*stop + o* *Group 1*	*flap + o* *Group 2*	*flap + ə* *Group 3*	*ə* *Group 4*	*Total* *No.*	*%*
I	16	22	38	12	88	63
you	7	7			14	10
we	2	6			8	6
they	1	3			4	3
lexical noun phrase	5			5	4	
pause	1			1	1	
adverb	2	2			4	3
neg. imper.	6			6	4	
interrogative	3	41	1		8	6
Total	43	44	39	12	138	100

noun phrases, but the pronoun that appears most often (*I*) also occurs with the most reduced forms of *don't*.

Why would these more frequent forms, the pronouns, condition flapping in *don't*? Flapping of a coronal occurs within a phonological unit when the coronal is both preceded by and followed by a syllabic unit and the syllable preceding the coronal has more stress than the one following it. Flapping rarely affects a word-initial /t/ or /d/. Apart from their all ending in vowels, there is nothing special about the pronouns that would cause them to condition flapping more than any other unit, except the frequency with which they occur with *don't*. Our hypothesis is that the pronoun and *don't* constitute a storage and processing unit that is gradually undergoing reduction due to frequency of use (Bybee 2000b).

The vowel reduction (o → ə) that is observed in this case is often regarded as conditioned by the relative lack of stress. We are not directly appealing to stress, however, because the main indicator of lack of stress is precisely the reduction of the vowel to schwa. In other words, the reduced vowel and the perception of reduced stress are not independent of each other.[4] Moreover, we would still have to explain why the vowel and stress reduction only occurs when *I* is the subject. Since there is nothing in the phonological properties of *I* that could condition the reduction, we must conclude that it is the frequency with which *I don't* is used as a sequence that is responsible for the reduction of the vowel.

In addition, as shown in table 14.2, the vowel reduction occurs primarily with the verbs that occur most frequently in the data: *know*, *think*, *have* (*to*), *want*, *like*. Vowel reduction also occurs with three other verbs, *mean*, *feel*, and *care*, which are not particularly frequent in our corpus, but of course all of these occur very commonly in the phrases *I don't mean*, *I don't feel*, and *I don't care* in American English. Again, there is nothing in the phonological properties of these verbs that would condition vowel reduction in a preceding word. It is rather the frequency with which these phrases—*I don't know*, *I don't think*, *I don't have* (*to*), *I don't want*, and *I don't care*—are used that has rendered them fused storage and processing units and has conditioned the loss of stress on the middle element and its consequent reduction.

4. Storage and processing

Almost all of these reduced phrases also occur in a nonreduced form in our data. The individual elements in these phrases, *I*, *don't*, and the verb, are storage and processing units as well, so that it is also possible to use them compositionally. In addition, each of these individual units is associated with the stored combinations, *I don't know*, *I don't mean*, and so on. In this way, these phrases are parallel cognitively to the morphological formations modeled in Bybee (1985, 1988) and Losiewicz (1992). In this associative network model of morphology, irregular forms are stored lexically and associated by lexical connections with similar forms. But lexical storage is highly affected by language use, such that high-frequency forms have stronger lexical representation than low-frequency forms. It is for this reason that low-frequency irregulars tend to regularize. But it also follows that high-frequency regular morphological formations would be represented in memory. So high-frequency past-tense/past-

TABLE 14.2. *Don't* Variants in Declarative Constructions by Following Expressions (n = 124)

Following	*stop + o* *Group 1*	*flap + o* *Group 2*	*flap + ə* *Group 3*	*ə* *Group 4*	*Total*
know	2	8	24	5	39
think	7	6	6	1	20
have	1	7	1		9
have to	1	2	1		4
want	1	1	3		5
see	3	1			4
like		1		1	2
get	1	2			3
mean				1	1
feel				1	1
care			1		1
play		1			1
meet		1			1
believe		2			2
inhale	2				2
work		1			1
support		1			1
give		1			1
need		1			1
intend	1				1
go out	1				1
make contact	1				1
follow through	1				1
do	1				1
code	1				1
eat	1				1
Total					
verb tokens	25	36	36	9	106
verb types	15	15	6	5	26
adverb/disc marker	3	3		1	7
pauses/break	6	1	2	2	11
Total	34	40	38	12	124

participle forms such as *supposed*, *wanted*, *liked*, and so on would also be strong storage and processing units (Stemberger and MacWhinney 1988; Baayen, Burani, and Schreuder 1996).

Losiewicz (1992) devised a test of this hypothesis based on earlier findings by Walsh and Parker (1983). In a series of experiments Walsh and Parker found that English /s/ in word-final position is longer in acoustic duration if it is the plural morpheme (*laps*) than if it is part of a monomorphemic word (*lapse*). Losiewicz (1992) found this same distinction applies to morphemic /d/ or /t/ (the past tense) as in *rapped* versus nonmorphemic final /d/ or /t/ as in *rapt*. Losiewicz further reasoned that if the difference in length is due to the non-morphemic segment being part of a lexical representation, while the morphemic one is added to the stem in processing, then the same difference in length should appear in low- versus high-frequency words with morphemic /d/, since the low-frequency words would be formed by using a schema and the high-frequency words would be accessed directly from the lexicon.

Losiewicz asked subjects to read sentences containing English past-tense forms that constituted rhyming pairs of high- and low-frequency verbs (*covered*, *hovered*; *needed*, *kneaded*). For all subjects and all pairs of verbs, the final past morpheme was longer in the low-frequency verb of the pair. The average duration difference was 7 ms, and this difference was highly significant and not due to overall differences in word length. It is not proposed that such a length difference is either perceptible or learnable from input, but rather that it reflects a difference in processing type. These results, then, can be taken to support the hypothesis that high-frequency inflected verbs are stored in the lexicon while low-frequency inflected forms are produced by combining a template for past -(*ɨ*)*d* to a base verb form.

Similarly, we can postulate two cognitive mechanisms by which a phrase such as *I don't know* can be produced. In one case, the expression is a construction and accessed whole from storage, and it thus includes reductions and coarticulations that have accumulated in its representation; in the other case it is put together from two (*I don't* and *know*) or three (*I*, *don't*, and *know*) elements, in which case the vowel of *don't* will not be reduced, though flapping can occur.[5] As with the past tense, these alternate means of access do not necessarily correspond to different degrees of semantic compositionality.

However, it has also been proposed in this model that stored units that are of high frequency of occurrence will be more autonomous from related units in storage, accounting for the fact that derivationally or inflectionally related words that are of high frequency are more prone to split off semantically from related words (Bybee 1985). For the phrases studied here, this means that with higher frequency of the unit as a whole, the function may deviate more from the function of the phrase as composed of three lexical units. This is precisely what is found in the data studied here. As Scheibman (2000) has shown, there is a high correspondence between the reduced version of *I don't know* and a special discourse function that is independent of the meaning of the phrase as a sum of its parts. A similar case can be made for the phrase *I don't think*, which is not used to literally mean that the speaker has failed to generate cognitive activity, and for *why don't you* when used in making suggestions.

5. Constituent structure

Let us look more closely now at what the data suggest about the constituent structure and storage of expressions containing *don't*. What we find is that where *don't* is used in a frequent phrase, it is more likely to occur in its most reduced form (with a flap and reduced vowel), and in these pronunciations it is often the case that the negative auxiliary does not contribute compositionally to the semantics of the expression. That is, in its reduced form in frequent collocations, *don't* does not always literally mean 'do not'; rather, the entire phrase can convey less analytically derived, discourse-dependent meanings. Such distinctions in function among phrases containing the negative auxiliary provide important evidence that constituent structure is subject to the same phonological, semantic, and functional usage factors that affect lexical items undergoing processes of grammaticization. In addition, the discovery of divergent functional-phonological correspondences for expressions whose elements reflect different degrees of bondedness is amenable to a processing and storage model that allows differential treatment of the *same* lexical item or phrase depending on its meaning, its pronunciation, and its linguistic and conversational contexts.

5.1. *I don't know*

Of the thirty-seven total tokens of *I don't know* in the corpus, only eight contain full-vowel variants of *don't* (two with a stop and six with a flap), whereas twenty-nine of these usages are schwa forms (twenty-four pronounced with a flap and five with just the reduced vowel). In other words, in 95% of *I don't know* tokens, *don't* is pronounced with either consonant or vowel reduction or both, and for 78%, the auxiliary is articulated with a reduced vowel.[6] In this frequent expression, then, there is consistent phonological reduction of the negative auxiliary.

With respect to their functions in conversation, all variants of *don't* in *I don't know* are able to express the compositionally transparent meaning of not having knowledge about an entity or a proposition, as illustrated in (1):

(1) (Sexual practices)
O: well but that's interesting,
because I was telling F,
I `don't `know ^any ^woman,
[dõt]
that I've discussed it with,
who hasn't tried it. (B2.434O)

Seven out of eight of the full-vowel variants participate in this function of verb negation, as do twelve out of the twenty-nine reduced-vowel forms. *I don't know*, then, pronounced in varying degrees of reduction may express a speaker's lack of knowledge about a given entity, event, or proposition. However, one out of the eight full-vowel variants of *don't* in the expression *I don't know*—a token pronounced with a flap—and seventeen out of the twenty-nine schwa variants in these constructions perform more pragmatic functions in discourse. Specifically, these uses of *I don't know* in conversation express a more subjectified meaning of *not knowing* by conveying a speaker's uncertainty toward a proposition or an extended stretch of talk.[7] For example, in (2), Z's *I don't know* is uttered rapidly as one unit (marked "allegro" in the transcript) after her narrative and right before a speaker change. In this context—and also in six out of seven similar usages of *I don't know* in the corpus that occur after usually lengthy stretches of talk—the speaker appears to be politely and deferentially opening up the floor to other participants. Östman (1981) notes a similar use of *I don't know* that expresses speakers' uncertainty about their contributions and can function as a floor-yielding device.

(2) (Getting old)

Z: well I talked to a guy that's thirty-four in my class,
and we were talking about the difference,
just in . . . physiology and how you feel,
and your best study hours,
and [the . . .] rest you need versus what you do,
O: [right]
Z: <A *I don't know* A>
[ɾə]
it's just—
O: right
5: you know what was the biggest giveaway for me?
I couldn't—
I can't get loaded anymore.
Z: uh-huh.
All: @@@@
S: I mean without paying. (G1.76Z)

Similarly in (3), O's *I don't know* does not convey a lack of knowledge about an entity or proposition; rather, her *knowing* or *not knowing* is dependent on her own future action, which has interactional consequences. As S's request—*will you finish the salad?*—favors an affirmative response from O (i.e., it is directed toward the salad's getting eaten up and not whether or not O actually wants more salad), the speaker's *I don't know*. *not now*. softens her negative response:

(3) (Dinner)

S: you don't get any more asparagus,
'cause I gave it all to Z.
N: that's okay.
O: @@@
S: how about salad?
N: nope I'm great.
S: no salad?
O will you finish the salad darling?
O: *I don't ^know.*
[ɾ̃o]
not now. (G2.464O)

In summary, though all variants of *don't* in *I don't know* convey the phrase's lexical sense of *not knowing*, only reduced forms (one token with a reduced consonant and seventeen tokens with reduced vowels) express the construction's more pragmatic functions of indicating speaker uncertainty and mitigating polite disagreement in conversation, as shown in table 14.3.

5.2. *I don't think*

The second-most frequent use of *don't* in the corpus is in the expression *I don't think* (nineteen occurrences of *I don't think* versus thirty-seven of *I don't know*); it is also

TABLE 14.3. Full-Vowel and Reduced-Vowel Variants of *Don't* in *I Don't Know* by Lexical versus Pragmatic Function

	Full Vowel	*Schwa*
Lexical sense	7	12
Pragmatic function	1	17

the second-most frequent site for reduced *don't*. As is the case for many tokens of *I don't know*, *I don't think* conveys a meaning in conversation that is noncompositional; that is, when speakers use this expression they are not expressing an inability to mentally formulate or ponder something. Unlike the case of *I don't know*, however, which may be used to indicate a lack of knowledge about something, speakers never use *I don't think* to convey the lexical sense of *not thinking*. This meaning is expressed by the verb *to think about*, as illustrated in (4):

(4) (Bad water)
O: I'll never drink out of the water fountains,
I mean they're so horrible,
the [stuff] tastes carcinogenic.
Z: [really]?
oh my god.
O: it's really horrible @@.
so . . . if I take this stuff,
I'll drink it but,
I don't `think about ^taking it. (G2.557O)
[ɾõn]

For all *I don't think* tokens, speakers convey an epistemic stance toward some elaborated or pro-complement (e.g., *I don't think so*), as illustrated in (5). And as is the case with pragmatic uses of *I don't know*, *I don't think* allows speakers to hedge and manage polite disagreement in conversation (Thompson and Mulac 1991):

(5) (Bed and Breakfasts)
S: you guys need some capital,
I keep saying.
F: we need the Mormon church behind us.
O: oh?
S: <A how'!! you get them A>?
F: how?
O: <@yeah *I don't think* they'll ^go for your ^fantasy @>.
[ɾ̃əɔ̈]
(B1.92O)

I don't think always has scope over an entire clause, so that what is negated is the following complement and not the verb in the construction. In other words, for *I* *don't think* + complement constructions, *don't* does not negate the verb *think* but rather the proposition in the subsequent clause. Though *I don't think* is a frequent site for reduced *don't*, unlike *I don't know*, there does not appear to be a form–function correspondence between variants of *don't* in *I don't think* and its meanings and uses. That is, though the meaning of *I don't think* is compositionally unanalyzable in conversation—indicating a more grammaticized unit—there is no consistent formal reduction concomitant with this functional shift as we saw for *I don't know* (i.e., we find in the data both full and reduced variants of *don't* in *I don't think*).

5.3. *Why don't* (*you*)

In the data, *don't* occurs eight times in interrogatives: twice in *don't you think*, in which it is not reduced, and six times in *why don't* (*you*). In this latter phrase *don't* is in its full form once and occurs with a flap four times and with a reduced vowel once. It is admittedly a small number of tokens, but the distribution of variants appears to be significant. Five of the instances of *why don't* (*you*) are not actually questions but rather suggestions, as seen in (6) and (7):

(6) (Dinner)
O: *why don't* you `take ^one,
[ɾõt]
before I pass it? (B1.53O)

(7) (Office hours)
S: she asked me a question,
I say,
no that's not ^one question,
so I started telling her how it's—
Z: hmm.
S: more than one question.
I said *why don't* you sit ^down,
[ɾ̃əʊ̈]
so that I can talk to you about it. (G3.9S)

All of the suggestion uses of *why don t* (*you*) have a flap in *don't* and one has a reduced vowel. Worth noting is that the one case of the full form in this context occurs in an actual question:

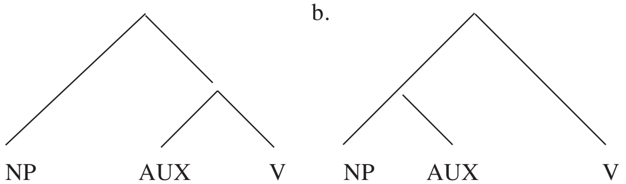
(8) (West Texas motel)
O: now *why don't* you think that's a fantasy?
[dõ]
S: I don't `know `what's ^fa-
[ɾəʊ̈]
I can't imagine what would . . . consti-
O: well just because you can't understand it
doesn't mean that it doesn't exist.
S: are you a little mad at me? (B1.179S)

In fact, it seems that if *don't* occurs in its full form after *why* it will be interpreted as asking a question rather than making a suggestion. Thus it appears that *why don't*, when it is used for suggestions, constitutes a processing and storage unit: it has a particular range of phonetic variation and it has a function not derivable directly from the sum of its parts. Furthermore, parallel to the case of *I don't know*, the reduced form of *why don't* (*you*) participates in a pragmatic function of making polite suggestions versus the full form's being involved in a more propositional task of asking for information.

6. Degrees of internal constituency

Within the three-unit phrase, *I don't know*, *think*, *want*, and so on, degrees of constituency exist.[8] The structure predicted by the syntactic structure would separate the subject from the verb phrase as the major constituent division, and the auxiliary would separate from the verb as a lower level division, as shown in (9a). However, we will argue from the phonological evidence that in these frequent phrases, and perhaps more generally in English, the constituent structure of NP–AUX–V is as in (9b):

(9) a. b.



The reduction of the vowel of *don't* is more dependent on the subject than it is on the following verb. The reduction only occurs with *I* (and in one instance with *why*), but it occurs with a variety of verbs: *know*, *think*, *have* (*to*), *want*, *like*, *mean*, *care*, and *feel*. Moreover, the deletion of the flap occurs only with *I* but with a variety of verbs: *know*, *think*, *like*, *mean*, and *feel* in this corpus.

Note also that an adverb intervening between the subject and *don't* blocks vowel reduction, as seen in examples (10a) (S's last utterance) and (10b), but an adverb between *don't* and the verb does not, as in example (11):

(10) a. (Bubbles)
N: mine was plain water,
and now it's carbonated.
I'm drinking somebody else's stuff.
O: no . . . you're not.
S: I don't ¤think ‘so.
[ɾ̃əʊ̈]
All: @@@
O: you're having [a hallucination].
S: [*I ¤really ¤don't `think* so]. (G2.540S)
[ɾõ]

b. (Sexual practices)
O: but *I also ‘don’t ‘know* any . . . one,
[rõ]
that I’ve discussed it with,
who enjoyed it. (B2.435O)

(11) (Coming of age)
Z: you had all the feelings,
but you just,
didn’t play,
you didn’t—
you just ignored it.
S: so you were just horny anyway,
so you might as well fuck men.
Z: yeah . . . well . . . I—
I—
I don’t even ^know,
[əɒ̈]
if I was `really all that ^horny. (G3384Zb)

The greater dependence of *don’t* on the subject is possibly part of the general tendency in English for auxiliaries to lean on the subject, as evidenced by contractions such as *’ll*, *’s*, *’ve*, *’d*, *’re*. This proclivity in English is in accordance with Halliday’s (e.g., 1994) positing of the mood element. In Hallidayan systemic grammar the mood is composed of subject plus auxiliary. This component of the clause is used by discourse participants to negotiate propositions with one another (e.g., A: *can she do it?* B: *yeah, she can*.) and may also express speaker stance, especially when the subject is first person singular or when the auxiliary is a modal.

However, given that traditionally the auxiliary syntactically and semantically belongs in the verb phrase—for example, grammatical categories such as person, number, tense, aspect, and mood typically occur with the verb—why does the phonological fusion unite the auxiliary with the subject?

Our data suggest an answer to this in terms of the number of types that occur in the position before the auxiliary versus the number that appear after the auxiliary. In the conversations analyzed here, which yielded 124 examples of *don’t* in declarative constructions, only twelve different items preceded *don’t* (three are adverb types), while thirty different items followed *don’t* (four are adverb types). This means in actual language use, the position following *don’t* is more than twice as flexible as the position preceding *don’t*.

(12)

	Subject	*don’t*	verb/adverb
type frequency	12		30
	(5 lexical NPs)		(16 verbs occurring only once)

It is also interesting to consider the transition probability in the sequence *I don't*. That is, what is the most common item to follow *I*? Krug (1998) has shown for a large corpus of British English that the most common auxiliaries to contract with *I* are also the ones that follow *I* most frequently. Therefore, we wanted to know what items follow *I* most commonly. Unfortunately, we did not have this information for the corpus studied for *don't* reduction because only utterances with *don't* in them were transcribed. However, it was possible to use another conversational corpus with 414 tokens of *I* to count the items following *I*. In this corpus we found that the most frequent item to follow *I* was *'m*, with forty-seven occurrences. The next most frequent item was *don't*, with forty-four occurrences. Krug has argued that conventionalized contractions are conditioned by frequency of cooccurrence. The reduction of *don't* after *I* could be regarded as a type of contraction and thus conditioned by the same factors that condition contraction of the auxiliary with the subject.

The item following *don't* with the highest token frequency was *know*, which occurred thirty-nine times; the item preceding *don't* with the highest token frequency was *I*, which occurred eighty-eight times. Fewer items with low token frequency preceded *don't* than followed it. Five lexical noun phrases, each occurring once, preceded *don't*, but sixteen verbs that occurred only once followed *don't*. That is, the class following *don't* is more open than the class preceding *don't*. The fewer types and the resultant higher token frequency of individual types creates a stronger syntagmatic relation in the storage unit. In other words, because of a higher token-to-type ratio, subject–auxiliary sequences are more entrenched than auxiliary–verb sequences.

This case is somewhat unusual because the frequency of co-occurrence does not correspond to semantic relations. Ordinarily, semantic relatedness would correspond to tighter constituent structure. Instead, in the case of reduced *don't*, we have elements joined into frequently used constructions because of their occurrence in discourse. Recall that all but one of the reduced variants of the negative auxiliary co-occur with *I*. Collocations composed of a first person singular pronoun plus a verb such as *know* or *think* may easily take on as a major part of their meaning the speaker's evaluation or construal of a proposition. The basic epistemic sense of *know* in the expression *I don't know* contributes to the construction's usefulness in discourse, not only as a coventionalized response to explicit or implied requests for information but also as an interactional softener with scope over both linguistic and social material. We propose, then, that functional relatedness, which leads to contiguity in expression, and variation in type and token frequency lead to the behavior attributed to abstract relations of constituent structure.

Traditional methods of determining constituent structure in competence-based models tend to consider distributional properties without considering type and token frequency. Thus pronouns are considered noun phrases because they occur in the same positions as full noun phrases: before a verb phrase, after a preposition, and so on. *I* is a pronoun because it shares properties with other pronouns: it takes no determiners or other modifiers and it has both subject and object forms. Our approach aims for a model of usage and performance, where constituents are processing units. Thus the evidence we consider here from phonological fusion and distribution in terms of type and token frequency suggests a different constituent structure for *I* plus *don't* and perhaps other auxiliaries than for full noun phrases plus a lexical verb or even for full noun phrases plus an auxiliary.

This proposal is not all that odd typologically. In many languages, pronouns affix to the verb while full noun phrases do not (e.g., Navajo [Young and Morgan 1980]). In other languages, person/number markers fuse with modality markers (e.g., Quileute [Andrade 1933]), or pronouns and auxiliaries of tense, aspect, and modality fuse into a clitic complex (e.g., Luiseño [Steele 1981]). In all cases, pronouns evince a different constituent structure than full noun phrases, and auxiliaries or other items that develop from verbs behave differently from lexical verbs.

Most recent analyses of cases where phonological processes operate across word boundaries only in certain contexts, some of which are not syntactically motivated, conclude that the relation between syntactic structure and phonological or prosodic units is at times arbitrary (Inkelas and Zec 1990). We suggest that such relations are not arbitrary, but that frequency of co-occurrence determines the application of phonological processes across word boundaries because it also determines the chunking of speech into processing units that we equate with constituents. Given the evidence we have offered here and the support from the structure of other languages, it seems to us that the claim that languages suffer a mismatch between syntactic structure and prosodic structure is the one that requires justification.

7. Conclusions

Following the many functionalists who have cited repetition as the mechanism that creates grammar (e.g., Givón 1979; Haiman 1994), we have shown that the probability that two items will be contiguous in naturally occurring speech determines their degree of fusion into constituents (see also Meillet 1912). Of course, since contiguity in discourse is determined by pragmatic and semantic factors, items that occur together will be relevant to one another. In the usual case, then, this principle will lead to the commonly occurring constituent relations—preposition with noun phrase, adjective with noun, auxiliary with verb, and so on. However, we have also seen that not all instances of a construction have the same status in storage and processing and not all semantic relations are expressed iconically in constituent structure. By examining cases in which phonological fusion and discourse function defy the usual constituent relations, we are able to reveal the effects of frequency of occurrence in creating new constituent relations.

Our study shows that a low type frequency for each slot in a construction coupled with a high token frequency of particular items in a construction creates the tightest constituent structure. It follows that the higher the type frequency of any particular position in a construction, the looser will be its constituent bonds to other parts of the construction. This hypothesis makes predictions that could be tested on the behavior of other constructions in naturalistic data. We believe that using naturally occurring discourse data to help determine the nature of cognitive representations will yield a better understanding of the set of relationships studied under the rubric of constituent structure.

Notes

This paper was presented at the Symposium on Constituency and Discourse held at the University of California, Santa Barbara, in 1997. We thank the participants in the symposium for their comments, especially Bill Croft, Mark Durie, and Sandra Thompson. Thanks also to John Haiman for comments on an earlier version of this chapter. Scheibman acknowledges the assistance of the Research, Project and Travel Grant awarded by the Office of Graduate Studies, University of New Mexico. Correspondence address: Department of Linguistics, Humanities 526, University of New Mexico, Albuquerque, NM 87131-1196. E-mail: jbybee@unm.edu.

1. But see below how the realities of language use affect the relative openness of these classes.

2. Kaisse (1985) claims that the presence or absence of a nasal consonant or vowel nasalization in a given pronunciation of *don't* is frequently predictable based on the initial consonant of the following word. However, the conversation data do not show such systematic variation.

3. Numbers following examples locate utterances in the database. Transcription is based on the DuBois et al. (1993) system. Stress is indicated only for the intonation units in which *don't* occurs. The following are the symbols used in this paper:

^	= primary stress
`	= secondary stress
.	= final transitional continuity
,	= continuing transitional continuity
?	= appeal transitional continuity
[]	= speech overlap
—	= truncated utterance
..	= pause
(0)	= latching
A	= allegro: rapid speech
@	= laughter
X	= indecipherable

4. Ladefoged (1975:99–102) distinguishes two types of unstressed syllables in English—those with a full vowel and those with a reduced vowel.

5. Kaisse (1985) also argued that *don't* with a reduced vowel has a separate representation from the one with a full vowel.

6. Furthermore, 64% of the total number of reduced forms of *don't* in all declarative utterances in the corpus—those that are pronounced with a flap and a reduced vowel and those pronounced just with a reduced vowel—occur in the expression *I don't know*.

7. See Scheibman (2000) for in-depth analyses of these usages.

8. We are assuming, of course, that *don't* is a single unit, the fusion of two etymological units.

References

Andrade, Manuel J. 1933. *Quileute*. New York: Columbia University Press.

Baayen, Harald, Cristina Burani, and Robert Schreuder. 1996. Effects of semantic markedness in the processing of regular nominal singulars and plurals in Italian. *Yearbook of morphology 1996*, 13–33.

Boyland, Joyce Tang. 1997. Morphosyntactic change in progress: A psycholinguistic approach. Ph.D. diss., University of California at Berkeley.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Philadelphia: John Benjamins.

———. 1988. Morphology as lexical organization. In *Theoretical morphology*, ed. M. Hammond and M. Noonan, 119–141. San Diego, Calif.: Academic Press.

———. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5), 425–455. (Reprinted in the volume as chapter 8.)

———. 2000a. Lexicalization of sound change and alternating environments. In *Papers in Laboratory Phonology*, vol. 5: *Acquisition and the lexicon*, ed. Michael Broe and Janet Pierrehumbert, 250–268. Cambridge: Cambridge University Press. (Reprinted in this volume as chapter 10.)

———. 2000b.The phonology of the lexicon: Evidence from lexical diffusion. In *Usage-based models of language*, ed. Michael Barlow and Suzanne Kemmer, 65–85. Stanford, Calif.: CSLI. (Reprinted in this volume as chapter 9.)

Bybee, Joan L., and Sandra Thompson. (1997). Three frequency effects in syntax. *Berkeley Linguistics Society* 23, 65–85. (Reprinted in this volume as chapter 12.)

DuBois, John W., Stephan Schuetze-Coburn, Susanna Cumming, and Danae Paolino. 1993. Outline of discourse transcription. In *Talking data: Transcription and coding in disourse research*, ed. Jane A. Edwards and Martin D. Lampert, 45–89. Hillsdale, N.J.: Erlbaum.

Givón, Talmy. 1979. *On understanding grammar*. New York: Academic Press.

Goldberg, Adele E. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: University of Chicago Press.

Haiman, John. 1994. Ritualization and the development of language. In *Perspectives on grammaticalization*, ed. William Pagliuca, 3–8. Amsterdam: John Benjamins.

Halliday, M. A. K. 1994. *An introduction to functional grammar*, 2nd ed. London: Edward Arnold.

Heine, Bernd. 1993. *Auxiliaries*. Oxford: Oxford University Press.

Heine, Bernd, Ulrike, Claudi, and Friederike Hünnemeyer. 1991. From cognition to grammar: Evidence from African languages. In *Approaches to grammaticalization*, vol. 1, ed. Elizabeth Traugott and Bernd Heine, 149–187. Amsterdam: John Benjamins.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphological change. In *Current progress in historical linguistics*, ed. William Christie, 96–105. Amsterdam: North Holland. (Reprinted in this volume as chapter 2.)

Hopper, Paul J. 1991. On some principles of grammaticalization. In *Approaches to grammaticalization*, vol. 1, ed. Elizabeth C. Traugott and Bernd Heine, 17–35. Amsterdam: John Benjamins.

Inkelas, Sharon, and Draga Zec. 1990. *The phonology–syntax connection*. Chicago: University of Chicago Press.

Kaisse, Ellen. 1985. *Connected speech: The interaction of syntax and phonology*. San Diego, Calif.: Academic Press.

Krug, Manfred. 1998. String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26, 286–320.

Ladefoged, Peter. 1975. *A course in phonetics*. New York: Harcourt Brace Jovanovich.

Langacker, Ron. 1995. A dynamic usage-based model. Paper presented at the Symposium on Usage-Based Models of Language, Rice University, March.

Losiewicz, B. L. 1992. The effect of frequency on linguistic morphology. Unpublished diss., University of Texas, Austin.

Meillet, Antoine. 1912. L'évolution des formes grammaticales. *Scientia* (*Revista di Scienza*) 12(1): 26, 6 (Reprinted in Antoine Mellet, [1958]. *Linguistique historique et linguistique général*, 130–148. Paris: H. Campion.)

Östman, Jan-Ola. 1981. *You know: A disourse-functional approach*. Amsterdam: John Benjamins.

Phillips, Betty S. 1984. Word frequency and the actuation of sound change. *Language* 60, 320–342.

Scheibman, Joanne. 2000. *I dunno but*. A usage-based account of the phonological reduction of *don't* in American English conversation. *Journal of Pragmatics* 32(1), 105–124.

Steele, Susan. 1981. *An encyclopedia of AUX: A study of cross-linguistic equivalence*. Linguistic Inquiry Monograph 5. Cambridge, Mass.: MIT Press.

Stemberger, Joseph Paul, and Brian MacWhinney. 1988. Are inflected forms stored in the lexicon? In *Theoretical morphology: Approaches in modern linguistics*, ed. Michael Hammond and Michael Noonan, 101–116. San Diego, Calif.: Academic Press.

Thompson, Sandra A., and Anthony Mulac. 1991. A quantitative perspective on the grammaticalization of epistemic parentheticals in English. In *Approaches to grammaticalization*, vol. 2, ed. Elizabeth Traugott and Bernd Heine, 313–329. Amsterdam: John Benjamins.

Tranel, Bernard. 1981. *Concreteness in generative phonology: Evidence from French*. Berkeley: University of California Press.

Walsh, T., and F. Parker. 1983. The duration of morphemic and non-morphemic /s/ in English. *Journal of Phonetics* 11, 201–206.

Young, Robert W., and William Morgan. 1980. *The Navajo language: A grammar and colloquial dictionary*. Albuquerque: University of New Mexico Press.

Chapter 15

Sequentiality as the Basis of Constituent Structure

1. Sequentiality and constituent structure

Recent investigations of the relation between grammar and language use, especially in the literature on grammaticization (Hopper and Traugott 1993; Bybee, Perkins, and Pagliuca 1994; among many others), have made it clear that grammar is not a fixed logical structure but rather is constantly changing and evolving. Haiman (1994) and I (Bybee 1998b) have argued that the basic mechanisms of the grammaticization process are neither domain-specific nor species-specific. Thus the hope has arisen that the highly abstract and heretofore mysterious properties of grammar might be explainable in more general terms. In my earlier work I have addressed the process by which grammatical morphemes develop, emphasizing the role that repetition plays in the process. This chapter addresses the mechanism behind the hierarchical arrangement of linguistic elements into constituents, once again emphasizing the role of language use and repetition.

The existence of constituent structure and the hierarchical organization resulting from it has always been taken by linguists as prime evidence that linguistic behavior does not consist merely of linear strings of elements. It is further believed that the hierarchical organization of sentences is one of the most basic aspects of language, indeed, a defining feature of human language.

Linguists rarely ask why natural language has constituent structure;[1] they merely assume that it does, just as they assume that all phonologies will be organized into segments and features. In the spirit of Lindblom, MacNeilage, and Studdert-Kennedy (1984), I submit that structure can be explained, that form is emergent from substance, and that the larger design of language can be interpreted as the indirect consequence of local behavior (Lindblom, MacNeilage, and Studdert-Kennedy 1984:186). My goal in this chapter is to present and defend the hypothesis that sequentiality is basic to language and constituent structure emerges from sequentiality because elements that are frequently used together bind together into constituents.

This point is likely to seem rather obvious to some readers and indeed follows directly from comments in Givón (1998, 2002). To my knowledge, however, it has not been directly argued for or demonstrated. In fact, some linguists seem to be convinced that constituent structure is basic to language and therefore might find the hypothesis surprising. Thus, toward the ultimate goal of identifying the mechanisms that create grammar and relating them as much as possible to domain-general cognitive processes, I undertake this demonstration.

2. Definition of constituent

A popular introductory textbook for linguistics, *Language Files* (Cipollone, Keiser, and Vasishth 1998), has a nice discussion of the notion of constituent. It begins with a discussion of linear order in sentences and then points out that in addition to linear order, there is hierarchical organization. The book says that the semantically coherent groupings of words in sentences are constituents, thereby identifying both semantics and linear order as important to constituent structure. It goes on to give three tests for identifying constituents:

1. Constituents can sensibly be used alone, as in the answers to questions. (*Who chewed up your shoe? My new puppy*.)
2. Constituents can be replaced by pro-forms. (E.g., *I saw <u>him</u> <u>do it</u>* where *him* replaces *my new puppy* and *do it* replaces *chew up my shoe*.)
3. Constituents can occur in various places in the sentence; for example, an NP can be the subject of the verb or object of the verb or a preposition; an NP can be extraposed (e.g., *My new puppy, he chewed up my shoe*).[2]

3. Explanations for constituent structure

Of course various explanations for why language is organized into constituents are possible. One could claim that specific types of phrases, such as noun phrases, verb phrases, and prepositional phrases, are innate and only need the presence of minimal input to trigger their language-specific organization. Or one could propose, as does Langacker (1987:310), that "hierarchy is fundamental to human cognition," making it a general cognitive attribute that can be applied to language. Certainly I agree with this view. However, we must also ask just what is organized hierarchically and how languages come to be organized in the particular way that they are.

Langacker's (1987, 1997) particular proposal is that constituency reflects semantic relations. Again, it is easy to agree with this position since many linguists have pointed out the iconic relation between the conventionalized structures of languages and their semantic relevance (e.g., Bybee 1985). However, as we will see later, constituency and meaning are not always in a strictly iconic relationship and the human mind can apparently find constituents in reoccurring meaningless syllables (Saffran, Aslin, and Newport 1996). Thus we search for an additional factor to fully explain the existence of constituent structure.

My hypothesis is that semantics, and, to some extent, pragmatics and our experience with the world, will determine what elements tend to occur together in sequences in an utterance, but repetition is the glue that binds constituents together. Thus I hypothesize that hierarchies of constituent structure are derivable from frequent sequential co-occurrence. In this view, the more often particular elements occur together, the tighter the constituent structure. Thus low-level constituents such as determiner, *the*, and a noun, such as *puppy*, frequently co-occur, while higher level constituents, such as an NP, *the puppy*, and verbs such as *ran*, *licked*, or *slept* occur together less often. Note that in this view constituent structure can be gradient and two constituents that seem to have the same structure may have different degrees of cohesion due to the differences in their co-occurrence patterns (Bybee and Scheibman 1999).

4. Knowledge of grammar is procedural knowledge

To understand the role that frequency or repetition plays in the creation of grammar it is important to recognize that language production is a neuromotor behavior based on procedural knowledge rather than propositional knowledge. Propositional knowledge is "knowing that" or knowing facts such as "Santa Fe is the capital of New Mexico." Procedural knowledge is "knowing how" and includes knowing how to tie shoelaces or how to drive a car. Propositional knowledge is conscious knowledge that is easy to report on. Procedural knowledge is usually below the level of conscious awareness, and while subjects can carry out the procedures, it is much more difficult for them to report what the procedure is. This distinction has an interesting parallel in the difference between lexical and grammatical knowledge. While speakers are often able to report on the meanings of words or phrases, it is much more difficult for untrained speakers to explain the meanings of grammatical morphemes or grammatical constructions. Thus we might conclude that lexical items involve at least some propositional knowledge, while grammatical constructions are largely procedural.

This conclusion is further supported by the manner in which procedural knowledge develops, as outlined by Anderson (1993) and Boyland (1996). Two properties of the development of procedural knowledge are important for our understanding of the way grammar develops. First, frequently used actions become fluent more quickly; that is, repetition increases fluency. For the purposes of language, our commonsense notion of fluency can be applied here: in fluent speech, words are strung together without inappropriate pauses. We can also go beyond this sense of fluency and note that with high levels of repetition, articulatory gestures can overlap one another and individual gestures can be reduced both in duration and in displacement. Thus grammaticizing constructions that undergo extreme frequency increases also undergo extreme phonological fusion and reduction, as, for example, when *going to* reduces to *gonna* (see Bybee, Pagliuca, and Perkins 1994; Bybee 2001). The second property is in a sense the mechanism that makes the first one possible: "recurring sequences of actions come to be represented at a higher level as single actions, thus increasing fluency" (Boyland 1996:10). Thus repeated sequences become fluent because they become automated into a single chunk that can be accessed and executed as a unit.

5. Linguistic evidence for chunking

Several types of evidence for the chunking of high-frequency sequences can be cited. First, one unit of a chunk primes or automates the other unit. In the plaza at Old Town in Albuquerque I watched a boy of about 5 years spot an antique cannon. He said, "Hey, Dad, there's a cannonball. Can I climb on the cannonball?" The father responded, "That's not a cannonball; it's a cannon." The boy insisted, "Dad, can I climb on the cannonball?" and the exchange repeated itself. The boy had learned *cannon* only in the context of the compound *cannonball*, and that was the only context in which he could access the word. Hearers have automated chunks as well, with analogous priming effects. In the United States, upon hearing *supreme*, one can expect *court* as the next word, or upon hearing *sesame* one can expect *street*.

Second, inside frequently used chunks, internal structure tends to be lost. Thus *gonna* no longer consists of the three morphemes *go*, *ing*, and *to*.

Third, the morphemes or words inside a chunk become autonomous from other instances. For example, speakers probably do not associate *go* in *gonna* with the lexical movement verb anymore. Sosa and MacFarlane (2002) show that subjects have difficulty identifying the word *of* when it occurs in frequent chunks such as *sort of* or *kind of*.

Fourth, the components of a chunk become compressed and reduced phonologically, as illustrated by *gonna*, but also by other verbs that are frequently followed by *to*, such as *wanna*, *hafta*, *gotta*, and so on.

Based on these facts, the hypothesis of this chapter is: items that are used together fuse together.

This could be called the Linear Fusion Hypothesis. The explanation for this phenomenon is twofold: first (and, in my mind, foremost) is the automation of production that is typical for procedures; second, the fact that items are predictable in context allows the speaker to unleash the reductive processes that increase fluency.

6. Further examples of the fusion of high-frequency combinations

Boyland (1996) argues that the sequence *would have* fuses due to the high frequency with which the two units occur together. The misanalysis apparent in the nonstandard spelling *would of* shows that the identity of the two units has been lost. Also, the context after a modal is the only place that *have* reduces down to [ə]. In the example of *be going to*, the most fusion is among the parts that are invariant and therefore occur together more often, that is, the *going to* part. The *be*-verb part, which varies among *am*, *is*, *are*, *was*, and *were*, reduces and fuses only in the high-frequency first person singular: *I'm going to* becomes not just *I'm gonna* but also [aimə̃ɾə̃] in casual speech. Krug (1998) studies the rate of contraction of English auxiliaries with the subject and finds that the more frequent a combination, the more likely it is to show contraction. For instance, the most frequently occurring contraction is *I'm* and that combination is the most frequent in the corpus. The basis for the high frequency of *I'm* is the strong subjectivity of speech, which makes first person singular the most frequent pronoun, and the fact that *am* cannot occur with any other subject (Scheibman 2002). Bybee and Scheibman (1999) study the reduction of *don't* in American English and find that it reduces most in the contexts in which it is used the most, that is, after *I* and before certain verbs such as *know*, *think*, *mean*, *care*, *feel*, and others.

While it is not necessarily true in the cases just cited, it is usually the case that high-frequency grammaticizing items that are used together are in the same constituent in the traditional sense. For example, a determiner will tend to show reduction and fusion that depends upon the noun with which it is in construction. Thus the English article *the* has two pronunciations, [ði] and [ðə], as does *an/a*. The second variant in each case developed as a reduction before the consonant in the noun or adjective of the same noun phrase. The French masculine definite article *le* loses its vowel before a vowel-initial noun: *le* + *ami* > *l'ami* 'the friend', and so on.

Postpositions fuse with the nouns they operate on to become case suffixes. Thus the Turkish word *değin* meaning 'as far as' has fused with the preceding noun to make a new case suffix, as in *köydek* 'as far as the village' (Kahr 1976). Auxiliaries that are part of tense and aspect constructions and that follow the verb fuse with the verb to become suffixes, as in the famous case in the Romance languages, where the infinitive + *habere* constructions yields the future tense: Old Spanish *cantar ha* becomes Modern Spanish *cantará* 'he will sing'. Examples such as these are abundant in the grammaticization literature.

7. Constituency with less frequent combinations

The examples just cited are the extreme cases, since they involve combinations of extremely high frequency. But more subtle effects can also be found in cases of co-occurrence that are less frequent, leading me to hypothesize that chunking and constituency relate directly to frequency of co-occurrence.

Gregory et al. (1999) examine three measures of shortening of content words later in a large corpus of conversation (the Switchboard corpus, see later). They find that the tapping of a word-final /t/ or /d/ is highly affected by the probabilistic variable "mutual information," which is a measure of the likelihood that two words will occur together. If the final /t/ or /d/ occurs before a word that is highly likely given the first word, the rate of tapping increases. For instance, tapping would be more common in the word pair *lot of* than in *out alone*. They write: "The effect of mutual information on tapping suggests that tapping is a process that may also preferentially apply internally to highly cohesive pairs" (Gregory et al. 1999:9). Word-final deletion of /t/ and /d/ and the duration of a word are also highly affected by mutual information as well as other measures of probability. Jurafsky et al. (2001) find similar effects of probability on function words.

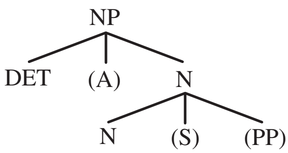
The usual interpretation of small phonological adjustments that occur between words is that they are due to on-line phonetic processes that apply as the words are strung together. However, if this is the case, it is difficult to explain why frequency of co-occurrence affects the probability that a phonological adjustment would take place. Part of the explanation is that the speaker, knowing that the next word has a high probability of occurring given the current word, allows reduction processes to occur. But for the speaker to know the probability of the occurrence of the next word, she or he must have a mental representation that includes knowledge about what words have occurred together in the past. Thus the occurrence of two words together strengthens the sequential link between them. It is this sequential link that is the local basis for the eventual emergence of constituent structure. That is, pairs of words with strong sequential links are regarded as occurring in the same constituent.

8. Sequentiality and constituency in the NP

To test the hypothesis that traditional notions of constituent structure correspond rather directly to frequency of co-occurrence, I elected to examine the English noun phrase in conversation. The noun phrase (especially one with a lexical noun) is a very good example of a constituent in English, as it has the same structure whether it is the subject or object of the verb or object of the preposition. Further, it is sometimes used independently in conversation (Ono and Thompson 1994).[3]

The basic structure of an English noun phrase was represented with the phrase structure tree in (1) in early models of generative grammar:

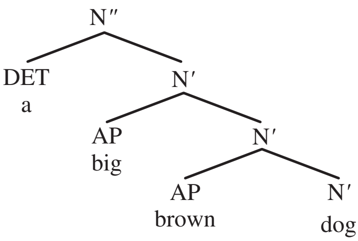
(1) NP structure:



The DET (determiner) and N are obligatorily present (although the DET can be a zero) and the other items are optional. The NP can contain an adjective that precedes the N, or a relative clause (S), or a prepositional phrase (PP), which follows.

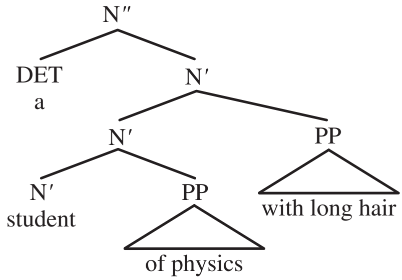
Another more recent representation of the noun phrase in X-bar notation provides an extra layer of constituency for each element added to the noun phrase (see Radford 1988). Thus (2) shows a noun phrase with two adjectives and (3) a noun phrase with postmodifiers:

(2)



What predictions are made about levels of constituency by these phrase structure trees? In (1) the DET (A) N are all under the same node, NP, so they will not be separated by any constituent boundaries except those that indicate that they are separate words. In (2) each adjective increases the distance between the determiner and the noun.

(3)



With post-modifiers in (1) or (3), there are two consequences for the constituent structure in the normal interpretation: first, the N will be separated from the DET (A) by an extra bracket, suggesting, perhaps erroneously, that an N with post-modifiers is not in the same constituent with the DET as the N is when it is final to the NP. Second, the post-modifier, being a constituent itself (as an S or PP), is separated from the N by several brackets, indicating a looser constituent structure.

While the second consequence seems fine, the first presents an odd prediction that a DET N combination has tighter constituent structure when no modifiers follow it then when they do. However, other mismatches between the bracketing generated by phrase structure rules and the surface constituency have been identified in the literature and a special component of rules, Readjustment Rules, was set up to fix these mismatches (Chomsky and Halle 1968). In addition, phonological phrasing or the domain of phonological rules is sometimes at odds with the generated phrase structure and needs adjusting. Thus it is generally clear that the structures predicted by trees such as (1), (2), and (3) cannot be taken in a completely literal way. However, certain elements of the general structure presented in these trees are well motivated, that is, that the NP is a constituent, and that these elements, DET, pre-modifying adjectives, and post-modifying PPs and Ss belong in the NP constituent. Thus I will compare this general structure to the quantitative facts of co-occurrence to determine whether or not generalizations about constituent structure can be said to be derived from multiple patterns of language use.

My hypothesis is that specific items that are used together frequently will form tighter bonds than items that occur together less often. The prediction is that items that occur within a traditional constituent are more likely to occur together in running discourse than are items that are in different constituents. One complication is the existence of optional items in a constituent. Strictly speaking, my hypothesis would predict that optional elements have weaker sequential links than obligatory elements and thus looser constituency. While this prediction may be correct, I will not pursue it further here. Instead, I will focus on testing, in a general way, the predictions about co-occurrence made by the noun phrase constituent. The question will simply be whether or not nouns occur more frequently with items inside or outside the noun phrase.

TABLE 15.1. Occurrence of Pauses before and after Eleven Nouns from the Switchboard Corpus (N = 7870)

Before the Noun	*After the Noun*
74 (1%)	2696 (34%)

To study the English noun phrase, I selected eleven lexical nouns from the most frequently occurring nouns in the Switchboard corpus. This corpus consists of over 2.4 million words from recorded telephone conversations. The nouns used were: *husband*, *mother*, *computer*, *movie*, *school*, *car*, *house*, *money*, *idea*, *class*, *problem*.[4] These nouns were studied only in their singular forms. The goal of the study was to discover which items most frequently preceded and most frequently followed these nouns in the conversations and to compare the frequency of preceding versus following items to see if the quantitative distributional facts correspond to traditional notions of constituency in the noun phrase.

First consider where pauses occur in relation to these eleven nouns. Table 15.1 shows that the probability of a pause following the noun is much greater than that of a pause preceding the noun. In fact, more than one-third of the tokens of these nouns were followed by a pause, while fewer than 1% were preceded by a pause.[5] This suggests a much weaker constituency bond to the item following the noun, either because it is in another constituent or because it is a more loosely joined part of the noun phrase. The reason for the low probability of a pause before the noun is presumably the obligatory presence of a determiner before the noun.

Table 15.2 shows the three most common linguistic items to precede and follow the noun. What is notable here is the greater predictability of the item preceding the noun compared to the one following it. In the Preceding column, *the* accounts for 17% of all tokens preceding the noun. Following the noun, *and* is the most common unit, but it only occurs after 7% of the tokens. Thus the co-occurrence patterns for

TABLE 15.2. The Three Most Frequent Linguistic Items Preceding and Following the Eleven Nouns

	Occurrences	
	Number	*%*
Preceding item		
the	1,348	17%
my	958	12%
a	767	10%
Following item		
and	562	7%
that	345	4%
is	270	3%

TABLE 15.3. Sets of Items Preceding the Eleven Nouns Studied

	Item	*Number of Occurrences before Nouns*	*Total for Category*
Articles	*the*	134	3,703
	a/an	767	
Possessives	*my*	958	
	your	147	
	our	127	
	his	48	
	their	31	
	's	19	
Other determiners	*that*	137	
	any	53	
	no	76	
	one	11	
Prepositions	*of*	260	592*
	to	178	
	in	154	
Adjectives	*good*	152	248
	new	80	
	whole	16	
Conjunction	*and*	6	6

X + N are stronger than for N + X, reflecting tighter constituency.[6] In terms of understanding the constituent structure of the noun phrase, it is instructive to view some of the most frequent items to precede and follow these eleven nouns. The items in table 15.3 were selected by first counting the ten most frequent items to precede each noun and then selecting all the items that occurred in more than one list. Table 15.4 was constructed in the same way with attention to the items that followed the noun. Tables 15.3 and 15.4 are organized by grammatical function. The totals in the last column are followed by asterisks if they present apparent counterexamples to the hypothesis that elements following or preceding the noun are in the same constituent. These cases are discussed later.

Tables 15.3 and 15.4 confirm that the constituency of the noun phrase as traditionally represented corresponds closely to the sequential co-occurrence patterns found in continuous speech. A large majority of the preceding and following items are part of the noun phrase.

The position preceding the noun most frequently has a member of the class of determiners, including the high-frequency articles, possessive pronouns (especially *my*), and quantifiers or demonstratives (a total of 47%). A few high-frequency adjectives represent the adjective position in the noun phrase. Finally, some items that are not in the noun phrase made the list: the three prepositions *of*, *to*, and *in* and the conjunction *and* with only six occurrences. A closer examination of the preposition + noun examples reveals that they occurred in particular phrases. *Of* occurred primarily with the mass noun *money* and was part of a quantifier phrase, for example,

TABLE 15.4. Sets of Items Following the Eleven Nouns Studied

	Item	*Number of Occurrences before Nouns*	*Total for Category*
Prepositions	*to*	120	445
	for	117	
	in	112	
	of	81	
	at	15	
Verbs	*is*	270	444*
	was	79	
	does	32	
	has	32	
	had	31	
Relative clause	*that*	345	409
	I	64	
N/NP Conjunction	*and*	342	342
S/VP Conjunction	*and*	286	286*
Possessive	*'s*	145	145

lot of occurred seventy-seven times, *amount*(*s*) *of* occurred forty times, and so on. Sinclair (1991:85) argues that in such phrases we do not have a head noun (*lot*) followed by prepositional phrase, but rather the second noun is the head of the phrase, and *lot of* or comparable phrases are modifiers to that noun. Under that analysis, which seems realistic for this data, *of* is not really a preposition and the phrase *lot of* is in the same constituent as the noun it modifies.

The preposition *to* occurs with *school* in 170 out of 178 examples and with *class* in the other 8 examples. *In* occurs with *school* in 139 out of 154 cases and with *class* and *computer* (e.g., *computer class*) in the other 15 cases. The usage suggests that *to/in school* and *in class* are locative adverbial rather than phrases involving full noun phrases. After all, it is quite idiosyncratic that these nouns lack a determiner in these expressions.

Even given that the prepositions *in* and *to* pose counterevidence to my hypothesis, 93% of the items preceding the nouns investigated are in the same noun phrase.

The position following the noun has no one category or item that is as strong as the articles or possessive pronouns in the preceding position. Prepositions, verbs, and relative clauses follow the noun with approximately equal frequency. Prepositions are expected in this position and are considered part of the noun phrase. Items that initiate relative clauses are also expected in this position, for example, *that*. *I* turns out to be common at the beginning of a relative clause (as in *the movie I saw*, *the class I'm taking*). While the relative clause is in the noun phrase, it also begins a new clause, so that it has a weaker constituent bond with the noun than any items that precede the noun. The rate of occurrence among these high-frequency classes—relative clauses, prepositional phrases, and high-frequency verbs—is about the same, even though the verbs are not in the same constituent with the noun phrase. This distribution suggests that the constituency bond between the noun and the modifying elements following the noun is weaker than with items preceding the noun. This relation is represented by additional constituency boundaries, as the relative clause would have a new S constituent and the prepositional phrase itself is a constituent.

The fact that the frequency with which the noun is followed by a verb, the first element in the other major constituent of the sentence, is comparable to that of the relative clause and prepositional phrase, which are part of the noun phrase, seems problematic for the hypothesis that frequency of co-occurrence corresponds to constituency. In particular, the high frequency of *is* would seem to be counter-evidence to this hypothesis. However, *is* and in some cases *has* and *had* are subject to contraction with the subject. Such cases of fusion where frequently co-occurring items are not in the same traditional constituents are the evidence I will use in sections 12 and 13 to argue that sequential co-occurrence is more basic than constituency.

None of the descriptions of the noun phrases consulted before examining the conversational data prepares us for the high frequency with which the conjunction *and* follows a noun. In table 15.4 I report a rough breakdown into those instances of *and* that conjoin another noun and those that introduce another clause or verb phrase. This breakdown was based on the item following *and*. Clause-conjunction was assumed if the following item was a conjunction, *then* or *so*, or an adverb. When a pronoun followed that was in the nominative case, the type of conjunction was assumed to be N conjunction if the preceding noun was animate (*husband and I*) but clause conjunction if the preceding noun was inanimate (*computer and we*). A following verb indicated verb or verb phrase conjunction. All other cases could be reasonably assumed to be noun conjunction. In a small number of cases these assumptions may have led to an erroneous assignment, but this breakdown allows us to get an idea of how many of these *and*s introduce a major constituent break.

Overall the data suggest the predicted correspondence between sequential co-occurrence and traditional notions of constituency within the noun phrase. They predict a stronger bond with the item preceding the noun than with the item following it, and most of the items found to precede or follow the noun are in the same traditional constituent. The primary counter-examples involve preceding prepositions, which seem to be restricted to certain high-frequency phrases, and common verbs or auxiliaries following the noun.

In the next two sections, I discuss the cognitive mechanisms that underlie this general correspondence, and then in the following sections I proceed to the argument that the hierarchical structure of language is derivable from the more basic sequential nature of language.

9. Fragments and networks

Language learning in a natural setting involves the storage and processing of repeated fragments of speech. Such fragments might include parts of words, words, or multiple words. Since human beings are sensitive to recurring sequences of stimuli and record them in memory, they learn to recognize repeated sequences of speech or speechlike stimuli (Saffran, Aslin, and Newport 1996; Gomez and Gerken 1999). Similarly the automated neuromotor sequences that correspond to these perceptual units are recorded in memory and strengthened with use. Since the number of fragments of speech that must be stored in memory for a mature user of a language is extremely large, a tight organization of these fragments is necessary. I suggest for multi-word sequences, as I have for morphologically complex words, that whole complex units may be stored in memory, but they are typically associated with other units that are similar phonologically and semantically (Bybee 1985, 1995, 1998a). Figure 15.1 shows a simplified, schematic representation of a possible organization for some of the noun phrases I discussed earlier. The connecting lines between words indicate a relation of phonological and semantic identity.[7] Figure 15.2 shows a set of relations among noun phrases centering on the determiners rather than the noun. The networks in figures 15.1 and 15.2 interlock.

In this model representation is affected by language use. Each token of use of a word or sequence of words strengthens its representation and makes it more easily accessed. In addition, each instance of use further automates and increases the fluency of the sequence, leading to fusion of the units. Thus, as in morphology, high-frequency combinations have a stronger representation. For instance, *my mother* (which occurred 182 times in the Switchboard corpus) has a stronger representation than *her house* (which occurred 22 times) or *this hospital* (which occurred only once). All of these noun phrases are interlocked in a network based on experience with language. From this network emerges the fact that any of these nouns, for example, *mother*, *house*, or *hospital*, can occur with any of the set of determiners. Since particular combinations are also represented, we do not lose the information that some of these nouns occur more often with certain determiners (*my* with kinship terms or *an* with *idea*) than others.

The hypothesis that very particular sequences of words from our experience are stored in lexical memory is supported by the findings of Erman and Warren (2000)

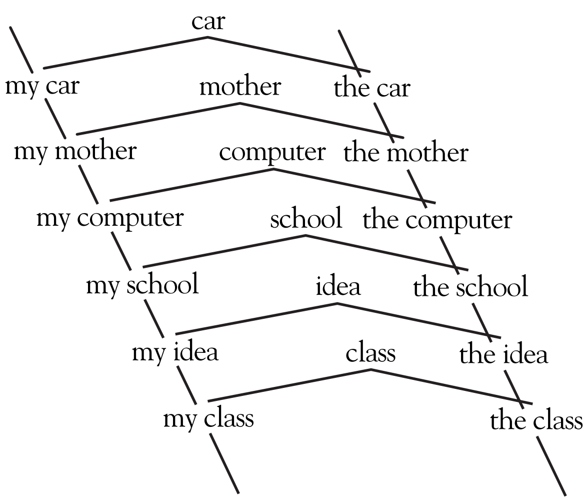


FIGURE 15.1. Possible organization of relations among NPs centering on frequently occurring nouns.

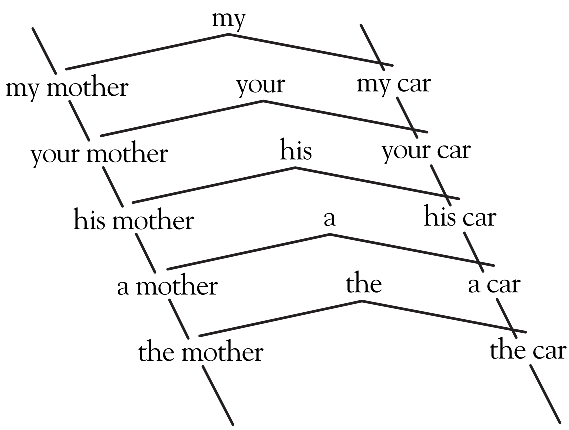


FIGURE 15.2. Possible organization of relations among NPs centering on determiners.

that in the spoken and written texts they examined, 55% of the texts consisted of prefabricated units, that is, sequences of words that are conventionalized. Prefabricated sequences occur more frequently and have more idiomatically determined meaning than sequences that are put together afresh each time. In other words, prefabricated sequences have been experienced before. Pawley and Syder (1983) point out that nativelike use of a language depends heavily upon knowing and using conventionalized word combinations.

10. Schemas emerge from the network

The organization apparent in figures 15.1 and 15.2 is only possible if categories are established for strings or parts of strings. A relation between *my car* and *your car* is possible only if a category is established for *car*. Such a category is based on the similarity between tokens of use, which includes information not just about phonology and semantics but also about contexts of use, both linguistic and non-linguistic. The linguistic context in which *car* is used will include information about what immediately precedes and follows it. Since, as we have seen, the preceding elements are more predictable, that is, the same ones occur more often, a category that includes the frequently occurring preceding element could be formed. The storage of multiword strings and the categorization of their component elements leads to the formation of constructions. Constructions are schematic generalizations over sequences with shared parts. I agree with Langacker (1987) that schemas are formed at various levels of abstraction. For instance, for NP in English, we might find the following levels of abstraction:

1. Very specific: *my mother*, *my computer*, *the car*, *a problem*, *an idea*
2. Partially general: [*my* + NOUN], [POSS PRO + *mother*]
3. More general: [POSSESSIVE + NOUN]
4. Fully general: [DETERMINER + NOUN]

These levels of abstraction involve categorization at differing levels of generality. *Mother* is a collection of all the exemplars of this word encountered so far. POSSESSIVE PRONOUN is a level of categorization that ranges over specific words such as *my*, *your*, *our*, *his*, and so on. NOUN obviously has an even broader range. Grammatical constructions arise, then, from the storage of frequently repeated sequences and the categorization of their parts at different levels of abstraction.

11. Sequentiality is more basic than hierarchy

In the previous sections I hope to have demonstrated that there is a correspondence between frequency of co-occurrence and traditionally established notions of constituent structure. Now the question arises as to the directionality of that correspondence: do elements occur together frequently because they are in the same constituents or are constituents derived from clusters of items that occur together frequently? In this section I argue that sequentiality is basic and that constituents and hierarchies arise because frequently used strings are chunked as single units.

First, observe that learning of sequences of behavior is part of implicit learning in other domains and that such sequences naturally fall into hierarchies with the more frequently used sequences recombined into larger units to form higher level constituents (Fentress 1983). Driving a car involves a number of automatic sequences such as shifting into reverse or drive, braking to slow down or stop, putting on a turn signal, turning right or left. On a frequently driven path, such as from one's home to one's office, these chunks of behavior are sequenced in a particular way and constitute a larger constituent that itself can be automated, as evidenced by the mistake of following this frequent path even when the destination is elsewhere. Also, any of the subunits of the whole sequence, such as braking, accelerating, turning left or right, can be recombined infinitely so that one can drive anywhere one wants to go. For the experienced driver, such recombinations occur with great ease and fluency, just as the native speaker recombines the automated chunks of language to produce a sentence s/he has never produced before. Thus the creation of hierarchy out of automated sequences of behavior is a domain-general cognitive process.

It is by now well established that Broca's area deals not just with language but also with sequential neuromotor behavior (Roland 1985; Greenfield 1991). In addition, many researchers have stressed the importance of left-hemisphere dominance for both motor control and language (Kimura 1979, 1993; Corballis 1989; Armstrong, Stokoe, and Wilcox 1995). Thus the hierarchy in automated motor activities and the hierarchy in grammar could stem from the same neurological source. Moreover, recent research into the perceptual processing of predictable visual stimuli suggests that Wernicke's area processes predictable events in time and may not be exclusively associated with language (Bischoff-Grethe et al. 2000). Since predictability is the perceptual side of sequentiality, it may turn out that the mechanisms behind the ability to perceive linguistic sequences and perhaps group them into constituents may also be domain-general.

Humans from 12 months to adulthood can learn repeated sequences of meaningless syllables, as shown by Saffran, Aslin, and Newport (1996) and Gomez and

Gerken (1999, 2000). Moreover, Gomez has recently shown that both babies and adults can learn sequences of two nonce words that are separated by a third "word" chosen from a large class (2001). Thus meaning is not necessarily involved in learning sequences, suggesting that the basis for constituent structure may be recurring sequences and not just semantics. (See also Santelmann and Jusczyk 1998.)

In addition, there is purely linguistic evidence for the dominance of sequentiality over hierarchy in cases where fusion between elements in different constituents occurs because the two elements occur frequently in sequence. Examples of such cases will be discussed in the next two sections (sections 12 and 13). Linguistic cases of the opposite type, that is, those that appear to show hierarchy dominating sequentiality, as when constituents are discontinuous, are discussed in section 14, where it is argued that the linear constraints on discontinuous constituents, such as the Heavy NP Constraint and the Distance Principle, demonstrate that sequential organization underlies even discontinuous constituents.

12. Chunking in violation of "constituent structure"

Linguistic units chunk together according to repeated sequences, sometimes in violation of usual notions of constituent structure (Bybee and Scheibman 1999). A very robust example is English auxiliary contraction, which occurs in *I'm*, *I've*, *I'd*, *I'll*, *he's*, *he'll*, *he'd*, and elsewhere. The auxiliary which is the first element in the verb phrase, contracts with the subject noun phrase, which is in the other major constituent of the clause. Thus contraction occurs across the primary constituent boundary of the clause. The reason for this is that the pronoun + auxiliary or noun phrase + auxiliary combinations are highly frequent. In a large corpus of British English, Krug (1998) finds that contraction occurs in the most frequently occurring combinations. The most common contraction is between *I* and *am*, and that is also the most frequent sequence.

But the auxiliary also occurs right before the verb, so why doesn't the auxiliary fuse with the following verb, where it belongs semantically and syntactically? In order to answer this question I examined the distribution of auxiliaries in spoken American English, using the Switchboard corpus again. In table 15.5 we see the token count of the ten most frequent items to precede and follow the auxiliary *will*, in its full and reduced version, *'ll*. What we see is that, indeed, the most frequent items to precede *will* or *'ll* are pronouns and the most frequent items to follow are verbs. What governs the contraction is the asymmetry in the frequency of the preceding versus the following items. Note that the most frequent pronoun (*I*) preceding *will* is twice as frequent as the most frequent verb (*be*) following *will*. A similar pattern is found for all contracted auxiliaries in the Switchboard corpus. Not only are the fused items in different major constituents, but also they have no semantic relevance to one another. The fusion seems due entirely to frequency of co-occurrence.

Contraction is recorded in Switchboard for all the items listed here as preceding *will* except for *people*. The only other preceding items showing contraction are *there*, *this*, and *who*. In other words, contraction only. occurs between the most frequent combinations in the case of *will*. Apparently contraction can generalize from the most

TABLE 15.5. Ten Most Frequent Items Occurring before and after *Will* and *'ll* (Switchboard Corpus)

	Number
Preceding	
I	918
they	471
we	368
it	256
you	200
that	183
he	122
she	53
,	47
people	38
Following	
be	466
,	244
have	199
get	130
go	119
do	103
probably	90
just	81
tell	75
.	42

frequent items to a general class of items, as demonstrated by the contraction of *has* to *'s*, which appears in the data with a large class of full noun phrases.

Other details of the distribution of auxiliaries support the Linear Fusion Hypothesis. For all the modal auxiliaries except *can* and *can't*, the most frequent items to follow are *be*, *have* and the negative.

1. When *have* is part of the Perfect construction, it contracts with the preceding modal, as in *could've*, *should've*, *would've*, *must've*, *might've*.
2. Of course the negative *not* also contracts in some cases, for example, *couldn't*, *shouldn't*, *wouldn't*.
3. Interestingly, *be* also forms a unit with some of these modals, but of a different sort. The combinations *could be* and *maybe* become lexicalized units that serve as independent epistemics, while *would be* has become an adjective.

Thus we have ample evidence from the auxiliaries that high-frequency sequences become chunked into units.

The observation that the combination of subject pronouns or nouns plus the auxiliary in English behaves like a constituent has been made before. Halliday and

Hasan (1976:197) call this complex the Modal Element and distinguish it from the Propositional Element (roughly the remainder of the clause). Halliday (1985) regards this collocation of units to be a constituent in the structure imposed on the clause by the interpersonal metafunction where again the Modal Element (subject + finite element of verb group) is distinguished from the Propositional Element. This level of constituent analysis coincides, then, with the frequency of co-occurrence found in conversational data and evidence from contraction that we have just considered.

The distribution of pronouns and auxiliaries and verbs, along with the Linear Fusion Hypothesis, explain why English auxiliaries contract with a preceding pronoun (and, in some cases, nouns) rather than becoming prefixes on the verb. If similar distributions occur in other languages it may explain why in languages where the auxiliary follows the verb (SOV languages) there is massive suffixation, while in languages where the auxiliary precedes the verb (VO languages) there is not a comparable trend toward prefixation (Bybee, Pagliuca, and Perkins 1990).

13. Other cases of non-constituents fusing

Another robust phenomenon demonstrable across a wide range of languages is the binding of a verb and preposition into a unit. Reh (1986) discusses this phenomenon in African languages as an explanation for why case affixes are almost always suffixes, that is, why prepositions do not tend to become case prefixes.

Reh (1986) points out that in several languages of the Southern Lwo family, a preposition following a verb becomes a suffix on the verb. This can be seen by examining the dative/benefactive preposition in Dhuluo, which is *nɪ* as shown in examples (4) and (5), where it is positioned between the verb and the noun phrase that is its object:

Dhuluo (Southern Lwo; Western Nilotic)

(4) Otieno o-kelo nɪ Odhiambo kitabu.
O. PERF-bring DAT/BEN O. book
'Otieno has brought a book to Odhiambd'.

(5) Onyangɔ tiyo nɪ japuonj.
O. IMPF:work DAT/BEN teacher
'Onyango works for the teacher'.

In Lango, the cognate preposition and the pronoun that was its object have fused with the verb to form "the benefactive stem." Thus *-kèlò* 'bring' + *nɪ* > *-kèllì* 'bring for someone', as shown in (6):

(6) Lango
Ò-kèllì dàkô.
3SG-bring-BEN woman
'She brought it for the woman'.

In this case, then, the former preposition has fused with the preceding verb rather than with the following noun. I propose that the explanation for this is that that particular prepositions would tend to occur with certain verbs, such as verbs meaning ‘bring’ or ‘give’, while the noun that follows is much less predictable, presumably being drawn from the entire class of human nouns and perhaps some non-human ones as well.

Analogous situations, but usually without affixation, can be found in European languages. Second-language learners of Spanish and French must learn lists of verb + preposition combinations, as particular verbs select particular prepositions. For instance, Spanish *pensar en* ‘to think about’, *acabar de* ‘to finish’, and *comenzar a* ‘to begin to’. Again, the verb + preposition sequence would be more frequent than any particular preposition + noun or preposition + infinitive sequence.

Another common fusion across constituent boundaries is the fusion of prepositions and determiners, European languages, for example, Spanish and French. For instance Spanish: *a* ‘to, at’ and *el* ‘the (MASC.SG) > *al*, *de* ‘of, from’ + *el* > *del*. In this case, as in the others, it is plausible to assume that the frequent co-occurrence of these grammatical items leads to their fusion. Note that there is no particular semantic relevance (in the sense of Bybee 1985) or semantic affinity between the meaning of a preposition and that of a determiner. This appears to be a case of pure sequentiality.

14. Discontinuous constituents and discontinuous dependencies

A major argument that the utterances of a language are not just linear strings of words is the fact that non-adjacent elements can be in the same construction or constituent. Thus in the English verb + particle combinations, the particle can be separated from its verb by a pronoun or short noun phrase, as in *look the number up*; *look it up*. However, the separation of the constituents does not necessarily mean that their connection is not still linear or sequential. Other types of neuromotor behavior can be suspended at certain points and then resumed with the association still being sequential. Furthermore, the predictability of the second element from the first can still be maintained across intervening items, as when the phrase *look the number* leads to the expectation of hearing *up*. This is analogous to waiting for the other shoe to drop.

A purely linguistic argument for the importance of sequentiality even in these cases is the well-documented existence of constraints on the material intervening between the two parts of the constituent. The Heavy NP Constraint describes the fact that, for instance, a noun phrase with a lot of modifiers does not do well between the verb and its particle. Chen (1986) shows that the separation of the verb and the particle is constrained by the number of syllables in the direct object. Separation of the verb and particle is practically non-existent in both spoken and written language for direct objects of more than five syllables. Thus (7) and (8) would be very rare or non-existent in discourse and have a very awkward feel:

(7) I need to look the number that I lost up.

(8) I need to look a word that I can’t remember how to spell up.

Other kinds of discontinuous dependencies also rely on sequentiality. For example, in French certain main clause verbs can have subjunctive verbs in their subordinate clauses. In Canadian French, the use of subjunctive or indicative is variable and apparently is not meaningful. Poplack (1992) has studied the variables that affect the mood choice and finds that certain high-frequency main clause verbs in combination with certain high-frequency subordinate clause verbs are more likely to have the subjunctive.

Examples (9) and (10) illustrate this variability. (*Faut* is a reduction of *il faut que*, which means 'it is necessary'.)

(9) Bien certain, faut qu'ils *aient* (*S*) une place eux-autres aussi pour vivre.
'Well, of course, they should have a place to live, too'.

(10) Faut j'*aille* (*S*) voir pour de l'ouvrage.
'I have to go look for a job'.

I would argue that such dependencies are basically sequential. In Poplack's analysis of Canadian French, it turns out that if a parenthetical expression intervenes between the main verb and the subordinate verb, the subjunctive verb form is much less likely to appear. In other words, intervening material can disrupt the dependency, suggesting that the use of subjunctive results from the selection of an automated chunk.

As mentioned earlier, infants are sensitive to sequential regularities in input even in the absence of semantic factors. Santelmann and Jusczyk (1998) found that 18-month-old infants can discriminate between grammatically correct English discontinuous constituents and ungrammatical ones. The construction in question was the progressive, in which the finite form of the verb *to be* forms a construction with the suffix *-ing* with an open class of verbs intervening. Thus the infants demonstrated Head Turn Preference for the natural passages, such as *Everyone is trying to be careful*, versus the unnatural one, *Everyone can trying to be careful*. When extra words were inserted between the discontinuous constituents, the infants no longer preferred the natural passages, suggesting that this is indeed a sequential effect that can be disrupted by intervening words.

Gomez (2001) also tested 18-month-olds using sequences of three nonce "words." The infants learned to discriminate sequences they had heard before even though other "words" from a large set always came between the first and third word. Since the nonce words are meaningless, this experiment demonstrates that sequential dependencies can be detected and learned in the absence of meaning.

The sequentiality hypothesis makes general predictions about center embedding, for example, that center embedding would be constrained by the length of the embedded unit. To my knowledge studies of the length of center-embedded clauses in natural discourse have not been undertaken.

15. Conclusion

Linguists have been accustomed to viewing language as emanating from a mental structure that is autonomous from actual usage events. A more explanatory view is afforded by recent functionalist views of language as highly contextualized and embodied (Fox 2001). Most tokens of language use are routine events that respond to the environment—both social and physical, that is, the people encountered and the utterances they produce. These responses are partially automatic, though they do involve an assessment of the environment and the choice of an appropriate response, as do other fine-tuned neuromotor behaviors. As with other neuromotor skills, language responds to practice. Perceptual skill also improves with repetition. Thus we have every reason to believe that repetition could be the main factor that builds up structure in language (Haiman 1994). Its importance in grammaticization has been well documented. Here I suggest that repetition of sequences of units is the main factor in the creation of linguistic patterns that have been identified as constituent structure.

My proposal is as follows. Constituents of the type proposed for generative grammar that are described by phrase structure trees do not exist. Instead, units of language (words or morphemes) are combined into chunks as a result of frequent repetition. Most of the time the units of these chunks bear a semantic and/or pragmatic relation to one another allowing them to fulfill the grammatical criteria for constituency: they can sensibly be used alone, as in the answers to questions; they can be replaced by a pro-form; and they can be used in various positions in a sentence (see the examples in section 2). In such cases, where frequency of co-occurrence corresponds to semantic relevance, we have traditional constituents. Indeed, the semantic coherence of such units may facilitate their establishment as chunks. However, other types of chunks also exist, as I have demonstrated in this chapter, showing that frequency of co-occurrence is an independent factor. Thus pronoun + auxiliary, preposition + determiner, and verb + preposition sequences can form chunks but are difficult to describe in traditional frameworks since they do not meet the criterion of semantic relevance. For this reason, too, they do not fulfill the grammatical criteria of occurring alone or being replaceable by a pro-form. Thus constituency in this view is the convergence of two other factors and is itself not a basic structure. It is an emergent property of language.

A second point is that this emergent constituency differs from traditional constituency in that it can be gradient, since the factors determining it are themselves gradient. *Gradience in constituency* refers to the fact that different items of the same putative category might fuse less with one another. If frequency of co-occurrence is a major determinant of emergent constituency, then the two units in *my mother* are in a tighter constituent bond than the two units in *my appointment*. This difference has no overt consequences in English, but in some languages it would be manifest as a difference between alienable and inalienable possession, where the latter always has a more fused expression (Nichols 1988). Other gradient differences in frequency of co-occurrence do have overt consequences, as seen in the fact that *I don't* is more fused than *they don't* and *hit'im* is more fused than *hit the ball*. These relations cannot be captured by using the same phrase structure tree for all instances of the same traditional constituent.

Considering now the evolution of language, the development of grammar may be easier to explain in terms of domain-general abilities than many researchers have supposed (e.g., Pinker and Bloom 1990). If constituent structure is epiphenomenal, then a theory of the evolution of language need not account for it directly, but rather by reference to the processes that underlie it and these appear to be domain-general. The abilities include (1) highly advanced motor skills, fine motor control, and the associated neurological capabilities including the ability to compose, store, and access relatively long sequences of behavior; (2) the ability to combine concepts into communicatively coherent sequences, which in turn is based on (3) extensive categorization capacity that is applied to both phonological form and meaning; and (4) the ability to store and categorize vast quantities of prefabricated sequences. These abilities interact in that one may facilitate the others. In particular, the semantic coherence of units in sequence may make it possible to compose longer sequences more fluently. And, as I emphasized throughout, the automation of lower level sequences makes the composition of hierarchically complex sequences possible. Thus abilities that are neither domain-specific nor species-specific interact in current language processing to create the apparent structure that is grammar. As these abilities evolved from a more primitive to a more advanced stage, language might also have evolved from a set of relatively short utterances consisting of first one, then two units, to much longer utterances with apparent hierarchical structure via the concatenation of preformed chunks.

Notes

1. Langacker (1997), which is a discussion of the iconic relations between conceptual structure and syntactic structure, is an exception.

2. Interestingly, the best examples of syntactic constituents in English seem to be noun phrases. Noun phrases also have the best support as constituents in the discourse literature (Ono and Thompson 1994). Verbs and verb phrases and prepositional phrases present certain problems, as we will see below.

3. On the difficulties of identifying the verb phrase as a constituent, see Givón (1995).

4. In selecting these nouns, there was a conscious bias toward count nouns. In addition, it was immediately obvious that kinship terms, such as *mother* and *husband*, had the special property of occurring more with possessive pronouns, and thus only two of them were included. Bleached or grammmaticized nouns such as *thing*, *couple*, and *stuff* were also passed over.

5. The Switchboard transcriptions distinguish several types of pauses, but I have added them all together for the purposes of table 15.1.

6. One might expect a higher percentage for *the* as the item to precede a noun. Table 15.3 shows that there are quite a number of frequently occurring determiners competing with *the*. In addition, some nouns frequently occur in compounds, so that the item preceding them is another noun. For instance, the item to most frequently precede *school* is *high*.

7. In the model developed in Bybee (1985), semantic and phonological connections can be made independently, but when they coincide a morphological relation is indicated.

References

Anderson, J. R. 1993. *Rules of the mind*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Armstrong, D., W. Stokoe, and S. Wilcox. 1995. *Gesture and the nature of language*. Cambridge: Cambridge University Press.

Bischoff-Grethe, A., S. M. Proper, H. Mao, K. A. Daniel, and G. Berns. 2000. Conscious and unconscious processing of nonverbal predictability in Wernicke's area. *Journal of Neuroscience* 20, 1975–1981.

Boyland, Joyce Tang. 1996. Morphosyntactic change in progress: A psycholinguistic approach. Ph.D. diss., University of California at Berkeley.

Bybee, J. L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.

———. 1995. Regular morphology and the lexicon. *Language and Cognitive Processes* 10(5), 425–455. (Reprinted in this volume as chapter 8.)

———. 1998a. The emergent lexicon. *CLS* 31, no. 1: *The Panels,* 421–435. Chicago: Chicago Linguistic Society. (Reprinted in this volume as chapter 13.)

———. 1998b. A functionalist approach to grammar and its evolution. *Evolution of Communication* 2, 249–278.

———. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

Bybee, J. L., William Pagliuca, and Revere D. Perkins. 1990. On the asymmetries in the affixation of grammatical material. In *Studies in diachronic typology for Joseph H. Greenberg*, ed. W. Croft, K. Denning, and S. Kemmer, 1–42. Amsterdam: John Benjamins.

Bybee, J. L., William Pagliuca, and Revere D. Perkins. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, Joan L., and Joanne Scheibman. 1999. The effect of usage on degree of constituency: The reduction of *don't* in American English. *Linguistics* 37, 575–96.

Chen, P. 1986. Discourse and particle movement in English. *Studies in Language* 10(1), 79–95.

Chomsky, Noam, and Morris Halle. 1968. *The sound pattern of English*. New York: Harper and Row.

Cipollone, N., S. H. Keiser, and S. Vasishth (eds). 1998. *Language files: Materials for an introduction to language and linguistics*, 7th ed. Columbus: Ohio State University Press.

Corballis, M. C. 1989. Laterality and human evolution. *Psychological Review* 96, 492–505.

Erman, Britt, and Beatrice Warren. 2000. The idiom principle and the open choice principle. *Text* 20, 29–62.

Fentress, J. C. 1983. Hierarchical motor control. In *The psychobiology of 1anguage*, ed. M. Studdert-Kennedy, 40–61. Cambridge, Mass.: MIT Press.

Fox, B. 2001. On the embodied nature of grammar: Embodied being-in-the-world. In *Complex Sentences in Grammar and Discourse: Essays in Honor of Sandra A. Thompson*, ed. J. L. Bybee and M. Noonan. Amsterdam: Benjamins.

Givón, T. 1995. *Functionalism and grammar*. Amsterdam: Benjamins.

———. 1998. Toward a neurology of 1anguage. *Behavioral and Brain Sciences* 21, 154–155.

———. The visual information-processing system as an evolutionary precursor of human language. In *The evolution of language out of pre-language*, ed. T. Givón and B. Malk, 3–50. Amsterdam: John Benjamins.

Gomez, R. L. 2001. Finding structure in language: Sometimes more variability is better. Presentation given at the annual meeting of the AAAS in the symposium: Tools Infants Might Use to Learn Language, San Francisco, February.

Gomez, R., and L. Gerken. 1999. Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition* 70, 109–135.

——— . 2000. Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences* 4, 178–186.

Greenfield, P. M. 1991. Language, tools and brain: The ontogeny and phylogeny of hierarchically organized sequential behavior. *Behavioral and Brain Sciences* 14, 531–595.

Gregory, Michelle L., William D. Raymond, Alan Bell, Eric Fosler-Lussier, and Daniel Jurafsky. 1999. The effects of collocational strength and contextual predictability in lexical production. *CLS* 99, 151–166.

Haiman, John. 1994. Ritualization and the development of language. In *Perspectives on grammaticalization*, ed. William Pagliuca, 3–28. Amsterdam: John Benjamins.

Halliday, M. A. K. 1985. *An introduction to functional grammar*. London: Edward Arnold.

Halliday, M. A. K., and R. Hasan. 1976. *Cohesion in English.* London: Longman.

Hopper, P., and E. Traugott. 1993. *Grammaticalization.* Cambridge: Cambridge University Press.

Jurafsky, Daniel, Alan Bell, Michelle Gregory, and William D. Raymond. 2001. Probabilistic relations between words: Evidence from reduction in lexical production. In *Frequency and the emergence of linguistic structure*, ed. Joan L. Bybee and Paul Hopper, 229–254. Amsterdam: John Benjamins.

Kahr, J. C. 1976. The renewal of case morphology: Sources and constraints. *Working Papers on Language Universals* 20, 107–151. Stanford, Calif.: Stanford University Press.

Kimura, D. 1979. Neuromotor mechanisms in the evolution of human communication. In *Neurobiology of social communication in primates*, ed. H. D. Steklis and M. J. Raleigh, 197–219. San Diego, Calif.: Academic Press.

———. 1993. *Neuromotor mechanisms in human communication.* Oxford: Oxford University Press.

Krug, Manfred. 1998. String frequency: A cognitive motivating factor in coalescence, language processing and linguistic change. *Journal of English Linguistics* 26, 286–320.

Langacker, R. 1987. *Foundations of cognitive grammar*, vol.1: *Theoretical prerequisites.* Stanford, Calif.: Stanford University Press.

———. 1997. Constituency, dependency and conceptual grouping. *Cognitive Linguistics* 88, 1–32.

Lindblom, Björn, Peter MacNeilage, and Michael Studdert-Kennedy. 1984. Self-organizing processes and the explanation of phonological universals. In *Explanations for language universals*, ed. Brian Butterworth, Bernard Comrie, and Östen Dahl, 181–203. Berlin: Mouton de Gruyter.

Nichols, J. 1988. On alienable and inalienable possession. In *In honor of Mary Haas: From the Haas Festival Conference on Native American Linguistics*, ed. W. Shipley, 557–609. Berlin: Mouton de Gruyter.

Ono, Y., and S. Thompson. 1994. Unattached NPs in English conversation. *BLS* 20.

Pawley, Andrew, and Frances Hodgetts Syder. 1983. Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In *Language and communication*, ed. Jack C. Richards and Richard W. Schmidt, 191–226. London: Longman.

Pinker, S., and P. Bloom. 1990. Natural language and natural selection. *Brain and Behavioral Sciences* 13, 707–784.

Poplack, S. 1992. The inherent variability of the French subjunctive. In *Theoretical Analyses in Romance Linguistics*, ed. C. Laeufer and T. A. Morgan, 235–263.

Radford, A. 1988. *Transformational grammar: A first course.* Cambridge: Cambridge University Press.

Reh, M. 1986. Where have all the prefixes gone? *Africanistische Arbeitspapiere* 5, 121–134.

Roland, P. E. 1985. Cortical organization of voluntary behavior in man. *Human Neurobiology* 4, 155–167.

Saffran, J. R., R. N. Aslin, and E. L. Newport. 1996. Statistical learning by 8-month-old infants. *Science* 274, 1926–1928.

Santelmann, L., and P. W. Jusczyk. 1998. Sensitivity to discontinuous dependencies in language learners: Evidence for limitations in processing space. *Cognition* 69, 105–134.

Scheibman, J. 2002. *Point of view and grammar: Structural patterns of subjectivity in American-English conversation.* Amsterdam: John Benjamins.

Sinclair, J. 1991. *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Sosa, A. V., and J. MacFarlane. 2002. Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word *of*. *Brain and Language* 83(2), 227–236.

Chapter 16

Mechanisms of Change in Grammaticization

The Role of Frequency

One of the most notable characteristics of grammatical morphemes (hereafter "grams"; see Bybee and Dahl 1989) and the constructions in which they occur is their extremely high text frequency as compared to typical lexical morphemes. Since grams commonly develop from lexical morphemes during the process of grammaticization, one striking feature of this process is a dramatic frequency increase. This increase comes about as a result of an increase in the number and types of contexts in which the gram is appropriate. Frequency is not just a result of grammaticization; it is also a primary contributor to the process, an active force in instigating the changes that occur in grammaticization. This chapter treats two topics: (1) the manner in which the extreme frequency increase occurs, which will be examined via a case study of *can* in Old and Middle English; and (2) those mechanisms of change associated with grammaticization that are attributable in some way to this dramatic frequency increase, including phonological, morphosyntactic, and semantic change. A third important theme of this chapter echoes that found in Traugott (2003): none of these changes can be studied except in the context of the construction in which the grammaticizing element occurs.

1. The grammaticization of constructions

The recent literature on grammaticization seems to agree that it is not enough to define grammaticization as the process by which a lexical item becomes a grammatical morpheme, but rather it is important to say that this process occurs in the context of a particular construction (see Heine 2003; Traugott 2003). In fact, it may be more accurate to say that a construction with particular lexical items in it becomes grammaticized, instead of saying that a lexical item becomes grammaticized. For instance, several movement verbs appropriately fit into the following constructional schema of English:

(1) [[movement verb + Progressive] + purpose clause (*to* + infinitive)]
E.g., *I am going to see the king*
I am traveling to see the king
I am riding to see the king

However, the only example of this construction that has grammaticized is the one with *go* in it. The particular example of this construction with *go* in it has undergone phonological, morphosyntactic, semantic, and pragmatic changes that have the effect of splitting the particular grammaticizing phrase off not only from other instances of *go* but also from other instances of this [movement verb + Progressive + purpose clause] construction.

2. The role of repetition

Also in the recent literature on grammaticization, we find extensive discussions of semantic change and its sources (see Traugott 1989; Heine, Claudi, and Hünnemeyer 1991; Bybee, Perkins, and Pagliuca 1994), but much less emphasis on the development of morphosyntactic and phonological properties of emerging grams. In an attempt to offer an integrated approach to the multiple changes that constitute grammaticization, I will focus in this chapter on the role that repetition plays in the various changes that a grammaticizing construction undergoes. The importance of repetition to grammaticization has been emphasized in Haiman's (1994) discussion of the parallels between the general cultural phenomenon of ritualization and the process of grammaticization in language and in Boyland's (1996) examination of the effects of repetition on the cognitive representation of grammaticizing constructions. Building on these works, I will argue for a new definition of grammaticization, one that recognizes the crucial role of repetition in grammaticization and characterizes it as the process by which a frequently used sequence of words or morphemes becomes automated as a single processing unit.

Haiman (1994) makes a case for regarding the process of grammaticization as ritualization, citing the following aspects of ritualization, all of which are the result of repetition: *habituation* results from repetition and depletes a cultural object or practice of its force and often its original significance as well; repetition leads to the *automatization* of a sequence of units and the reanalysis of the sequence as a single processing chunk, with formerly separate units losing their individual meaning; repetition also leads to the *reduction of form* through the weakening of the individual gestures comprising the act and through the reorganization of a series of formerly separate gestures into one automated unit; and *emancipation* occurs as the original, more instrumental function of the practice gives way to a more symbolic function inferred from the context in which it occurs.

Applying these aspects of ritualization to the grammaticization process in particular, I will argue that frequent repetition plays an important role in the following changes that take place in grammaticization:

1. Frequency of use leads to weakening of semantic force by habituation—the process by which an organism ceases to respond at the same level to a repeated stimulus (section 4).
2. Phonological changes of reduction and fusion of grammaticizing constructions are conditioned by their high frequency and their use in the portions of the utterance containing old or backgrounded information (section 5).
3. Increased frequency conditions a greater autonomy for a construction, which means that the individual components of the construction (such as *go*, *to*, or *-ing* in the *be going to* example of [1]) weaken or lose their association with other instances of the same item (as the phrase reduces to *gonna*) (section 6).
4. The loss of semantic transparency accompanying the rift between the components of the grammaticizing construction and their lexical congeners allows the use of the phrase in new contexts with new pragmatic associations, leading to semantic change (section 7).
5. Autonomy of a frequent phrase makes it more entrenched in the language and often conditions the preservation of otherwise obsolete morphosyntactic characteristics (section 8).

Before moving to an expanded discussion of each of these aspects of grammaticization, I will discuss the two ways of counting frequency in section 3 and demonstrate in section 4, with a case study of the development of *can* in English, how a grammaticizing construction increases its frequency.

3. Type and token frequency

Two methods of counting frequency are relevant for linguistic studies: one method yields token frequency and the other type frequency. Token or text frequency is the frequency of occurrence of a unit, usually a word or morpheme, in running text. For instance, *broke* (the past tense of *break*) occurs sixty-six times per million in Francis and Kučera (1982), while the past-tense verb *damaged* occurs five times in the same corpus. The token frequency of *broke* is much higher than that of *damaged*. We can also count the token frequency of a grammaticizing construction, such as *be going to*, by counting just those occurrences of *be going to* that are used with a following verb (rather than a noun).

Type frequency refers to the dictionary frequency of a particular pattern, such as a stress pattern, an affix, and so on. For instance, English past tense is expressed in several different ways, but the expression with the highest type frequency is the suffix *-ed*, as in *damaged*, which occurs on thousands of verbs. The pattern found in *broke* has a much lower type frequency, occurring with only a handful of verbs (depending upon how you count them: *spoke*, *wrote*, *rode*, etc.).

The notion of type frequency can also be applied to grammaticizing constructions by counting the different lexical items with which a construction can be used: for instance, when in Shakespeare's English *be going to* had its literal meaning of a subject traveling to a location in order to do something, the subject position could only be occupied by a noun phrase denoting an animate, mobile entity, and the verb following the phrase would have to be a dynamic verb. As the phrase grammaticized and changed its meaning the number of different types appropriate for subject position expanded to include non-animate and non-mobile entities and the verb position expanded to include a broader range of predicates (e.g., current usage allows *The tree is going to lose its leaves*; *I'm going to be ready at nine*; etc.). A grammaticizing phrase is thus said to increase in generality (Bybee 1985) as the contexts in which it is appropriate move from very specific to more general.

A much-noted property of grammaticizing constructions is this increase in type frequency of co-occurring lexical items. As a consequence, the token frequency of units such as *going to* or *gonna* also increases dramatically. As important as the increase in type frequency or generality is, it is the high token frequency of grammaticizing phrases that provides the triggering device for many of the changes that occur in the form and function of the grammaticizing construction. High token frequency triggers many changes because it affects the nature of the cognitive representations in ways that will be explained as I proceed. First, however, I turn to the issue of the increase in token frequency of grammaticizing constructions, using the English modal auxiliary *can* as a case study.

4. How does frequency increase? A case study of *can*

4.1. Generalization of meaning

One of the earliest-mentioned mechanisms of semantic change in grammaticization is bleaching or generalization, the process by which specific features of meaning are lost, with an associated increase in the contexts in which the gram may be appropriately used (Meillet 1912; Lehmann 1982).[1] in fact, generalization seems to characterize the entire grammaticization continuum—we note that as the process unfolds, grams always become more general and more abstract in their meaning, more widely applicable, and more frequently used.[2] The mechanism behind bleaching is habituation: a stimulus loses its impact if it occurs very frequently.

Grammaticizing expressions have inherent meaning derivable from the meanings of their component parts. It is this inherent meaning that is said to be bleached as grammaticization proceeds. In some cases (though certainly not all), a neat diagram may be constructed showing which parts of the original meaning are lost along the way. For instance, Modern English *can*, derived from an Old English main verb, *cunnan* 'to know', can be charted as going through the stages in table 16.1[3]

TABLE 16.1. Stages of Development for *Can*

Stage	*Meaning*	
Mental ability	Mental	enabling conditions exist in the agent
Ability	______	enabling conditions exist in the agent
Root possibility	______	enabling conditions exist __________

(cf. Bybee 1988 on *may*). At each stage, *can* is used in a wider range of contexts (table 16.2).

Ability and mental ability are self-explanatory; root possibility asserts that enabling conditions exist in general. They include the inherent abilities of the agent but also factors in the external world that create enabling conditions. Examples follow:

(2) *Mental ability*:
Ful wys is he that kan hymselven knowe!
'Completely wise is one who knows himself!' (B. Mk. 3329)[4]

(3) *Skill*:
Ther seen men who kan juste and who kan ryde
'Men are seen there who can (i.e., know how to) joust and who can ride'. (A. Kn. 2604)

(4) *Ability*:
But I wol passe as lightly as I kan
'I will pass by as lightly as I can' (B. NP. 4129)

(5) *Root possibility*:
Thou cannest not haue of Phocion a frende and a flaterer both to gether
'You cannot (It is not possible to) have of Phocion both a friend and a flatterer both' (UDALL Erasm. Apoph. 299a)

Tables 16.1 and 16.2 show what is meant by generalization or bleaching: specific features of meaning drop off, leaving a semantic core. The classes of main

TABLE 16.2. Contexts of Use of *Can*

Stage	*Subject*	*Main Verbs*
Mental ability	Human agents	Intellectual states and activities Communicating Skills
Ability	Human agents	All of the above Overt actions and activities
Root possibility	Human agents Passive subjects Inanimate subjects	All of the above

verbs with which the auxiliary *can* is used generalize, as does the range of possible subjects of *can*. However, this is not all there is to the story. It must be remembered that both specific and general meanings of a gram can coexist; old uses may be retained in certain contexts (Bybee and Pagliuca 1987; Hopper 1991). Furthermore, tables 16.1 and 16.2 are just schematic summaries; they do not actually inform us of how the changes took place. The result is generalization of meaning and contexts of use, but what exactly was the mechanism by which this generalization occurred?

4.2. From noun phrase complement to verb phrase complement (Old English)

The ancestor of the modern auxiliary *can* is the main verb *cunnan*, which expresses various types of knowing.[5] With a noun phrase complement denoting a person, a skill, or a language, the sense of knowing is acquaintance or acquired skill or knowledge (Goosens 1992). *Cunnan* is also used in the sense of understanding, as in "knowing the holy writings":

(6) Ge dweliað and ne cunnon halige gewritu
'You are led into error and do not know the holy writings' (Ags. Gospel of Matthew xxii)

In order for a main verb such as *cunnan* to begin its development into an auxiliary, it must expand its syntactic distribution to take verb phrase objects. *Cunnan* had very limited use with infinitive objects in the Old English period, so that studying the specific contexts in which it was used with an infinitive can give us some idea of how the development may have taken place. The infinitives used with *cunnan* in Old English mostly fit into three semantic classes of main verbs: verbs of mental state or activity, verbs of communication, and verbs describing skills. Table 16.3 shows the thirteen examples listed in the OED of *cunnan* used with an infinitive before 1100, plus the additional items listed by Goosens from his sample.

Goosens takes the mental state class as central and describes the other classes as related to this class more or less directly. There is no doubt that the mental state class is important, but when we consider how *cunnan* might have come to be used with infinitives, it seems likely that there were distinct motivations for the different semantic classes of verbs.

Since use of *cunnan* with mental state verbs is clearly important in Old English (and it continues to be important in Middle English), let us consider what would motivate the use of *cunnan* with verbs having a meaning that is so closely related to its own meaning. Indeed, meanings such as 'be able to know' and 'know how to understand' seem rather redundant. However, it is important to bear in mind that as a main verb, *cunnan* was already fairly frequent and thus would have begun to lose some of its semantic force and specificity. I suggest that mental state infinitives then appear to be added in to bolster the meaning of *cunnan*, to flesh out the specific sort of knowledge intended in the context:[6]

TABLE 16.3. Verb Classes Used with *Cunnan* in Old English

OED		*Additional Items Listed by Goosens (1992)*	
Mental states or activities			
understandan		*geþencean*	‘to comprehend’
ongietan (2)	‘to understand’	*behabban*	‘to comprehend’
tocnawan	‘to distinguish, discern’	*wurðian*	‘to esteem’
		gecnawan	‘to perceive, know’
Communication			
secgan (3)	‘to say’	*sprecan*	‘to speak’
geandettan	‘to confess’	*tæcan*	‘to teach’
		læran	‘to teach’
Skills			
gretan hearpan	‘to touch a harp’		
huntian	‘to hunt’		
wunda sniðan	‘to cut a wound’		
Other			
afandian	‘to prove, try’		
bebeorgan	‘to defend oneself’		

(7) He ne con ongitan forhwy swylc God geþafað
he not *con* understand why such God allows

‘He does not understand why God allows such as that’ (950 Alfred’s Boeth. xxxix)

The near synonymy of ‘can’t understand’ and ‘doesn’t understand’ supports the idea that *con ongitan* is a harmonic phrase that means about the same as either component alone would mean. Perhaps *cunnan* is beginning to bleach and grow too weak to stand alone in such contexts.

In other cases, it is not clear whether the form of *cunnan* means ‘know how to’ or is expressing a meaning similar to the main verb:

(8) Nu cunne ge tocnawan heofones hiw
Now *cunne* 2p distinguish-INF heaven-GEN hue

‘Now you can distinguish/interpret heaven’s hue’ (Ags. Gospel of Matthew xvi.3)

In this passage the speaker is pointing out to the addressees that they know how to and do in fact interpret the color of the sky at sunset and dawn to predict the weather.

Thus it appears that one avenue by which *cunnan* begins to grammaticize as an auxiliary is determined by the fact that it was already frequent and had already undergone some weakening of its semantic content. Of course, the use of *cunnan* with infinitives whose meaning is a more specific version than that covered by *cunnan* results in the further weakening of its semantic content.

With verbs of communication and instruction, *cunnan* is used in contexts in which it retains its “knowledge” interpretation: it is not used in quotative contexts but rather where the content of what is said is asserted to be based on accurate knowledge of facts:

(9) þæt hi andsware secgan cunnen
that they answer say-INF *cunn*-PL

‘That they can say the answer’ (*c*. 1000 Elena 374)

(10) Weras þa me soðlice secgan cunnon
man-PL then 1s-DAT truly say-INF *cunn*-PL

‘Then men can truly say to me’ (*c*. 1000 Elena 317)

The third verb class used with *cunnan* in Old English consists of those denoting skills, particularly those with a strong intellectual component, such as reading, writing, or singing, but not excluding the more physical, such as hunting. This verb class corresponds to a set of nominal objects frequently used with *cunnan*. The most frequent nominal or pronominal objects of *cunnan* refer to people, but the second most frequent set comprises objects that refer to intellectual skills.[7] Thus we find ‘know the holy writing’, ‘know songs’, ‘know book-learning’, ‘know 1etters’. An infinitive construction could arise in these contexts by adding in the infinitive to an object that already consists of a noun phrase. Consider the following example, where the first instance of *cunnan* is followed by a noun phrase while the second has both a noun and an infinitive:

(11) ðy læs ðe him con leoða worn, oððe mid hondum con hearpan
unless 3s-DAT *con* song-PL many or with hand-PDat *con* harp
grētan, hafaþ him his gliwes giefe
touch-INF have-3s 3s-DAT 3s-GEN glee-GEN gift

‘unless he knows many songs, or can (knows how to) touch the harp with his hands, has his gift of glee’ (*c*. 1000 Versus Gnom. 172)

In this case, it appears that the infinitive complement could develop directly out of the noun phrase complement as a vehicle for adding in more specific information about the skill being described.

The uses of these three verb classes in infinitival form with *cunnan*, then, appear to arise for different reasons and perhaps are simultaneous developments. They are not necessarily totally independent, however. While each developing class of *cunnan* plus infinitive is a separate construction, it is plausible to assume that some more abstract generalization emerges from the similarities among these constructions.

Since the uses of *cunnan* are highly constrained lexically, they are appropriately described in a cognitive grammar (Langacker 1987) or associative network framework (Bybee 1985, 1998) in which phrases or constructions are stored in the lexicon and generalizations are abstracted from these stored units. In this framework there is no strict separation of lexicon and grammar, but rather units of varying lengths and degrees of complexity may be stored lexically with the following properties: (1) the degree of strength or entrenchment of stored units is based on their text frequency; (2) connections or associations of both a phonological and semantic nature are made among items, based on similarity or identity; and (3) schemas of varying degrees of generality emerge from these representations.

A description of *cunnan* in Old English would require three quite specific schemas, one for each verb class, as shown in (12), and a more abstract schema, as in (13):

(12) a. *cunnan* + V-infinitive
'know' {mental state, activity}

b. *cunnan* + V-infinitive
'know to' {communicate, instruct}

c. *cunnan* + NP (V-infinitive)
'know how to' {skill} ({do a skill})

(13) *cunnan* + V-infinitive
'know (how) (to)' {activity involving mental capacity}

4.3. Expansion to auxiliary status (Middle English)

Since we are interested in how bleaching and generalization to new contexts take place, an appropriate time period to focus on is the Middle English period. Once more our concern will be the semantic classes of verbs that appear in infinitival form with *can*. To determine the relative text frequency of the verb classes and individual members of these classes, it is necessary to examine all instances of *can* + infinitive in a stretch of text. For this purpose I have chosen the works of Geoffrey Chaucer and have examined the first three hundred uses of *can* + infinitive listed in *A Concordance to the Complete Works of Geoffrey Chaucer* (Tatlock and Kennedy 1927), which includes all of the *Canterbury Tales*, most of *Troilus and Criseyde*, and several shorter poems.[8]

First we observe that the three verb classes that appeared with *cunnan* in Old English are still strongly represented in Middle English (in the following, the verbs are rendered in their Modern English spelling, unless that distorts the meaning or shape of the verb radically):

(14) Intellectual states or activities (52 tokens, 18 types):
deem, believe, see, know, guess, understand, *espy* (discover), judge, construe, imagine, comprehend, conclude, bethink, remember, find a difference, find a reason, shape a remedy, (wit) suffice

(15) Communication (102 tokens, 31 types):
clepen (name), *devyce* (describe), thank, say, tell (or count), express, expound, make mention, make a description, make by argument, answer, cry, bewail, speak, report, swear, lie, preach, *reherce* (describe), declare, reckon, amend, beguile, portray sorrow, assure, describe, write, complain, record, define, *distreyne* (urge), treat

(16) Skills ('know how to') (26 tokens, 18 types):
read, gloss, form, paint, *counterfete* (imitate), shape, do craft, do craftily, delve in herbs, work in philosophy, sing, dance, joust, play an instrument, play, entune, sound, make a beard

The frequency increase of *can* from Old to Middle English is manifested both in the use of *can* with a larger number of verbs of each class (i.e., type frequency) and in the development of a high token frequency for some combinations in the intellectual state and communication classes. Both kinds of frequency contribute to the bleaching of the meaning of an element.

Because of certain commonly used fixed phrases, the token-to-type ratio in the "intellectual states and activities" class and the "communication" class is approximately three to one. Here are the most commonly used main verbs:

(17) Communication class:

High token frequency:	tell	30
	say	29
	devyce (describe)	8
Type frequency:	31 distinct verbs	

(18) Intellectual states or activities:

High token frequency:	see	12
	deem	6
	understand	6
	espy (discover)	5
Type frequency:	18 distinct verbs	

In the associative network or cognitive framework described earlier, type frequency corresponds to the generality of the schema, which in turn corresponds to a higher degree of grammaticization. High token frequency corresponds to a local schema that is very strong or highly entrenched, such as *can say*, *can tell*, or *can see*. Increases in frequency of both types lead to the continued bleaching of the meaning of *can*.

Actually, the phrases listed earlier are abstractions from larger ritualized phrases found frequently in the Chaucer texts, phrases such as the following:

(19) I kan say yow no ferre (farther) (A. Kn. 2060)
I kan say you namoore (B. ML. 175; B. NP. 4159; G. CY. 651)

(20) more than I kan telle (B. ML. 1120)
mo than I kan make of mencioun (A. Kn. 1935)
more than I kan yow devyse (describe) (B. ML. 429)

(21) I kan nat seen (that) (B. Mel. 2735; TC II 794; TC IV 1365)

Each of these can be viewed as a construction with varying degrees of generality and varying degrees of entrenchment.

The Chaucer texts also reveal that the use of *can* with infinitives has expanded to other semantic classes of verbs, that is:

1. verbs denoting states of mind that are not strictly intellectual, such as *love*, *suffer*, *have patience*, and so on;
2. verbs denoting states that are not mental or emotional: *be wrye* (twisted), *be rotten*, and so on;
3. verbs indicating a change of state in another person (these are probably related to verbs of instruction of Old English): *teach*, *heal*, *comfort*, *disturb*, and so on;
4. verbs indicating an overt action: *ride*, *go*, *send*, *climb*, *steal*, and so on.

It is plausible to assume that these verb classes arose out of the earlier three classes gradually, since lines between semantic classes of verbs are not discrete (cf. the study of Kemmer 1995). I propose the following developments:

(22) 'know' > 'experience'
main predicates: Intellectual states > States of mind > States

(23) 'know to tell' > 'know how to' > 'be able to'
main predicates: Instruction > Change of state (transitive)

(24) 'know a skill' > 'be able to'
main predicates: Mental skills > Physical skills > Overt action

By the time these developments have occurred, there are very few predicates that cannot be used with *can*. Despite the generality with main predicates, *can* does not yet express root possibility with any regularity, since use with inanimate subjects is extremely rare: only twelve examples are found in the corpus of three hundred and all but two of these are metonymic expressions for humans, that is, "inanimate" objects such as the eyes, the heart, wit, foolishness, and beauty. Two other inanimate objects that can tell or hide (the truth) are a book and the gossip or prattle (of women):

(25) As ferforth as my wit kan comprehend
'As far as my wit can comprehend' (TC IV 891)

(26) Swich vanyte ne kan don hire non ese
'Such foolishness cannot please her' (TC IV 703)

It appears that the most general schema for *can* in Chaucer's English is centered on human subjects and is only occasionally extended beyond humans to aspects of their behavior or metonymic uses of mind-body parts (such as eyes and wit). The most general schema, (27), does not have restrictions on the type of main predicate *can* occurs with. The tendency to use *can* commonly with certain semantic classes of verbs is captured in more specific schemas referring to the verb types listed in (22) through (24) or covering very specific constructions, such as those represented in (19) through (21):

(27) {human subject} *can* + infinitive
{know how to}
{experience}
{be able to}

At this period, *kan* has generalized to expressing human ability of all types, but it has not yet generalized beyond ability to include root possibility.[9]

4.4. Further generalization: Root possibility

General ability is very closely related semantically and functionally to root possibility. While ability applies only to properties internal to an agent, root possibility includes both internal and external enabling conditions. It is paraphrasable as "it is possible for *x* to *y*." Thus in one of the few examples of root possibility in the Chaucer texts, we can see how this paraphrase would apply:

(28) Til we be roten, kan we nat be rype
'Until we are rotten, it is not possible for us to be ripe' (A. Rv. 3875)

The close relation of ability to root possibility is due to practical considerations in the real world: the ability to do something often depends on both internal and external conditions. Thus in this example, again from Chaucer, it is difficult to tell if the conditions are internal, external, or both:

(29) Allas! kan they nat flee the fires heete? (G. CY. 1408)

Later in the Middle English period, examples interpretable as root possibility become much more common, and the syntactic conditions under which *can* is used continue to shift. In particular, the root possibility reading makes the use of *can* with stative predicates and in passive clauses possible, as the following two examples show:

(30) No worldely thyng can be wythout stryfe. (1509 Hawes Past. Pleas. xvi.xlix)

(31) Gij, But and thou array thy body sumptuously thou canst not be excused as chast in mind. (1540, Hyrde Vives' Instr. Chr. Worn. 1592)

Also examples of *can* expressing capacity, a use close to root possibility, appear in the sixteenth century:

(32) There is great number that fayne would aborde
our ship can holde no more. (Barclay Ship of Fooles 1570)

In this use, *can* begins to replace *may*, which, as noted earlier (n. 9), was much used in the root sense in Chaucer's works. *May* is undergoing its own development, however, and beginning to be used more often in the epistemic sense ('it is possible that').

This brief survey of the development of *can* from Old to Middle English illustrates how the sharp frequency increase takes place: (1) the grammatical construction (*can* + infinitive) gradually extends to use with more and more types of verbs and then subjects; this extension is based on semantic similarity among the verbs in question, but its result is a generalization or bleaching of the meaning of *can*; (2) certain phrases have a high token frequency, which also serves to bleach the meaning of their component parts. The result is a major change from the meaning of *cunnan*: in these root possibility readings of (30), (31), and even (32), no hint of the meaning of *cunnan* as 'know' remains.

5. Phonological changes

A recognized concomitant of grammaticization is reduction in phonological form. In a large cross-linguistic sample, Bybee, Perkins, and Pagliuca (1991, 1994) demonstrate a significant association between degree of semantic grammaticization and phonological reduction, particularly in the length of the grams in question but also in the degree of fusion of the gram with surrounding material.

The previous section illustrated in some detail the way increases in token and type frequency occur over time. In this section I will examine the link between frequency, phonological reduction, and fusion of grammaticizing phrases. The example of *can* is less useful here, since it is a single monosyllable, so other examples will be taken up. It should not be concluded, however, that *can* has undergone no phonological reduction just because its orthographic shape is fairly constant. Since the Old English period it has lost the final inflectional syllable that occurred in many forms (*cunnan*, *cann*, *canst*, *cunnon*, *cunne*) as have other verbs, and furthermore, in Modern English *can* is phonetically reduced to [kŋ] or [ŋ] in high-frequency contexts, such as after the pronoun *I*.

5.1. Phonological reduction

Recent studies of the lexical diffusion of regular sound changes have shown that in many cases high-frequency words undergo sound change at a faster rate than low-frequency words. The effects of frequency have been shown for vowel reduction and deletion in English (Fidelholtz 1975; Hooper 1976), and for the raising of /a/ to /o/ before nasals in Old English (Phillips 1980), for various changes in Ethiopian languages (Leslau 1969), for the weakening of stops in American English and vowel change in the Cologne dialect of German (Johnson 1983), for ongoing vowel changes in San Francisco English (Moonwomon 1992), for tensing of short *a* in Philadelphia (Labov 1994:506–507), and for t/d deletion in American English (Bybee 2001:23ff, 112ff, 151ff).

Pagliuca and Mowrey (1987) argue that when one views articulation in terms of sets of overlapping gestures, all sound change can be classified as due to Substantive

Reduction, the reduction in the magnitude of a gesture, or Temporal Reduction, the reduction in the duration of a constellation of gestures, resulting in the shortening of individual gestures or the increase in the overlap of gestures. This hypothesis is meant to explain the dominance of weakening and assimilation in attested sound changes. Browman and Goldstein (1990, 1992) make a very similar claim for casual speech processes (which I take to include the same range of phenomena as the category "sound change"). Browman and Goldstein hypothesize that all casual speech processes result from either the reduction in magnitude of a gesture or the increase in the overlap of gestures.

These hypotheses await further investigation, but even if they turn out to have some counterexamples, the fact will remain that a large proportion of phonological changes are reductive in nature. Thus it is reasonable to ask why reductive changes would affect high-frequency words or phrases earlier and at a faster rate than low-frequency words and phrases. Several factors can be identified.

First, Fowler and Housum (1987) found that the second repetition of the same word in a single discourse was significantly shorter than the first token of the word. The speaker can be less explicit about the articulation of a word if it has already been used, because it will be easier for the listener to access if it has just been activated. Furthermore, Fowler and Housum point out that the reduction can actually be a signal to the listener that the word being used is just the same as one used earlier rather than a new and different word. It would follow then that words or phrases that are often repeated in the same discourse (high-frequency and grammaticizing phrases) would be in position to be shortened more often than words and phrases of low frequency.

Second, D'Introno and Sosa (1986) point out that frequency effects in the spread of a sound change are better viewed as familiarity effects: their position is that it is not so much the frequency of a word but rather its use in casual or familiar social situations that allows it to reduce or undergo change at a faster rate. Since the changes in question occur more often in casual speech, words that are used more often in casual speech will be more often subjected to the change.

Other factors might be involved as well, especially for grammaticizing constructions: as meaning generalizes, the informational contribution of the grammaticized elements decreases and along with that the intonational and rhythmic emphasis. Such prosodic reduction will have an effect on the segmental properties of the phrase as well.

For all of these reasons (and perhaps others), increasing frequency of use of grammaticizing constructions leads to phonological reduction. While the reduction is extreme in many cases, it usually follows patterns that are also seen in ongoing or future sound changes, suggesting that it is the frequency of use that hastens the changes. For instance, in Old Spanish the second person plural suffix was *-des* (from Latin *-tis*) and was preceded by a stressed vowel: *ádes*, *-édes*, or *-ídes*. In Old Spanish this medial *d* (pronounced [ð]) was gradually deleted, so that in Modern Spanish (in the dialects that use it) the forms are *-áis*, *-éis*, and *-ís*. Currently in most dialects of Spanish other instances of medial [ð] are gradually deleting. What is interesting is that this earlier morpheme-specific change was an instance of a more general change that would be current many centuries later.

Other instances of phonological reduction in grammaticization seem more extraordinary, but even most of these can be analyzed into steps that reflect the general reduction patterns of the language. For example, *going to* [goɪŋtuwl reducing to *gonna* [gə̃nə̃] or even further, as in *I'm gonna* reducing to [aimə̃ɾ̃ə̃], involves the following: (1) the reduction of full [o] to schwa; (2) change of the velar to alveolar nasal; (3) vowel nasalization; and (4) flapping, all of which occur in other words as well. However, certain aspects of this reduction are extraordinary: (1) reduction of the diphthong [oi]; (2) flapping of [nt]; and (3) deletion of [g] in [aimə̃ɾ̃ə̃].

5.2. Phonological fusion

Besides the reduction of the consonants and vowels within words, grammaticization often involves the phonological fusion of words or morphemes that formerly were separate. Here frequency is at work as well. Combinations of words and morphemes that occur together very frequently come to be stored and processed in one chunk. Boyland (1997) points out that as high-frequency sequences of units come to be processed as single units, their gestural representation changes: what were previously multiple gestures come to be reorganized into single gestures and along with this reorganization comes reduction and increased overlap of gestures.

In Bybee and Scheibman (1999), we have shown that the reduction of the auxiliary *don't* in English is most extreme in precisely the phrases in which it most commonly occurs. Out of 138 occurrences of *don't* from spontaneous conversation, 87 were after the first singular pronoun *I*, making this the most common element to precede *don't*. There were fifty-one tokens in which the vowel was reduced to schwa and fifty of these occurred with *I*. (The other token was in the phrase *why don't you*, used to make a suggestion.) The reduction to schwa was also influenced by the following verb. The most common verb to follow *don't* was *know*, and twenty-nine of the tokens with a schwa occurred with this verb. In fact, twenty-nine out of thirty-nine cases of *don't know* were reduced and all of these were in the phrase *I don't know*. The second most common verb to be used with *don't* was *think*, and seven out of nineteen of these cases, again all with *I*, were reduced to schwa. Other phrases in which *don't* was reduced were *I don't have* (*to*), *I don't want*, *I don't like*, *I don't mean*, *I don't feel* and *I don't care*. The reduction did not occur with any other pronouns with the twenty other verb types found in the conversations.

We concluded that neither phonological nor syntactic conditioning is responsible for the reduction of *don't*, but rather that this reduction occurs inside of automated processing units, chunks that are automated primarily because they occur with high frequency. As *I don't know* comes to be produced as a single unit, the medial syllable loses its stress, allowing the vowel of *don't* to reduce.

6. Autonomy

Another consequence of a high frequency of use of a word or phrase consisting of multiple morphemes is a growing autonomy from other instances of these same morphemes. I have argued that token frequency is an important determinant of semantic split among derivationally related words (Bybee 1985). That is, derived words that are of relatively high frequency (compared to their base form) are more likely to be semantically opaque and to have additional meanings or nuances not present in the base form. The reason for this is that high-frequency words are present enough in the input to have strong representations of their own; they do not have to be understood in terms of other related words.

The same process applies to grammaticizing phrases—they gradually grow increasingly independent of their composite morphemes and other instances of the same construction. Thus the phrase (*be*) *going to* is becoming less and less associated with the individual morphemes, *go*, *-ing*, and *to*, until a point may well come when speakers are surprised to find out what its etymological source is. Similarly, but on a different plane, (*be*) *going to* has disassociated itself from other instances of the construction, as given in (1). Such dissociations are phonological, semantic, and morphosyntactic.

Dissociations due to growing autonomy of grammaticizing phrases account for the splits that are often found between a morpheme in a grammaticizing phrase and its lexical source (Heine and Reh 1984; Hopper 1991). French *pas* in the negative phrase *ne . . . pas* is no longer associated with its etymological source, the noun *pas* meaning 'step'. The forms of *avoir* in French are still used for possession but are also found in the construction of the *passé composé* (*j'ai chanté*, *tu as chanté*, *il a chanté*, etc.) and in the formation of the future (*je chanterai*, *tu chanteras*, *il chantera*, etc.). In these three uses, despite similarities of phonology, these forms are best analyzed as autonomous from one another; they occur in different constructions and their meanings are in no way transparently related across these constructions.

7. New pragmatic associations

The autonomy of grammaticizing phrases and their growing opacity of internal structure makes it possible for new pragmatic functions to be assigned to them. Such new functions originate in the contexts in which the expressions are frequently used.

As an example, consider the phrase *I don't know* as used in colloquial American English. As mentioned earlier, this phrase can reduce to [aiɾə̃nə̃] or [aiə̃nə̃]. While it can be used with its literal meaning as an answer to a question, it can also be used in conversation to mitigate an assertion or to politely disagree or refuse something being offered (Scheibman 2000). In these cases, *I don't know* is a single processing unit that is losing its association with the words from which it was derived. Due to its growing autonomy, it is capable of taking on new discourse functions that arise from the contexts in which it is commonly used.

8. Entrenchment: The evolving morphosyntactic properties of English auxiliaries

Another effect of high token frequency on complex forms is their maintenance of conservative structure despite the pressure of productive patterns (Bybee 1985).

High token frequency explains why some English verbs (*ate*, *broke*, *wrote*) retain their irregular vowel changes despite the extreme productivity of the *-ed* affix for expressing past tense. High-frequency constructions can also retain conservative morphosyntactic characteristics even in the face of new productive morphosyntactic patterns. Bybee and Thompson (1997) argue that even morphosyntactic constructions can exhibit this type of entrenchment due to the strength of the representation of the construction.

It is well known that English modal auxiliaries (*can*, *could*, *may*, *might*, *will*, *would*, *shall*, *should*, and *must*) have a set of syntactic properties that distinguish them from main verbs: the use of a bare infinitive, subject inversion in questions and other contexts, and the placement of the negative immediately following the auxiliary. How did these properties develop? Space is not available here for a detailed treatment of these properties, but the basic answer is that these properties were once variable properties of all verbs, but they have become conventionalized in these high-frequency verbs, while all other verbs changed their properties in accordance with the changing syntax of the English language.

Consider first the use of the bare infinitive rather than the *to* infinitive. In Old English, the infinitive was formed by adding *-an* to the verb stem. Thus verb + infinitive constructions in Old English had no intervening *to*. With general reduction of final syllables and the loss of inflections in verbs and nouns, the infinitive suffix gradually disappeared. Long before this suffix was lost, however, a new infinitive marker began to develop in the form of the preposition *to*. Haspelmath (1991) has shown that the primary source of infinitive markers cross-linguistically is allative or dative markers, which are first used in purpose clauses and subsequently generalized to other infinitival uses. This is exactly what happened in English: *to* with the infinitive (an erstwhile verbal noun) inflected in the dative was first used in purpose clauses and gradually extended to general use as an infinitive marker. During the Middle English period there was still some variation in the use of infinitives with and without *to*.[10]

Modals such as *can* have very consistently occurred throughout their history in constructions without *to*. The reason for this is that these constructions were first created and apparently entrenched before *to* developed as the infinitive marker. Since constructions with the modal auxiliaries were of high frquency and thus highly entrenched, they were not reformulated after the *to* infinitive generalized in the language. The same is true of other verb + infinitive constructions that have survived from the Old English period. For instance, *go* + infinitive and *see* + infinitive constructions use bare infinitives even today: *Let's go see*; *I saw him do it*. More recent formations with functions similar to those of the modal auxiliaries, such as *want to*, *be going to*, *have to*, use the newer infinitive construction that was established before these constructions became entrenched. Thus it is the fact that the constructions with *can* + infinitive arose before the *to* infinitive and the fact that they were of high frequency that together explain why *can* uses a bare infinitive.

Another striking characteristic of the class of auxiliaries to which *can* belongs today is that they invert with the subject in certain constructions, primarily questions, but also (perhaps archaically) in conditional protases lacking *if*, and in clauses with fronted negative elements. In the Middle English period this verb–subject order in these contexts was a variable property of all verbs; it was not restricted to auxiliaries (Mossé 1952:126–128). Consider these examples:

(33) Gaf ye the chyld any thyng?
'Did you give the child anything?'

(34) Ne sunge ich hom never so longe,
Mi songe were i-spild ech del

'Even if I sang to them ever so long,
My song would be entirely lost (on them}'

Since the modal auxiliaries and *be* and *have* as auxiliaries were becoming increasingly frequent in this period, they would commonly occur before the subject in these contexts. While other verbs eventually ceased to appear in this position, taking instead the position after the subject, which eventually became obligatory, the auxiliaries, including the newly developed pro-verb, *do*, remained in inverted positions in these special constructions. Again it is their high frequency that accounts for their conservative behavior. The constructions with inverted auxiliaries were highly entrenched and thus not prone to revision despite the other syntactic changes occurring in English.

The position of the negative *not* after *can* and other auxiliaries has a similar diachronic explanation. The sentence negation particle in Old English, *ne*, occurred before the verb, but in Middle English it was reinforced by another negative *nought*, *not*, which derived from *ne* + *wiht* (literally: 'not a creature'). *Not* occurred after the verb in Middle English and became the normal negative marker as the preposed *ne* was lost. It occurred after simple finite main verbs as well as after the auxiliary (Mossé 1952:112):

(35) My wyfe rose nott

(36) cry not so

The position of the negative after *can* and other auxiliaries is the preservation for this high-frequency group of the order that once applied to all verbs. While other verbs require the use of *do*-support, the auxiliaries have simply continued to participate in the highly entrenched construction that was established in the fourteenth century.

Thus it can be said that the special properties of the auxiliaries in English are the retention of older morphosyntactic properties that were once general to English verbs. These modal auxiliaries and the other auxiliaries, *be*, *have*, and *do*, have retained these properties because of their high frequency: due to repetition their participation in certain constructions is highly entrenched and not likely to change. By the same token, modal constructions developing more recently will reflect the morphosyntax of the period in which they develop and are highly unlikely to fall in with the older modals and take on their characteristics, such as using the bare infinitive and occurring before the subject in questions.

This preservation of older morphosyntactic characteristics in high-frequency constructions can be attributed to the same mechanism as the preservation of irregularities (older morphological properties) in inflected forms. While analogical change generally operates to level or regularize morphophonemic alternations (e.g., as *wept* becomes *weeped*), forms with high token frequency tend to resist such change (e.g., *kept* is not becoming *keeped*); see Anttila 2003; Dressler 2003; Hock 2003).

9. The effects of repetition

This survey of the changes that occur in grammaticization has revealed that repetition affects semantics and phonology by promoting change, in particular, reductive change, and that repetition affects morphosyntax by ensuring the retention of older characteristics. It might seem contradictory that repetition could both encourage innovation in one domain and enhance conservatism in another. This paradox is also found in the lexical diffusion of phonetic versus morphophonemic change. In Hooper (1976), I pointed out that sound change affects high-frequency items first, while analogical leveling affects low-frequency items first. The substantive properties of words or phrases, their meaning and phonetic shape, are modified, usually reduced, with use. The ritualization or automatization process has an online effect of compressing and reducing; this is a processing effect. In contrast, the structural properties of words and phrases—that is, the morphological structure of words and the syntactic properties of constructions—are preserved by repetition; this is a storage effect. Frequently used words and phrases are highly entrenched and more likely to be accessed as whole units and less likely to be re-formed online. Thus their general structure—the morphological irregularity of high-frequency nouns and verbs, or the structure of high-frequency construction—will tend to be preserved. We can say, then, that repetition has a reductive effect on-line but a conserving effect in storage.

Repetition is universal to the grammaticization process. Repetition and its consequences for cognitive representation are major factors in the creation of grammar. The conventionalized aspects of language provide the framework for manipulation of our thoughts into objects of communication. Repetition alone, however, cannot account for the universals of grammaticization. The fact that the same paths of change are followed in unrelated languages has multiple causes. It is not just the fact of repetition that is important but in addition *what* is repeated that determines the universal paths. The explanation for the content of what is repeated requires reference to the kinds of things human beings talk about and the way they choose to structure their communications.

Notes

I wish to thank John Haiman, Barbara Need, Sandra Thompson, and Elizabeth Traugott for helpful comments on an earlier draft of this chapter.

1. Care must be taken here to distinguish between meaning and use: as a gram loses specific features of meaning, it appears to take on more uses. Being used in a wide range of contexts does not mean that the gram has *more* inherent meaning.

2. At the end of the grammaticization process, an old gram may be restricted in use by newer grams that replace some of the uses of the older one. The consequent addition of contextual meanings to the old gram may appear to make meanings more restrictive.

3. The permission use of *can* is not treated here. In Bybee and Pagliuca (1985) and in Bybee, Perkins, and Pagliuca (1994), we argue that the permission use of grams originally expressing ability develops out of the root possibility sense. Root possibility expresses a highly generalized set of enabling conditions, which include the social conditions that govern permission.

4. Abbreviations for examples from Chaucer are: B. Mk. (Monk's Tale); A. Kn. (Knight's Tale); B. NP. (Nun's Priest's Tale); B. ML. (Man of Law's Tale); G. CY. (Canon's Yeoman's Tale); B. Mel. (Tale of Melibeus); TC II (Troilus and Criseyde, book 2); TC IV (Troilus and Criseyde, book 4); A. Rv. (Reeve's Tale). All other abbreviations are from the OED and follow the OED's format for dates and details for locating the example in the text: date of publication, author/title of work, chapter, page number, etc.

5. The past tense of *cunnan*, OE *cuþ*, which gives Modern English *could*, will not be treated here. See Bybee (1995) for the development of past-tense modals in English.

6. Lyons (1977) refers to cases in which two modals of similar meaning co-occur in a clause without increasing or decreasing the degree of modality as modal harmony. *Cunnan* 'to know' plus a verb of knowledge could be regarded as a harmonic expression.

7. This is true of Goosen's sample and the small OED sample consulted.

8. There are several varieties subsumed under Old English and even more under the designation Middle English, so it cannot be assumed that there is necessarily a direct developmental relation between the languages represented in the texts used here. Still it is clear that in some general sense a type of diachronic relation exists.

9. In Chaucer's English, root possibility is expressed by *may*, which derived from a verb expressing physical power or ability. *May* is more grammaticized semantically than *can*: in the Middle English period it is used frequently with inanimate and generic subjects to express root possibility. It is also commonly used in subordinate clauses and is even beginning to express epistemic possibility in some contexts (see Bybee 1988).

10. Indeed in Modern English there is still variation between the bare infinitive and the infinitive with *to*, as in *help someone* (*to*) *do something*.

References

Anttila, Raimo. 2003. Analogy: The warp and woof of cognition. In *Handbook of historical linguistics*, ed. R. Janda and B. Joseph, 425–440. Oxford: Basil Blackwell.

Boyland, Joyce Tang. 1997. Morphosyntactic change in progress: A psycholinguistic approach. Ph.D. diss., University of California at Berkeley.

Browman, Catherine P., and Louis M. Goldstein. 1990. Tiers in articulatory phonology, with some implications for casual speech. In *Papers in laboratory phonology*, vol. 1: *Between the grammar and physics of speech*, ed. John C. Kingston and Mary E. Beckman, 341–376. Cambridge: Cambridge University Press.

———. 1992. Articulatory phonology: An overview. *Phonetica* 49, 155–180.

Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins.

———. 1988. Semantic substance vs. contrast in the development of grammatical meaning. *BLS* 14, 247–264.

———. 1995. The development of past tense modals in English. In *Modality in grammar and discourse*, ed. Joan L. Bybee and Suzanne Fleischman, 503–517. Amsterdam: John Benjamins.

———. 1998. The emergent lexicon. *CLS 31*, number one: *The Panels*, 421–435. Chicago: Chicago Linguistic Society. (Reprinted in this volume as chapter 13.)

———. 2001. *Phonology and language use*. Cambridge: Cambridge University Press.

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 William Pagliuca. 1985. Cross-linguistic comparison and the development of grammatical meaning. In *Historical Semantics, Historical Word Formation*, ed. J. Fisiak, 59–83. The Hague: Mouton.

———. 1987. The evolution of future meaning. In *Papers from the VIIth International Conference on Historical Linguistics*, ed. A. G. Ramat, O. Carruba, and G. Bernini, 109–122. Amsterdam: John Benjamins.

Bybee, Joan L., Revere D. Perkins, and William Pagliuca. 1991. Back to the future. In *Approaches to grammaticalization*, ed. Elizabeth C. Traugott and Bernd Heine, 17–58. Amsterdam: John Benjamins.

———. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.

Bybee, Joan L., and Joanne Scheibman. 1999. The effect of usage on degrees of constituency: The reduction of *don't* in American English. *Linguistics* 37(4): 575–596. (Reprinted in this volume as chapter 14.)

Bybee, Joan L., and Sandra A. Thompson. 1997. Three frequency effects in syntax. *BLS* 23, 65–85. (Reprinted in this volume as chapter 12.)

D'Introno, F., and J. M. Sosa. 1986. Elisión de la /d/ en el español de Caracas: Aspectos sociolingüísticos e implicaciones teóricas. In *Estudios sobre la fonología del español del Caribe*, ed. Rafael A. Núñez Cedeño, Iraset Páez Urdaneta, and Jorge M. Guitart, 135–163. Caracas: Ediciones La Casa de Bello.

Dressler, Wolfgang U. 2003. Naturalness and morphological change. In *Handbook of historical linguistics*, ed. R. Janda and B. Joseph, 461–471. Oxford: Basil Blackwell.

Fidelholtz, James L. 1975. Word frequency and vowel reduction in English. In *Papers from the eleventh annual meeting of the Chicago Linguistic Society*, 200–213. Chicago: Chicago Linguistic Society.

Fowler, Carol A., and Jonathan Housum. 1987. Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language* 26(5), 489–504.

Francis, W. Nelson, and Henry Kučera. 1982. *Frequency analysis of English usage*. Boston, Mass.: Houghton Mifflin.

Goosens, Louis. 1992. *Cunnan*, *conne*(*n*), *can*: The development of a radial category. In *Diachrony within synchrony: Language history and cognition*, ed. G. Kellermann and M. D. Morrissey, 377–394. Frankfurt am Main: Peter Lang.

Haiman, John. 1994. Ritualization and the development of language. In *Perspective on grammaticalization*, ed. William Pagliuca, 3–28. Amsterdam: John Benjamins.

Haspelmath, Martin. 1991. From purposive to infinitive: A universal path of grammaticization. *Folia Linguistic Historica* 10(1/2), 287–310.

Heine, Bernd. 2003. Grammaticalization. In *Handbook of Historical Linguistics*, ed. R. Janda and B. Joseph, 575–601. Oxford: Basil Blackwell.

Heine, Bernd, Ulrike Claudi, and Friederike Hünnemeyer. 1991. *Grammaticalization: A conceptual framework*. Chicago: University of Chicago Press.

Heine, Bernd, and Mechthild Reh. 1984. *Grammaticalization and reanalysis in African languages*. Hamburg: Buske.

Hock, Hans Henrich. 2003. Analogical change. In *Handbook of historical linguistics*, ed. R. Janda and B. Joseph, 441–460. Oxford: Basil Blackwell.

Hooper, Joan B. 1976. Word frequency in lexical diffusion and the source of morphophonological change. In *Current progress in historical linguistic*, ed. William Christie, 96–105. Amsterdam: North Holland. (Reprinted in this volume as chapter 2.)

Hopper, Paul. 1991. On some properties of grammaticization. In *Approaches to grammaticalization*, vol. 1, ed. Elizabeth Traugott and Bernd Heine, 17–35. Amsterdam: John Benjamins.

Johnson, Theodore. 1983. Phonological free variation, word frequency, and lexical diffusion. Ph.D. diss., University of Washington, Seattle.

Kemmer, Suzanne. 1995. An analogical model of syntactic change. Paper presented at the Conference on Functional Approaches to Grammar, University of New Mexico, July.

Labov, William. 1994. *Principles of linguistic change: Internal factors*. Oxford: Basil Blackwell.

Langacker, Ronald. 1987. *Foundations of cognitive grammar*, volume one. Stanford, Calif.: Stanford University Press.

Lehmann, Christian. 1982. *Thoughts on grammaticalization: A programmatic sketch*, vol. 1. Cologne: Arbeiten des Kölner Universalien-Projekts Nr. 48.)

Leslau, Wolf. 1969. Frequency as a determinant of linguistic change in the Ethiopian languages. *Word* 25, 180–189.

Lyons, J. 1977. *Semantics*. Cambridge: Cambridge University Press.

Meillet, Antoine. 1912. L'evolution des formes grammaticales. *Scientia* (*Revista di Scienza*) 12(26): 6. (Reprinted in Antoine Meillet 1948, 130–148.)

Moonwomon, Birch. 1992. The mechanism of lexical diffusion. Paper presented at the Annual Meeting of the Linguistic Society of America, Philadelphia, January.

Mossé, Fernand. 1952. *A handbook of Middle English*. Baltimore: Johns Hopkins Press.

Pagliuca, William, and Richard Mowrey. 1987. Articulatory evolution. In *Papers from the 7th International Conference on Historical Linguistics*, ed. A. G. Ramat, O. Carruba, and G. Bernini, 459–472. Amsterdam: John Benjamins.

Phillips, Betty S. 1980. Old English *an-on*: A new appraisal. *Journal of English Linguistics* 14, 20–23.

Scheibman, Joanne. 2000. *I dunno but*: A usage-based account of the phonological reduction of *don't* in American English conversation. *Journal of Pragmatics* 32(1), 105–124.

Tatlock, John S. P., and Arthur G. Kennedy. 1927. *A concordance to the complete works of Geoffrey Chaucer*. Washington, D.C.: Carnegie Institute.

Traugott, E. C. 1989. On the rise of epistemic meaning: An example of subjectification in semantic change. *Language* 65, 31–55.

———. 2003. Constructions in grammaticalization. In *Handbook of historical linguistics*, ed. R. Janda and B. Joseph, 624–647. Oxford: Basil Blackwell.

This page intentionally left blank

INDEX