INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
A processing model of phonological rule application

Myers, James Tomlinson, Ph.D.

The University of Arizona, 1993
A PROCESSING MODEL OF PHONOLOGICAL RULE APPLICATION

by

James Tomlinson Myers

A Dissertation Submitted to the Faculty of the DEPARTMENT OF LINGUISTICS
In Partial Fulfillment of the Requirements For the Degree of DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

1993
As members of the Final Examination Committee, we certify that we have read the dissertation prepared by James Tomlinson Myers entitled A PROCESSING MODEL OF PHONOLOGICAL RULE APPLICATION and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Michael Hammond
Richard Oehrle
Merrill Garrett

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Dissertation Director Michael Hammond

July 27, 1992
STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: [Signature]
To my parents.
ACKNOWLEDGMENTS

I am solely responsible for the shortcomings of this pamphlet. If any good comes out of it, it is due to the help of other people.

First off, I must thank the folks of the National Science Foundation, who paid my bills during the bulk of the research. I would not have had the NSF Fellowship at all if it hadn't been for Mike Flynn, Mike Hammond, Doug Saddy, Kelly Sloan and Sue Steele.

Then there is my advisor, Mike Hammond. If nothing else, this dissertation gave me a chance to get to know him better. He both does good linguistics and presents a good model for how to live the linguist's life. To say any more would make us both embarrassed.

In addition to Mike, and my other committee members Merrill Garrett and Dick Oehrle, fate brought me in contact with many other smart people in Tucson who set me thinking; the ones who helped most with my dissertation include Jean Ann, Diana Archangeli, David Basilico, Paul Bloom, Tom Bourgeois, Jo Calder, Dick Demers, Lee Fulmer, Chip Gerfen, Chris Golston, Kerry Green, Sung-Hoon Hong, Janet Nicol, Pat Perez, Curt Rice, Jane Tsay and Malcah Yaeger-Dror.

I also owe thanks to numerous other people who were involved at specific stages of the work. In addition to my committee members, Oliver Myers and Jane Tsay performed the monumental task of reading my whole dissertation. Fellow dissertator Jean Ann complained a lot, which helped me feel better. Anne Cutler, Ken Forster, Victoria Fromkin, Chris Golston, Jeri Jaeger, Janet Nicol, Stephanie Shattuck-Hufnagel and Joe Stemberger provided insights very useful to the discussion in Chapter 3. Greg Iverson, Samuel Wang and the audience at the presentation of Myers (1992) at BLS 18 provided good feedback on ideas that ended up in Chapter 4; Jack Chambers, Jonathan Kaye, Curt Rice and Malcah Yaeger-Dror provided comments helpful specifically with Canadian Raising. Jeff Bowers, Betty Glisky, and Steven Zepp were the speakers of Canadian English I checked some hypotheses with. Communications with Jennifer Cole were helpful in thinking about the general topics in Chapter 5; communications with Greg Guy and Malcah Yaeger-Dror helped in my discussion of Coronal Deletion in that chapter.

Above all, I would like to express my unending gratitude to my colleague and friend Jane Tsay. Without her inexhaustible patience, common sense and kindness, I would have gone insane much sooner than I did.
# TABLE OF CONTENTS

**Abstract** ........................................................................................................................ 12

**Chapter 1. An introduction to Double Lookup** ........................................................ 14

1.0 Introduction ........................................................................................................... 14

1.1 The relation between linguistics and psycholinguistics .................................. 15

1.2 Double Lookup .................................................................................................. 19

1.2.1 Lexical Lookup and Rule Lookup ............................................................ 19

1.2.2 Gradient Retrievability and Gradient Productivity .................................. 20

1.2.3 Redundancy of Patterning ......................................................................... 27

1.2.4 The Productivity Hypothesis ..................................................................... 38

1.3 Chapter summary ............................................................................................... 51

**Chapter 2. Partially productive rules in English** .............................................. 55

2.0 Introduction ........................................................................................................ 55

2.1 Syllabification ................................................................................................... 59

2.2 Rules involving vowel length and vowel height ............................................ 64

2.2.1 The Vowel Shift alternations ...................................................................... 66

2.2.2 Cluster Shortening ..................................................................................... 70

2.2.3 Trisyllabic Shortening ................................................................................ 78

2.2.4 -ic Shortening .............................................................................................. 81

2.2.5 i-Shortening ............................................................................................... 82
2.2.6 CiV Lengthening

2.2.7 Prenasal g-Deletion and compensatory lengthening

2.2.8 Vowel length rules and syllabification

2.2.9 Vowels: Summary

2.3 Rules involving coronal consonants

2.3.1 s-Voicing

2.3.1.1 s-Voicing before CiV Lengthening

2.3.1.2 s-Voicing before i-Shortening

2.3.1.3 s-Voicing with prefixes

2.3.1.4 s-Voicing: Summary

2.3.2 Velar Softening

2.3.3 Spirantization, y-Insertion and Palatalization

2.3.4 Coronals: Summary

2.4 Rules involving nasals

2.4.1 Nasal Assimilation

2.4.2 Noncoronal Deletion

2.4.3 Nasals: Summary

2.5 Chapter summary

Chapter 3. Evidence for productivity

3.0 Introduction
3.1 The concept of productivity.................................................................172
3.2 Linguistic evidence for productivity..................................................176
  3.2.1 Generalizations across real words..............................................177
  3.2.2 Novel forms in linguistics...........................................................188
  3.2.3 Linguistic evidence for productivity: Summary.......................192
3.3 Psycholinguistic evidence for productivity......................................192
  3.3.1 Experiments.................................................................................193
    3.3.1.1 The production task............................................................194
    3.3.1.2 The judgment task...............................................................196
    3.3.1.3 The memory task.................................................................199
    3.3.1.4 The perception task.............................................................202
    3.3.1.5 The concept formation task.................................................205
    3.3.1.6 Experiments: Summary.......................................................213
  3.3.2 Speech error accommodations.................................................214
    3.3.2.1 Vowel length and Vowel Shift...........................................217
    3.3.2.2 s-Voicing and Velar Softening............................................220
    3.3.2.3 Spirantization, y-Insertion and Palatalization....................221
    3.3.2.4 Nasal Assimilation and g-Deletion.......................................227
    3.3.2.5 Evidence from speech error accommodations: Summary.........231
3.3.3 Malapropisms........................................................................................................232
  3.3.3.1 Vowel Shift....................................................................................................238
  3.3.3.2 s-Voicing.....................................................................................................241
  3.3.3.3 Velar Softening............................................................................................242
  3.3.3.4 Spirantization...............................................................................................242
  3.3.3.5 y-Insertion....................................................................................................244
  3.3.3.6 Palatalization...............................................................................................247
  3.3.3.7 Nasal Assimilation and Postnasal g-Deletion...........................................248
  3.3.3.8 Evidence from malapropisms: Summary................................................250

3.4 Chapter summary......................................................................................................250

Chapter 4. Fully productive rules in English................................................................254

4.0 Introduction.............................................................................................................254

4.1 Automatic Feeding..................................................................................................256
  4.1.1 Palatalization and y-Deletion...........................................................................257
  4.1.2 Coronal Deletion and two assimilation rules.................................................261
  4.1.3 Automatic Feeding: Summary.........................................................................266

4.2 Prosody First............................................................................................................266
  4.2.1 Resyllabification and Flapping.........................................................................269
  4.2.2 Resyllabification and Coronal Deletion..........................................................275
4.2.3 Prosody First: Summary.................................................................277

4.3 Word-internal rules precede word-external rules..........................278
  4.3.1 Raising and Flapping in Canadian English...............................279
  4.3.2 Regressive Nasalization and Flapping.................................288
  4.3.3 Prevelar Raising and Nasal Assimilation...............................293
  4.3.4 Word-external rules last: Summary.........................................294

4.4 Word-internal rules and the Productivity Hypothesis...............295
  4.4.1 Fortition precedes lenition.....................................................296
  4.4.2 Word-internal rules: Summary.............................................300

4.5 Chapter summary...........................................................................300

Chapter 5. Prepatterning, productivity and cyclicity...............303

5.0 Introduction..................................................................................303

5.1 Derived environment effects........................................................308

5.2 Rule ordering paradoxes.............................................................311
  5.2.1 Real cyclicity and fake cyclicity............................................315
    5.2.1.1 Nasal Assimilation and Palatalization in English.............318
    5.2.1.2 Syncope and Umlaut in Icelandic.................................325
    5.2.1.3 o-Lowering and Umlaut in Schaffhausen....................332
    5.2.1.4 Fake cyclicity: Summary.............................................334
  5.2.2 Idiosyncratic ordering............................................................335
5.2.2.1 Contraction and i-Spirantization in Finnish
5.2.2.2 a-Rounding and Umlaut in Kesswil
5.2.2.3 Idiosyncratic ordering: summary
5.2.3 Rule ordering paradoxes: summary
5.3 Direct evidence for cyclic rule application
5.3.1 Prepatterned fully productive rules
5.3.2 Secondary stress in English
5.3.3 Coronal Deletion in English
5.3.4 Direct evidence: Summary
5.4 Chapter summary

Chapter 6. Theories of rule ordering
6.0 Introduction
6.1 The Standard Model
6.2 Simultaneous rule application
6.3 Predictable rule ordering
6.4 Rule typology
6.5 Chapter summary

References
ABSTRACT

This dissertation proposes a formal model of phonological performance, Double Lookup, that also has empirical consequences for theories of phonological competence. The most significant of these is the Productivity Hypothesis, the claim that the ordering of rules derives from their relative productivity.

According to Double Lookup, the use of phonological knowledge during speech production occurs in two steps. First, forms are retrieved from memory; second, phonological rules are retrieved from memory and applied, if appropriate, to the retrieved forms. Phonological patterns may be applied during speech in this way or be prepatterned (stored as patterns across lexical items in memory). The productivity of a rule is defined to be the likelihood of its being retrieved and applied during speech production. In general, less productive rules are more likely to be prepatterned than more productive rules. The Productivity Hypothesis then follows: Because prepatterned forms are retrieved before rules are retrieved and applied, less productive rules will be ordered before more productive rules.

Double Lookup and the Productivity Hypothesis are tested in several ways. First it is shown that the ordering of partially productive rules in English, as determined using standard linguistic methods, corresponds with their ranking in productivity, as determined through experiments described in the literature and through original surveys of speech errors.
The application of fully productive rules in English is also shown to be consistent with the Productivity Hypothesis; fully productive rules do not apply in a linear sequence, but rather interact in accordance with universal principles. All apparent counterexamples actually involve less than fully productive rules.

Next it is shown that the phenomenon referred to in the literature as cyclicity is correctly predicted to arise under certain well-defined circumstances, as when a rule is both prepatterned and very productive. In addition, it is shown that there are large categories of examples that cannot be handled by the notion of cyclicity at all, but find a simple account within Double Lookup.

Finally, evidence for the model is summarized by comparing it with other models of rule ordering which face conceptual and empirical problems Double Lookup avoids.
CHAPTER 1
AN INTRODUCTION TO DOUBLE LOOKUP

1.0 Introduction

The goal of this dissertation is to propose and defend a formal model of phonological performance that is specifically designed to have empirical consequences for theories of phonological competence as well. The model concerns the way the knowledge of phonological patterns is used in the actual production of linguistic forms, but is couched in terms that make predictions about this knowledge itself. Consequently, I will intentionally use the term application of phonological rules with a systematic ambiguity, to refer both to the knowledge of phonological patterns and to the use of this knowledge in speech processing.

The model makes claims that are testable both through psycholinguistic studies as well as traditional linguistic ones. In fact, some of its predictions are only testable by comparing results from both sorts of studies. One such claim, the major focus of this dissertation, is that the ordering of phonological rules, as determined by the standard methods of generative phonology, follows from their relative productivity, as determined by psycholinguistic experiments and speech error studies.

A model of this sort naturally raises questions about the appropriateness of the separation between psychological and linguistic theories. In 1.1 I discuss some of these questions, noting that the debate is very old and far from resolved. I conclude that although the existence of a single model that accounts for both
competence and performance in some linguistic domain is not required \textit{a priori}, evidence for such a model would nevertheless be most welcome. The remainder of the dissertation is then devoted to providing some of this evidence for a specific model of phonological rule application.

In 1.2 I introduce this model, which I call \textit{Double Lookup}. After describing each aspect of the model, I motivate it by citing arguments found in earlier works that have made similar claims. In Section 1.3 I list the aspects of Double Lookup I will focus on in the rest of the dissertation. This leads directly into an outline of the remaining chapters of this dissertation.

1.1 The relation between linguistics and psycholinguistics

It has long been recognized that theoretical linguistics, assumed by most of its practitioners to be a branch of cognitive psychology, is in practice distinct from other psychological disciplines in its aims and methods. According to the standard stereotype, theoretical linguists find the results of experimental psycholinguistics trivial and unsophisticated, while psychologists find the theorizing of linguists to be little more than a form of hubris, given the unreliability of the empirical foundations. Miller (1990) characterizes the fundamental difference as follows:

What is holding up the free flow of ideas back and forth between linguists and psychologists? For what it is worth, my own view is that linguists and psychologists subscribe to different theories of explanation. Linguists tend to accept simplifications as explanations.... To an experimental psychologist, [a linguistic theory like] X-bar
theory is not an explanation; rather, if it is true, it is something to be explained. (Miller 1990:321)

Although this split between psychology and linguistics has a long history (Derwing and Baker 1978), it has been perhaps strengthened in the generative era by the influential proposal that the study of language be divided up into the study of linguistic competence, or the knowledge of language, and the study of linguistic performance, or the use of language (Chomsky 1965). Linguistic competence is studied by using the traditional methods of structural linguists; for phonologists this would include the search for sound patterns in words in a dictionary transcribed in some standard orthography and in intuitions about the pronunciation of sentences. Linguistic performance, on the other hand, is what psycholinguists study; within the domain of phonology, this would include experiments testing the productivity of phonological rules and the examination of speech errors affecting phonological forms. A finding in one area does not necessarily affect the theorizing of the other, since linguists and psycholinguists are thought to be studying different things.

There is much that is useful in this distinction. There is no logical necessity that an entity that plays a role in one's knowledge of language is actually used during the processing of the language, and there may well be entities that come into being solely to enable language processing and have no parallel in competence at all.

However, there is a fundamental problem with maintaining the competence/performance distinction too strictly: the investigation of
competence can only be carried out through the study of performance. In order to know what people know about their language, one must observe some aspect of how they use it; the knowledge itself is not directly accessible. This means that data that bear on competence issues is just one kind of performance data. Indeed, some researchers have questioned whether the traditional sources of performance data used in the study of competence (dictionary work and introspection, in the case of phonology) are really more reliable than many other sources that could be envisioned (see Zwicky 1975 and Ohala 1986a for lists). Shortly we will see that there are in fact some basic questions of interest to investigators of linguistic competence that cannot be answered using these methods.

This is as far into the metatheoretical debate over the linguistics-psycholinguistics split as I dare go; for further discussion, see Stampe (1973:43ff), Derwing (1979), Chomsky (1980:203-205), Stemberger (1982:1-14), McCawley (1986), Bagemihl (1988:18-26) and Ohala (1990), among numerous other works. All I wish to add here is one small point, which is, however, the primary point of this dissertation. Recognition of the possibility of cases where theories of competence and performance (specifically, processing) fail to overlap completely in a given linguistic domain means only that there is no logical necessity for a single theory of competence and performance to exist. There is, however, no logical necessity for such a theory not to exist. It is simply an empirical question. Moreover, the
motivation for looking for such a theory is of the familiar Occam's razor type: If such a single theory for competence and processing were found, it would be preferred to a model where two separate theories are required.

One recent example of such a theory in the phonological literature is found in Golston (1991). The theory presented there overlaps in many ways with models argued for in Garrett (1980) and other works, whereby lexical stems are argued to be inserted into sentences before function words and inflectional affixes. One of the novel things about Golston (1991), however, is the source of evidence used to support this model. Whereas Garrett (1980) relies solely on psycholinguistic evidence such as speech errors, much of the evidence presented in Golston (1991) is of a sort very familiar to linguists: the distribution of affixes, patterns in phrasal phonology, restriction of phonological rules to different lexical categories (in this case, closed class versus open class morphemes), and so forth. Golston (1991) has succeeded in demonstrating the fruitfulness of a model of linguistic competence that closely mirrors a well-motivated model of linguistic performance.

This dissertation aims to do something very similar. As noted above, Double Lookup is intended to be a model comprehensible only if standard linguistic evidence and data produced by psycholinguistic methods are considered equally valid sources of information about the same object of inquiry. In this case, the object of inquiry is the ordering of phonological rules.
1.2 Double Lookup

In this section I describe the various tenets of the Double Lookup model. As we will see, every one of these tenets has precedents in the literature. The key aspects of Double Lookup that I will discuss are Lexical Lookup and Rule Lookup, Gradient Retrievability and Gradient Productivity, and Redundancy of Patterning. Together these will lead to the Productivity Hypothesis, the claim that the ordering of a pair of rules follows from the rules' relative productivity.

1.2.1 Lexical Lookup and Rule Lookup

The heart of the Double Lookup model is the claim that the production of phonological forms is accomplished in two separate but equally important steps. In Lexical Lookup, phonological forms are sought and retrieved from memory. In Rule Lookup, rules are sought in memory and then applied, if appropriate, to the forms that were retrieved during Lexical Lookup. Both processes occur on-line, that is, during the production of actual spoken utterances.

In considering Lexical Lookup and Rule Lookup to be separate, ordered steps, Double Lookup is thus no different from any theory that posits that the knowledge of phonological patterns (ie, rules) is distinct from the knowledge of phonological forms. There are, of course, current linguistic theories that deny the separation between knowledge of forms and knowledge of rules, for example analogical models (eg, Skousen 1989), connectionist models (eg, Rumelhart and McClelland 1987), and others (eg, Bybee 1988). The question of the
merits of Double Lookup relative to such ruleless models is far beyond the scope of this dissertation. Instead, I will concentrate my energies on comparing Double Lookup with other models accepting the existence of linguistic rules.

1.2.2 Gradient Retrievability and Gradient Productivity

It has long been known that forms stored in memory vary gradiently in their ease of retrieval, as measured, among other things, by how long it takes to call them up from memory (see Forster 1989 for a review); factors influencing a lexical item's ease of retrieval include most notably its frequency of occurrence and the context (phonological, syntactic, semantic, extralinguistic) in which the item is to be used. It is thus not quite accurate to say that a given lexical item is or is not part of one's knowledge; often the most one can say is that the item is more or less easy to access. I will call this observation Gradient Retrievability.

If phonological rules are also to be stored in memory, as any theory claiming psychological reality for rules as distinct from forms must do, it seems reasonable to suppose that rules are subject to Gradient Retrievability as well. This means that the likelihood that a given rule will apply on-line is at least partly due to the ease with which the rule can be retrieved from memory.

If we define the productivity of a rule as the likelihood with which the rule will apply on-line, then we are forced to conclude that rule productivity can be gradient too. I will call the doctrine of
Gradient Retrievability as applied to rule application \textit{Gradient Productivity}.

Two aspects of this conception of productivity are important to emphasize. First, productivity is defined as the likelihood of \textit{on-line} application. The reliability of a rule in describing patterns found in real words is thus not necessarily of any help in determining the rule's productivity, since it is possible that the words have simply been memorized with the pattern already present. Second, productivity is defined as a \textit{gradient} property. Rules need not be fully productive or fully unproductive; many, if not most, fall somewhere in between.

These two assumptions underlying my conception of productivity are commonly made in the literature. For example, the idea that the productivity of a rule cannot be reliably measured by studying real words finds a parallel in Bromberger and Halle's (1989) argument that phonology and syntax are different. One major aspect of this difference lies in the fact that phonological competence, much more than syntactic competence, is dependent on memory. Syntax involves patterns across sentences, and since they are potentially infinite in number, it is impossible to memorize all of the sentences of one's language. Therefore any syntactic patterns seen across sentences must be part of one's knowledge of language as well. By contrast, it is quite possible to memorize all of a language's lexical items. In fact, by definition the only thing distinguishing between a real word and a nonreal but possible word is that the former is
memorized while the latter is not. Thus there is no way to guarantee that a phonological pattern found across such memorized items is not simply memorized with them.

The key difference of interest to me is of course not that between sentence-building and word-building, but between forms that may be memorized and forms that are novel and therefore cannot be. Syntactic rules do not always create novel sentences and morphological rules are not always associated with nonnovel ("real") words. Thus in practice, as Zwicky (1992) has pointed out, people tend to use the same syntactic frames over and over, allowing for the possibility that such frames are not derived on-line but simply memorized. Similarly, rules of inflectional morphology typically seem to generate words on-line (Marcus et al. 1990, Stemberger and MacWhinney 1988); the same is presumably also true of compounding in English, and word-formation rules of various sorts in languages such as Yup'ik and Turkish, all of which allow for the formation of novel forms. The difference between sentence-building and word-building with regard to the production of novel forms is more a matter of degree. My point is that finding a phonological pattern across words does not in and of itself preclude the possibility that the pattern is simply memorized with the words.

Recognition of this problem has led many researchers to reject definitions of productivity that rely on the behavior of patterns across real words. Aronoff (1976:37), for example, defines the productivity of a morphological rule as the "likelihood of [a form
derived by application of the rule] being a word of the speaker's active vocabulary." He notes that attempting to determine the productivity of such a rule by examining lists of real words misses a fundamental insight, namely that if a word is derivable by a very productive rule, there is no need to list it at all. The question of whether *Ciceronically* is a real word cannot be answered by looking through a published word list, since it simply will not be there; yet the intuition is clear that if *Ciceronical* is a real word, then so is *Ciceronically*, simply because -ly suffixation is very productive.

Moreover, since likelihood is a gradient property, Aronoff (1976) consistently refers to productivity in a way that emphasizes its gradient nature, as for instance in his long discussion of the relative productivity of -ity suffixation and -ness suffixation.

Kiparsky (1975) is even more explicit in treating rule productivity as a gradient property that must not be defined in terms of real ("listed") items. He makes this point in his response to critics who, motivated by the failure of many experiments to confirm the productivity of phonological rules proposed by generativists, have denied their existence entirely:

...what many recent discussions... ignore is that productivity [sic] is traditionally and correctly viewed as a gradient phenomenon. The strongest sense in which one might speak of a rule being productive is that any new word or formation which meets the structural analysis of the rule *must* undergo it. Another, weaker, sense is that in which we speak of a rule as being productive if new words or formations *can* become subject to it, that is, if the scope of the rule is being extended in the language.
shall distinguish the two cases as *full* and *partial* productivity, respectively. Thus, in my terms there are degrees of partial productivity depending on how strongly the rule spreads. (Notice that since applicability to *new* items is made the criterion, even fully productive rules need not be *automatic* in the usual sense of that term, for they can clearly have exceptions in the old vocabulary.) [Kiparsky 1975:195]

Since productivity is defined in terms of on-line rule application, the easiest way to measure it is to examine forms that display patterns that must have been put there on-line. Forms that meet this criterion I will call *novel forms*. This term covers all phonological forms, word-sized or otherwise, that could not have been memorized, so that any phonological patterns found in them must be derived through on-line rule application. Novel forms thus include borrowings, coinages, experimental nonce forms, nonwords produced by speech errors, and novel concatenations of words into phrases.

The use of novel forms appears to be crucial in demonstrating that phonological rules are greater than fully unproductive. Consequently, when phonologists find themselves in the position of having to argue that phonological knowledge is rule-based, they must make reference to novel forms.

One clear example of this is found in the discussion in Kenstowicz and Kisseberth (1979:26-31) of what they call the *Null Hypothesis*, the claim that language has no phonological component at all. According to this hypothesis, syntax and morphology arrange morphemes that are extracted directly from the lexicon with their
pronunciations already fully specified. The simplest way to argue against this, as Kenstowicz and Kisseberth (1979) demonstrate, is to show that there are cases where phonological patterns could not have been memorized with the forms they appear on. Kenstowicz and Kisseberth (1979) give four examples of this; naturally all of them involve what I am calling novel forms.

First, they point out that phonological rules can apply to novel morphemes, such as voicing assimilation in the pronunciation of *two fips* [fɪps] and *two mags* [mægz].\(^1\) Halle (1962) makes a similar argument using the morpheme structure constraints that allow *plast* but rule out *ptak* as possible words of English. Second, phonological rules can be triggered by syntactic processes, as vowel harmony is triggered by noun incorporation in Chukchee, or as phrasal phonology is triggered by syntactic concatenation in almost any language. Since syntax can create novel sentences, any pattern triggered by it cannot have simply been memorized. Third, speakers often continue to apply phonological rules from their native tongue to forms in the second language. And finally, phonological rules can be triggered by speech errors.

The conclusion they derive from these examples is that as phonological patterns could not have spread to these nonmemorized forms if phonological rules did not live a life at least partly independent of lexical items, the Null Hypothesis is therefore false.

\(^1\)Kenstowicz and Kisseberth's (1979) use of *mags* as an example is of course somewhat unfortunate, since it is in fact a real word ("magazines").
There is, however, another important lesson to be learned from their arguments.

As noted above, the simplest way to argue against the Null Hypothesis is to examine what happens to novel forms. What this means is that any phonologist who is interested in dealing with such a hypothesis must go beyond word lists, often to what is called "external" evidence such as speech errors. In fact, except for rules that apply across word boundaries, there is no way to argue against the Null Hypothesis with respect to many types of rules unless external evidence (including psycholinguistic evidence) is taken into account. This follows directly from the difference between phonology and syntax noted earlier. Syntactic claims are standardly tested against novel sentences; at least some phonological claims, such as the Null Hypothesis, must be tested against novel phonological forms. As Bromberger and Halle (1989) observe, however, phonology is different, and this is because, unlike with syntax, novel forms triggering many phonological rules of interest are created only rarely in natural speech. Thus evidence for the productivity of such rules must be sought in experiments (e.g., checking people's judgments on novel forms such as ptak or asking them to produce the plural of nonwords such as fip), in speech error studies, in studies of foreign accents -- in short, in psycholinguistics.

The various aspects of Double Lookup that we have seen so far, namely Lexical Lookup, Rule Lookup, Gradient Retrievability and Gradient Productivity, find close parallels in the generative
phonological literature. I now move into somewhat more unorthodox territory, though, as I hope to show, the move is very natural given where we have been up to this point.

1.2.3 Redundancy of Patterning

We have seen that for rules that do not apply across word boundaries there is no way to be certain, without the use of external evidence, that the patterns they capture are not simply stored with forms in memory. Thus far I have only presented this as a logical possibility. I now claim that such cases of what I will call prepatterened phonological rules do occur, in fact quite commonly. This doctrine, that surface phonological generalizations can result both from the on-line application of phonological rules as well as from the generalizations' being prepatterned across forms stored in memory, I will call Redundancy of Patterning.

This concept, like the other concepts underlying Double Lookup, is hardly without precedent in the literature. Standardly, however, the redundancy is thought to be limited to a specific well-defined subclass of rules. That is, when ideas like Redundancy of Patterning appear in the literature, they typically take the following form: Some surface phonological generalizations are always produced by on-line rule application, while others are always prepatterned, though they still may be part of a speaker's phonological knowledge. I will illustrate this conception with brief descriptions of three distinct, though strikingly similar, theories of rule typology (discussed further in Chapter 6).
Sapir (1925) proposes a typology based on the distinction between rules governing the distribution of phonemes and rules deriving allophones from these phonemes. As McCawley (1986:30) notes in his commentary, these rule types differ in an important way for this discussion. Namely, only allophonic rules can be thought of as deriving one representation from another, in that they change the phonemic representation into the surface representation; morphophonemic rules do not derive forms, but show relations between forms in memory. Thus this typology seems to imply that while both morphophonemic and allophonic rules are part of a speaker's knowledge, only the latter are productive in the sense that they apply on-line.

The model of Natural Phonology (eg, Stampe 1973, Donegan and Stampe 1979) has a similar division of rule types. In this model, what are called rules (roughly equivalent to morphophonemic rules), which do not necessarily apply on-line, are distinguished from processes, which are explicitly claimed to apply on-line (Stampe 1973:43-45). The evidence used to support this distinction in productivity depends as usual on the behavior of rules and processes with respect to novel forms:

Processes apply to tongue-slips..., to Pig Latins, to foreign words, etc.... Rules do not ordinarily apply in these cases. This leaves open the question of whether rules in fact apply in speech production. [Donegan and Stampe 1979:144]
Finally, this conception of Redundancy of Patterning shows up again in the model of Lexical Phonology (e.g., Kiparsky 1982, Mohanan 1982, Kaisse and Shaw 1985). Here what are called lexical rules (again roughly equivalent to morphophonemic rules) are contrasted with postlexical rules. Unlike Natural Phonology, however, Lexical Phonology does not claim to be a model of speech processing. Nevertheless, while postlexical rules are claimed to have all the properties of fully productive rules, applying freely to novel forms such as phrases, lexical rules have the hallmarks of less productive rules, such as lexical exceptions.

Some workers within Lexical Phonology have been more explicit in claiming that lexical rules and postlexical rules have a different psychological status. Mohanan (1982) claims that postlexical rules may be triggered by speech error forms, while lexical rules may not be. This concept comes out clearly in the following passage from Shattuck-Hufnagel (1986):

...the output of earlier strata [posited by Lexical Phonology] might correspond to the forms stored in the production lexicon that is accessed during phonological planning, while the operations described in later strata might correspond more closely to processing mechanisms. [Shattuck-Hufnagel 1986:145]

The special nature of the knowledge of lexical rules as essentially nonproductive (in my sense) is described by Kiparsky (1982) as follows, where he considers what it means to speak of the "application" of a lexical rule like Velar Softening, the rule relating
pairs such as *criti[s]ize-*criti[k]al (to be discussed further in the following two chapters):

This does not imply that the speaker or hearer need in any way mentally "derive" the words he says or hears by means of such rules as Velar Softening. What it does mean is that the alternations they govern belong to the regular phonological pattern of English, while for example a hypothetical $k\sim s$ alternation in the reverse context, such as *criti[k]ize-*criti[s]al, would be irregular. The claim made is that someone who knows English implicitly knows that pattern, and will under appropriate circumstances recognize the difference between regular and irregular alternations, though he may not be able, even after reflection, to verbalize the rules that underlie it. [Kiparsky 1982:34-35]

In short, what all these various models agree on is this: There are two fundamentally different types of phonological rules. One type can be said to be part of a speaker's phonological knowledge by virtue of its productivity; rules of this type always apply on-line. The second type, by contrast, is used by speakers in a very different way. Such rules virtually never apply on-line, but instead are used to relate forms stored in memory.

To my mind, there are two basic flaws with typologies such as this. The first is that they neglect the reality of Gradient Productivity. Such typologies assume that rules are either fully productive or essentially not productive at all, but there is abundant evidence against there being such a clear-cut dichotomy. In Chapter 3, for example, we will consider speech error and experimental evidence suggesting that what Lexical Phonology would treat as
lexical rules vary greatly in their relative productivity; while Velar Softening does in fact seem to be very unproductive, other rules with no parallel in the postlexical stratum, such as y-Insertion, appear to be relatively productive. It might be countered that the lexical/postlexical distinction divides between fully productive rules (postlexical) and all others (lexical), but even this division will not work; in Chapter 4 I will show that many if not most apparently postlexical rules are really less than fully productive, raising the question of whether there really are any rules that always apply online.

The second problem with such typologies is the supposition that rules can be part of a speaker's knowledge of phonology without being productive. I commented earlier that Gradient Retrievability implies that it makes no sense to speak of "knowing" a lexical item in terms other than its ease of retrievability. The same thing should be true of rules. If rules can be known and used in recognizing patterns across forms stored in memory, as the passage from Kiparsky (1982) suggests, then it seems strange that such rules cannot also be retrieved from memory and applied to novel forms. As it happens, this dilemma is immediately resolved by the facts; as noted above, lexical rules do in fact apply to novel forms, just at a lower rate than postlexical rules. And yet if this fact is granted, it does not seem possible to consider lexical rules psychologically distinct from postlexical rules. The difference resides less in the rules themselves than in the fact that lexical rules, being less productive, are
associated with more prepatterned forms than the more productive postlexical rules. That postlexical rules do not typically "relate forms in memory" is not a fundamental defining property of such rules; it simply follows from the fact that postlexical rules, being highly productive, typically do not have prepatterned forms to relate.

Considerations raised by Gradient Retrievability and Gradient Productivity thus lead to a slightly broader interpretation of Redundancy of Patterning than is standardly found in the literature. By rejecting a rule typology where there are two different types of rules, one of which is always prepatterned and one of which always applies on-line, Double Lookup assumes that any given surface phonological generalization may arise either through prepatterning or through on-line rule application. All else being equal, the more productive the rule, the more likely the generalization arises through on-line rule application. We are forced to this conclusion because we cannot say that some rules are fully productive while others are fully unproductive. If most rules fall somewhere between, this means that for any given form, sometimes the rule applies and sometimes it doesn't. If the rule is nevertheless associated with a descriptively accurate surface generalization, then it must be the case that sometimes the generalization results from prepatterning and sometimes from on-line rule application.

If we allow both that a rule may sometimes change a form and other times not, and that nevertheless in both cases the same surface representation may result, we must suppose that lexical items may
be represented in memory by different allomorphs, some prepatterned according to the rule and others not. These allomorphs would presumably be stored as distinct entities in memory, with Gradient Retrievability being applicable to each one individually. Thus in the production of a single phonological form subject to a single phonological rule, Lexical Lookup may have two allomorphs to choose between. In Rule Lookup, each of the allomorphs may then either undergo the rule or not, depending on the rule's productivity.

Even this concept of allomorphy is not without precedent in the literature. Aronoff (1976) argues for the necessity of a class of rules called allomorphy rules that select between different forms of morphemes stored separately in the lexicon; these forms are then subject to phonological rules across the whole range of productivity. Hayes (1990), working within the Lexical Phonology model, is even more explicit. He notes that the existence of rules that have some properties of postlexical rules (in particular, cross-word application) as well as some properties of lexical rules (eg, the tolerance of lexical exceptions) implies that it is possible for certain rules to be what he calls "precompiled" in the form of multiple allomorphs stored separately in the lexicon. These allomorphs are chosen between during the insertion of lexical items into utterances, thus allowing for there to be influences from information outside the word boundary. This class of rules will be discussed more fully in Chapter 6, where we see that it supports the Double Lookup model much better than Lexical Phonology. Hayes (1990:94) himself is agnostic on the
question of the psychological nature of his allomorph analysis. Double Lookup of course is not; it claims that allomorphs really are stored separately in memory and that the selection between them occurs on-line during Lexical Lookup.

In spite of the relative indeterminacy of this conception of Redundancy of Patterning, there is a general principle that should hold: If a surface generalization does not come from on-line rule application, then it must come from prepatterning, and vice versa. This means that if a rule is not prepatterned at all, it must be fully productive, and if a rule is not productive at all, it must be prepatterned. It is true that a very productive rule may nevertheless be prepatterned as well (we will come across this situation several times during the course of this dissertation, most notably in Chapter 5), but in general we should expect productivity and prepatterning to be complementary. Thus the more productive a rule is, the less prepatterned it should be, and the less productive a rule is, the more prepatterned it should be.

Consider for example the very productive English rule of Coronal Deletion, to be discussed in detail in Chapter 5, which deletes word-final /t,d/. The optional application of this rule allows for two pronunciations of the word hand, namely [hænd] and [hæn]. According to Redundancy of Patterning, each of the forms could be stored as different allomorphs, separately retrievable in Lexical Lookup. In Rule Lookup, each of these allomorphs could then either
undergo the rule or fail to do so. This results in four possible derivations for these forms, as illustrated below in (1).

(1) Derivations of *hand* in Double Lookup.

Because Coronal Deletion is a rather productive rule (as we will see in later chapters), and because of the complementary nature of productivity and prepatterning, we should expect that it is typically the left-hand box, that is the non-prepatterned form, that is available to be chosen in Lexical Lookup. If both forms are available, it is of course solely their relative frequency that determines which will be chosen, in accordance with Gradient Retrievability. This relative frequency, however, is not the frequency with which a given allomorph appears on the surface, but the frequency with which a given allomorph has been retrieved in the past. This frequency is dependent on the productivity of the associated rule; if the rule is very productive, the prepatterned form will thus be less frequently retrieved.
Likewise, we would expect that it is typically the right-hand box that is chosen in Rule Lookup. Hence the path of derivation during speech production should typically follow the dark lines in (1).

Again, as noted above, it is quite possible for even a fully productive rule to be prepatterned; we will see in Chapter 5 that the factors determining whether a rule is prepatterned go beyond its productivity. In fact, Double Lookup, as a theory of rule application, does not make any specific claims about prepatterning, which should more properly be included in a theory of how forms are stored in memory. Thus in my discussion I merely illustrate the ideal case, where the application of an on-line rule is assumed to change one form into another; the assumption of this change requires the assumption that the form retrieved by Lexical Lookup is distinct from the surface form, and hence is not prepatterned.

Let's compare this with the much less productive rule of Vowel Shift (evidence of the relative unproductivity of this rule is discussed in Chapter 3). This rule accounts for alternations such as [iy]-[ε] seen in pairs like sleep-slept; at this point the rule can be thought of as changing [iy] to [ε] when followed by a consonant cluster; a much more thorough description of this pattern is found in Chapter 2. The possible derivations of slept, according to Double Lookup, appear as in (2) below.
Since Vowel Shift is a less productive rule, the on-line derivation of *slept* will tend to follow the dark lines in (2), with the prepatterned form being retrieved in Lexical Lookup and the rule not being retrieved and applied during Rule Lookup. As with the derivation of *hand*, the other paths in (2) are possible but unlikely; when they are taken, the output may be considered a speech error.

In this section, then, we have seen how Redundancy of Patterning follows from Gradient Retrievability and Gradient Productivity. As with these others aspects of Double Lookup, parallels to Redundancy of Patterning have appeared in the literature for a long time, though perhaps not in quite as strong a form as is adopted here. Together these principles allow us to see how Double Lookup treats phonological patterns in on-line processing, as represented in the figures in (1) and (2). In the next section I will show how they also allow us to derive a very strong claim about the ordering of phonological rules.
1.2.4 The Productivity Hypothesis

Let us now consider a consequence of the previous discussions for a topic of central importance in generative phonology: rule ordering. Consider two rules, A and B, where A is very unproductive (ie, rarely applies on-line) and B is very productive (ie, usually applies on-line). Given what we have said about Lexical Lookup, Rule Lookup and the rest, what could we say about the relative ordering of A and B?

Since A is less productive than B, we would expect it to be more prepatterned. That is, the surface generalization captured by rule A typically arises through the choosing of prepatterned allomorphs in Lexical Lookup. By contrast, the surface generalization captured by rule B typically arises through the on-line application of B during Rule Lookup.

The result of A being associated with Lexical Lookup and B being associated with Rule Lookup is that rule A will appear to be ordered before B. More precisely, it will appear that rule B applies to forms that have already undergone rule A.

The idea that relative productivity of rules can lead to their apparent ordering also has a long history in the literature. For example, in the three models discussed above, namely Sapir's (1925) morphophonemic/allophonic model, Natural Phonology and Lexical Phonology, the ordering of rules of the unproductive type before rules of the productive type does not need to be stipulated. For example, if lexical rules never apply on-line but postlexical rules
always do, then lexical rules must necessarily precede postlexical rules.

Lexical Phonology recognizes another more general, but weaker, connection between ordering and productivity in addition to the ordering of lexical and postlexical rules. In order to account for the interaction between phonological and morphological rules, Lexical Phonology supposes that the rule block of lexical rules is further subdivided into morphologically defined rule blocks, called levels or strata. One property of these levels that has often been remarked upon is that the earlier levels contain less productive rules than the later levels. Thus -ity suffixation, which triggers Level 1 rules, can be shown to be much less productive than -ness suffixation, which triggers Level 2 rules (see Chapter 3). Borowsky (1986) provides another example. In responding to charges that certain patterns analyzable with Velar Softening are restricted to so few forms that there is no need to suppose that the pattern is associated with a rule at all (see Chapter 2), she comments:

Within our framework, level 1 phonology is characteristically unproductive and often limited, yet nonetheless phonologically explainable, so I consider the velar softening cases interesting evidence. [Borowsky 1986:125fn36]

Regarding the relative productivity of rules in the earlier and later levels in English, Goldsmith (1990) makes an interesting observation. He points out that prosodic rules such as rules of syllabification and stress assignment tend to treat forms derived by
Level 1 morphology as if they were monomorphemic, while they treat forms derived by Level 2 morphology as polymorphemic. He gives the example of the Level 1 suffix -ism, meaning "philosophical position," which does not tolerate a preceding schwa, changing Buddha+ism into Buddhism, just as the sequence [əɪ] is not allowed in monomorphemic forms. By contrast, its Level 2 namesake -ism, meaning "peculiar trait," has no trouble with this, changing Indiana+ism into Indianaism. The same holds for stress rules: Cathóli[s]ism, the philosophical position, has a legal monomorphemic stress pattern, while Cát'holi[k]ism, the peculiar trait, does not. In other words, morphologically derived forms in Level 1 tend to be treated by prosodic rules as if they were not derived. Befitting the nondeterministic nature of Gradient Productivity, however, this is only a tendency; in the succeeding chapters we will encounter much evidence that words derived by less productive affixation processes may nevertheless be treated as polymorphemic by prosodic and other rules.

Notice that in the view of Lexical Phonology, the relation between productivity and ordering is not necessarily conceived of as causal. Whereas the ordering of lexical before postlexical rules can be thought of as following from the fact lexical rules are unproductive and postlexical rules are productive, the ordering of Level 1 before Level 2 is not considered to be derivable from the fact that Level 1 rules are much less productive. Moreover, the ordering
of rules within each level is thought to be entirely arbitrary, at least with relation to the rule's relative productivity.

This is where I break into new ground. The close relation between productivity and ordering, especially given what we have seen about the rest of the Double Lookup model, suggests the possibility that all rule ordering is derivable from differences in productivity. I will call this bold claim the *Productivity Hypothesis*.

(3) **Productivity Hypothesis** (preliminary version)

Rules are ordered solely in accordance with their relative productivity, with less productive rules appearing earlier than more productive rules.

Before going on, it is important to note what sort of ordering principle this is. The ordering of a less productive rule before a more productive one is in a sense *intrinsic*, since it follows automatically from intrinsic properties of the rules through application of a principle. In practice, however, the ordering is really *extrinsic*, or arbitrary, since there may well be no simple way to predict whether or not a given rule is productive purely from its form. It remains true, however, that according the the Productivity Hypothesis it is unnecessary to know both the productivity of a rule and its ordering relative to other rules, since the properties are derivable from each other.

The closest direct parallel in the literature to this view of the causal relation between productivity and ordering is the Elsewhere
Condition, discussed in Kiparsky (1973a), among other places.\textsuperscript{2} In its many slightly different versions, this condition stipulates that if the conditioning environment of a rule A is contained within that of another, more general rule B, then rule A must be ordered before B. One major difference between the Elsewhere Condition and the Productivity Hypothesis is that while the latter is intended to be a statement about phonological processing, the former is a purely formal principle; given just the form of two rules, the Elsewhere Condition can determine whether or not their ordering is predictable. Another difference is that according to the Elsewhere Condition, the more specific rule A is supposed not only to precede the more general rule B but also to preempt it; if A applies, B will not, even if its structural description is still met. By contrast, the Productivity Hypothesis makes no such claim about disjunctivity. Although the Elsewhere Condition cannot therefore be subsumed under the Productivity Hypothesis, their parallels are still striking: all else being equal, a rule with a more restricted environment than another, as in the Elsewhere Condition, will have fewer chances to apply online, and will therefore be less productive. It thus may be possible to reduce these two principles to one, though I will not attempt this here.\textsuperscript{3}

As an example of the use of the Productivity Hypothesis, consider the two rules discussed in the previous section, Vowel Shift

\textsuperscript{2}This parallel was suggested to me by Chris Golston.
\textsuperscript{3}See Janda (1987) for a discussion of other principles similar to the Elsewhere Condition that nevertheless appear to be distinct from it.
and Coronal Deletion. Given Redundancy of Patterning, any given lexical item can potentially be stored as prepatterned according to either or both of these rules. Likewise, this lexical item can undergo either or both of these rules on-line. Since Vowel Shift is much less productive than Coronal Deletion, we would expect that most of the time Vowel Shift will be prepatterned and Coronal Deletion will apply on-line. This will result in Vowel Shift appearing to be ordered before Coronal Deletion. Using conventional notation, forms will be expected to undergo a derivation as in (4a) rather than (4b).

\[(4)\]

\[a.\]

\[
\begin{array}{l}
\text{UR} \\
\text{Vowel Shift} \\
\text{Coronal Deletion}
\end{array}
\]

\[
\begin{array}{l}
/sliypt/ \\
/slept/ \\
[sl\text{e}p]
\end{array}
\]

\[b.\]

\[
\begin{array}{l}
\text{Coronal Deletion} \\
\text{Vowel Shift} \\
\text{Output}
\end{array}
\]

\[
\begin{array}{l}
/sliypt/ \\
/sl\text{iyp} \\
[sl\text{iyp}]
\end{array}
\]

Two questions arise with regard to this example. First, what is the allomorph of \textit{slept} that is prepatterned according to both Vowel Shift and Coronal Deletion? The first thing to note is that it is not possible to speak of just one such allomorph. The way two interacting rules are prepatterned depends on the way they interacted before they were both prepatterned. That is, if at some point in time rule A were prepatterned and rule B were not, resulting in the apparent ordering of A before B, then if at a later stage rule B becomes prepatterned as well, the prepatterning will
continue to reflect this ordering. Thus the form prepatterned according to both Coronal Deletion and Vowel Shift may be either /sliyp/, where Coronal Deletion appears to have applied before Vowel Shift, or /slep/, where Vowel Shift appears to have applied before Coronal Deletion.

The question posed above is therefore just a special case of a more general question: Where do prepatterned forms come from? A reasonable hypothesis seems to be as follows. Prepatterned forms are more typical of less productive rules because with the knowledge of the rule weak, a learner is liable to store a surface form, with its pattern and all, as an underived form in memory. If prepatterned forms can be associated with more productive rules as well, as I will show, then the same thing should be true of them; prepatterned forms are the stored representations of surface forms. If the patterns seen in these surface forms are rule-governed at all, it must be that at some point in the past the patterns were put there through on-line rule application, but knowledge of this history behind the pattern is irrelevant to the storing of the prepatterned form.

The hypothesized stored allomorphs of slept can be expected to have a similar origin. If [sliyp] and [slep] appear on the surface, they may be stored as /sliyp/ and /slep/ by a language learner. Note that given the relative productivity of the two rules, the form that is more likely to appear on the surface is [slep]. As a reasonable first guess, we might therefore suppose that this is the form that is more likely to be stored in memory. The form
prepatterned according to both Coronal Deletion and Vowel Shift is therefore the one where it appears that Vowel Shift has applied first.

The second question that arises is parallel. We have seen that if Lexical Lookup retrieves a form that is prepatterned according to both Vowel Shift and Coronal Deletion, this form will tend to be prepatterned so that it appears that Vowel Shift has applied before Coronal Deletion. What happens, then, if neither Vowel Shift or Coronal Deletion are prepatterned, so that both rules apply on-line?

It is clear that if all rule ordering follows from productivity, then fully productive rules cannot be ordered at all. The problem resides in the proper way to understand what it means for rules to apply and not be ordered. Two possibilities suggest themselves.

The simplest hypothesis is that two on-line rules apply at the same time. Unfortunately, as I will show in Chapter 4, this hypothesis cannot be maintained. There are rule pairs that must be considered as applying on-line where simultaneous rule application will not explain their interaction.

A slightly more complex hypothesis, one consistent with the facts, is the claim that on-line rules interact in a principled fashion so that all are able to apply. As Koutsoudas, Sanders and Noll (1974), Donegan and Stampe (1979) and others have pointed out, this can be achieved by allowing the rules to apply and reapply until no more applications are possible. Although the precise ordering the rules have actually taken in any particular instance is not possible to reconstruct, the effect of this sort of rule application "free-for-all" is
that rule pairs will always appear to be *feeding* and/or *counterbleeding*. That is, if a rule A creates the environment allowing a rule B to apply, A will be ordered before B (a feeding order); if A destroys the environment allowing B to apply, B will be ordered before A (a counterbleeding order).

I will refer to this hypothesis, that on-line rules interact in such a way that all have a chance to apply, by the mnemonic *Automatic Feeding*. I test this in Chapter 4 by examining the relative ordering of apparently fully productive rules. Naturally, in the case of Vowel Shift and Coronal Deletion, this situation will rarely if ever arise, since Vowel Shift is so much less productive than Coronal Deletion.

The relation between productivity and ordering can thus be expressed as in the following fuller formulation of the Productivity Hypothesis.

(5) **Productivity Hypothesis** (final version)

The ordering of the segmental rules A and B is determined in the following way:

a. If A is less productive than B, then A is ordered before B.

b. If A and B are synchronically so unproductive so as to be primarily prepatterned, then the ordering of A and B follows from the relative productivity of A' and B' according to (5a), where A' and B' are the diachronic predecessors of A and B, respectively.
c. If A and B are both fully productive, then they are not ordered. Instead they apply in accordance with Automatic Feeding.

What this principle means, then, is that all ordering follows solely from productivity. The specific way this works depends on the relative productivity of the rules in question.

First, (5a) means that it is impossible to find a pair of rules where the less productive rule follows the more productive one. This is the central prediction of the Productivity Hypothesis.

Second, (5b) means that equally or almost equally productive rules can only be ordered if synchronically both rules are very unproductive and if diachronically they are derived from rules that differed in productivity in accordance with (5a). This follows from the discussion above concerning the origins of prepatterning. If two rules ordered by (5a) later come to be equally unproductive, this means that forms conforming to these rules will typically surface through allomorph selection in Lexical Lookup. Since such allomorphs presumably originate as surface forms that language learners have stored “as is” in memory, they will continue to encode the ordering pattern that had earlier been actively derived by (5a).

Further, we expect that if the earlier of the rules in such a pair were somehow to increase in productivity, the historically determined ordering relation could not be maintained. Either a new rule corresponding to the historically earlier rule would be added to the end of the derivation, or else the rules themselves would switch
order. Both such situations have been found in real languages, as I will show in Chapter 5.

Finally, (5c) means that it is impossible to find a pair of fully productive rules which are ordered other than in an apparent feeding or counterfeeding order. As I show in Chapter 4, segmental rules that are ordered any other way are not equally productive (for the behavior of prosodic rules see below).

Because the Productivity Hypothesis is a derived concept and not a fundamental axiom, there are principled exceptions to it. Specifically, the Productivity Hypothesis is derived from the claim that ordering results from an earlier rule being prepatterned while a later rule need not be. Since, as noted above, rules can be prepatterned even if they are productive, it is possible for a more productive rule to be ordered before a less productive one or for two equally productive rules to be ordered if in both cases the earlier rule is prepatterned.

Such exceptions only arise in two specific cases, however. First, as we will see in Chapter 5, it is possible for particular lexical items to show prepatterning associated with a rule ordering that is no longer synchronically relevant. That is, if rule A was once less productive than rule B, and so was ordered before it, but since then the relative productivities have changed and with them the ordering, it is still possible for there to be some lexical items that preserve the older ordering of A before B. I will call this idiosyncratic ordering. Idiosyncratic ordering can only arise if two factors hold. First,
neither rule can be productive enough to regularize these lexical items into the current pattern. Second, the lexical items preserving the older ordering must be so easy to retrieve during Lexical Lookup that there is no need to derive the relevant aspects of their phonological form on-line in Rule Lookup. This second factor means that items that show the idiosyncratic ordering tend to be more frequent than items that show the synchronic ordering; as I noted above, items are easier to retrieve if they are more frequent. Thus in these cases the lexical items must have both rules A and B prepatterned. Such cases of idiosyncratic ordering, to be discussed in Chapter 5, are a natural consequence of Redundancy of Patterning, in particular the role Lexical Lookup plays in it.

A second, more interesting sort of exception to the Productivity Hypothesis concerns the relative ordering of prosodic and segmental rules. As it happens, prosodic rules appear to be ordered before segmental rules even if both are equally productive. This pattern will be seen throughout the discussion of English rules in this dissertation: whether the rules are partially productive or fully productive, prosodic rules always precede segmental rules. I will call this observation the Prosody First principle.

The necessity of a principle like this has long been recognized; researchers as varied as Donegan and Stampe (1979) and Kaisse (1985) both argue that during on-line rule application, prosodic rules apply first. In this dissertation I will assume the attitude towards this phenomenon expressed in the following passage:
The application of prosodic processes is the most important factor in the living phonological pattern of a language and its long-range phonological 'drift'; the selection of segmental processes is largely determined, even in childhood, by the way segmental representations are mapped onto prosodic structure in speech.... However, since the remainder of our discussion is mostly concerned with segmental issues, we must turn to the processes which govern segments. [Donegan and Stampe 1979:142]

That is, I assume that the prosodic constituency of the speech signal is basic to the application of all other segmental processes; consequently, prosodic structure must either be present or be built before segmental structure can be built or changed. This may have to do with the apparent fact, encoded in concepts such as prosodic licensing (Ito 1986), that segmental processes can only take place within the framework provided by prosodic structures. If no framework exists underlyingly, as within phrases, it must be built before any segmental process can take place. Following the lead of Donegan and Stampe (1979), however, I will not discuss the principle of Prosody First any further than to show in Chapter 4 that it is correct, working to order even fully productive on-line rules. For most of this dissertation I will focus instead on segmental rules.

By way of summary, I now give a figure showing what the Productivity Hypothesis predicts about the relative ordering of Vowel Shift and Coronal Deletion. The figure is given below, with the most likely derivational path indicated by the dark lines. Numerous additional possibilities have been left out due to lack of space.
Derivations of *slept* in Double Lookup.

[VS = Vowel Shift, CD = Coronal Deletion]

Notice that even though this figure represents the behavior of two different rules, there is only one instance of rule application during Rule Lookup. Moreover, the different allomorphs chosen between during Lexical Lookup are not ordered either, but rather are stored "side by side" at the same level in memory. The only ordering that occurs is that between Lexical Lookup and Rule Lookup. The Productivity Hypothesis claims that all examples of rule ordering follow from this universal ordering.

This concludes the introduction to Double Lookup. The remainder of the dissertation is devoted to providing empirical arguments for aspects of this model, in particular the Productivity Hypothesis, using both psycholinguistic and linguistic evidence. I outline these arguments in the final section of this chapter.

1.3 Chapter summary

In this chapter I have had the luxury of introducing the Double Lookup model without providing much in the way of empirical...
support. In the remainder of this dissertation I will attempt to provide this support, discussing other interesting conceptual issues raised by this model as they come up.

In Chapters 2 and 3 I demonstrate a direct method of arguing for the Productivity Hypothesis by comparing the ordering of several partially productive phonological rules in English, as determined by standard linguistic arguments, with their relative productivity, as determined by psycholinguistic methods such as experiments and speech error studies. The reason for focusing on English is simple: it has by far the best studied set of phonological rules, both from a traditional generativist point of view as well as from a psycholinguistic perspective. In Chapter 2 I will introduce and motivate the rules and argue for their ordering, while in Chapter 3 I will argue for the relative productivity of the rules on the basis of psycholinguistic evidence. We will see that though the currently available psycholinguistic evidence for the relative productivity of these rules is less than fully satisfying, the evidence we do have strongly suggests that the Productivity Hypothesis is on the right track.

In Chapter 4 I will examine the question of what happens with rules that are identical in productivity, in particular fully productive rules, focusing again on rules in English. The primary criterion used in this chapter to indicate full productivity is the ability to apply across word boundaries, since word concatenations constitute novel forms (see above). What I find is that rule pairs where both rules
apply across word boundaries always conform to Automatic Feeding and Prosody First. If, however, one of the rules applies solely within word boundaries, implying that it is less productive, it is ordered earlier. All examples given in the literature of rules that are purported to apply on-line but which are ordered in unpredictable ways involve rule pairs where both rules apply solely within word boundaries and thus need not apply on-line. Such examples thus do not constitute counterevidence to the Productivity Hypothesis unless independent evidence can be given for the relative productivity of the rules, as I do in Chapters 2 and 3 for many of the word-internal rules of English proposed in Chomsky and Halle (1968). Indeed, the behavior of truly on-line interacting rules is always in accordance with the principles of Double Lookup.

In Chapter 5 I discuss rule cyclicity, where a rule appears to be extrinsically ordered in more than one place in a derivation. Such ordering patterns are disallowed by the Productivity Hypothesis, since a rule cannot be more or less productive than itself and so can only appear once in a derivation. On the basis of examples in Icelandic, dialects of German, and other languages discussed in the literature, I show that apparent cases of cyclicity are the natural result of allowing a rule both to be prepatterned as well as to apply on-line. In contrast to the claim made by the cyclicity hypothesis, therefore, the Double Lookup analysis supposes that apparent multiple applications of a single rule are actually the result of theoretically distinct ways of giving rise to the same pattern. In
some cases the prepatterned version of the rule is historically older, with the on-line version being added later through a reanalysis of the surface pattern, while in other cases a historically older on-line version has become prepatterned although the on-line version has remained as well. This analysis subsumes cases credited to cyclicity in the literature, and also allows for other cases that cannot be analyzed with cyclicity, namely where the prepatterned and on-line rules are not identical, though too similar to dismiss as a chance occurrence. In addition, I show that there are other cases ascribable to cyclicity that actually involve what I above called idiosyncratic ordering. Thus apparent cases of cyclicity actually argue instead for the fundamental correctness of the Double Lookup model.

Finally, in Chapter 6 I sum up the arguments made for Double Lookup by comparing this model with other models making claims about rule ordering, including the ones I have briefly mentioned in this chapter. The strengths and weaknesses of these other approaches will be examined; Double Lookup will be seen to have all of the strengths and none of the weaknesses.

The Double Lookup model, claiming as it does to be an all-encompassing theory of rule application in both performance as well as in competence, will take much more motivation than I have been able to provide in this one short chapter. Let us therefore waste no time in turning to the first step in the complex set of arguments to come.
CHAPTER 2
PARTIALLY PRODUCTIVE RULES
IN ENGLISH

2.0 Introduction

The Productivity Hypothesis claims that the relative productivity of rules predicts their ordering, with less productive rules ordered before more productive rules. As we saw in Chapter 1, this is uncontroversially the case with the ordering of partially productive rules like Vowel Shift before fully productive rules like Coronal Deletion. The Productivity Hypothesis goes further, however, claiming that even partially productive rules are ordered according to their relative productivity, while fully productive rules are not ordered at all. In this chapter and the next I deal with the ordering of partially productive rules in English; the ordering of (apparently) fully productive rules in English is left until Chapter 4.

Experimental and speech error evidence that the rules discussed in this chapter are less than fully productive will be given in Chapter 3, but their lesser productivity is also suggested by two linguistic properties they share. First, all of the rules discussed here apply solely within words.1 Since word sequences are a kind of novel form, rules that apply across word boundaries must be very productive; there is no such guarantee for rules that do not apply across word boundaries. Second, many of the rules discussed in this

---

1Two of the word-internal rules discussed in this chapter, Palatalization and Nasal Assimilation, have been claimed to apply across word boundaries as well, but I will argue in Chapter 5 that the word-external rules are in fact distinct from the word-internal rules.
chapter are sensitive to morphological information. All else being equal, a rule that is sensitive to morphological information will be less productive than one that is not, since its structural description will be more specific, thereby allowing it to apply under fewer circumstances. Moreover, the morphological processes associated with the rules discussed in this chapter tend not to be used to create novel words (see discussion of -ity and -ness suffixation in Chapter 3). Hence the forms that show the phonological patterns discussed here are not novel forms, meaning the rules may well be prepatterned; in Chapter 3, I will show that this is in fact the case.

The first step needed to test the Productivity Hypothesis with regard to partially productive rules is to determine what phonological patterns need to be accounted for with rules at all and which, if any, of these rules need to be ordered.

This exercise will have two other subsidiary purposes. First, it will allow me to demonstrate that rule ordering can be a descriptively adequate device. I will show that in at least some of the cases I discuss rule ordering is the simplest way to account for the way rules interact. Theories that totally disallow rule ordering (discussed more in Chapter 6) thus face an empirical challenge.

2 Recall the parallels between the Productivity Principle and Elsewhere Condition mentioned in Chapter 1.
3 Computer searches were conducted with the use of the Moby Pronunciator 1.01 database of approximately 110,000 English words. This database was accessed with the Grep/Word Count 1.1 program on a Macintosh LC. In some cases pronunciations were double-checked against those in Kenyon and Knott (1953) and Woolf et al. (1981).
Second, I will show that some rules for English proposed in Chomsky and Halle (1968) and which continue to be assumed in Halle and Mohanan (1985), Borowsky (1986) and other recent works do not in fact describe valid generalizations about English.

One thing I wish to emphasize very strongly is that the methodologies employed in this chapter to argue against the necessity of certain rules are entirely orthodox within generative phonology. Some of the same rules have also been argued against on the basis of experimental or other psycholinguistic findings, to be reviewed in the next chapter; I show in this chapter that in many cases such studies are not even necessary to argue against these rules. Often one needs nothing more than the standard methodologies of generative phonology to demonstrate that a proposed rule is not in fact an adequate description of distributional patterns and allomorphic variations.

I divide the rules I will examine into three groups. In 2.2 I examine rules affecting vowels, specifically rules affecting vowel height and rules shortening or lengthening vowels. I show that the evidence does indeed support the ordering of the vowel length rules before the vowel height rules, as is standardly assumed. However, it appears that the vowel length rules cannot be accounted for in terms of automatic syllabification processes, contra S. Myers (1987).

In 2.3 I examine rules that in some way involve the coronal obstruents /t, d, s, z/, specifically the rules referred to in Halle and Mohanan (1985) as s-Voicing, Velar Softening, Spirantization, y-
Insertion and Palatalization. Of these rules, I show that only Spirantization, y-Insertion and Palatalization behave essentially as described in the literature. First, I show that s-Voicing is much more restricted than commonly described, applying regularly only with the rather small class of forms with a latinate prefix. This means that evidence for its interaction with vowel lengthening rules, Velar Softening, and Spirantization is much weaker than standardly assumed. Second, I note that Velar Softening, too, is much more restricted in its domain of application than often assumed. However, neither of these observations will have much affect on the primary focus of interest in this chapter, which is these rules' relative orderings.

In 2.4 I discuss two rules involved in the realization of nasal consonants, namely the rules Halle and Mohanan (1985) call Nasal Assimilation and Noncoronal Deletion. I show, partly following Borowsky (1986), that there must be two separate applications of Nasal Assimilation, one obligatory and one optional; in Chapter 5 I show that these are actually two separate rules. I argue that the literature appears to be right in that in some cases [ŋ] can indeed be described as being derived from /Ng/, where /N/ represents an underlying nasal consonant unspecified for place. However, in other cases [ŋ] does not seem to be derived, suggesting that it is an underlying phoneme in English, though like [ʒ], a marginal one.

A common theme running through all three sections is the importance of syllabification in the application of these rules. Hence
in 2.1 I begin by briefly discussing two important syllabification processes of English. Syllabification processes will be seen to precede each of the segmental rules discussed in the chapter.⁴

In 2.5 I conclude with a list of rules of partially productive rules in English whose descriptive adequacy has been checked as thoroughly as possible, arranged so that all of the necessary ordering relationships are clear. The orderings shown in this list will then be compared in Chapter 3 with the information about the productivity of these rules obtainable from available speech error corpora and published experiments. We will find that the available data, while not conclusive, are strongly suggestive of the correctness of the Productivity Hypothesis.

2.1 Syllabification

As noted above, many of the rules to be examined in this chapter depend for their application on syllabic information: some rules only apply within syllables, while others only apply across syllable boundaries. In this section I therefore focus on two important aspects of the syllabification process, namely the Onset Principle and stress-sensitive Resyllabification.

One traditional source of evidence for such processes in English involves the characterization of aspiration. Examples such as those

⁴The issue of whether syllabic patterns arise through application of rules, the meeting of conditions (cf eg S. Myers 1991) or some other system will not arise in my discussion. I will refer to generalizations describing syllabic patterns such as Resyllabification and the Onset Principle as "rules" and "processes," though nothing I say about them precludes the possibility that they are conditions instead.
below in (1a) show that voiceless obstruents in English are aspirated if syllable-initial; if not syllable-initial, as in (1b), they are not aspirated.

(1) [after Selkirk 1983:363-365]

a. **Aspirated:**

Toronto, pathetic, calamity
repair, recant
atrocious, apply, accretion, improve, betwixt, acquaint

b. **Unaspirated:**

stir, spinach, skate
act, actor, activity, Atkins, napkin, ichthyosis
cat, pack, map

Forms like repair and recant, which have aspirated segments, suggest that a consonant between two vowels is syllabified as the onset of a syllable containing the second vowel as nucleus. This principle, long recognized in the literature (see Selkirk 1983 and Itô 1986 for reviews), is formalized in Itô (1989) as the **Onset Principle**. The idea is that in the absence of other information, an intervocalic consonant will syllabify as the onset of the following syllable (indicated with "σ") rather than as the coda of the preceding one.

---

5There are of course numerous other sources of evidence for this principle, such as the fact that the first syllable in a CVCV sequence is light for stress rules, but it would detour the discussion much more than necessary to go into these here.
(2) **Onset Principle**

\[ \text{[= (3) in Itö 1989:223]} \]

\[ \text{Avoid } \sigma[v] \]

There is a well-defined class of exceptions to this principle, however. In contrast to forms like *repair* and *recant*, where the vowel following the aspirated obstruent is stressed, forms like those given in (3) show that an intervocalic consonant is not aspirated when the following vowel is unstressed. This indicates that under these circumstances it is not a regular syllable onset.

(3) [after Selkirk 1983:364]

wacky, attitude, happen, interpolate, elliptical, actor, wimpy

Kahn (1976) argues that under the stress circumstances illustrated in (3), the relevant consonants become ambisyllabic, being both onset of the following syllable as well as coda of the previous one. By contrast, Selkirk (1983) argues that the influence of stress is to resyllabify the relevant consonants solely as the coda of the previous syllable. Since the precise formulation of this process is not crucial to the discussion in this and later chapters, I will simply adopt a version of Kahn's (1976) version of *Resyllabification* without argument. I give this version below.\(^6\)

---

\(^6\)In addition to these analyses implicating Resyllabification, Kiparsky (1979) argues that the relevant prosodic level in this process is not the syllable but the foot; see Hammond (1982) for critical discussion of this alternative. Again, all that is relevant here is that prosody, in particular the stress pattern, is crucial.
The extent to which this rule applies depends on two factors. The first is the relative sonority of the originally final segment of the first syllable. The onset consonant is obligatorily resyllabified if this final segment is very sonorous, such as a vowel; it is only optionally resyllabified if the final segment of the first syllable is an obstruent, nasal or /l/. This is illustrated with the examples below. If an obstruent is preceded by a stressed syllable ending in a vowel, aspirating it sounds very unnatural, as in the forms in (5a); if it is preceded by a stressed syllable ending in an obstruent, nasal or /l/ as in the forms in (5b), however, aspiration is optional.

(5) [after Selkirk 1983:366]

a. **Vowel:**

happy, mightiest, accolade, beaker, goiter, wacky

b. **Obstruent, nasal, /l:**

elliptical, aptitude, actor, after, napkin, Rifkin contemplate, pantomime, winter, anchor, lanky filter, altitude, poltergeist, alcohol

The claim that segments can vary in their sonority and that this has effects on syllabification has been widely studied. The *sonority hierarchy* for English (ie, the ranking of segments from most to least sonorous) is generally taken to be as given in (6).
(6) **Sonority hierarchy**
[after Selkirk 1983 and Goldsmith 1990:111]

vowels
- low vowels
- mid vowels
- high vowels
glides
liquids
nasals
obstruents
- voiced fricatives
- voiceless fricatives
- voiced stops
- voiceless stops

The second factor determining the application of Resyllabification is the presence or absence of a word boundary. Even under the appropriate stressing and sonority conditions, Resyllabification appears to apply across word boundaries less readily than it will word internally. Contrast the forms in (7a), where aspiration is virtually not allowed at all, with the examples in (7b), where aspiration is possible. In both sets of examples, the relevant consonant is immediately preceded by a stressed vowel and is followed by an unstressed vowel, so any difference must be ascribed solely to the presence or absence of the word boundary.

(7) a. vóter, béisér, híker, súper
b. gó tomarrow, sée today, hígh catástrophe, súe podiatrists

Kaisse (1985) assumes that this means that Resyllabification is marked so as to be restricted to the word-internal domain, making it essentially a lexical rule in the terms of Lexical Phonology. Hammond (1982), however, shows that this assumption is not
necessary. We may assume instead that Resyllabification applies prior to destressing of the initial syllable in the words in (7b). At the stage when Resyllabification would apply, then, the syllable following the relevant consonant is not in fact stressless, thus blocking the application of Resyllabification.⁷

In any event, the important point is that regardless of the reasons, a word-initial consonant is not readily resyllabified as the coda of the final syllable in the previous word. This fact will play an important role in some of the discussions in Chapter 4, where we will see evidence that segmental rules that are restricted by their dependence on Resyllabification to apply solely within words are less productive than rules that are not dependent on Resyllabification in the same way and thus can apply across word boundaries.

In this section, then, I have introduced the concepts of the Onset Principle and a stress-sensitive rule of Resyllabification. Both will appear repeatedly in the discussions of partially productive rules to follow.

2.2 Rules involving vowel length and vowel height

In this section I examine the evidence for rules affecting vowel length and vowel height in English. I find that there is good evidence supporting the claim that the vowel length rules precede the vowel

---

⁷Hammond (1982) supposes that forms like tomorrow, toady allow Resyllabification more readily than otherwise phonologically similar forms like tomato because they may be analyzed as beginning with the inherently stressless particle to; alternatively, we might suppose that the high frequency of these words has caused the effects of initial destressing to become prepatterned. See further discussions in Chapters 4 and 5.
height rules. Moreover, I find new support for the claim in S. Myers (1987) that the vowel length rules follow at least to some extent from more general principles of syllabification, including application of Resyllabification. However, neither the vowel length nor the vowel height rules will be reducible to a single rule or principle, contrary to what has been claimed in the literature.

Another notable finding is the close intertwining of the vowel length and vowel height rules. It is very rare to find alternations motivating one set of rules independently of the other. Because of this, I find that I must discuss these rules together. I will call the alternations that these rules together work to derive the Vowel Shift alternations.

I begin in 2.2.1 with a discussion of these alternations and the proper formulation of the vowel height rules. I then turn to the vowel length rules. In 2.2.2 I consider Cluster Shortening, in 2.2.3 Trisyllabic Shortening, in 2.2.4 -ic Shortening, in 2.2.5 i-Shortening, in 2.2.6 CiV Lengthening, and in 2.2.7 compensatory lengthening triggered by Prenasal g-Deletion. In 2.2.8 I explore the suggestion in S. Myers (1987) that the various vowel shortening rules can be reduced to a single process of Closed Syllable Shortening, triggered in some cases by Resyllabification. Although I find suggestive evidence that seems to support a syllable-based analysis of the vowel length rules, the specific way vowel height interacts with these rules will be seen not to be derivable from syllabically relevant factors like sonority. I thus conclude in 2.2.9 that there are several distinct rules
responsible for the vowel length and vowel height patterns involved in the Vowel Shift alternations.

2.2.1 The Vowel Shift alternations

Since Chomsky and Halle (1968), numerous theoretical phonologists (e.g., Halle 1977, Rubach 1984, Halle and Mohanan 1985 and S. Myers 1987) have been interested in vowel alternations like those illustrated below in (8).\(^8\)

(8) [after (2) in Jaeger 1986:78]

\[
\begin{array}{ll}
[ay]\text{-}[i] & \text{divine-divinity} \\
[iy]\text{-}[e] & \text{serene-serenity} \\
[ey]\text{-}[æ] & \text{sane-sanity} \\
[aw]\text{-}[ʌ] & \text{profound-profundity} \\
[uw]\text{-}[ʌ] & \text{reduce-reduction} \\
[oy]\text{-}[ʌ] & \text{destroy-destruction} \\
[uw]\text{-}[a/ə] & \text{lose-lost} \\
[ow]\text{-}[a/ə] & \text{verbose-verbosity}
\end{array}
\]

Chomsky and Halle (1968) note that such alternations involve differences in tenseness, rounding, backness, vowel height and diphthongization. Halle (1977), Halle and Mohanan (1985) and S. Myers (1987) argue that the fundamental distinction is vowel length, with the vowel quality differences resulting from this prosodic dimension. That is, in deriving, say, *divine* and *divinity* from the same underlying stem, rules apply that result in a difference in vowel length in the second syllable; this length difference then

---

\(^8\)The precise set of alternations considered worthy of examination varies from author to author. In (8) I list all that have been mentioned.
triggers the application of later rules that result in the vowel quality differences. I will follow this proposal throughout my discussion.

The diphthongization, backing, rounding and tensing are handled by rules I will not discuss here. I will instead focus on the vowel height alternations, which are claimed to be handled by means of the unitary rule of Vowel Shift, which is given in Halle and Mohanan (1985) essentially as in (9). This rule affects long vowels in the following manner: long low vowels become mid vowels, long mid vowels become high vowels, and long high vowels become low vowels (eventually surfacing as [ay] or [aw]). Vowel length is represented in (9) by doubly linking the vowel segment to two timing units (specifically, moras, represented with "μ").

(9) Vowel Shift
[after (61) in Halle and Mohanan 1985:78]

\[
\begin{align*}
\{[\alpha_{\text{high}}] \to [-\alpha_{\text{high}}] \} & \quad \text{μ μ} \\
\{[-\text{low}] \to [-\text{high}] \} & \quad \text{\textbackslash /}
\end{align*}
\]

It is well known that there are numerous problems with this rule. One major problem is that although the rule as written should affect back vowels the same way as front vowels, this is simply not the case, as can be seen in the alternations in (8). While the front vowels show a simple alternation in vowel height (with diphthongization and backing of the low vowel), characterization of the back vowels requires numerous additional "readjustment rules"
(Chomsky and Halle 1968, Halle 1977, Halle and Mohanan 1985).\(^9\) In order to restrict the discussion of the Vowel Shift alternations to a reasonable length, I will ignore the complex issue of the back vowels through the remainder of this chapter. The vowel alternations I will consider, therefore, are just those illustrated below in (10).

(10) \[\text{[ay]-[t]} \quad \text{divine-divinity} \]
\[\text{[iy]-[e]} \quad \text{serene-serenity} \]
\[\text{[ey]-[æ]} \quad \text{sane-sanity} \]

Other major problems with the rule as given in (9) are formal in nature. Consider first the fact that the two subparts of this rule must apply in order. The intention behind the rule is to change low vowels into mid vowels, mid vowels into high vowels, and high vowels into low vowels. This is accomplished by having the first subpart exchange mid vowels and high vowels and the second subpart exchange low vowels and mid vowels. The conversion of high vowels into low vowels crucially requires the two subparts to be ordered, since the first subpart changes high vowels into mid vowels which the second subpart then converts into low vowels. The rule of Vowel Shift is thus seen actually to consist of two rules, extrinsically ordered with respect to each other.

Second, each of the subparts is itself problematic. Each subpart has the effect of exchanging feature values, the first one, for example, converting \([+\text{high}]\) to \([-\text{high}]\) and \([-\text{high}]\) to \([+\text{high}]\)

\(^9\) As we will see in the next chapter, experimental studies have also consistently shown more problems with people's knowledge of the back vowel alternations than with the front vowel alternations.
simultaneously. As Anderson and Browne (1973) demonstrate, if Chomsky and Halle's (1968) analysis is correct, the Vowel Shift rule represents the sole example of a purely phonological exchange rule known in any language. Chances seem good, then, that Chomsky and Halle's (1968) analysis is not correct. However, Anderson and Browne's (1973) suggested alternative seems just as unlikely:

In particular, if the high vowels become centralized diphthongs before lowering, and the mid and low vowels are simply raised one step, no exchange is involved, and the ordering problem [ie, the requirement that the switches involved in an exchange rule must apply simultaneously] does not arise. [Anderson and Browne 1973:447]

What this suggestion means is that mid vowels become high, low vowels become mid, and high vowels become mid diphthongs which then later become low diphthongs. Vowel Shift thus becomes four separate noncollapsible processes, two of which need to be ordered with respect to each other.

At this point my purpose is not to determine the proper formalism for Vowel Shift, but rather merely to prepare the reader for the systematic deconstruction of the vowel shift alternations that is to come. Hence throughout the remainder of my discussion I will merely assume the formalism given above in (9).

In the remaining subsections of section 2.2, I will examine the rules shortening and lengthening vowels. It is the evidence for these rules that really provide evidence for the Vowel Shift rule given in (9). This is because there is no morphological process in English
affecting vowel length directly. Hence all Vowel Shift alternations must involve other phonological rules, triggered by morphological processes like suffixation, that affect vowel length.

2.2.2 Cluster Shortening

In this section I review the evidence for the rule Halle and Mohanan (1985) call Cluster Shortening. I find that application of this rule is dependent on vowel quality in a hitherto unnoticed way.

Following Chomsky and Halle (1968), Halle and Mohanan (1985) propose this rule to account for examples like those given below. In all cases the underlined vowels become short when they appear before a consonant cluster.

(12) [from S. Myers 1987:489-490]

a. [ay]-[ɨ]

describe-descriptive, description
five-fifteen, fifty, fifth
scribe-scripture
wide-width
wise-wisdom

b. [iy]-[ɛ]

clean-cleanse
deep-depth
intervene-intervention
perceive-perceptive, perception
steal-stealth
thief-theft

10 In Chapter 3 we will consider speech error evidence that has been suggested as possible examples of the application of the Vowel Shift rules in the absence of the application of vowel length rules.
This rule may be formalized as given below.

(13) Cluster Shortening [first version]
[based on (56b) in Halle and Mohanan 1985:77]

\[
\begin{align*}
\mu & \rightarrow \mu \\
V & \rightarrow \emptyset \\
V & \rightarrow V / \_CC
\end{align*}
\]

As argued by Halle and Mohanan (1985), this rule feeds Vowel Shift. This can be seen in the fact that vowels that are shortened by this rule alternate with vowels that differ in vowel height in precisely the way predicted by the rule in (9). Specifically, long [iy] and [ay] alternate with short [ε] and [i], respectively. Thus if we suppose that the vowel in a word like perceive is underlingly /e:/ and that in describe is underlingly /i:/, we can account for the Vowel Shift alternations illustrated above through an ordered application of Cluster Shortening before Vowel Shift, as in (14).

(14) a. perceive perception
UR /e:/ /ε:/
Cluster Shortening --- e
Vowel Shift iy ---
Output [iy] [ε]

b. describe description
UR /i:z/ /i:z/
Cluster Shortening --- i
Vowel Shift ay ---
Output [ay] [i]

One strange thing about the examples discussed so far is the absence of the Vowel Shift alternation [εy]-[ε]. This is not an
oversight; there are in fact no examples of this alternation being triggered by Cluster Shortening. What we find instead is the alternation [ey]-[e], as illustrated by the examples below.

(15) abstain-abstention { *abst[ae]ntion }
detain-detention { *det[ae]ntion }
retain-retention { *ret[ae]ntion }

In the discussion of these facts in S. Myers (1987), it is suggested that the forms in (15) contain /e:/ underlyingly but are marked as lexical exceptions to Vowel Shift. This explains why the vowels are appropriately shortened before the consonant cluster (with the long vowel diphthongized and tensed) even though they do not undergo Vowel Shift. However, marking these particular forms as lexical exceptions does not avoid the generalization that Cluster Shortening never triggers the [ey]-[ae] alternation. That is, there appear to be no morphemes containing underlying /ae:/ that undergo Cluster Shortening.

I illustrate the generality of this by examining in detail cases of Cluster Shortening and Vowel Shift in the past tense forms of strong verbs. In fact, we will see that in this subset of cases illustrating Cluster Shortening, only the [iy]-[e] alternation is robustly represented. The [ey]-[ae] alternation is not found at all because there are no strong verbs in English that take the suffix -t where the present tense form contains [ey].

Halle and Mohanan (1985) and S. Myers (1987) argue that vowel alternations seen in past tense forms of strong verbs match the Vowel Shift alternations when the past tense forms are derived
through suffixation of the irregular past tense marker -t. This suggests that suffixation of -t triggers Cluster Shortening, which then gives rise to the Vowel Shift alternations. This is true even if the cluster never shows up on the surface, as when the stem already ends in /t/ or /d/. The short vowels derived in this way then surface in accordance with Vowel Shift. Sample alternations are shown in (16a), with derivations in (16b). I include the pair hit-hit as an example of a verb that already contains a short vowel in the present tense, and so is correctly predicted not to change in the past tense.

(16) a. leave-left  
meet-met  
bite-bit  
hit-hit  

b.  

| Verb | /leːv/ | /leːv + PAST 
leaves | levvt |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Shortening</td>
<td>--</td>
</tr>
<tr>
<td>Vowel Shift</td>
<td>liːv</td>
</tr>
<tr>
<td>Other rules</td>
<td>[liːv]</td>
</tr>
</tbody>
</table>

| Verb | /meːt/ | /meːt + PAST 
meets | mett |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Shortening</td>
<td>--</td>
</tr>
<tr>
<td>Vowel Shift</td>
<td>miːt</td>
</tr>
<tr>
<td>Other rules</td>
<td>[miːt]</td>
</tr>
</tbody>
</table>

| Verb | /biːt/ | /biːt + PAST 
bites | bitt |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Shortening</td>
<td>--</td>
</tr>
<tr>
<td>Vowel Shift</td>
<td>bayt</td>
</tr>
<tr>
<td>Other rules</td>
<td>[bayt]</td>
</tr>
</tbody>
</table>
The pattern is actually somewhat more complicated. First, as noted above, there are no strong verbs taking the -t suffix in the past tense whose present tense form contains [ey]; the only strong [ey] verb pairs are bear-bore and swear-sware. Hence there is no evidence for Cluster Shortening triggering the [ey]-[æ] Vowel Shift alternation. Second, the most robust alternation within the group of strong verbs taking the -t suffix involves strong verbs that end in [iyC], where C represents a consonant other than [t] or [d]. All examples of this type are given in (17).

(17) Consistent with Vowel Shift Inconsistent with Vowel Shift?
bereave-bereft hear-heard [iy]-[æ]
cleave-cleft
creep-crept
deal-dealt
dream-dreamt
feel-felt
keep-kept
kneel-knelt
lean-leant
leap-leapt
leave-left
mean-meant
sleep-slept
sweep-swept

By contrast, other strong verbs taking the -t suffix conform to the Vowel Shift pattern less regularly. I give all relevant examples below. The examples in (18) show strong [iy] verbs ending in [t] or
[d]. Two out of the nine verbs (22%) fail to shorten in accordance with Cluster Shortening, and one shows a non-Vowel Shift vowel quality alternation.

(18) **Consistent with Vowel Shift** | **Inconsistent with Vowel Shift**
---|---
bleed-bled | beat-beat [iy]-[iy]
breed-bred | 
feed-fed | 
et-ate (some dialects) | eat-ate (some dialects) [iy]-[ey]
lead-led | 
meet-met | 
read-read | 
speed-sped | 

The worst violations of the predictions of Cluster Shortening and Vowel Shift involve strong verbs whose present tense form contains [ay]. We expect that when these forms undergo -t suffixation, the vowel will become [i]. This is true in only five out of the fourteen relevant forms (36%).

All strong [ay] forms end in [t] or [d]. Of those that end in [ayt] or [ayd], only half show Cluster Shortening, triggering the appropriate Vowel Shift alternations, as illustrated below.

(19) **Consistent with Vowel Shift** | **Inconsistent with Vowel Shift**
---|---
bite-bit | fight-fought [ay]-[o]
chide-chid | ride-rode [ay]-[ow]
hide-hid | smite-smote [ay]-[ow]
light-lit | stride-strode [ay]-[ow]
slide-slid | write-wrote [ay]-[ow]

The remaining strong [ay] verbs end in [nd]. None of them show the Vowel Shift alternations.
To summarize the pattern in strong verbs: There are no strong [ey] verbs that undergo -t suffixation, so there is no evidence of Vowel Shift in these verbs. Strong [ay] verbs, all of which end in [t] or [d], do not show a robust Vowel Shift pattern with -t suffixation. Finally, strong [iy] verbs do show a robust Vowel Shift pattern with -t suffixation, but primarily if they do not already end in [t] or [d]. I claim that this pattern shows that only the [iy]-[e] Vowel Shift alternation is robustly supported in strong verbs.

It might be countered that the correlation of robust Vowel Shift alternations with the suffixation of -t to stems not ending in [t] and [d] suggests instead that Cluster Shortening actually reacts to surface clusters, like that in leave-left, rather than the nonsurfacing clusters hypothesized for pairs like beat-beat. The lack of [ey]-[æ] alternations is due to there being no appropriate strong [ey] verbs, and the nonrobustness of the [ay]-[ɪ] alternations is due to [ay] verbs always ending in [t] or [d].

This suggestion has a problematic consequence, however. Namely, the vowel alternations that do show up in verbs ending in [t] or [d], such as in meet-met and bite-bit, must now be considered as arbitrary as non-Vowel Shift alternations like hold-held. Once we allow for this possibility, the claim that phonological rules like
Cluster Shortening or Vowel Shift are necessary in the strong verbs at all becomes questionable.

Wherever we decide to have Cluster Shortening apply, then, the fact remains that the strong verbs only robustly support the [iy]–[e] Vowel Shift alternation. This is true even in the [iyt] and [iyd] verbs, which show the appropriate Vowel Shift alternation more often than [ayt] and [ayd] verbs (78% as opposed to 50%).

I summarize the findings of this entire subsection in (21). We see that the application of Cluster Shortening appears to be general except with the strong verbs derived with -t suffixation; within these verbs, only [iy] verbs (ie, verbs with underlying /e:/) undergo Cluster Shortening reliably. None of the cases of Cluster Shortening, whether involving strong verbs or not, show the [ey]–[æ] alternation (ie, show evidence that /æ/ undergoes Vowel Shift as it should). In the table, "S" indicates that a shortening rule occurs in the given environment with the given underlying vowel, while "VS" indicates the same thing for Vowel Shift.

(21)  

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/e/</th>
<th>/æ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Shortening</td>
<td>S</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>in general</td>
<td>VS</td>
<td>VS</td>
<td></td>
</tr>
<tr>
<td>Strong verbs with -t</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-t suffixation</td>
<td></td>
<td>VS</td>
<td></td>
</tr>
</tbody>
</table>

Leaving aside the special problems posed by strong verbs, I therefore suggest that the rule of Cluster Shortening be amended as below, so that it fails to apply with underlyingly [+low] vowels. As
we will see, this sort of restriction is hardly unique to Cluster Shortening; virtually all of the vowel length rules will be restricted to certain vowel classes.

(22) **Cluster Shortening** [final version]

\[
\begin{array}{c}
\mu \mu \\
/ \rightarrow \\
V / -CC
\end{array}
\]

2.2.3 **Trisyllabic Shortening**

Following Chomsky and Halle (1968), Halle and Mohanan (1985) suggest a rule of *Trisyllabic Shortening* to account for examples like those given below. In all cases the underlined vowels become short when they appear in the antepenultimate syllable of the word. As with Cluster Shortening, this rule feeds Vowel Shift as well.

(23) [from (13) in S. Myers 1987:494-495]

a.  [ay]-[i]

<table>
<thead>
<tr>
<th>prime</th>
<th>primitive</th>
</tr>
</thead>
<tbody>
<tr>
<td>oblige</td>
<td>obligatory</td>
</tr>
<tr>
<td>divide</td>
<td>divisible</td>
</tr>
<tr>
<td>vile</td>
<td>vilify</td>
</tr>
<tr>
<td>rite</td>
<td>ritual</td>
</tr>
</tbody>
</table>

---

11 Below in 2.2.8 I will explore the suggestion of S. Myers (1987) that Cluster Shortening is actually a special case of what he calls Closed Syllable Shortening. Although a syllable-based analysis of shortening like his or like that given in Archangeli (1991) is certainly possible here, I remain agnostic on this point for reasons discussed in 2.2.8. I consider the notation given in my final version of this rule adequate for my purposes.
b. [iy]-[e]

sincere-sincrity
saline-salinity
compete- competitive
repeat-repetitive
creed-credible
clear-clarify
supreme-supremacy

c. [ey]-[æ]

cave-cavity
opaque-opacity
profane-profanity
exclaim-exclamative
inflame-inflamatory
flame-flammable
saline-salify
grateful-gratitude

This rule may be formalized as given below.

(24) **Trisyllabic Shortening**
[based on (56a) in Halle and Mohanan 1985:77]

\[
\mu \mu \quad \mu
\]
\[
\nu \to \]
\[
\nu \quad \nu / \quad - \sigma \sigma \text{Word}
\]

Unlike Cluster Shortening, then, Trisyllabic Shortening does not appear to be sensitive to vowel quality. There are of course exceptions to Trisyllabic Shortening, but they appear to be idiosyncratic rather than systematic. For example, the list of anomalous pairs given in (25a) represents all exceptions to Trisyllabic Shortening or Vowel Shift (found through a computer search) involving a front vowel and the suffix -ity. One shows the
alternation [iy]-[i], indicating the application of Trisyllabic Shortening but not Vowel Shift, while the other shows the non-alternation [iy]-[iy], indicating the nonapplication of Trisyllabic Shortening. The derivations in (25b) show that chlorine may be thought of as containing an underlying /i:/, while obese has an underlying /æ:/.

The small number of examples makes it unclear whether the absence of an exception with /æ:/ is systematic or coincidental.

(25) a. chlorine-chlorinity  [iy]-[i]
obese-obesity  [iy]-[iy]

b. 

<table>
<thead>
<tr>
<th></th>
<th>chlorine</th>
<th>chlorinity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
<td>/i:/</td>
<td>/i:/</td>
</tr>
<tr>
<td>Trisyllabic Shortening</td>
<td>---</td>
<td>i</td>
</tr>
<tr>
<td>Vowel Shift rules</td>
<td>{exception}</td>
<td>---</td>
</tr>
<tr>
<td>Output</td>
<td>[iy]</td>
<td>[i]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>obese</th>
<th>obesity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
<td>/æ:/</td>
<td>/æ:/</td>
</tr>
<tr>
<td>Trisyllabic Shortening</td>
<td>{exception}</td>
<td>{exception}</td>
</tr>
<tr>
<td>Vowel Shift rules</td>
<td>iy</td>
<td>iy</td>
</tr>
<tr>
<td>Output</td>
<td>[iy]</td>
<td>[iy]</td>
</tr>
</tbody>
</table>

This generality is represented schematically below, where "S" and "VS" represent the same thing they did in the previous section.

(26) 

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/æ/</th>
<th>/e/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trisyllabic Shortening</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>VS</td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
</tr>
</tbody>
</table>
2.2.4 -ic Shortening

Again following Chomsky and Halle (1968), Halle and Mohanan (1985) suggest a rule of -ic Shortening to account for examples like those given below. In all cases the underlined vowels become short when they appear in the syllable preceding the suffix -ic or -id. Once again, this rule feeds Vowel Shift, and like Trisyllabic Shortening it does not appear to have any systematic exceptions.

(27) [from S. Myers 1987:499]

a. [ay]-[i]

Semite-Semitic
parasite-parasitic
mime-mimic
satire-satiric
pyrite-pyritic
type-typical
cycle-cyclic
arthritiS-arthritic
Cushite-Cushitic

b. [iy]-[e]

meter-metric
HelLene-Hellenic
diabetes-diabetic
hygiene-hygienic

c. [ey]-[æ]

mage-magic
state-static
(28) -ic Shortening

[after (56c) in Halle and Mohanan 1985:77]

\[
\begin{array}{cccc}
\mu & \mu & \mu \\
V & \rightarrow & 1 \\
V & V / -ic, -id, etc
\end{array}
\]

Thus we see that with both Trisyllabic Shortening and -ic Shortening, low, mid and high vowels are all equally affected by Vowel Shift. I express this schematically below.

(29)

<table>
<thead>
<tr>
<th>/i/</th>
<th>/e/</th>
<th>/æ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>VS</td>
<td>VS</td>
<td>VS</td>
</tr>
</tbody>
</table>

These two rules will turn out to be the only vowel length rules with this property. Later I will argue that this similarity is not a coincidence, but rather follows from the fact that both rules are collapsible into a single one triggered by Resyllabification, following S. Myers (1987).12

2.2.5 i-Shortening

In this section I discuss yet another vowel length rule that makes reference to vowel quality, namely the rule Halle and Mohanan (1985) call i-Shortening. Like Cluster Shortening, it has systematic exceptions.

---

12 In addition to S. Myers (1987), there are at least two other reanalyses of Trisyllabic Shortening and -ic Shortening that I will not discuss in detail. As I observe in a later footnote, Halle and Vergnaud (1987) reanalyze these patterns as involving a foot-based process. Yip (1987) argues that they can be collapsed with Cluster Shortening on the basis of her independently motivated claim that the /i/ found in shortening-triggering affixes like -ity and -ic is in fact epenthetic; see Stemberger (1992) for counterarguments to this claim.
Chomsky and Halle (1968), assuming that the suffix *-ion* is underlingly /ion/, predicted that the vowels in a stem-final syllable would be shortened under *-ion* suffixation through Trisyllabic Shortening. Rubach (1984) and Halle and Mohanan (1985) point out, however, that while there are many cases of the [a:y]-[i:] alternation with *-ion* suffixation, as illustrated in (30a), there are fewer cases of the [i:y]-[e] alternation, as illustrated in (30b); note that most involve the same root, *cede/ceed*, and that there are also at least as many exceptions as expected forms. Moreover, there are no cases of the [e:y]-[æ] alternation at all; as illustrated in (30c), *-ion* has no effect on /æ/. Based on this and other evidence, Rubach (1984) and Halle and Mohanan (1985) decide that *-ion* is actually underlingly /yon/, meaning that the shortening that it triggers cannot be due to Trisyllabic Shortening.\(^{13}\)

\(^{13}\)The alternation that is seen pairs like *discreet-discretion* may thus be considered completely suppletive (parallel to idiosyncratic vowel alternations such as that found in the pair *dr[a]ma-dr[æ]matist* in my speech). Alternatively, such words may be thought of as marked to undergo a special rule that mirrors Vowel Shift but only affects words specially marked to undergo it (parallel to ablaut in *sing-sung* but not *bring-*brung*).
(30) a. [from (7e) in S. Myers 1987:487]

[ay]-[i]

decide-decision
precise-precision
circumcise-circumcision
collide-collision
revise-revision
supervise-supervision
divide-division
provide-provision
extradite-extradition
expedite-expedition
ignite-ignition
erudite-erudition
contrite-contrition
recognize-recognition

b. [partly from (7e) in S. Myers 1987:489 and (43b) in Rubach 1984:43]

Shortening applies  Shortening doesn't apply

[iy]-[ɛ]  [iy]-[iy]
discreet-discretion  delgate-deletion
succeed-succession  excrete-excretion
recede-recession  cohesive-cohesion
concede-concession  complete-completion
proceed-procession

c. [ey] - [ey] (no alternation):

abrade - abrasion
evade - evasion
relate - relation
Rubach (1984) and Halle and Mohanan (1985) therefore conclude that this pattern is derived through a special shortening rule that only affects /iː/. Halle and Mohanan (1985) formulate this rule of *i-Shortening* essentially as given below.

(31) **i-Shortening**  
[after (110) in Halle and Mohanan 1985:98]

\[ \mu \mu \rightarrow \mu \]  
\[ V \rightarrow | \]  
\[ V \quad V / - CyV \]

[+high]

Again, then, we see that a vowel length rule is able to single out vowels of a particular quality. This is expressed schematically below, using the standard notation.

(32)  

\[ /i/ \quad /e/ \quad /æ/ \]

i-Shortening  
S  
VS

### 2.2.6 CiV Lengthening

Vowel length can be affected by lengthening rules and not just shortening rules. The first lengthening rule I discuss is what Halle and Mohanan (1985) call *CiV Lengthening*. Again, this rule will be seen to affect vowels differently depending on their vowel quality.

As discussed in the literature, the rule is intended to account for the length differences (with the accompanying Vowel Shift alternations) seen in pairs such as the following.
(33) [from (34) in Rubach 1984:39 and (58) in Halle and Mohanan 1985:78]

a. [e]-[iy]
   remedy-remedial
   comedy-comedian
   college-collegian

b. [æ]-[ey]
   Canada-Canadian
   Panama-Panamanian
   Arab-Arabian
   Jordan-Jordanian
   Caucasus-Caucasian
   regal-regalia
   mendacity-mendacious

Assuming that the categorization given above is correct, then, the generalization seems to be that a nonhigh vowel is lengthened in a consonant-final stem before a suffix of the form [iV], where "V" represents a vowel. This vowel length rule, like Cluster Shortening and i-Shortening, is sensitive to vowel height features. Not only must the suffix begin with /i/, but the affected vowel must be either /e/ or /æ/, that is, any vowel except /i/. As seen in (34), the behavior of CiV Lengthening is highly unpredictable when the stem ends in a syllable containing /i/. The forms in (34a) seem to undergo CiV Lengthening but not Vowel Shift, while the forms in (34b) do not undergo either rule in the derived form.

14The relevant vowels in remedy, Canada, and so forth, being unstressed, are reduced to schwa. Evidence that they were in fact of the relevant quality before being reduced comes from their treatment as short (lax) vowels by the stress rules; see Chomsky and Halle (1968) for discussion.
(34) a. [iy]-[iy]

Parisian
Venetian

b. [a]-[i] [from Chomsky and Halle 1968:182]

Darwinian
reptilian
precise-precisian

Given this pattern, Halle and Mohanan (1985) formalize CiV Lengthening as in (35).

(35) CiV Lengthening
[after (57) in Halle and Mohanan 1985:78]

\[
\begin{array}{c}
\mu & \mu & \mu \\
I & & \\
V & \rightarrow & \backslash /
\end{array}
\]

[-high]

Once again, then, we see that a vowel length rule can be sensitive to vowel quality. This sensitivity is represented schematically below, where "L" represents vowel lengthening.

(36)

\[
\begin{array}{ccc}
/i/ & /e/ & /æ/ \\
\text{CiV Lengthening} & L & L \\
& VS & VS
\end{array}
\]

2.2.7 Prenasal g-Deletion and compensatory lengthening

I now move on to a final rule affecting vowel length, namely what Halle and Mohanan (1985) call Prenasal g-Deletion, also

---

15This is not a typographical error. The words precisian and precision are distinct.
discussed in Chomsky and Halle (1968), Borowsky (1986) and S. Myers (1987). This rule, too, will be seen to be sensitive to vowel height features.

The rule of Prenasal g-Deletion is motivated on the basis of alternations like those listed in (37), where /g/ appears before a nasal consonant only if the nasal is followed by a vowel. This list was compiled through a computer search, and I believe it is exhaustive.  

(37) a. [ay]-[ɪɡ]
    malign - malignant/malignancy
    sign - signature/signatory/signal/signify
    resign - resignation
    consign - consignation
    design - designate/designation/designatory/designee
    assign - assignment
    benign - benignant/benignancy/benignity
    indignant - indignant/indignity/indignation
    condign - condignity
    paradigm - paradigmatic

b. [ɛ]-[ɛɡ]
    apothegm - apothegmatic
    phlegm - phlegmatic/phlegmatize

c. [æ]-[æɡ]
    diaphragm - diaphragmatic

16There are a few other alternations claimed to involve this rule. Chomsky and Halle (1968) cite the examples repugn - repugnant, impugn - pugnacious/pugnacity, which do not involve front vowel alternations. Borowsky (1986:238) also includes the pairs knowledge-acknowledge and gnostic-agnostic, which do not involve vowel length alternations.
As Borowsky (1986) points out, the fact that /g/ remains only if the following nasal is followed by a vowel suggests that the driving force behind this process is syllabification. That is, since the sequence /gN/, where N represents a nasal consonant, is not allowed syllable-finally as it violates the sonority hierarchy, such a sequence can only surface if the N is resyllabified as the onset of the following syllable, which is what happens if a vowel follows. If this N cannot be resyllabified, the sequence /gN/ cannot surface, forcing the application of the Prenasal g-Deletion rule.

Construing this process as syllable-driven also makes sense of another property seen in the alternations in (37a): the lengthening of the preceding vowel, with the subsequent application of Vowel Shift. As Chomsky and Halle (1968) and all subsequent workers have noted, this finds a natural explanation as a case of postvocalic deletion with compensatory lengthening. Autosegmental models of vowel length can handle this quite easily, as I show in (38) with a moraic model of the syllable (cf Hayes 1989).

(38) Prenasal g-Deletion with compensatory lengthening

<table>
<thead>
<tr>
<th>UR</th>
<th>sign</th>
<th>signature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>/sɪɡn/</td>
<td>/sɪɡn + atur/</td>
</tr>
<tr>
<td></td>
<td>σ</td>
<td>σ</td>
</tr>
<tr>
<td></td>
<td>/\</td>
<td>/\</td>
</tr>
<tr>
<td>Syllabification</td>
<td>/ ɛɪmɛ /</td>
<td>/ ɛɪmɛ / ɛɪm ...</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>s ɪɡ n</td>
<td>s ɪɡ n+atur</td>
</tr>
</tbody>
</table>
Halle and Mohanan (1985) thus formulate the rule of Prenasal g-Deletion essentially as given below, writing it as if the deletion of /g/ and the subsequent compensatory lengthening were part of the same process.

(39) Prenasal g-Deletion [preliminary version]
[after (106a) in Halle and Mohanan 1985:96]

\[
\begin{array}{c}
\sigma \\
/\backslash/
\end{array}
\begin{array}{c}
\mu \\
/\backslash/
\end{array}
\begin{array}{c}
\mu \\
/\backslash/
\end{array}
\begin{array}{c}
\mu \\
/\backslash/
\end{array}
\mu

\begin{array}{c}
| | \\
/ | | \\
\mu | \\
/ | | \\
\mu | \\
/ | | \\
\mu | \\
/ | | \\
\mu | \\
/ | | \\
\mu | \\
/ | | \\
\mu | \\
/ | | \\
\mu | \\
/ | | \\
\mu |
\end{array}

\begin{array}{c}
V \\
/ | | \\
g | \\
/ | | \\
V |
\end{array}
/ - N

This cannot be right, however, as can be seen by comparing the examples in (37a) with those in (37b,c). Although /g/ consistently deletes in the environment given in (39), the preceding vowel lengthens only if it is [+high]. That is, the rule should actually be represented as below. If the preceding vowel is [+high], the deletion of /g/ triggers compensatory lengthening; otherwise it does not.
Prenasal g-Deletion

a. \[ \mu\mu\mu\mu\mu \]
   \[ \rightarrow \]
   \[ V\ g\ V\ /\_\ N \]
   \[ [+\text{high}]\ [+\text{high}] \]

b. \[ \mu\mu\mu\mu\mu \]
   \[ \rightarrow \]
   \[ V\ g\ V\ /\_\ N \]
   \[ [-\text{high}]\ [-\text{high}] \]

Notice that because of the sensitivity of compensatory lengthening to the deletion specifically of /g/, we cannot describe the pattern using a general g-deletion rule followed by a more specific rule of lengthening. It is not the case that /i/ is the only vowel that ever lengthens in English; after all, above we saw that with CiV Lengthening, /i/ is the only vowel that fails to lengthen. Rather, /i/ is the only vowel that lengthens \textit{when a following /g/ is deleted}. Hence we are forced to represent these rules roughly as Halle and Mohanan (1985) have done, where the deletion of /g/ and the compensatory lengthening happen as part of a single process.\(^{17}\)

The conclusion is strange, but hardly unexpected. As we have seen, with the exception of Trisyllabic Shortening and -ic Shortening,

\(^{17}\)An alternative analysis would be to posit two rules of Prenasal g-Deletion in addition to a single rule of compensatory lengthening. One g-Deletion rule would apply after a [+high] vowel, deleting the featural material of a prenasal /g/, while the other, applying after all other vowels, would delete both the /g/ and the mora it is attached to. Compensatory lengthening would then be able to apply only in the [+high] case. As compensatory lengthening applies in no other situation in English, I will retain my slightly simpler formulation.
all the vowel length rules have been seen to be dependent on vowel height features, and in apparently all different ways. Below I represent schematically the restrictions shown with Prenasal g-Deletion.

\[ (41) \]

\[
\begin{array}{c|c|c|c}
\text{Compensatory Lengthening} & /i/ & /e/ & /æ/ \\
\text{with Prenasal g-Deletion} & \text{L} & \text{VS} & \\
\end{array}
\]

In the next section I will consider the question of whether these idiosyncrasies can nevertheless be unified in a model where all vowel length rules are triggered by automatic syllabification processes. I will conclude that for the most part they cannot.

2.2.8 Vowel length rules and syllabification

S. Myers (1987) has made the interesting suggestion that most of the shortening rules discussed above, namely Cluster Shortening, Trisyllabic Shortening and -ic-Shortening, can be collapsed into a single generalization he calls Closed Syllable Shortening. Following his line of reasoning, I will indicate how the lengthening rules may also be thought of as triggered by independently motivated processes of syllabification. It will become clear, however, that the specific patterns illustrated above will not reduce to a single rule.

S. Myers (1987) conceives of Closed Syllable Shortening not as a rule, but as a constraint on syllable structure, as represented schematically below. The idea is that the rime of a syllable can either consist of a long vowel or a short vowel followed by a consonant. As can be seen in his formalization, Closed Syllable
Shortening makes no reference to the featural level. This will become significant shortly.\textsuperscript{18}

\text{(42) \[(=32) \text{ in S. Myers 1987:511}\]

\[\sigma \rightarrow \text{C}^* \text{V} (X) \quad \text{(where X is C or V)}\]

In terms of the moraic representation of the syllable, this constraint could be represented as in (43). That is, no mora may be linked to more than one Root node. This ensures that only syllables of the forms illustrated in (44a,b) will be allowed; one like that in (44c) will not.

\text{(43) Closed Syllable Shortening}

\[\begin{array}{c}
*\mu \\
/ \ \\
\cdot \cdot \cdot \text{Root}
\end{array}\]

\text{(44) a. } /\text{at}/ \quad \sigma \\
\quad |\backslash \\
\quad \mu \mu \\
\quad | | \\
\quad \text{a t}

\text{b. } /\text{aː/} \quad \sigma \\
\quad |\backslash \\
\quad \mu \mu \\
\quad |/ \\
\quad \text{a}

\textsuperscript{18}Archangeli (1991) achieves a similar result through the use of Weight-by-Position, adopted from Hayes (1989). Although with her formalism it is possible for shortening to be ordered after segmental processes that are sensitive to syllable structure, her analysis of shortening itself still makes no reference to the featural level.
Of course long vowels do appear in closed syllables, but as S. Myers (1987) notes, following Hayes (1982) and Selkirk (1983), only under certain circumstances. Specifically, the syllable must be word-final. This suggests that word-final closed syllables get around Closed Syllable Shortening by virtue of the fact that word-final consonants are extraprosodic. That is, word-final consonants are marked to be invisible to prosodic structure such as syllables. Closed syllables with long vowels may thus be represented as below, where the "< >" enclose the extraprosodic word-final consonant.

\[
\begin{array}{c}
\text{Word-finally} /a:t/ \\
\sigma \text{ jWord} \\
\mu \mu \\
/ \\
a < t
\end{array}
\]

S. Myers (1987) notes further, following Selkirk (1983), that word-final coronal consonant clusters appear to be made extrametrical as a whole, since such clusters do not trigger Closed Syllable Shortening, as illustrated by the examples in (46a). Such forms would be represented as in (46b).19

---

19 As Mike Hammond has pointed out to me, this analysis thereby confronts examples like mean-meaned, which inexplicably show the vowel shift alternation [iy]-[e].
Given this background, S. Myers (1987) suggests that the cases of shortening discussed above can be reduced to the effects of Closed Syllable Shortening alone. This is especially clear in the case of Cluster Shortening. Because word-final consonants are extraprosodic, word-final syllables with a single consonant in the coda can contain long vowels. However, no syllable should be able to contain a long vowel if it ends in a consonant cluster (unless the consonants in this cluster are both coronals, as noted above). According to this view, then, Cluster Shortening is just a special case of Closed Syllable Shortening. A sample derivation illustrating this is given below, where the resyllabification of the onset consonant /t/ as a coda is represented by moving it before the "." used to separate syllables.

(47)    intervene intervention
UR      /ve:< n >/   /ve:< n > + t + yən/
Closed Syllable Shortening   --     vent.yən
Vowel Shift      vi< n >   --
Other rules        [viyn]     [vənʃən]

As S. Myers (1987) shows, however, Closed Syllable Shortening can also be used to account for cases of shortening before vowel-initial suffixes, specifically those involved in Trisyllabic Shortening and -ic Shortening. This is because in words with these suffixes, the
affected vowel is always before an unstressed syllable, which meets the conditions necessary for the application of Resyllabification. In these words, the stem-final consonant is resyllabified as the coda of the stressed syllable, thus making it into a closed syllable, as seen in (48). Closed Syllable Shortening is thus relevant, preventing the application of Vowel Shift in these syllables.

(48) a.  

<table>
<thead>
<tr>
<th></th>
<th>sane</th>
<th>sanity</th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
<td>/sæ:&lt; n &gt;/</td>
<td>/sæ:&lt; n &gt; + ˌti/</td>
</tr>
<tr>
<td>Stress and resyllabification</td>
<td>ˌsæ:&lt; n &gt;</td>
<td>ˌsæ:nˌti</td>
</tr>
<tr>
<td>Closed Syllable Shortening</td>
<td>--</td>
<td>ˌsæ:nˌti</td>
</tr>
<tr>
<td>Vowel Shift</td>
<td>ˌséy&lt;n &gt;</td>
<td>--</td>
</tr>
<tr>
<td>Other rules</td>
<td>[ˌséyn]</td>
<td>[ˌsæ:nˌti]</td>
</tr>
</tbody>
</table>

b.  

<table>
<thead>
<tr>
<th></th>
<th>satire</th>
<th>satiric</th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
<td>/sæt:&lt; r &gt;/</td>
<td>/sæt:&lt; r &gt; + ˌik/</td>
</tr>
<tr>
<td>Stress and resyllabification</td>
<td>ˌsæt.i:&lt; r &gt;</td>
<td>ˌsæ.tˌirˌik</td>
</tr>
<tr>
<td>Closed Syllable Shortening</td>
<td>--</td>
<td>ˌsæ.tˌirˌik</td>
</tr>
<tr>
<td>Vowel Shift</td>
<td>ˌsæt.ay&lt;r &gt;</td>
<td>--</td>
</tr>
<tr>
<td>Other rules</td>
<td>[ˌsætˌayr]</td>
<td>ˌsæ.tˌirˌik</td>
</tr>
</tbody>
</table>

S. Myers (1987:507) observes that constraints like Closed Syllable Shortening are extremely common in the world's languages, and that it is "natural" in the sense of being "a straight phonologization of a universal fact of phonetic implementation: that vowels in closed syllables have a significantly shorter duration than corresponding vowels in open syllables". Later, however, he goes a bit too far, claiming that "[c]losed syllable shortening is not a language-particular rule and need not be stipulated in the grammar"
What about the lengthening rules, namely CiV Lengthening and the compensatory lengthening triggered by Prenasal g-Deletion? Although S. Myers (1987) does not mention these, it seems rather trivial to extend his ideas to these as well. This has already been illustrated with Prenasal g-Deletion, where the compensatory lengthening triggered by the deletion of /g/ can be understood as a natural consequence of syllabification as well. A similar view of CiV Lengthening is also possible, since the addition of a vowel-initial suffix should allow the stem-final coda consonant to be resyllabified as an onset of the new syllable now following, in accordance with the Onset Principle. This will then cause the stem-final syllable to become open, allowing the vowel to lengthen.

In spite of the appeal of a purely prosodic analysis of the shortening and lengthening rules, however, we have already seen sufficient evidence above suggesting that these are not in fact triggered automatically through syllabification processes. The problem specifically is this: each of the vowel length rules is sensitive to vowel height features in a different way. I illustrate this below schematically by collecting each of the summary figures from the above sections in a single chart, given below in (49).

---

20See Archangeli (1991) for further arguments that the syllabification subprocesses involved in shortening may be extrinsically ordered with respect to rules of the segmental phonology.
(49)

<table>
<thead>
<tr>
<th></th>
<th>/i/</th>
<th>/e/</th>
<th>/æ/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cluster Shortening</td>
<td>S</td>
<td>S</td>
<td></td>
</tr>
<tr>
<td>in general</td>
<td>VS</td>
<td>VS</td>
<td></td>
</tr>
<tr>
<td>Strong verbs with</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-t suffixation</td>
<td>S</td>
<td>VS</td>
<td></td>
</tr>
<tr>
<td>Trisyllabic Shortening</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
</tr>
<tr>
<td>-ic Shortening</td>
<td>S</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td>VS</td>
<td>VS</td>
</tr>
<tr>
<td>i-Shortening</td>
<td>S</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CiV Lengthening</td>
<td>L</td>
<td></td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>VS</td>
<td></td>
<td>VS</td>
</tr>
<tr>
<td>Compensatory Lengthening</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>with Prenasal g-Deletion</td>
<td>VS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

It is true that this pattern is not entirely random. Notably, the only two cases where all three Vowel Shift rules can apply, namely with Trisyllabic Shortening and -ic Shortening, are also the only two cases where S. Myers (1987) claims Resyllabification applies. This suggests that this aspect of S. Myers's (1987) really is on the right track.

There appears to be a way to reduce the apparent randomness of the other cases as well.21 Recall from the discussion of sonority earlier in this chapter that vowel height is one of the relevant features: high vowels are less sonorant than mid vowels, which in

21This was suggested to me by Mike Hammond.
turn are less sonorant than low vowels. The fact that vowel height plays a role in determining the application of the vowel length rules might then be understood as an automatic consequence of the relevance of vowel height features in syllabification.

Unfortunately the idea will not work. First of all, consider the role the relative sonority of segments plays in the application of Resyllabification. As noted above, Resyllabification applies more readily the more sonorous the final segment of the previous syllable is. Thus we would expect that a consonant will be more likely to resyllabify as a coda if preceded by a more sonorous vowel like /æ/ than by a less sonorous vowel like /i/. This predicts that in cases involving Resyllabification, /æ/ should be more susceptible to shortening than /i/. As we saw above, though, in the only cases where Resyllabification seems relevant, namely Trisyllabic Shortening and -ic Shortening, all three vowel heights show equal ability to shorten.

Even if we propose something like the otherwise unmotivated hypothesis that more sonorous vowels are inherently "longer" than less sonorous vowels, the idiosyncrasies reflected above cannot be eliminated. This hypothesis would predict that /æ/ should tend to resist shortening and favor lengthening; /i/, by contrast, should favor shortening and resist lengthening. Part of the first tendency is in fact seen: /æ/ does not shorten with Cluster Shortening or i-Shortening, while /i/ shortens with both. This is as successful as the hypothesis gets, though. It cannot explain why /e/ is shortened with
Cluster Shortening and not i-Shortening, nor can it explain why /i/ resists lengthening in CiV Lengthening while in the compensatory lengthening triggered by Prenasal g-Deletion it is only /i/ that lengthens.

Given these idiosyncrasies, I conclude that as appealing as it is, a purely prosodic analysis cannot be adopted. The various vowel length rules must remain separate, just as Chomsky and Halle (1968) originally argued. Except for the shortening rules triggered by Resyllabification, the forms of these rules must make reference to exactly the same vowel height features that are manipulated by Vowel Shift.

I thus suggest that the rules of Cluster Shortening, i-Shortening, CiV Lengthening and Prenasal g-Deletion with compensatory lengthening be left as they were formalized above. The rules triggered by Resyllabification, namely Trisyllabic Shortening and -ic Shortening, which do appear to be collapsible into a single process, must nevertheless be handled slightly differently from the way S. Myers (1987) does. This is because of the evidence given above that Cluster Shortening affects a slightly different set of vowels from those affected by the rules triggered by Resyllabification. Hence it cannot be the case that what is causing shortening in the Resyllabification examples is the fact that the vowel is now in a closed syllable; otherwise the vowels before consonant clusters would be shortened just as consistently.
The consequence of this is that we must now suppose that the shortening in these cases occurs in a closed syllable only if the syllable became closed through Resyllabification. The formulation of the rule collapsing Trisyllabic and -ic Shortening, given below, must therefore overlap somewhat with Resyllabification.22

\[ \begin{align*}
\mu \mu & \quad \mu \\
\sqrt{\text{V}} & \rightarrow \quad - \\
V & V \quad / \quad C \check{V}
\end{align*} \]

(50) **Resyllabification Shortening**

2.2.9 **Vowels: Summary**

In this section I have argued that the vowel height rules involved in the Vowel Shift pattern are not reducible to a single process. The only truly systematic aspect involves what Halle and Mohanan (1985) call Trisyllabic Shortening and -ic Shortening, which I reanalyze as involving Resyllabification Shortening. Cluster Shortening, i-Shortening, CiV Lengthening, and compensatory lengthening triggered by Prenasal g-Deletion are all sensitive to idiosyncratic subsets of the vowel height features and must therefore be kept as separate rules.

---

22 Interestingly, the grouping of the shortening rules suggested by the Vowel Shift data presented here, where Trisyllabic Shortening and -ic Shortening are due to one rule while Cluster Shortening is due to another, mirrors the analysis independently argued for by Halle and Vergnaud (1987:251-254). There the effects of Trisyllabic Shortening and -ic Shortening are ascribed to a single rule that shortens a vowel in the head syllable of a binary foot, whereas the effects of Cluster Shortening are ascribed to a rule that shortens a vowel in a closed syllable.
Since all of the various rules affecting vowel length are mutually exclusive (that is, each apply in different environments), there is no evidence for their relative ordering. All of them, however, must be ordered before Vowel Shift in order to derive the Vowel Shift alternations in the proper way. The picture of ordered rules thus ends up looking roughly as given below, where all the rules in (51a) are ordered before the rule in (51b).

(51) a. Resyllabification Shortening
    Cluster Shortening
    i-Shortening
    CiV Lengthening
    Prenasal g-Deletion with compensatory lengthening

b. Vowel Shift

2.3 Rules involving coronal consonants

The next collection of interacting rules I consider are rules that either produce or affect coronal consonants, specifically the rules I will call (after Halle and Mohanan 1985) s-Voicing, Velar Softening, Spirantization, y-Insertion and Palatalization. As I noted in the introduction, I will not have much new to say about the last three rules; I find that the evidence essentially supports the analysis argued for in Chomsky and Halle (1968), Rubach (1984), Halle and Mohanan (1985) and Borowsky (1986), at least as far as their ordering is concerned. By contrast, I will show that the evidence for s-Voicing and Velar Softening is much weaker than often assumed.

2.3.1 s-Voicing

The first rule I examine here is s-Voicing, discussed in Chomsky and Halle (1968), Rubach (1984), Halle and Mohanan
(1985) and Borowsky (1986). We will see that this rule appears to apply robustly only in a very restricted morphological domain.

Halle and Mohanan (1985), following Rubach (1984), formalize s-Voicing roughly as follows. What the rule is intended to do is simple: voice /s/ when followed by a vowel and preceded by a long vowel. 23

\[ s \rightarrow z / V \_ V \]

(52) **s-Voicing**  [after (111) in Halle and Mohanan 1985:98]

Chomsky and Halle (1968) and Halle and Mohanan (1985) give evidence from three morphological domains. In two of these domains, I will show, the evidence for s-Voicing is extremely weak; that is, in these areas s-Voicing is extremely unproductive. I will discuss them first and then focus the rest of my attention on the remaining source of evidence.

2.3.1.1 s-Voicing before CiV Lengthening

On the basis of the examples shown below, Rubach (1984) concludes that s-Voicing must apply after CiV Lengthening. This is because the vowel preceding the /s/ in these forms is not

---

23 In addition, Chomsky and Halle (1968:228-9) note that a rule is also needed to voice /s/ before a stressed syllable and after /k/, as in words such as *exist, examine, auxiliary, and exasperate* (cf nonapplication in *axis* and *maxillary*). I will not discuss this pattern here. One aspect of this rule that appears peculiar from the hindsight offered by years of discussion of the Strict Cycle Condition (see Chapter 5) is the fact, duly noted by Chomsky and Halle (1968:229), that this rule only applies within morphemes; cf its nonapplication in *hex+de+meter, tox+le+ly, annex+ation.*
underlyingly long, so that in order for s-Voicing to apply, CiV Lengthening must apply first to lengthen it.

(53) [after (38) in Rubach 1984:412]

\[ \text{[Vs]} - \text{[V:z]} \]

Cáucasisus - Caucásian
Málthús - Malthúsian
gymnástics - gymnásium

Since it is not clear that the [s] in gymnastics is underlyingly voiceless or derived from /z/ by the voicing assimilation that occurs within consonant clusters, both of the clear examples cited in the literature for this claimed ordered of CiV Lengthening before s-Voicing, namely Caucasus-Caucasian and Malthus-Malthusian, involve the suffix -ian. Hence I decided to conduct a computer search of all words with this suffix. Below I list all such forms associated with a stem ending in /s/ including the two in (53). The columns indicate whether the stem has a long vowel, short vowel or consonant before /s/; the rows indicate what changes take place when -ian is suffixed. The forms that conform to s-Voicing (ie, the derived form shows [z] after a long vowel, [s] otherwise) are underlined.
(54) [pronunciations checked with Woolf et al. 1981 and native speakers]

<table>
<thead>
<tr>
<th>stem:</th>
<th>/Vs/</th>
<th>/Vːs/</th>
<th>/Cz/</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Vːz]</td>
<td>Caucasian-Caucasian</td>
<td>Dionysus-Dionysian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Malthus-Malthusian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Paris-Parisian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Venus-Venusian</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[Vz]</td>
<td>Dionysus-Dionysian</td>
<td></td>
<td>precise-precisian</td>
</tr>
<tr>
<td>[Cs]</td>
<td></td>
<td>Mars-Martian</td>
<td></td>
</tr>
<tr>
<td>[Vs]</td>
<td>Hesse-Hessian</td>
<td>Dionysus-Dionysian</td>
<td></td>
</tr>
<tr>
<td>[Vːs]</td>
<td>Horace-Horatian</td>
<td>Dionysus-Dionysian</td>
<td>Thrace-Thracian</td>
</tr>
<tr>
<td>Venice-Venetian</td>
<td></td>
<td>Venus-Venusian</td>
<td></td>
</tr>
</tbody>
</table>

As can be seen, when the complete data set is considered, evidence for s-Voicing triggered by CiV Lengthening becomes much less convincing. Out of the ten distinct word pairs, only four (Caucasus-Caucasian, Malthus-Malthusian, Paris-Parisian, precise-precisian) show evidence of actually having undergone the rule obligatorily. The other six either properly fail to undergo the rule (Hesse-Hessian, which is an exception to CiV Lengthening, and Mars-Martian) or else are exceptions to s-Voicing, obligatorily (Thrace-Thracian, Venice-Venetian, Horace-Horatian) or optionally (Venus-Venusian, Dionysus-Dionysian).24

24 Note also the voicing of [z] in Mars, which appears actually to be lost with -ian suffixation in Martian. We will see other occasional examples in this chapter of [z] appearing after [r] when [s] would be expected, suggesting the existence of a rather unproductive and idiosyncratic rule we may term Post-r s-Voicing.
Thus in this morphological domain there are only six word pairs that show evidence of s-Voicing (four obligatorily, two optionally). The doctrine of Gradient Productivity forces us to conclude nevertheless that s-Voicing is a descriptively adequate rule of English for this class, although extremely unproductive. I will thus assume that CiV Lengthening is in fact ordered before s-Voicing. The reader should not forget, however, that the evidence for this conclusion is extremely weak.

2.3.1.2 s-Voicing before i-Shortening

The second claimed source of evidence for s-Voicing is the putative voicing of /s/ after long vowels that are then shortened before suffixes like -ion. The two examples that have been put forth in the literature in support of this ordering of s-Voicing before the shortening rules are listed below.

(55) [from Rubach 1984:44]

[ays] - [i3]

concise - concision
precise - precision

The shortness of this list, coupled with the fact that both forms contain the same root (-cise), makes one suspect the necessity of s-Voicing in these cases. The suspicion receives support from the results of a thorough computer search of all words ending in -ion, which should cause shortening of the last vowel in the stem by i-Shortening (see discussion earlier in this chapter).
A total of approximately sixty of these words end in [ʒən], ie with a voiced alveolar fricative. If s-Voicing were applicable here, we might expect many of them to be derived from words ending with an underlying /s/ preceded by a long vowel. This is simply not the case.

As illustrated in (56), [ʒən] appears in eight sorts of cases. The only cases where s-Voicing may potentially be applicable, namely (56a) and (56b), where the stems end in a long vowel followed by a voiceless consonant, consist of only four forms, one of which has an optional pronunciation where the final consonant is already voiced as [z]. In the remaining cases, [ʒ] is either derived from an underlyingly voiced consonant, making s-Voicing superfluous, or else is derived from a voiceless consonant following /t/.²⁵

(56) Words ending in [ʒən]

a. /V:s/
   precise-precision
   concise-concision
   recluse-reclusion  {also reclu[z]}

b. /V:t/
   equate-equation

²⁵Note that again /s/ appears to be voiced after /t/ for some morphemes (cf footnote 24).
c. /V:d/

abrade-abrasion
allude-allusion
collude-collusion
conclude-conclusion
corrade-corrasion
corrode-corrosion
delude-delusion
detrude-detruision
dissuade-dissuasion
elude-elusion
erode-erosion
evade-evasion
exclude-exclusion
explode-explosion
extrude-extrusion
implode-implosion
include-inclusion
intrude-intrusion
include-inclusion
invade-invasion
persuade-persuasion
protrude-protrusion
seclude-seclusion
suade-suasion

d. /V:z/

affuse-affusion
circumfuse-circumfusion
confuse-confusion
contuse-contusion
defuse-defusion
diffuse-diffusion
effuse-effusion
fuse-fusion
infuse-infusion
superfuse-superfusion
transfuse-transfusion
e. /rs/
   asperse-aspersion
   averse-aversion
   disperse-dispersion
   immerse-immersion
   intersperse-interspersion
   obverse-obversion
   coerce-coercion {also coer[f]ion}

f. /rt/
   convert-conversion
   divert-diversion
   evert-eversion
   extravert-extraverion
   extrovert-extroversion
   introvert-introversion
   invert-inversion
   pervert-perversion
   retrovert-retroversion
   revert-reversion
   subvert-subversion

g. /V:r/
   adhere-adhesion
   cohere-cohesion
   inhere-inhesion

h. /rʒ /
   emerge-emersion
   submerge-submersion

One of the examples that might be derived using s-Voicing, namely that in (56b) (equate-equation) must actually be removed from the list of supporting examples. In order for s-Voicing to apply in this pair, the underlying /t/ must first be changed into /s/. Later
we will see that such a rule, Spirantization, does seem to be independently required. Unfortunately, its interaction with s-Voicing in all other cases shows that s-Voicing would have to be thought of as applying first. Thus, for example, in the pair *relate-relation*, the alternation is between [t] and [ʃ], not [ʒ], which is standardly taken as evidence that s-Voicing gets its sole chance to apply to these forms, and is therefore inapplicable, *before* Spirantization applies.26

We are thus left with the original two pairs, *concise-concison* and *precise-precision*, as constituting the sole evidence that s-Voicing applies before i-Shortening. The fact that both contain the same root thus becomes significant. It is well known that individual roots can show allomorphy that show up in no other forms (eg, Aronoff 1976). For example, an alternation between [ʤ] and [ʒ] shows up under -ion suffixation only with the root *-merge*, as in *emerge-emersion*, *submerge-submersion*, while an alternation between [r] and [ʒ] shows up only with the root *-here*, as in *adhere-adhesion*, *cohere-cohesion*, *inhere-inhesion*. It would thus not be unreasonable to suggest that the root *-cise* allows for an alternation between [s] and [z] that is not phonologically governed but is merely allomorphy.

Nevertheless, to avoid taking the discussion in a direction not crucial to my main arguments, I will simply follow the analysis found

---

26 An alternative account of the alternation in *equate-equation*, then, would be to suppose that this item is specially marked to undergo Spirantization before s-Voicing. Such cases of special marking of rule orderings do seem to exist, both in English and in other languages. What this means for Double Lookup is discussed in Chapter 5.
in the literature and suppose that s-Voicing is in fact ordered before i-Shortening.\textsuperscript{27} As with the cases involving -\textit{ian} suffixation, I merely observe that the evidence for this conclusion is extremely weak.

\textbf{2.3.1.3 s-Voicing with prefixes}

Thus we are left with the final source of evidence to argue for the existence of s-Voicing, namely forms containing prefixes affixed to roots that seem to begin with underlying /s/. The generalization is claimed to be as follows: when prefixes ending in long vowels, such as \textit{de-}, \textit{re-}, and \textit{pre-} are suffixed to roots beginning with /s/ followed by a vowel, /s/ surfaces as [z]. Otherwise, that is, when the root is preceded by a consonant-final prefix like \textit{per-}, \textit{dis-}, \textit{in-} or \textit{con-}, is preceded by a prefix with a short vowel like \textit{a(d)-}, or appears word-initially, /s/ surfaces as [s]. In (57) I list all examples of alternations I have found with help of a computer search.\textsuperscript{28}

\begin{center}
\begin{tabular}{ccc}
\textit{semb}le & [s] & \textit{sembl}ence [z] \\
\textit{assemble} & res\textit{émbl}e & \textit{diss\textit{é}mbl}e \\
\textit{sent} & ass\textit{é}nt & res\textit{é}nt
\end{tabular}
\end{center}

\textsuperscript{27}Note that by transitivity, this would force us to suppose that CiV Lengthening, which is ordered before s-Voicing, is ordered before i-Shortening. This would then be more evidence that the vowel length rules cannot be thought of as due solely to automatic syllabification processes.

\textsuperscript{28}"*" indicates that a pronunciation with [z] is also possible. "**" indicates that a pronunciation with [s] is also possible. The lack of s-Voicing in forms such as \textit{resent} ("sent again") is evidence for this rule's dependence on morphological information, as the \textit{re-} in this word is not the same as that in \textit{resent} ("feel annoyance at").
<table>
<thead>
<tr>
<th>[s]</th>
<th>[z]</th>
</tr>
</thead>
<tbody>
<tr>
<td>serve</td>
<td>desérve</td>
</tr>
<tr>
<td>serve</td>
<td>presérve</td>
</tr>
<tr>
<td>serve</td>
<td>resérve</td>
</tr>
<tr>
<td>sign</td>
<td>désign</td>
</tr>
<tr>
<td>assign</td>
<td>resign</td>
</tr>
<tr>
<td>consign</td>
<td>resign</td>
</tr>
<tr>
<td>sist</td>
<td>desist**</td>
</tr>
<tr>
<td>sist</td>
<td>resist</td>
</tr>
<tr>
<td>insist</td>
<td>persist*</td>
</tr>
<tr>
<td>persist</td>
<td></td>
</tr>
<tr>
<td>insist</td>
<td></td>
</tr>
<tr>
<td>persisit*</td>
<td></td>
</tr>
<tr>
<td>sol</td>
<td>désolâte**</td>
</tr>
<tr>
<td>solitary</td>
<td></td>
</tr>
<tr>
<td>solve</td>
<td>dissólve</td>
</tr>
<tr>
<td>solve</td>
<td>resólve</td>
</tr>
<tr>
<td>dissolve</td>
<td>absólve</td>
</tr>
<tr>
<td>son</td>
<td>résonant</td>
</tr>
<tr>
<td>assonant</td>
<td></td>
</tr>
<tr>
<td>dissolútion</td>
<td></td>
</tr>
<tr>
<td>dissonant</td>
<td></td>
</tr>
<tr>
<td>sorb</td>
<td>desórb**</td>
</tr>
<tr>
<td>sorb</td>
<td>resórb**</td>
</tr>
<tr>
<td>absorb*</td>
<td></td>
</tr>
<tr>
<td>adsorb*</td>
<td></td>
</tr>
<tr>
<td>sort</td>
<td>resórt</td>
</tr>
<tr>
<td>sort</td>
<td></td>
</tr>
<tr>
<td>consórt</td>
<td></td>
</tr>
<tr>
<td>assórt</td>
<td></td>
</tr>
<tr>
<td>source</td>
<td>résource**</td>
</tr>
<tr>
<td>source</td>
<td></td>
</tr>
<tr>
<td>consúlte</td>
<td>désultôry**</td>
</tr>
<tr>
<td>insúlte</td>
<td>resúlt</td>
</tr>
<tr>
<td>sume</td>
<td>presúme</td>
</tr>
<tr>
<td>consume</td>
<td>resúme</td>
</tr>
</tbody>
</table>
It is clear that the generalization given above is not without its problems when applied to all of the forms in (57). One problem is that out of all the words with long-vowel prefixes, six (desultory, resorb, desist, desorb, desolate, resource) allow an alternative pronunciation with [s]. According to Kenyon and Knott (1953), in fact, desultory, desolate, and resource only have the [s] pronunciation in American English (resorb and desorb are not listed). We therefore must change the generalization given above from "/s/ surfaces as [z] after long-vowel prefixes" to "/s/ may surface as [z] after long-vowel prefixes." This would account for why [z] may (optionally) appear after de-, re-, and pre-, but never after in-, con-, and a(d)-.

We may attempt to reduce the problem (the optional nature of s-Voicing after long-vowel prefixes) by supposing that the vowel in de-, re-, and pre- is sometimes underlyingly short, but this move is not supported by any independent evidence. It is true that these suffixes virtually always surface with the short vowel [ɪ] (resource, resorb and desorb, with their atypical word-initial stress, are the sole exceptions), but as Halle and Mohanan (1985) observe, this is typical of unstressed tense vowels. What evidence we do have suggests that these prefixes really are underlyingly long. As seen in the forms in (58a), these suffixes surface with the vowel [ɛ] when they appear in a position where they should undergo the Resyllabification Shortening rule discussed above, namely in a stressed syllable followed by an unstressed syllable. In other words, the vowel in these suffixes appears to be /e:/ underlyingly. When shortened, as
in (58a), the vowel surfaces as [e]; when it is not shortened, as in (57), it is raised by Vowel Shift to /i:/, which surfaces as [i:] if stressed and as [l] if unstressed. By contrast, the vowel in the prefix a(d)- is not raised by Vowel shift, surfacing as [ə] in most of the forms in (57) and as [æ] in the form in (58b), implying it is an underlyingly short /æ/.

(58) a. [l]-[e]

resign - résignation
design - désignate
prèsérve - préservation

b. [ə]-[æ]

assign - assignation

It appears, then, that the prefixes de-, re-, and pre- really do end in underlyingly long vowels while the prefixes in-, con-, and a(d)- do not. It therefore seems we must assume that s-Voicing is optional for certain words.

Two of the (optional) exceptions to s-Voicing share a curious property in common, however: désolâtre and désultory are both unstressed on the second syllable. This immediately suggests an explanation for why they may fail to undergo s-Voicing. As we saw above, Resyllabification acts to resyllabify a consonant followed by a stressless vowel as the coda of the preceding syllable. Forms with this pattern, then, namely désolâtre and désultory, should undergo Resyllabification Shortening in the first syllable. The elimination of the underlying long vowel will then remove part of the conditioning
environment for s-Voicing, which will then be unable to apply. This concept is illustrated with the derivations below.

(59) desultory desolation

<table>
<thead>
<tr>
<th></th>
<th>des:Al-to:ri</th>
<th>de:so:la:tyan</th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
<td>/de:+sAltori/</td>
<td>/de:+so:la:tyan/</td>
</tr>
<tr>
<td>Stress</td>
<td>dé:sAltòri</td>
<td>dé:solàtyan</td>
</tr>
<tr>
<td>Resyllabification</td>
<td>dé:s.Àl.tòr.i</td>
<td>dé:s.o.là:tyan</td>
</tr>
<tr>
<td>Shortening</td>
<td>dé:s.Àl.tòr.i</td>
<td>dé:s.o.là:tyan</td>
</tr>
<tr>
<td>s-Voicing</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Vowel Shift, etc</td>
<td>[dèsÀltòri]</td>
<td>[dèsoléyfàn]</td>
</tr>
</tbody>
</table>

This analysis allows us to connect the unusual stressing pattern with the (optional) lack of s-Voicing, but it raises many problems itself. The first problem is that it only works properly for the two words shown in (59): the second syllable in resource, which also fails to undergo s-Voicing, is stressed and so this word cannot undergo Resyllabification. Apparently, then, the (optional) blocking of s-Voicing in this word must arise another way.

Another major problem concerns the interaction of s-Voicing with Shortening. Observe that in the form resonant, which is stressed like the cases given in (59), the /s/ always surfaces as [z]. We could decide that this word is a lexically marked exception to Resyllabification. However, note that the prefix vowel in resonant is [e], i.e. the short form of the underlying /e:/, not the expected [i] or [i:]. It appears, then, that Resyllabification Shortening has applied here, implying that the /s/ is in fact ambisyllabic. The /s/ should then surface as [s], since s-Voicing will no longer be applicable.

In summary, then, as illustrated in (60) below, Shortening and s-Voicing can interact in all four logically possible ways.
The interaction of s-Voicing and Shortening in prefixed words

<table>
<thead>
<tr>
<th>s-Voicing:</th>
<th>yes</th>
<th>no</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shortening yes</td>
<td>résonânt</td>
<td>désolâte</td>
</tr>
<tr>
<td></td>
<td>désultory</td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>resúlt</td>
<td></td>
</tr>
<tr>
<td></td>
<td>{etc}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>résorbe</td>
<td></td>
</tr>
<tr>
<td></td>
<td>désorbe</td>
<td></td>
</tr>
</tbody>
</table>

This makes it impossible to avoid having idiosyncratic problem cases. As we saw above, if we ignore ambisyllabicity, we would correctly predict the [z] in resonant, but desolate, desultory, resource, resorb and desorb would be incorrectly predicted to have obligatory [z] as well. If we take ambisyllabicity into account, we still incorrectly predict resource, resorb and desorb to have [z], since in these words the second syllable is not unstressed and so ambisyllabicity should not occur. If we somehow allowed ambisyllabicity to apply in these cases, we would incorrectly predict the shortening of the prefix vowel. Finally, if we take ambisyllabicity into account, we now incorrectly predict [s] in resonant; as noted above, if we didn't employ ambisyllabicity in this word, we would incorrectly predict a long high vowel in the prefix.

It is clear, then, s-Voicing does not describe a very consistent generalization even in the small set of forms with vowel-final prefixes. What about consonant-final prefixes? Here we see that in- and con- are the only consonant-final prefixes where /s/ consistently surfaces as [s]. After ad- and ab- it is not too surprising that [z] sometimes appears (as in adsorb and absorb), since these prefixes
end in voiced obstruents, and clusters of obstruents tend to agree in voicing in English; the only exceptions (eg, *Aztec*) involve clusters split by a syllable boundary, as these words are (allowing for *ad[s]orb* and *ab[s]orb*). The fact that /s/ sometimes surfaces as [z] after *per-*-, as in an alternate pronunciation of *persist*, and sometimes as [s] (as in *persuade*; the root *suade* also appears in *dissuade*, but unfortunately never with long-vowel prefixes) is likewise not necessarily a problem with s-Voicing. As we saw above, /s/ appears to be voiced occasionally after /ʃ/ in several cases, as in *Martian*, *disperse- dispersion*, *coerce-coercion*. The optionality of this [z], unlike the optional [z] in *resource* and so forth, has nothing whatsoever to do with the optionality of resyllabification. Not only is the stress pattern inappropriate for ambisyllabicity, but there is no dialect split; according to Kenyon and Knott (1953), [z] may appear in *persist* in American speech as well.

A stranger example, though, involves the prefix *dis-*. In most cases /s/ surfaces as [s] following this prefix, as in *dissuade*, *dissemble*, *dissolution*. In *dissolve*, however, the [z] is obligatory. What makes this particularly strange is that the noun formed from this verb, *dissolution*, contains an obligatory [s]. One possible way of handling *dissolve* might be to suppose that this word has somehow become reanalyzed as containing the prefix *de-* [dɪ], which allows s-Voicing. This is plausible, since *de-* and *dis-* are semantically very similar. Such a solution, however, would force us to the conclusion that *dissolution* is not in fact derived from *dissolve*.
It appears, then, that in prefixed forms there is some evidence for s-Voicing, though not much. Its application appears to be sensitive to syllabification, in particular syllable structure after the application of Resyllabification. Nevertheless, idiosyncrasies abound; for example, s-Voicing is obligatory in *resist* but only optional in *desist*, even though both show precisely the same stress pattern.

2.3.1.4 s-Voicing: Summary

In all three cases, namely prefixed forms, forms involving CiV Lengthening, and forms involving i-Shortening, there is very little evidence for the necessity of a rule of s-Voicing. In fact, the total number of forms where s-Voicing obligatorily applies is small enough to list all in a single figure, as I do below.

(61) a. Caucus - Caucasian  
    Malthus - Malthusian  
    Paris - Parisian  
    precise - precisian  

b. precise - precision  
    concise - concision  
    recluse - reclusion
I conclude, therefore, that although there is evidence for the existence of this rule, there is not much. In Chapter 3 I will discuss the implications of this fact for its relative productivity.

2.3.2 Velar Softening

The next rule I discuss is Velar Softening (Chomsky and Halle 1968, Rubach 1984, Halle and Mohanan 1985, Borowsky 1986, and others). I argue that the evidence supports the traditional analysis of this rule for the most part. As with s-Voicing, however, we will see that it is much more restricted than is commonly assumed. This will have consequences for the discussion of its relative productivity in the next chapter.

The alternations this rule is intended to account for are illustrated in (62). As can be seen, most of the cited cases involve the suffix -ic, which alternates in form between [ik] and [is] or [if] when followed by a nonlow front vowel.

a. [k] - [s]

-ic: critic - criticize, criticism
  public - publicity, publicize
  classic - classicism, classicist
  Catholic - Catholicism
  toxic - toxicity
  mediate - medicine
  Stoic - Stoicism
  lyric - lyricist
  syllabic - syllabicity
  electric - electricity
  phonetic - phoneticity

other: matrix - matrices
  reciprocal - reciprocity
  opaque - opacity

b. [g] - [ʤ]

  rigor - rigid
  larynx - larynges
  fungus - fungi, fungicide {also fun[g]i, fun[g]icide}
  analogue - analogy, analogous, analogize
  intellect - intelligent
  pedagog - pedagogy
  regal - regicide
  prodigal - prodigy
  dialog - dialogist {also dialo[g]ist}

²⁹Chomsky and Halle (1968:219) also include the pair allegation - allege.
Since Chomsky and Halle (1968), phonologists have assumed that the pattern seen in (62a,b) results from the application of the rule of Velar Softening, which changes /k/ to /s/ and /g/ to [ç]. The pattern in (62c) shows that Velar Softening precedes the rule of Palatalization (to be discussed below), which changes /s/ to /ʃ/.

Halle and Mohanan (1985) give the rule of Velar Softening essentially as follows.

(63) Velar Softening
[after (64) in Halle and Mohanan 1985:79]

\[ \{ k \rightarrow s \}, \quad [\text{-low}] \]
\[ \{ g \rightarrow ç \}, \quad [\text{-back}] \]

There are numerous theoretical and empirical problems with this rule. As Chomsky and Halle (1968) first observed, there is no way to change /k/ into /s/ by changing one feature; at the very least, both the Place feature and the value for [continuant] must change. Chomsky and Halle (1968:224) therefore accomplish the change of /k/ into /s/ in two steps. First, /k/ becomes a [+coronal, +strident, +anterior] segment, that is, the dental affricate /ç/. A later rule (a special case of Spirantization) then changes /ç/ to /s/. This
second step is not wholly implausible, since /c/ is not a distinctive sound in English and we may expect there to be automatic marking conventions/redundancy rules to convert it into a phoneme. Borowsky (1986:131-132) adopts a similar two-step analysis, whereby /k/ first becomes [-back], that is the palatal affricate /ʃ/, and markedness conventions then convert this into /s/. Her analysis is slightly less plausible in this already credulity-stretching area, since /ʃ/ already is a phoneme of English. Finally, Rubach (1984:27,39) rejects the two-step derivation of /s/ from /k/, preferring instead a rule that changes two features (namely, [+coronal] and [+continuant]) at the same time. His reason seems to be that for him Spirantization cannot apply after Velar Softening (he considers Spirantization to be cyclic and Velar Softening to be postcyclic; see Chapter 5 for discussion of such concepts), and so it cannot be the case that /k/ becomes /c/ by Velar Softening and then /c/ becomes /s/ by Spirantization.

Under all three analyses, the conversion of /g/ to /ʒ/ is direct. Chomsky and Halle (1968) and Rubach (1984) stipulate that Velar Softening makes /g/ [+coronal] and [+strident], but unlike the case with /k/, not [+anterior], thereby producing the voiced palatal affricate /ʒ/. Borowsky's (1986) analysis has the advantage here, as by applying the same rule she supposes for /k/, namely the insertion of [-back], /g/ will become /ʒ/. In both cases, there is no need for further rules, since /ʒ/ is an English phoneme.
In other words, either the change from /k/ to /s/ can be made a relatively reasonable one- or two-step process but the change from /g/ to /ç/ requires a slightly different rule, or the changes from /k/ to /s/ and from /g/ to /ç/ can be covered by the same rule that nevertheless necessitates an unmotivated clean-up process to make the /s/ surface properly. In two versions, an intermediate form (/c/ for Chomsky and Halle, /ʃ/ for Borowsky) is derived which never surfaces or has any effect whatsoever. Hence most authors, such as Kiparsky (1982) and Halle and Mohanan (1985), ignore the issue entirely and give the alternations simply as /k/ to /s/ and /g/ to /ç/, as in (63), without trying to reduce the rule to features.

Since we've recognized that under any analysis the /k/ to /s/ alternations in (62a,c) must be treated differently from the /g/ to /ç/ alternations in (62b), I will henceforth focus solely on the [k] cases. In any event the cases where [g] alternates seem much rarer and involve rather idiosyncratic processes (eg, the suffixation of -id in rigid and -t in intellect and the unexplained change in voicing in larynx-larynges and intellect-intelligent).

As can be seen from the way I arranged the forms in (62), most cases where [k] alternates (nineteen out of twenty-three examples cited in the literature) involve the same morpheme, namely the suffix -ic. As with the question in earlier sections about the necessity of s-Voicing in precise-precision, concise-concision, it would be preferable if Velar Softening affected other morphemes regularly as well.
The suffixes that should trigger Velar Softening consist of all those beginning with /i/ or /e/. This includes -ine, -i, -y, -ist, -ism, -ize, -ian, -ity and the special plural suffix -es (as in matrices). Below I have collected all examples (from a computer search) where a stem-final /k/ becomes [s] or [ʃ] before these suffixes, and where the /k/ is not part of the suffix -ic. To aid in the later discussion, I've divided the examples into two groups: those where the affected stem ends in the submorphemic sequence [ɪk], and all others.

(64) All Velar Softening alternations not involving the suffix -ic:

With [ɪk]

a. [k] - [s]
colchicum - colchicine
duplicate - duplicity
lubricate - lubricity
medicate - medicine
mendicant - mendicity
vortical - vorticity

b. [k] - [s]
calix - calices
calyx - calyces
cervix - cervicis
coadjutrix - coadjutrices
curatrix - curatrices
cyclix - cyclices
directrix - directrices
executrix - executrices
generatrix - generatrices
helix - helices
initiatrix - initiatrices
matrix - matrices
radix - radices
spadix - spadices
tractrix - tractrices
victrix - victrices
Without [\(k\)]

c. [k] - [s]  
caduous - cadugity  
opaque - opacity  
reciprocal - reciprocity  
abacus - abaci  
cergus - cergi  
coccus - cocci  
focus - foci  
logus - loci  
abacus - abacist  
pharmacology - pharmacist, pharmacy  
Spartagus - Spartacist  
Greek - Gregism  
Australopithecus - Australopithegine  
Greek - Gregize  

d. [k] - [s]  
apex - apices  
cimex - cimiges  
codex - codices  
harusped - haruspices  
ibex - ibiges  
index - indices  
lutex - latiges  
pontifex - pontifices  
scolex - scoliges  
vortex - vortiges  
bora - boraces  
crux - cruces  
feal - feges  
hallux - halluces  
hyra - hyrages  
lux - leges  

e. [k] - [\(\phi\)]  
Greek - Gregian  
Markus - Margian  

By far the largest morphological class of such examples involves the special plural suffix -es. The -ity class, for instance, involves only seven morphemes. It is true that there are numerous other instances of -ity appearing after [s] (I list a sampling below), but in all of these words [s] either does not alternate at all, or else only with [\(\phi\)] (due to the application of Palatalization, discussed below). Since [s] does not alternate with [k] in these cases, there is no reason to suppose that Velar Softening applies here.
The connection between the presence of -ic and Velar Softening with -ity is further illustrated by the existence of many pairs where it seems that -ic and -ity have been added together, as if they formed the single suffix [tstti]; that is, for these morphemes there is no independent form with -ic [tk] alone.30

(66) -icity

impudent - impudicity
multiple - multiplicity
simple - simplicity
triple - triplicity

30Ohala's (1986a) comments about the pair witty-witticism are also relevant here. As he notes, it appears that in the coining of witticism, the suffix was perceived to be not -ism but -icism.
Given the difficulty in finding large classes of forms showing Velar Softening but not involving the suffix -ic, the alternations with -es become quite important. Unfortunately the -es cases do not present straightforward evidence for Velar Softening either. On examining the examples in (64c), two things become clear. First, as these are all the cases in my database where suffixation of -es triggers Velar Softening, it is curious that this suffix may only be used if the singular form ends with /ks/, never, for example, just with /k/ alone. Thus given that the stems end in /ks/ and not /k/, Velar Softening should not be able to apply at all. In order to allow the vowel of -es (presumably underlyingly /e:/, since it surfaces as [iy] after Vowel Shift) to affect the /k/, the intervening /s/ must first be deleted. Alternatively, we could suppose that the stems underlyingly end with /k/, with /s/ added word-finally; this would account for alternations like those shown below. No matter how this problem is overcome, however, the special /s/ rule must be marked to apply only with stems that are also marked to take the -es plural.

(67) apex - apical
    helix - helical
    radix - radical
    vortex - vortical

Second, it is striking how often [ː] appears before the /s/ ostensibly derived by Velar Softening. As can be seen in (64c), it appears in most of the -es examples; sixteen out of the thirty-two examples end in [ːk] before -es suffixation, and an additional ten more end in [ːk] afterwards. Note that the string [ːk] also appears
in four out of the seven -ity cases not involving the suffix -ic
(namely, duplicate-duplicity, lubricate-lubricity, mendicant-
mendicity, vortical-vorticity). The presence of /l/ before /k/ is not
necessary for the application of Velar Softening, as can be seen from
the (rare) example like opaque-opacity, but the widespreadness of
such cases, coupled with the fact that most cases of Velar Softening
involve -ic, makes one suspect that this rule actually consists in fact
of nothing more than a set of allomorphy rules (in the sense of
Aronoff 1976) that have evolved in analogy to -ic allomorphy.

I will, however, not explore this possibility further here.
Rather, I now turn to the question of the ordering of Velar Softening.
The only rule Velar Softening has been claimed to follow is s-Voicing.
The argument runs as follows.

Recall that in my discussion of s-Voicing, a root-initial /s/ may
be pronounced [z] in at least one dialect if preceded by a long-vowel
prefix. There are, however, exceptional cases where s-Voicing never
applies in any dialect. Examples are listed below.

(68) [includes all cases from Chomsky and Halle 1968:121; Halle and
Mohanan 1985:99; Borowsky 1986:128]

C[s] - V:[s]
i ncite, excite - re cite
c ongede - re cede
c ongeive - rece ive
e xcess - r esc ess

What all these roots have in common is that they contain a
vowel, namely /i/ or /e/, before which Velar Softening is supposed
to apply. Thus it is possible that these roots begin with an underlying /k/, which then becomes /s/ by Velar Softening. We can see by the (unique) example *Grecian* that the /s/ produced by Velar Softening is not subject to s-Voicing. This can be easily accommodated if we simply assume that s-Voicing applies before Velar Softening. Hence the underlying /k/ in the roots in (68) do not become /s/ by Velar Softening until after s-Voicing has had its chance to apply.

There is additional evidence for the application of Velar Softening inside roots. Namely, when the prefix *a(d)-* is added to roots like those shown in (69), a [k] unaccountably appears.

(69) [includes all examples from Chomsky and Halle 1968:222; Halle and Mohanan 1985:99; Borowsky 1986]

\[
[s] - [ks]
\]

incident - accident
recede - accede
recess - access
degenerate - accelerate

A similar pattern is seen in words with the prefix *su(s)-*, where a [k] or a [g] unaccountably surfaces.\(^{31}\)

---

\(^{31}\)Borowsky (1986:122) also includes the words *accent* and *succinct* (also pronounced without [k]), even though to my knowledge the roots contained in these words do not appear anywhere else.
(70) [includes all examples from Chomsky and Halle 1968:222; Halle and Mohanan 1985:99; Borowsky 1986:122,129]

recess - success
recede - succeed
congest - suggest

This suggests that the prefixes \textit{a(d)-} and \textit{su(s)-} preserve the underlying root-initial /k/ or /g/, specifically by allowing them to geminate before Velar Softening applies. In autosegmental terms this is easy to represent, as Halle and Mohanan (1985) and Borowsky (1986) demonstrate, by supposing that these suffixes come with an empty prosodic slot that becomes filled by a following consonant. The right half of the geminate is then alone affected by Velar Softening.

\begin{align*}
\text{(71) } & X X & X X X & X X X X X \\
& \text{æ} & \text{k e s} & \text{æ} \text{ k e s}
\end{align*}

Note that this analysis forces us to say something concerning \textit{Geminate Inalterability}, the observation that universally, rules cannot affect one half of a geminate (Hayes 1986a, 1986b, Steriade 1982, Steriade and Schein 1986). According to this observation, Velar Softening should not be able to apply in the structure shown in (71), and yet it apparently does. The solution adopted by Steriade and Schein (1986) and Borowsky (1986) is to conclude that gemination after \textit{a(d)-} and \textit{su(s)-} involves not spreading but copying. The geminate thus formed is not doubly linked, and so Geminate Inalterability does not apply to it.
This seems ad hoc. It is true that in some cases copying appears to be preferable to spreading as a description of harmony systems (see, eg, Archangeli and Pulleyblank forthcoming), but all such cases involve nonlocal harmonies. The special copying rule here can be avoided if we simply assume instead that prefixes like a(d)- come in different context-sensitive allomorphs, specifically /æk/, /æd/ and /æ/, exactly parallel to the context-sensitive allomorphs of the English indefinite article. As Aronoff (1976) emphasizes, allomorphy in English affixes is not at all unusual. This can be seen quite clearly with the suffix -ion, which shows up as -ion in rebellion, -ation in information, -ition in repetition, and -tion in deduction (see also Chapter 3). I will briefly return to this concept later. Nevertheless, given the conservative stance I am forced to adopt in this chapter to avoid obscuring my major interests, I will merely note the problem posed by the gemination-as-copying analysis, mention a possible solution involving allomorphy, and move on.33

32 More on the significance of cases like a/an allomorphy will appear in Chapter 6.
33 One datum relevant in any discussion of this problem is the pair coccus-cocci, which shows an alternation between [k] and [ks]. The appearance of the [s] could be ascribed to Velar Softening triggered by the suffix -i. The presence of [k] in cocci is problematic, however. Unlike the case with a(d)- and su(s)-, we cannot say that the appearance of [ks] rather than [s] is due to the suffix; cf the [k] - [s] alternation in locus-loci, focus-foci. By analogy with the gemination-as-copying idea mentioned above, we could suppose that the stem coc- contains an underlying geminate, represented as two adjacent [k] segments; however, the Obligatory Contour Principle (see eg McCarthy 1986, Yip 1988) disallows just such an analysis. Also relevant is the pair ascend-descend, in which the root-initial consonant remains [s] in both words; this case is not amenable to an analysis involving root-internal application of
In conclusion, then, I accept that Velar Softening is a descriptively adequate rule, albeit rather a restricted one. Specifically, the vast majority of cases showing alternations between [k] and [s] involve either the suffix -ic or a morpheme ending in the sequence [lk]. I also accept the traditional ordering of s-Voicing before Velar Softening, although again the evidence for this is not strong. The weakness of the argument for the ordering comes from its crucial reference to the nonapplication of s-Voicing in certain cases, while we saw above that s-Voicing often fails to apply even where there is no other rule to ascribe the exceptions to.

2.3.3 Spirantization, y-Insertion and Palatalization

The rules we have considered so far in this section, namely s-Voicing and Velar Softening, have turned out to be rather exceptional-ridden. As we move to rules that are thought to apply later in the derivation, however, we begin to find rules with more robustness. The connection between the ordering of these rules and their reliability as generalizations will be discussed in Chapter 3.

The primary pattern I am interested in in this subsection is the alternation between the alveolar stops and fricatives [t,d,s,z] and the palatal affricates and fricatives [ʃ,ʒ,ʃ,ʒ]. Which alveolar is paired with which palatal is rather complex, as illustrated by the examples below. One important thing to keep in mind is that this time, this list is only selective; the pattern illustrated here is too widespread to

Velar Softening without positing an excessively abstract underlying form for send, specifically /skend/.
make an exhaustive list. For example, of over twelve thousand words where -ion is suffixed to a stem ending in /t,d,s,z/, every single one conforms to the appropriate patterns illustrated below.

(72) [includes pairs from Rubach 1984, Halle and Mohanan 1985]

a. \(/s/ \rightarrow [ʃ] / -ion, -ian, -ial, -ious, -ual, -uous\)
   
   confess - confession
   Venice - Venetian
   face - facial
   office - official
   space - spacious
   sense - sensual
   sex - sexual
   sense - sensuous

b. \(/z/ \rightarrow [ʒ] / -ion, -ian, -ial, -ious, -ual, -uous\)
   
   confuse - confusion
   use - usual

c. \(/t/ \rightarrow [ʧ] / s -ion, -ian, -ial, -ious\)
   
   digest - digestion
   Christ - Christian
   beast - bestial

d. \(/t/ \rightarrow [ʃ] / X -ion, -ian, -ial, -ious, where X \neq s\)
   
   extinct - extinction
   contort - contortion
   delete - deletion
   invent - invention
   resident - residential
   torrent - torrential
   part - partial
   infect - infectious
e. /d/ \rightarrow [ʒ] / -ion, -ian, -ial, -ious 
    conclude - conclusion

f. /d/ \rightarrow [ʃ] / n -ion, -ian, -ial, -ious 
    extend - extension 
    expand - expansion

g. /l/ \rightarrow [ʒ] / -ual, -uous 
    concept - conceptual 
    event - eventual 
    habit - habitual 
    tumult - tumultuous 
    tempest - tempestuous 
    contempt - contemptuous

h. /d/ \rightarrow [ʒ] / -ual, -uous 
    grade - gradual 
    reside - residual

The first thing to notice is that all of the suffixes that cause palatalization begin with /y/, as can be seen when they appear in words where palatalization doesn't apply.

(73) rebel - rebellion [ˈrebɛlɪən] 
    bile - bilious [ˈbɪlys] 
    manipulate - manual [ˈmænjuəl]

By contrast, palatalization does not occur with suffixes that begin with [i] rather than [y], as can be seen by the examples in (74).
(74) a. *ive:
transit - transitive
permit - permissive
gerund - gerundive
divide - divisive
abuse - abusive

b. *ible:
contempt - contemptible
permit - permissible
auditory - audible
sense - sensible

c. *ity:
quantify - quantity
scarcity - scarcity
atrocious - atrocity

If the cause of the palatalization in the words in (72) is a /yl, the question arises about why this /yl does not surface. Chomsky and Halle (1968), with Rubach (1984) and Halle and Mohanan (1985) following, propose that it is removed after palatalization by a rule that deletes /yl after palatal segments.34 This rule of y-Deletion is easily expressed as in (75). I follow Pulleyblank (1989) and Wiswall (1991) in assuming palatals are marked [+front], which depends from the Coronal node (cf Sagey 1986). The rule of Palatalization can then be formalized as the spread of [+front], as below in (76).35

34Borowsky (1986) suggests, alternatively, that the /yl is deleted automatically as part of the palatalization process. I see much to recommend this analysis, but will not pursue this line here since this question is not relevant to my primary concerns.

35Later we will see why the presence of the moraic slot following /yl is necessary.
(75) **y-Deletion**

\[ y \rightarrow \emptyset / [+\text{front}] \_ \]

(76) **Palatalization**

\[
\begin{array}{c|c|c|c|c}
\sigma & \sigma \\
\hline
\mu & \mu & \mu & \mu \\
\hline
\text{Cor} & \text{Cor} & \rightarrow & \text{Cor} & \text{Cor} \\
\hline
\text{[+front]} & \text{[+front]} \\
\end{array}
\]

So far we know that palatalization is triggered by /y/ and that this /y/ is then deleted. I now return to the alternations in (72).

One aspect of the pattern seen there is that the stops /t, d/ are not just palatalized but are also changed to fricatives (rather than, say, the palatal affricates [ʃ, ç]). That is, they change in more than one feature. Beginning with Chomsky and Halle (1968), this has been suggested to be due to the application of another rule, **Spirantization**, which converts /t, d/ to /s, z/, which are then changed to [ʃ, ç] by Palatalization.

One claimed source of independent evidence for Spirantization comes from the [t]-[s] alternations seen with the suffix -y. Some examples are given below.
Although this pattern is quite general for -y, there are some problems worth noting. First of all, unlike the alternations seen in (72), /d/ is not affected; Chomsky and Halle (1968:229) cite as support for this claim the example remedies (cf remedial). Second, as Chomsky and Halle (1968:229) also note, some cases of stem-final /t/ do not change into /sl/: difficult-difficulty, modest-modesty. They comment that this might be handled by supposing that the suffix in these cases is actually -ty, which applies only with adjectives (cf royal-royalty, loyal-loyalty). However, the spirantizing -y also occurs quite commonly with adjectives, eg accurate-accuracy, brilliant-brilliancy, decent-decency, private-privacy. This implies that difficult and modest are simply exceptions to the rule.

Spirantization is also claimed to play an important role in handling the data in (72), however, specifically the difference in output produced by suffixes like -ion, -ian, -ial, -ious (what I'll call the YV suffixes) as opposed to suffixes like -ual and -uous (what I'll call the YU suffixes). The argument, due to Chomsky and Halle (1968) and followed by all other workers since, runs approximately as follows.
First note that in all cases except those in (72c), where /t/ is preceded by /s/, the YV suffixes produce palatals that are fricatives, namely [ʃ,ʒ]. By contrast, the YU suffixes only produce fricative palatals if the underlying segment already was a fricative; alveolar stops become the affricate palatals [tʃ,çʃ].

Now observe that the palatalization process triggered by YU suffixes is easy to capture with a single feature change, namely that of place: alveolars become palatals. The fact that alveolar stops become affricates is an automatic consequence of the fact that English has no palatal stops, only palatal affricates and fricatives. By contrast, the change triggered by YV suffixes is more complex. Not only do the alveolars change place to become palatals, the stops also change manner, becoming full-fledged fricatives. This suggests, as noted above, that this change is done in two steps: first, /t,d/ become /s,z/; then /s,z/ become /ʃ,ʒ/.

This assumption has several advantages in addition to allowing rules that only change one feature at a time. First, it allows there to be only one palatalization rule, which changes /t/, /d/, /s/ and /z/ to /ʃ/, /çʃ/, /ʃ/ and /ʒ/, respectively. This rule would apply after the /t,d/-to-/s,z/ rule. Second, there is already independent motivation for such a /t,d/-to-/s,z/ rule, namely the evidence for Spirantization discussed above. Third, the fact that /st/ becomes [ʃʃ] after YV suffixes, and not [ʃʃ] as expected, can be attributed to a reasonable constraint on Spirantization, not on Palatalization. This is because, as Borowsky (1986) has observed, in order to derive [ʃʃ] under the
current analysis, there would have to be a stage where Spirantization
would create two adjacent /s/s, a situation that could be ruled out by

Following Chomsky and Halle (1968) and later workers, I
assume that the suffix -y is underlyingly /y/; the independent
evidence for this will not be discussed here. The rule of
Spirantization may thus be expressed as follows (assuming it may be
blocked by the OCP under the appropriate circumstances).

(78) Spirantization
[after (85) in Halle and Mohanan 1985:87]

\[
\begin{array}{c|c}
\sigma & \sigma \\
| & | \\
\mu & \mu \\
& / \ \backslash \\
\text{Cor} & \text{Cor} \rightarrow \text{Cor} [+\text{cont}] & \text{Cor} \\
& | & | \\
& [+\text{front}] & [+\text{front}] \\
\end{array}
\]

One apparently strange consequence of this analysis is that it
predicts that the /y/ in YU suffixes like -ual and -uous is not present
at the time Spirantization applies. But this strange prediction does in
fact seem to be supported by the evidence. It appears as if the [y]
that appears before [u] in English is not underlying at all, but is
inserted by rule.

Simplifying somewhat, the observation that supports this
conclusion is that in English, [y] may not follow a consonant unless
the following vowel is [u]. Hence, for example, we find the pattern
with initial [k] illustrated below.
The rule that inserts this [y], called $y$-Insertion in Halle and Mohanan (1985), is somewhat difficult to express. The first problem is that the rule cannot be given as simply as "insert /y/ before /u/" because then the rule has literally thousands of exceptions. Some examples are listed below.

Typically, therefore, it is assumed that forms must be marked in some way to undergo the rule. Chomsky and Halle (1968), Rubach (1984) and Halle and Mohanan (1985) accomplish this by supposing that the vowel in a $y$-inserting word like cube is underlingly different from that of a non-$y$-inserting word like cool, even though both end up surfacing precisely the same. Borowsky (1986), similarly, suggests that words that undergo $y$-Insertion already contain a /y/ of a sort in their underlying representation; for her, $y$-Insertion is replaced by a set of syllabic conditions that allow the underlying /y/ to emerge in the proper cases.
Unfortunately, both analyses are ad hoc. The first requires the phonetic [u] in cool to be represented in quite a different way from the phonetic [u] in cube, making it a coincidence that some words with [u] vary freely in whether they insert /y/. This is especially problematic given that the underlying forms must also undergo Vowel Shift. Hence in Chomsky and Halle (1968) and Halle and Mohanan (1985) the underlying vowel in cube is argued to be the high back unrounded vowel /iː/, while that in cool is /oː/. The situation gets even more complex when short vowels are considered.

Borowsky's (1986) analysis sets up the underlying form of a word like cube as illustrated in (81a), leaving it a mystery why underlying forms such as those in (81b) are not allowed.

\[
\begin{align*}
\text{(81) a. } & \quad \text{[based on (24) in Borowsky 1986:281]} \\
& \quad \begin{array}{c} X \ X \ X \\
| & | | \\
\text{k I U b} \\
\end{array} \\
\text{b. } & \quad \begin{array}{c} *X \ X \ X \\
| & | | \\
\text{k U I b} \\
\end{array} \\
& \quad \begin{array}{c} *X \ X \ X \\
| & | | \\
\text{k I U b} \\
\end{array} \\
& \quad \begin{array}{c} *X \ X \ X \\
| & | | \\
\text{k U I b} \\
\end{array}
\end{align*}
\]

There is a simple way out of this dilemma, however. As Kiparsky (1982) observes, y-Insertion is actually quite a general process in a subset of English vocabulary: the Latinate, or more precisely, the non-Germanic forms. Very few of the relevant non-Germanic forms are exceptions to y-Insertion (eg, acoustic, coupon in some dialects); few Germanic forms (eg, few) undergo the rule. The diacritic feature [+Germanic] plays an important role in much of
English morphology and phonology; in the next chapter we will see psycholinguistic evidence for such a diacritic as well.\footnote{Note that this analysis of y-Insertion correctly predicts that it is not possible for y-Insertion to apply in one part of a word but not in another, if its conditioning environment is met in both places. That is, there are no words containing strings like *CyuCu or *CuCyu.}

Thus I propose that the vowels in words like \textit{cube} are underlyingly the same as vowels in words like \textit{cool}. The difference is that the former words are diacritically marked to undergo y-Insertion by virtue of its being [-Germanic] and the latter are not. Non-Germanic forms that fail to undergo y-Insertion, like \textit{acoustic}, are simply marked as exceptions. Germanic forms that undergo the rule are specially marked to undergo the rule as well. I formalize the rule as follows.

\begin{equation}
\emptyset \rightarrow y / _u \quad \text{in [-Germanic] and other marked forms}
\end{equation}

In spite of its apparent strangeness, then, I adopt the essentials of the Chomsky and Halle (1968) analysis of the pattern illustrated in (62). Spirantization applies first in (62c,e,f). The rule of y-Insertion then inserts /y/ on the suffixes in (62g,h). Finally, Palatalization applies in all of the relevant suffixed forms in (62); y-Deletion removes the triggering /y/.

\footnote{Borowsky (1986) includes a very interesting discussion of the Palatalization that is triggered by the /y/ inserted by y-Insertion and how it is sensitive to stress in different ways in different dialects. I will not discuss this here, but interested readers can find more evidence there of the suggestion I made above on the basis of dialect differences in the application of s-Voicing and Flapping that it is actually dialect differences in Resyllabification that cause the variation.}

\textit{y-Insertion} \footnote{Note that this analysis of y-Insertion correctly predicts that it is not possible for y-Insertion to apply in one part of a word but not in another, if its conditioning environment is met in both places. That is, there are no words containing strings like *CyuCu or *CuCyu.}
2.3.4 Coronals: Summary

In this section I have discussed rules that affect or produce coronals, namely s-Voicing, Velar Softening, Spirantization, y-Insertion, Palatalization and y-Deletion. They are standardly ordered in the way I have listed them, and I have found no strong reason to dispute this standard analysis. I have, however, found evidence for supposing that s-Voicing and Velar Softening are rather unproductive rules, that Palatalization is rather productive, while the productivity of y-Insertion, applying as it does only in specially marked word classes, falls somewhere in between. The fact that Spirantization only applies independently of Palatalization with the suffix -y also suggests that it is not as productive as Palatalization. These facts about the rules' relative productivity will be discussed further in Chapter 3. For now I just give the ordering of these rules in the following chart.

(83) a. s-Voicing [before b, c]
    b. Velar Softening [before e]
    c. Spirantization [before d, e]
    d. y-Insertion [before e]
    e. Palatalization [before f]
    f. y-Deletion

2.4 Rules involving nasals

The final set of rules I discuss consists of the rules Halle and Mohanan (1985) call Nasal Assimilation and Noncoronal Deletion. I provide additional evidence supporting the claim in Borowsky (1986) that there are two distinct applications of Nasal Assimilation, although my formulation of these rules will differ somewhat from
hers. The question of whether these are two applications of the same rule or two different rules is left to Chapter 5. I then show that while aspects of the distribution of [ŋ] indicate that in some cases it is derived from /Ng/ (ie, a nasal unspecified for place followed by a /g/ that later deletes by Noncoronal Deletion), we must nevertheless assume that /ŋ/ is also a phoneme of English, contrary to Borowsky (1986) and earlier workers, although a marginal phoneme like /ʒ/.

2.4.1 Nasal Assimilation

In this section I argue that there are two rules assimilating nasal consonants in place to the following segment in English, which I will call Obligatory Nasal Assimilation and Optional Nasal Assimilation. These rules differ in four ways. First, as the names imply, one is obligatory while the other is optional. Second, the optional rule applies more readily if trigger and target are "temporally close"; the obligatory rule has no such constraint. The concept of temporal closeness will be discussed below.

Third, the obligatory rule is a feature fill-in rule (ie, it does not apply if the nasal consonant is already specified for place), while the optional rule is a feature-changing rule (ie, it applies even if the nasal consonant is already specified for place). Finally, the obligatory rule applies before a dialect rule I will call Prevelar Raising, while the optional rule follows it.

I formulate these two rules as in (84a) and (84b), respectively, assuming some version of feature geometry (cf Sagey 1986, McCarthy 1988). The idea is that Obligatory Nasal Assimilation only

---

38 The concept of temporal closeness will be discussed below.
affects nasal consonants not already specified for place, while Optional Nasal Assimilation only affects nasal consonants specified for place. Moreover, Optional Nasal Assimilation applies more readily if target and trigger are temporally close.

(84) a. **Obligatory Nasal Assimilation**

```
[+cons] [+nas] [+cons]
    \\
Root     Root
```

b. **Optional Nasal Assimilation**

```
[+cons] [+nas] [+cons]
    \\
Root₁     Root₂
```

In (85) below I give examples showing Obligatory Nasal Assimilation. Note that for these words assimilation is obligatory. Moreover, stress, which as we'll see plays a role in determining temporal closeness, has no effect here.
(85) Obligatory Nasal Assimilation.
[pronunciations checked with Woolf et al. 1981]

a. /N/ → [n]:
   Honda         Hondúras

b. /N/ → [m]:
   crumble      umbrëlla

c. /N/ → [ŋ]:
   linguist     linguistics
   Húngary      Hungárian
   lñger        lñguñine

The examples in (86) show that when the nasal is already specified for place, Obligatory Nasal Assimilation does not apply; the assimilation that optionally occurs must be ascribed to Optional Nasal Assimilation. Derivations illustrating the contrast between Obligatory and Optional Nasal Assimilation in this respect are given in (87).

(86) Obligatory Nasal Assimilation is not feature-changing:
[pronunciations checked with Woolf et al. 1981]

a. /m/ → [m] or [n]:
   Camden

b. /n/ → [n] or [m]:
   Cánberrá  lóganbërry     fnpùt
c. /n/ → [n] or [ŋ]:

<table>
<thead>
<tr>
<th>Obligatory Nasal Assim</th>
<th>Optional Nasal Assim</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>Móngol</td>
<td>Mongól</td>
<td>Mongól</td>
</tr>
<tr>
<td>tránquil</td>
<td>tranquility</td>
<td>tranquility</td>
</tr>
</tbody>
</table>

By contrast to Obligatory Nasal Assimilation, Optional Nasal Assimilation is sensitive to stress. Specifically, a nasal is more likely to assimilate through Optional Nasal Assimilation if preceded by a stressed syllable than by a nonstressed syllable, as seen in the contrast between Canberra, where [m] is quite likely, and loganbêrry, where [m] is not as likely. However, it cannot be the case that a particular stress pattern is a necessary condition for the application of the rule, since nasal assimilation is not fully blocked in forms like loganbêrry. Note also that in forms such as Ínpút, which consists of two feet, nasal assimilation nevertheless occurs quite readily. Hence I follow Borowsky (1986:95) in assuming that what is crucial is not prosodic structure per se that is relevant, but what she calls temporal closeness. The nasals in Canberra and Ínpút are "closer" to the following consonants as an indirect result of the stress pattern, where the nasal is preceded by a greater degree of stress than that which appears on the following syllable. By contrast, the nasal in loganbêrry is not as "close" to the following consonant because it is preceded by a lesser degree of stress (namely no stress) than that
which appears on the following syllable. Other factors influencing this temporal closeness include speaking rate; thus Optional Nasal Assimilation is more likely to apply in faster speaking rates than in slower.

Obligatory Nasal Assimilation does not apply across word boundaries. The examples below show that word-external nasal assimilation is always optional, even if the stress environment is designed to make target and trigger temporally close.

(88) Obligatory Nasal Assimilation does not apply across word boundaries:

\[ /n/ \rightarrow [n] \text{ or } [m]: \]

téñ bûcks
in Pittsburgh
Don't lôse both games -- wîn both!

One peculiarity that arises in this picture of Nasal Assimilation, as observed by Borowsky (1986), is the fact that the nasals in nasal-final prefixes such as con-, in-, and syn- obligatorily assimilate only in Coronal and Labial, as seen in (89a). By contrast, assimilation of Dorsal is optional and stress-sensitive, as illustrated in (89b); "%" indicates that the marked word may optionally be pronounced with the unassimilated [n] rather than the assimilated [ŋ].
(89) Prefixes and Dorsal assimilation:
[after Borowsky 1986:89, 92; pronunciations checked with Woolf et al. 1981]
a. /N/ → [n]:

condemnátion  condémn
intonátion  intóne

/N/ → [m]:

còmposition  compósé
cómpetent  compártement
sýmbol  symbólic
sýmphony  symphónic

b. /N/ → [ŋ] or [n]:

cóngress  %congrésional
cóngruous  %congrúity
cónquer  %concúr
%cóncord  %cóncórdance
%sýnchrony  %sýnchrónic
%incubate  %inclúde
%cóncrete  %concrétion

Based on this asymmetry, Borowsky (1986) concludes that Obligatory Nasal Assimilation does not apply to spread Dorsal across a prefix boundary, though tautomorphemically Dorsal can be spread by this rule, as seen by examples like those in (85c), where an unassimilated [n] is not possible. This then is another difference between the obligatory and optional versions of this rule. The apparently special nature of [ŋ] will be discussed further in the next section.
The final piece of evidence showing that Obligatory and Optional Nasal Assimilation are distinct (at least distinct applications) is the fact that Obligatory Nasal Assimilation feeds the dialect rule that I will call Prevelar Raising while Optional Nasal Assimilation does not.

This rule, discussed in Donegan and Stampe (1979), appears in different dialects in different forms, while some dialects do not have it at all. For some speakers (including myself), the vowel in a word like *rang* sounds more like the vowel in a word like *rain* (ie, [ey]) than in *ran* (ie, [æ]). Evidence we will consider in the next section suggests that even for these speakers, the vowel in *rang* at some level actually is /æ/. This implies that such speakers have a rule, Prevelar Raising, which raises /æ/ to [ey] before [ŋ]. Note that this rule is phonetically natural, since [ŋ] is [+high]. This same process is seen in these speakers' pronunciation of a word like *length* with the vowel [i], not the standard [e].

As Donegan and Stampe (1979) note, in some dialects this process applies in even broader environments, so that in some [ɪ] appears as [ɪŋ] and in others [e] appears as [eŋ]. Examples from these various dialects are given below.

(90) a. [from my own speech]

<table>
<thead>
<tr>
<th>Word</th>
<th>Pronunciation</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>rang</td>
<td>[reyn]</td>
<td>cf rain [reyn]</td>
</tr>
<tr>
<td>strength</td>
<td>[struŋθ]</td>
<td>cf string [struŋ]</td>
</tr>
</tbody>
</table>
b. [after Donegan and Stampe 1979:149]

fish [fɪɻ]  
leg [leɻg]

Like Obligatory Nasal Assimilation, stress is irrelevant for the application of Prevelar Raising, at least in my own speech. This is illustrated below in (91).  

(91) [from my own speech]

a. [æŋ] in other dialects always appears as [ɛŋ] in mine:

ángér Angóla  
tránquil tranquility

b. [ɛŋ] in other dialects always appears as [ɪŋ] in mine:

stréŋth encóurage  
léŋth encóunter

Hence I formulate the rule as it appears in my own dialect as follows. In my dialect the rule does not affect high vowels, so sing contains [ɪ].

(92) Prevelar Raising

\[
\begin{cases}
  [+\text{low}] \rightarrow [-\text{low}] \\
  [-\text{high}] \rightarrow [+\text{high}] 
\end{cases}
\]

As Donegan and Stampe (1979) show, Prevelar Raising is ordered before Optional Nasal Assimilation, but not Obligatory Nasal Assimilation. Prevelar Raising must apply after Obligatory Nasal Assimilation.

---

39 I have found some words where I agree with other dialects in saying [æŋ], namely the words angóra and Rangóon. Here I follow a suggestion of Mike Hammond and assume that this has more to do with these words being foreign borrowings (of low frequency) than with their stress patterns.
Assimilation because it is triggered by obligatory derived velar nasals in monomorphemic forms, such as that in *anger*. However, it must apply before Optional Nasal Assimilation because when [ŋ] is derived through assimilation to the initial consonant of another word, Prevelar Raising is not triggered. Examples showing various stress patterns are given in (93), with sample derivations in (94).

(93) mankind [mæŋkænd] *[məŋkænd]
    send comics [sɛŋ kâmiks] *[sɪŋ kâmiks]
    tan canœ [tæŋ kənûw] *[tɛŋ kənûw]

(94) tan canœ /tæŋkənûw/ anger /æŋɡər/
    Obligatory Nasal Assimilation --- æŋɡər
    Prevelar Raising --- éyŋɡər
    Optional Nasal Assimilation tankənûw ---
    Output [tæŋkənûw] [ɛyŋɡər]

In this section, then, I have argued that nasal assimilation in English occurs through the application of Obligatory Nasal Assimilation and Optional Nasal Assimilation. These rules are distinguished in various ways, including in their ordering relative to Prevelar Raising; the question of whether they can nevertheless be collapsed into one rule is left to Chapter 5.

2.4.2 Noncoronal Deletion

In this section I consider the rule Halle and Mohanan (1985) call *Noncoronal Deletion*, which deletes /g/ and /b/ after tautosyllabic nasal consonants. Like Prevelar Raising, this rule is ordered after Obligatory Nasal Assimilation and before Optional Nasal
Assimilation. I show that though this rule does seem to be accurate in describing some distributional patterns, it cannot be used to justify claims that /ŋ/ is not a phoneme of English, as is done in Borowsky (1986).

The primary purpose of this rule is to allow for an analysis of [ŋ] as underlyingly /Ng/. If Nasal Assimilation applies before Noncoronal Deletion, /Ng/ should surface as [ŋ] if /N/ and /g/ are tautosyllabic, as for example when /Ng/ is word-final. If /g/ forms the onset of another syllable, Noncoronal Deletion will be blocked and /Ng/ should surface as [ŋɡ], as for example when /Ng/ is word-internal.

The rule is also argued to delete /b/ after tautosyllabic nasals, changing underlying word-final /Nb/ into [m] and word-internal /Nb/ to [mb]. While [ŋ] is claimed always to be derived, either through assimilation to a /g/ (or some other segment) which later deletes or, as in the cases discussed in the previous section, through assimilation to a velar that does not delete, [m] is claimed only to be derived from /Nb/ in certain words.

Given this, it is clear that Noncoronal Deletion must follow some version of Nasal Assimilation in order to spread the Dorsal feature from /g/ before it is deleted. This spreading rule is obligatory; speakers who pronounce words like fighting with a word-final [n] instead of [ŋ] simply have a different lexical representation for this suffix (see Labov 1989 for arguments for this). Hence it appears that Obligatory Nasal Assimilation is ordered before Noncoronal Deletion.
I provisionally formulate Noncoronal Deletion as in (95). Note that this formulation, based on those found in the literature, requires reference to something like an "anti-place" feature, something explicitly rejected as impossible in feature geometry theories (see, eg, Sagey 1986, McCarthy 1988). Following Halle and Mohanan (1985), I use the feature specification [-coronal].\(^{40}\) Shortly, however, we will see a way to get around having to use this sort of notation.

(95) **Noncoronal Deletion** (penultimate version)
[after (106c) in Halle and Mohanan 1985:96 and (11) in Borowsky 1986:73]

\[
\begin{align*}
\text{[-son]} & \rightarrow \emptyset / [+\text{nusal}] \sigma \\
\text{[+vcd]} \\
\text{[-cor]} &
\end{align*}
\]

As it happens, the evidence for this rule is primarily indirect, as there is very little evidence from alternations supporting either the claim that word-final /g/ is deleted or that word-final /b/ is deleted. The results of an exhaustive computer search are given below. The forms in (96a) show alternations involving /Ng/, and those in (96b) show alternations involving /Nb/.

(96) a. \([\eta]-[\eta g]\)

long - longer, longest, elongate, prolongate
strong - stronger, strongest
young - younger, youngest

---

\(^{40}\)Borowsky (1986:73) avoids the problem entirely by formulating the rule as \(g/b \rightarrow \emptyset / N \sigma\), that is, without specifying features.
b. [m]-[mb]

bomb - bombard, bombardier, bombardon
(chor)amb - (chor)ambus, (chor)ambic
crumb, crumby - crumble
plumb, plumber, plummet - plumbic, plumbiferous,
plumbism, plumbum
rhomb - rhomboid, rhombus
thumb, thumbkin - Thumbelina

In addition to the shortness of this list, two factors make the alternation evidence for Noncoronal Deletion very poor. First, in the case of the [ŋ]-[ŋg] alternation, the only morphemes showing this alternation, namely long, strong, and young, are also unique in allowing the /g/ to surface with the suffixes -er and -est (compare wronger, wrongest, where /g/ does not surface). That is, it may well be that these cases of the [ŋ]-[ŋg] alternation are due to idiosyncratic properties of these morphemes and not to a general rule.

Second, with regard to the [m]-[mb] alternations, there are problems as well. Note that the suffixes that supposedly bleed the /b/-deletion rule show up in virtually no other word, implying that many of the forms on the right in (96b) are essentially monomorphemic; this rules out bombard and its derivatives, crumble, and Thumbelina. Other suffixes appear in other words, but in these words result in strange shifts in meaning; eg, plumb has to do with measuring depth by using a weight, historically made of lead, while plumbic, etc, have nothing to do with measuring, only with lead. It might be argued that the reason why [b] surfaces in
these forms is precisely because these suffixes are so unproductive; according to Lexical Phonology, unproductive suffixes are affixed early in the derivation, and so could be added before /b/-deletion. The trouble is, unproductive suffixes appear on the left side of (96) as well, presumably after /b/ deletion has already applied: plummet (drop like a plumb), thumbkin.

When we turn to distributional evidence, the case for Noncoronal Deletion becomes a bit stronger. However, of the three distributional arguments Borowsky (1986) presents for the claim that [ŋ] is derived from /Ng/ by the ordered application of Nasal Assimilation and Noncoronal Deletion, only one of them seems to be truly valid.

The first argument is that [ŋ] never appears syllable initially. If [ŋ] is underlyingly /Ng/, this is understandable, since such a sequence violates the sonority hierarchy and so cannot form a legal onset.

The problem is that /Ng/ and /Nb/ apparently also form illegal codas. This is seen in the fact that unlike [nd], which always may appear word-finally (unless Coronal Deletion, discussed in Chapters 4 and 5, optionally removes the /d/), in most dialects [ŋɡ] and [mb] may never appear word-finally. The reason /b/ and /g/ are targeted for deletion while /d/ may stay seems to be related to the

---

41 In some dialects, word-final /Ng/ is always pronounced [ŋɡ]. Similarly, in some dialects some subset of the words iamb, choriamb, gamb, rhomb and rhumb may be pronounced with a word-final [mb] (according to Woolf et al. 1981).
fact, noted earlier in this chapter during the discussion of Closed Syllable Shortening, that a word-final coronal is always extraprosodic. Thus for instance syllable-final [nd] does not triggerClosed Syllable Shortening, while [mp] and [ŋk] do. In addition, as Selkirk (1983) points out, coda obstruent clusters are disallowed unless the final consonant is a coronal; thus forms like rift and six are possible, but not *rifk or *sipf. As has often been noted (see, eg, S. Myers 1987), such facts suggest a rule like that given in (97), which assigns extraprosodicity to word-final coronals.

(97) **Coronal Extraprosodicity**

\[
\text{Coronal} \rightarrow [\text{+extraprosodic}] / \text{Word}
\]

With the prior application of this rule, Noncoronal Deletion can thus be rewritten as in (98). This rule will act to delete all syllable-final voiced obstruents. It will not affect /d/ because the prior application of the rule in (97) makes /d/ invisible to the rule in (98).

(98) **Noncoronal Deletion** (final version)

\[
[\text{-son}] \rightarrow \emptyset / [\text{+nasal}] \_\sigma
\]

Not only does this formulation remove the necessity of referring to an "anti-place" feature, but the rule may now be understood as motivated primarily by considerations of proper syllable structure. That is, Noncoronal Deletion can be seen, like the rule of Prenasal g-Deletion discussed earlier in this chapter, as removing illegal consonant clusters from syllables. Similar to the
deletion of the /g/ in a form like /sign/, which must be removed because /g/ is less sonorous than /n/, we can understand the deletion of /b/ and /g/ after /N/ as being motivated by the fact that nasals and voiced stops are relatively close in sonority, closer, anyway, than nasals and voiceless stops (see the sonority hierarchy given near the beginning of this chapter).

If this is so, however, we face a strange dilemma. On the one hand, /Ng/ makes a bad coda and so must be corrected by rule. On the other hand, /Ng/ makes a bad onset, but instead of being corrected by rule, this representation is simply disallowed. This difference in where the syllable structure constraints apply is illustrated schematically below.

(99) Where do syllable-structure constraints apply?

\[
\begin{array}{c}
\text{UR} & */NgV/ & */VNg/ \\
\downarrow & & \downarrow \\
\text{Surface} & *[ŋgV] & [Vŋ] \\
\end{array}
\]

The dilemma can be resolved if we simply assume a syllable-structure constraint like that given in (100a), which specifically rules out syllable-initial /ŋ/ (ie, /ŋ/ linked directly to the syllable node, following the convention of Hayes 1989). Such a constraint is parallel to the constraint in (100b), which rules out syllable-final /ŋ/ (more generally, any /ŋ/ not linked directly to the syllable node).

(100) a. *σ

\[
\begin{array}{c}
| \\
ŋ \\
\end{array}
\]
Thus I claim there is no evidence for the derivation of $[\eta]$ from $/Ng/$ from the nonappearance of $/\eta/$ in onset position.

The second of Borowsky's (1986) three arguments concerns the claimed nonappearance of $[\eta]$ morpheme-internally where no velar consonant follows. If this segment really is underlyingly $/Ng/$, we'd expect it to always appear as $[\eta\g]$ wherever the $/g/$ may be syllabified as the onset of a syllable. It is true that examples of morpheme-internal $[\eta]$ not preceding a velar are rather rare; in (101) I list all the examples found through a computer search. However, these examples are sufficient to show that $/\eta/$ is a free-standing phoneme in some words.

(101) [pronunciations verified with Woolf et al. 1981 where possible]

a. $[\etaC]$, where C is not a velar

angma
angstrom
anxiety
anxious {also with [k]}
Brobdingnag
dingbat
Ingmar
Langdon
Langley
Langtry
Shanghai
tungsten
whangdoodle
Borowsky (1986) attempts to eliminate these cases by claiming that the [ŋ] is derived through assimilation to a consonant that later deletes. She argues that in the forms like those in (101a) where [ŋ] appears before a consonant, an underlying velar /g/ is deleted because the following consonant makes it impossible to syllabify it as an onset.\footnote{Borowsky (1986) includes the word orangoutan in this list, presumably because for her the second syllable is [yuw], and so resyllabification of the /g/ is ruled out since *(gy) is not a legal onset cluster in her dialect. However, [gy] does appear in the optional pronunciation of some words, such as gubernatorial. Moreover, in the standard American pronunciation of orangoutan, the second syllable is [ə], not [yuw]; hence I place this word in the list in (101b) where [ŋ] appears before a vowel.} This analysis is needlessly abstract; there is no surface evidence for the prior existence of this /g/.\footnote{The optional [k] in anxious appears to be epenthetic, as it appears between a nasal and a voiceless fricative [ʃ] (cf epenthetic [t] between [n] and [s] in prince and [p] between [m] and [θ] in warmth).} Moreover, in at least one case the analysis will not work at all. In Langley, a word
Borowsky does not discuss, the hypothesized /g/ is followed by /l/, which can in fact form a legal onset with it (cf *gangly*, with obligatory [g]). Nevertheless, in this word the /g/ never surfaces. Since there is no principled way to delete it, we must assume that it was never there.

Borowsky's (1986) explanation for forms like those in (101b) where [ŋ] appears before a vowel is also needlessly abstract; a non-abstract analysis explains the facts just as well, if not better. She claims that such forms actually contain an underlying /h/ after the /N/, to which the /N/ assimilates; this /h/ then deletes by the same rule that deletes /h/ in all unstressed syllables (cf [h] in *a history book* but not in *an histórian*). As support for her argument, she notes that this /h/ appears on the surface in forms like *Bírmínghàm* where the syllable containing it may optionally be stressed.

This analysis captures the fact that [ŋ] can (generally) only appear before a vowel if this vowel is unstressed, but it can't be right. First, in *Klíngòn*, a word Borowsky does not discuss, the vowel following [ŋ] has some stress; in fact, the word has the same stress pattern as in the word *cóhört*, where the /h/ does not delete.

Second, some speakers (such as myself) only apply the h-Deletion rule in fast speech; for them, *histórian* does not lose its /h/ obligatorily, and yet [h] never surfaces in the appropriately stressed forms in (101b). Third, /h/ is not a velar segment. In fact, in current feature geometry models (e.g., Sagey 1986), /h/ is thought to have no place specification at all. Hence /N/ cannot acquire its Dorsal
place specification from it. Finally, except for cases like Birmingham, no alternate pronunciations of the words in (101b) contain [h]; instead, they contain [k] or [g].

A very simple non-abstract analysis of the forms in (101) is available: /j/ is a phoneme of English, though a marginal one. That is, like /z/, /j/ is not always derivable from other segments by rule, but the number of forms where it is not derivable is very small. A computer search shows that in the vast majority of instances of [3], this segment is derivable from /z/ by Palatalization, as in please-pleasure, video-vigual. The number of words with non-derivable [3] is less than thirty. These are listed below.

(102) /3/ as a marginal phoneme

azure
beige
bijou
bourgeois
ciaesura  {also with [z]}
camouflage  {also with [ʧ]}
cashmere  {also with [ʃ]}
caguayist
collage
corsage  {also with [ʤ]}
entourage
espionage  {also with [ʤ]}
garage  {also with [ʤ]}
genre

44 Ironically, the original version of this analysis in Chomsky and Halle (1968:234), works better precisely because it is more abstract. There an underlying voiceless velar fricative /x/ is used in place of Borowsky's (1986) /h/; naturally rules must then conspire to ensure that this segment never appears on the surface.
/ʒ/ as a marginal phoneme (continued)

Hoosier  
jabot  {also with [ɕ]}  
lingerie  {also with [ɕ]}  
loge  
luge  
massage  {also with [ɕ]}  
mirage  
montage  
regime  {also with [ɕ]}  
rouge  {also with [ɕ]}  

This sort of situation is hardly unique. For example, Katzner (1984) lists only sixteen words in Russian beginning with [ɕ], all of them appearing to be relatively recent borrowings, and McCarthy and Prince (1986:32) observe that in Mokilese, "[w] is arguable [sic] phonemic in only a few words."

Borowsky's (1986) observation that [ŋ] (almost) never appears before a stressed vowel can be easily explained without postulating abstract segments. The constraint given above in (100a) prevents /ŋ/ being syllabified as an onset. Thus /ŋ/ can only appear intervocally if resyllabified as a coda by Resyllabification. This rule is triggered in the stress environment illustrated in (101b), where [ŋ] is followed by an unstressed syllable.45

45The word K'lingdon, where the second syllable is stressed, thus poses problems not only for Borowsky's (1986) analysis, but for mine as well. It is possible, as suggested by Mike Hammond, that this word is in fact considered polymorphemic, with -on being treated as a suffix by analogy with words like xenon. Further examples of this sort are needed to determine what is in fact the appropriate analysis of this word.
I thus conclude that like the argument from onsets, the argument from independent morpheme-internal [ŋ] does not work either. The constraint against the appearance of /ŋ/ in onset position, combined with the status of /ŋ/ as a marginal phoneme, explain both patterns without requiring an underlying representation as /Ng/.

Borowsky's (1986) final argument for the derivation of [ŋ] from /Ng/ is more convincing, however. This concerns the length of vowels that appear before [ŋ]. If this segment is underlyingly /Ng/, ie a consonant cluster, we would expect Closed Syllable Shortening to prevent the appearance of a long vowel before it. This prediction turns out to be correct; [ŋ] never appears after a long vowel.

The situation is actually a bit more complex than described in the last paragraph. As the chart in (103) illustrates, [n] and [m] are also somewhat restricted in what vowels they may follow. Moreover, [ŋ] does not follow all short vowels and may follow long vowels in restricted cases (see footnotes). The only systematic gap, however, is that represented by the cooccurrence of [ŋ] with long vowels. Thus I conclude that there is good reason to believe that word-final [ŋ] is derived from underlying /Ng/ by the ordered application of Nasal Assimilation before Noncoronal Deletion.46

---

46 As pointed out by Mike Hammond, if this analysis is correct, we will require a constraint of the form [ŋ]Word to prevent the marginal phoneme /ŋ/ from appearing word-finally. Note that this is parallel to the restriction against the marginal phoneme /ʒ/ from appearing word-initially; as seen in the list in (102), /ʒ/ only appears word-initially in genre and jabot.
Thus I conclude that though [ŋ] is in fact derived in most cases by Nasal Assimilation, and that the derivation of word-final [ŋ] also requires the application of Noncoronal Deletion, this segment is nevertheless a phoneme of English. This conclusion is not unprecedented. Thus Borowsky (1986) admits that [ŋ] must be

<table>
<thead>
<tr>
<th></th>
<th>-n</th>
<th>-m</th>
<th>-ŋ</th>
</tr>
</thead>
<tbody>
<tr>
<td>iy</td>
<td>scene</td>
<td>seem</td>
<td>--</td>
</tr>
<tr>
<td>ey</td>
<td>sane</td>
<td>same</td>
<td>--.47</td>
</tr>
<tr>
<td>uw</td>
<td>soon</td>
<td>room</td>
<td>--</td>
</tr>
<tr>
<td>ow</td>
<td>phone</td>
<td>home</td>
<td>--</td>
</tr>
<tr>
<td>oy</td>
<td>coin</td>
<td>--</td>
<td>--.48</td>
</tr>
<tr>
<td>ay</td>
<td>fine</td>
<td>dime</td>
<td>--</td>
</tr>
<tr>
<td>aw</td>
<td>brown</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>i</td>
<td>sin</td>
<td>gym</td>
<td>sing</td>
</tr>
<tr>
<td>e</td>
<td>hen</td>
<td>hem</td>
<td>--.49</td>
</tr>
<tr>
<td>u</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>o</td>
<td>lawn</td>
<td>mom.50</td>
<td>song.51</td>
</tr>
<tr>
<td>ae</td>
<td>ban</td>
<td>Sam</td>
<td>bang</td>
</tr>
<tr>
<td>a</td>
<td>bun</td>
<td>rum</td>
<td>rung</td>
</tr>
<tr>
<td>A</td>
<td>Hahn</td>
<td>palm</td>
<td>--</td>
</tr>
</tbody>
</table>

47For speakers with Prevelar Raising, [eyŋ] is in fact possible. This indicates that Noncoronal Deletion applies before Prevelar Raising, which presumably could not lengthen a vowel before a consonant cluster. Note that this ordering is also historically plausible, as Noncoronal Deletion is a much older rule.

48As Borowsky (1986) observes, there are two instances of [ŋ] after [oy], namely boing and oink, both of which are onomatopoeic.

49There are virtually no cases of word-final [eŋ] in English words. After a computer search, the only native words I could find were length(en) and strength(en); the only other relatively common word is the borrowed ginseng.

50As is well known, the pronunciation of [ə] versus [ɔ] varies from dialect to dialect.

51There is some question about whether [ɔ] is a short vowel. McCarthy and Prince (1986) show that the minimal word in English is bimoraic; hence only function words can be monosyllabic and end in a short vowel. This rules out possible words like *[pə, pə, pə, pə, pə]*. Notice, however, that [ɔ] may appear in this situation, as in paw, law, raw, draw, etc ([ə] also appears in the word spa).
allowed to appear in the output of the lexical strata (i.e., at the phonemic level) due to its obligatory, stress-insensitive appearance in monomorphemic pairs such as linguist-linguistics. However, she only allows [ŋ] to appear lexically if its place features are shared with an adjacent segment. The existence of cases with free-standing [ŋ], such as Klingon, suggests that her claim must actually be extended to free-standing /ŋ/ as well.

The sole remaining apparent problem with /ŋ/ being a phoneme is thus the difference between monomorphemic and prefixed forms noted in the previous section. While in monomorphemic forms the derivation of [ŋ] may be obligatory, in prefixed forms (as illustrated above in (89)) a prefix-final nasal obligatorily assimilates in place to the following segment only if this segment is labial or coronal. By contrast, in prefixed forms assimilation of Dorsal is optional in many cases and also appears to be stress-sensitive. Borowsky (1986) argues that this follows from the non-phonemic status of [ŋ], which can only be derived in prefixes by the postlexical stress-sensitive rule of Optional Nasal Assimilation.

There is no simple way to get this analysis to work properly, however. Because of the existence of obligatory [ŋ] in monomorphemic forms like linguistics, Borowsky (1986) cannot claim that the optionality of [ŋ] in prefixed forms like congréssional is due solely to properties of /ŋ/, such as its proposed nonphonemic status. Instead, something about the prefix itself must be relevant as
well. Borowsky's specific proposal is that prefixes appear on a
different plane from roots, and that across morpheme boundaries
Nasal Assimilation involves copying, not spreading, of place features;
the prohibition against [ŋ] appearing in the lexical strata unless its
place features are linked to an adjacent segment thus means that
Dorsal cannot be assimilated in the same way. The figure below
explicates this concept. The form co[ŋ]gressional is not derivable
within the lexicon (ie, through Obligatory Nasal Assimilation) because
the proposed copying nature of Nasal Assimilation in prefixed forms
would create an unlinked [ŋ].

(104)  [cf (39) in Borowsky 1986:102]

This analysis cannot be entirely right, since as we saw above,
free-standing [ŋ] is in fact legal, though rare. In spite of this and the
cumbersome formalism of this analysis, however, I think that Borowsky captures the necessary insight. The device of having Nasal Assimilation in prefixed forms involve copying and not spreading parallels the same use of this device noted in the discussion above in 2.3.2 of the gemination rule claimed to operate in forms like _accept_. There I noted briefly that such cases could be handled equally as well if we supposed that prefixes come in different allomorphs, a device that is clearly necessary with other affixes in English, in particular -ion.

Precisely the same thing can be seen here. Specifically, if we assume that the prefix _con_ - has the allomorphs /kan/ and /kam/ but not /kaŋ/, we can account for the data even better than Borowsky. As in Borowsky's analysis, in this view the way [m] comes to appear before [p] in a prefixed form like _compass_ is different from the way [m] comes to appear before [p] in a monomorphemic word like _campus_; specifically, in _campus_ the [m] is derived through Obligatory Nasal Assimilation while in _compass_ it is underlying. The difference is that the obligatory appearance of [kaŋ] is blocked simply by the fact that there is no such allomorph. This is explained in turn by the fact that /ŋ/ is a marginal phoneme, appearing underlingly in only a very few morphemes. That [ŋ] nevertheless still appears independently in some morphemes is thus consistent with this
analysis in a way it is not with Borowsky's, which rules out all free-standing cases of [ŋ].52

From the perspective of the goals of this work, then, the key point is that the marginal phonemic status of /ŋ/ has no consequences for how Nasal Assimilation applies. Obligatory Nasal Assimilation is in fact free to derive [ŋ], as in linguistics, just as Optional Nasal Assimilation can derive it in congressional. Obligatory and Optional Nasal Assimilation thus do not differ with regard to the creation of this phoneme, contrary to what has been claimed. As we will see in Chapter 5, however, these rules do differ in other unpredictable ways, implying that they are in fact two separate rules, not one that may appear twice in the derivation as Borowsky (1986) assumes.

2.4.3 Nasals: Summary

As usual, I conclude with a figure showing the relative ordering of the rules discussed in this section.

(105) a. Obligatory Nasal Assimilation [before c, d]
b. Coronal Extraprosodicity [before c]
c. Noncoronal Deletion [before d, e]
d. Prenasal Raising [before e]
e. Optional Nasal Assimilation

52 It may be that the allomorph /kaŋ/ does show up in some words, such as congress, congruous and conquer, where [ŋ] is obligatory, not optional as it is in the similarly stressed concord and concrete. The fact that [ŋ] is again optional in the derived forms congressional and congruity, implying the prefix here is /kan/, not /kaŋ/, is fundamentally no stranger than de- being the prefix in dissolve while dis- is the prefix in the derived form dissolution (judging from the behavior of s-Voicing; see 2.3.1.3 above).
2.5 Chapter summary

In the following chart I list the rules I have argued are required for an adequate description of English phonology. Rules that are crucially ordered are indicated as such. As promised, the ordering I give here does not differ from what is argued for in works such as Chomsky and Halle (1968), Halle and Vergnaud (1985) and Borowsky (1986).

(106)   a. Lengthening rules [before b, d]  
b. s-Voicing [before c, d, f]  
c. Shortening rules [before d]  
d. Velar Softening [before e]  
e. Vowel Shift  
f. Spirantization [before g, h]  
g. y-Insertion [before h]  
h. Palatalization [before i]  
i. y-Deletion  
j. Obligatory Nasal Assimilation [before l, m]  
k. Coronal Extraprosodicity [before l]  
l. Noncoronal Deletion [before m, n]  
m. Prevelar Raising [before n]  
n. Optional Nasal Assimilation

In the next chapter, I show that the orderings illustrated in this figure are consistent with the Productivity Hypothesis: less productive rules are ordered before more productive rules.
CHAPTER 3
EVIDENCE FOR PRODUCTIVITY

3.0 Introduction

The Productivity Hypothesis claims that rules are ordered if they differ in productivity, with the less productive rules ordered before the more productive ones. The concept of rule productivity is thus fundamental to this principle and, indeed, to the model of Double Lookup as a whole.

In this chapter I explicate the concept of rule productivity. I begin in 3.1 by comparing and contrasting this concept in Double Lookup with other conceptions found elsewhere in the literature. Then in 3.2 and 3.3 I discuss what sorts of evidence can be used to measure a rule's productivity. Since one of the primary goals in this section is to rank the rules of English discussed in the last chapter according to their productivity, I will focus on evidence relating to these rules in particular.

In 3.2 I examine linguistic evidence for rule productivity, namely the reliability of generalizations across real words (3.2.1) and studies of novel linguistic forms (3.2.2).

In 3.3 I discuss psycholinguistic evidence for rule productivity. In 3.3.1 I review some experimental evidence for the relative productivity of rules. Next in 3.3.2 I examine a set of speech error accommodations, speech errors where novel forms are erroneously generated and then appear to be subject to rules of the grammar; the proportion of such forms which are accommodated in the appropriate way should give some indication of the productivity of the various
rules. In 3.3.3 I look at a set of malapropisms, or erroneous sound-based word substitutions, a different sort of speech error which seems to shed light on the representation of forms in memory. This sort of evidence for the real-time use of rules will be seen to be indirect at best; we will see that several rules which other tests show are very productive turn out to be associated with stored representations. That is, it is possible for a rule to be both very productive and prepatterned. This finding will play an important role in the discussion of apparent rule cyclicity in Chapter 5.

Finally, I end the chapter by unifying the information gathered from the previous sections in a general statement about the relative productivity of rules in English. I conclude that the facts of English discussed here are fully in conformity with the Productivity Hypothesis.

### 3.1 The concept of productivity

In Chapter 1 I defined the *productivity of a rule* as the likelihood that the rule would apply in real time during speech production. As noted there, this differs from other common views on productivity in at least two ways.

First, it presupposes that rules can vary gradiently in productivity. It is sometimes thought that rules either are fully productive or fully nonproductive. Throughout this chapter, however, we will see evidence that rules really do vary gradiently in productivity. It is in fact rare to find rules that are fully productive according to every test or that are fully unproductive. The source of
this gradient productivity will not be investigated here, but presumably it is related in some way to the reliability of the rule as a generalization across utterances. Consequences of this assumption will be seen in Chapter 5.

It may be objected that all rules really are either fully productive or fully unproductive for any given speaker; it is only when individual competences are averaged across speakers that it appears as if a given rule only has partial productivity. This view makes a prediction that is very easy to prove false: any given person is always consistent in how she pronounces words. The vast body of literature on naturally occurring speech, pioneered primarily by William Labov in the sixties and seventies, proves this false beyond any reasonable doubt. For many rules, it is often not possible to predict whether a given word spoken by a given speaker at a given time will conform to the rule or not. For example, Guy (1980:11) writes of the rule of Coronal Deletion (to be discussed in Chapters 4 and 5) that "variation is inherent, and cannot be scrubbed out of our linguistic description by ever-finer subdivisions of the data." This is even true of rules we might want to call "lexical". Thus, as S. Myers (1987) notes, the rule of Closed Syllable Shortening is optional in words like *economic*, the first vowel of which can be either [iy] or [e]. It is true that there are speakers who consistently use only the one or the other form, but there are also speakers who vary somewhat more freely between them. Such a case may be thought of as an instance of Gradient Productivity or of Gradient Retrievability;
in either case, we must assume that individual speakers allow gradient behavior in their use of phonology.

The second way my definition of productivity differs from many is that since productivity is understood as the likelihood a rule will apply on-line, we are forced to test it by examining novel forms. As we will see in this chapter, tests involving real words do not necessarily bring the same results. For example, we may find that a rule associated with many lexical exceptions is actually very productive when tested with novel forms. It seems plausible to assume that if a rule is very productive, it tends not to be associated with prepatterned forms; that is, with a productive rule, most forms derive their surface representation by the on-line application of this rule. We will find, however, that under certain circumstances a productive rule may be associated with prepatterned forms. This fact will play a crucial role in my discussion of apparent cyclicity in Chapter 5.

The relation between a rule's productivity and whether or not it has prepatterned forms is complex, but some generalizations can be made. First, if we find evidence that a rule is fully unproductive, then any surface generalizations we might ascribe to this rule must result from prepatterning. Note that as observed above, this does not work the other way: a rule that is not unproductive may nevertheless be associated with prepatterned forms as well. Second, if we find evidence that a rule is not prepatterned, then any surface generalizations we might ascribe to this rule must logically result
from the on-line application of the rule. Again, it does not work the other way: a prepatterned rule may nevertheless be highly productive. Because of these complexities, it is always safer to test a rule's productivity directly, that is, by testing if it applies to novel forms.

Defining productivity in terms of real forms rather than novel forms creates a conceptual problem as well. Consider the /k/ copying rule mentioned in the last chapter that is presumed in accounts of alternations like recess-access. As Borowsky (1986) notes, this rule only applies in at most six forms, along with their derivatives: access, accident, accept, accelerate, accede, accent. Within this small set, however, the rule has no exceptions; there are no cases where a- is prefixed to a root beginning /ke/ or /ki/ where the /k/ copying rule fails to apply. Nevertheless, the intuition is clear that this rule is essentially unproductive; Borowsky (1986) herself exploits this intuition in suggesting that the low productivity of this rule is related to its applying in a very early lexical stratum.

The /k/ copying rule with a(d) - affixation is but one example of a phonological rule triggered by specific morphemes. The question of the productivity of such rules is thus related to the productivity of the rules affixing these morphemes. If we could find a way of defining morphological productivity in terms of real words, we would then be able to determine the productivity of phonological rules with real words as well. As Aronoff (1976) shows, however, even in the
morphological domain it will not do to define productivity in terms of real words.

First, given a pair of affixes, it may be that one affix appears more commonly in a given morphological domain than the other, while in another domain it is the other affix that is associated with more words. For example, Aronoff (1976) points out that while there are far more words ending in -ive that take -ness than -ity, more words ending in -ile take -ity than -ness.

Second, the more productive an affix is, the less likely it is that any given word formed with the affix will already be in one's lexicon. Thus many published word lists (e.g., dictionaries) severely restrict the number of words formed with productive affixes, meaning that the relative productivity of affixes does not necessarily correlate with lists of real words.

Aronoff (1976) does discuss methods that may be interpreted as providing a means of measuring relative productivity of morphological rules by examining real words (see next section). However, such methods must be considered indirect evidence at best.

3.2 Linguistic evidence for productivity

Given the above discussion, there are two ways to test the productivity of a rule. On the one hand, it may be tested directly by examining its behavior in novel forms. On the other hand, it may be tested indirectly by examining its behavior in real words. As we will see, both methods may be used within either a linguistic or a psycholinguistic framework. In this section I discuss linguistic
evidence for productivity. In 3.2.1 I see what evidence about productivity can be gained indirectly from an examination of real words; in 3.2.2 I briefly discuss how novel forms may be used within the linguistic framework to investigate productivity.

3.2.1 Generalizations across real words

By their very nature as word-internal rules, many of the rules I discussed in the last chapter depend for their application on morphological information. Thus the relative productivity of these phonological rules may be estimated by measuring the productivity of the associated morphological rules, as noted above. Below I list the rules I discussed in Chapter 2 along with the affixes that seem to be required for their application. Several of the rules apply both across and within morphemes; I've indicated these with the word "general".

(1) Rule | Affixes
---|---
Vowel length rules | general
s-Voicing | de-, re-, pre
Velar Softening | -ic, -es
Vowel Shift | general
Spirantization | -y, -ion, -ian, -ial
y-Insertion | general
Palatalization | -ion, -ian, -ial, -ious, -ual, -uous
Nasal Assimilation | general
Noncoronal Deletion | general

1 As we saw in Chapter 2, the evidence that s-Voicing applies with other affixes is extremely limited; I will therefore focus solely on the prefixes here.

2 As the reader will recall from Chapter 2, these two affixes play different roles in Velar Softening: the suffix -ic is the primary target of Velar Softening (when followed by suffixes like -ion, -ian, -ity, -ious); -es is a reliable trigger of Velar Softening in a restricted set of cases.
How do we determine the relative productivity of these affixation processes? The simplest way is of course simply to count the number of word pairs that show evidence for the various rules. Using the computer-generated lists discussed in Chapter 2, I have done this for the rules of s-Voicing, Velar Softening and Palatalization, all of which are triggered by specific sets of morphemes. The results of this word count is given below. Note that there is a clear ranking of these rules in terms of number of forms affected, whether we count the number of words or just the number of different morphemes (ie, counting all examples of Velar Softening affecting \(-ic\) as one, and all examples of Palatalization affecting \(-ate\) as one). This ranking is exactly that predicted by the Productivity Hypothesis, since as we saw in Chapter 2, s-Voicing is ordered before Velar Softening, which in turn is ordered before Palatalization. Significantly, this ordering cannot be ascribed merely to differences in productivity across the lexical strata of Lexical Phonology, since at least two of these rules must be in the same stratum. Thus Halle and Mohanan (1985) place both Velar Softening and Palatalization in Level 2; Borowsky (1986) holds that all three rules may apply in Level 1.

(2) number of number of
\hspace{1cm} \text{word pairs} \hspace{1cm} \text{morphemes}
\begin{align*}
a. \text{ s-Voicing:} & \quad 21 & 21 \\
b. \text{ Velar Softening:} & \quad 320 & 62 \\
c. \text{ Palatalization:} & \quad \text{over 877} & \text{over 384}
\end{align*}
In addition to this crude measure of productivity, there are more subtle methods that can be used with real words. In particular, Aronoff (1976) mentions several properties that seem to correlate reliably with the productivity of an affix.

First, although it is true that what we'd like to call less productive affixes may occur with many words in specific morphological environments, this is not true across morphological environments in general. Thus Aronoff (1976:36fn3) notes that while there are domains where -ity applies readily but -ness doesn't, -ness applies to a much broader range of domains than -ity.

The second observation Aronoff (1976) makes about the productivity of affixes is this: the more productive an affix is, the more transparent is its contribution to a word's meaning. Thus adding the suffix -ness to an adjective X produces a word that means something like "the quality of being X"; callousness roughly means "the quality of being callous". In other words, the affixation of -ness performs a relatively well-defined function on the semantics of the stem. Words derived by suffixing -ity do not necessarily have any meaning like this, and may even have other meanings as well. Hence variety can mean "the quality of being various", but it can also mean "type", as in "How many varieties of fish are there in the pond?" (Aronoff 1976:38). Aronoff (1976) notes that similar semantic complexity holds between the members of the pairs notorious-notoriety, curious-curiosity, porous-porosity, monstrous-monstrosity, continuous-continuity, and discontinuous-discontinuity.
Third, the less productive an affix is, the more likely it is to make idiosyncratic changes in the forms of the stems it affixes to. The suffix -ness simply affixes on to adjectives ending in -ous with no change. With -ity, however, whether or not the -ous remains or is dropped is to some degree unpredictable. Hence it remains (shortened by Closed Syllable Shortening) in fabulous-fabulosity, nebulous-nebulosity, cellulous-cellulosity, and scrupulous-scrupulosity but is lost in credulous-credulity, incredulous-incredulity, sedulous-sedulity and garrulous-garrulity (Aronoff 1976:42).

Aronoff (1976) argues that all three of these properties, namely that less productive affixes are more restricted in their domain of application, perform complex and unpredictable changes in semantics, and may be associated with arbitrary form changes, can be understood if we suppose that forms derived by less productive affixes are stored in the lexicon. Such affixes would then be less able to spread to new morphological domains, and since the words containing them would be stored as wholes, the words could be expected, over time, to have less of a semantic and phonological relation with the parts they are made out of than words that are derived from pieces stored separately in the lexicon and then joined together. Note that Aronoff (1976) comes to this conclusion for purely linguistic reasons.

Once again, though, the argument only runs one way. Since -ity is less productive than -ness, words formed from it may be
anomalous in the ways discussed above, due to their being listed, but it does not follow that since -ness is more productive than -ity, it cannot have some anomalous semantic and phonological properties as well. For example, likeness does not mean "the quality of being like"; rather, it means something that is like (something else), specifically an imitation. Similarly, cleanliness is derived from clean, since there is no adjective *cleanly *[klenliy], implying that the stem has undergone the idiosyncratic insertion of -ly (triggering Shortening); compare mean-meanness. What this means, continuing with Aronoff's (1976) argument, is that even words derived with the productive suffix -ness may be stored in memory.3

What does all this imply for the rules listed above in (1)? Four of these rules, namely s-Voicing, Velar Softening, Spirantization and Palatalization are clearly restricted morphologically. Let's see how they compare on Aronoff's dimensions.

As we saw in the last chapter, s-Voicing only seems to apply, if at all, with the prefixes de-, re-, pre-. Unlike other prefixes (eg, in- and un-), these prefixes attach to roots, not stems. The number of roots, which by definition are not derived, is much smaller than the number of stems, which may be derived or not derived. Hence affixation of these prefixes appears to be rather restricted.

3Given the discussion of the origin of prepatterning in Chapter 1, we may expect such words to have other special properties marking them as stored. In particular, we would expect them in general to be more frequent than words that do not have such idiosyncrasies.
Many if not most of the roots these prefixes attach to do not have transparent meanings, so it is difficult to say how the prefixes contribute to the meaning of the words as a whole. Thus there appears to be no consistent meaning that carries over for all the words in any column or row in the chart below.

<table>
<thead>
<tr>
<th></th>
<th>re</th>
<th>de</th>
<th>pre</th>
<th>con</th>
</tr>
</thead>
<tbody>
<tr>
<td>sist</td>
<td>resist</td>
<td>desist</td>
<td>consist</td>
<td></td>
</tr>
<tr>
<td>sign</td>
<td>resign</td>
<td>design</td>
<td>consign</td>
<td></td>
</tr>
<tr>
<td>sult</td>
<td>result</td>
<td>consult</td>
<td></td>
<td></td>
</tr>
<tr>
<td>sume</td>
<td>resume</td>
<td>presume</td>
<td>consume</td>
<td></td>
</tr>
</tbody>
</table>

There is some evidence that the prefixes trigger idiosyncratic form changes as well. One possible example is *dissolve*, which inexplicably undergoes voicing (while the derived form *dissolution* does not). Consider also the fact that *semble*, which must take the suffix *-ance* if not prefixed (*semble-semblance*), need only take it optionally if prefixed by *re-* (*resemble-resemblance*) but may not take it at all if prefixed by *dis-* (*dissemble-*(dissemblance)*).

Together this evidence suggests that the prefixes *de-*-, *re-*-, and so forth are not very productive according to Aronoff's (1976) criteria. Consequently, the phonological rule of *s-Voicing*, which is triggered specifically by these affixes, cannot be very productive either.

*Velar Softening* doesn't fare much better. We saw in the last chapter that virtually all cases of the application of this rule involve changes in the suffix *-ic*, namely from *[k]* to *[s]*; most of the remaining cases are triggered by the special plural suffix *-es*. This
suffix appears to be rather unproductive according to Aronoff's criteria. As we saw, it is restricted to apply only with the twenty or so nouns ending in [ks]. Indeed, in many of these cases, the regular plural suffix -s is actually preferred (eg, crux-cruxes, not cruces). Although it does not seem to be associated with any semantic anomalies, it does, as we saw, cause strange form changes; specifically, the deletion (or insertion) of the stem-final [s] requires at the very least a phonological rule that is marked to apply only with these specific stems.

As we saw, the role of -ic in Velar Softening acted to restrict application of this rule as well. Recall that most examples of Velar Softening triggered by -ity either involve words ending in the morpheme -ic or in the phonological sequence [lk]. In addition, most of the cases of Velar Softening triggered by the plural suffix -es also involved stems ending in [lk] (or [lks], if the stem-final [s] is deleted and not epenthetic). This means that Velar Softening is restricted in a very special way: it applies most readily to the morpheme -ic and to forms that sound like -ic.

Nevertheless, Velar Softening shows signs of being more productive than s-Voicing. Forms derived with -ic and -es show fewer semantic anomalies than those derived from prefixes, for the simple reason that the suffixes attach to stems and not roots. The suffix -ic (and -ity and -ian, the Velar Softening-triggering suffixes that attach to it) is also much less restricted in its application than the prefixes, and for the same reason: it attaches to stems and not
roots. Both sets of affixes, however, are associated with anomalous phonological changes, however, so we wouldn't expect Velar Softening to be too much more productive than s-Voicing.

I now turn to Spirantization. As we saw in the last chapter, this rule virtually never applies independently of Palatalization. In fact, the only time Spirantization applies by itself is with the suffix -y. What is the productivity of this suffix?

The morphological domain of application of this suffix is relatively broad, deriving nouns both from nouns (mendicant-mendicancy) and from adjectives (accurate-accuracy), meaning it is more productive than the prefixes associated with s-Voicing, which attach only to roots. However, the semantic changes it confers are somewhat anomalous in many cases. Thus democracy is related only to what is often called "small-d" democrat, not the much more common meaning found in Democrat (the political party). Moreover, even ignoring phonologically strange cases like difficult-difficulty, normal-normalcy, there remain cases like lunacy, which are related not to the nonword *lunate but to lunatic (cf mendicant-mendicancy). This implies that -y suffixation is roughly equal in productivity to the -ic and -ian suffixation associated with Velar Softening. Recall, though, that Velar Softening is also triggered by the more unproductive -es suffix, and that its application is restricted primarily to stems containing [ík]. Thus it appears that due to its dependence on -y suffixation, Spirantization is more productive than s-Voicing and Velar Softening, but is still not fully productive.
Palatalization appears to be more productive by Aronoff's criteria. First, like -y the suffixes associated with Palatalization are semantically transparent compared with the prefixes. Second, unlike s-Voicing it is associated with a wide variety of affixes, including -ion, -ian, -ial, -ious, -ual, -uous, each of which has a relatively wide domain of application as well. Unlike Spirantization, which has exceptions (cf permit-permissive, where it applies, and transit-transitive, where it doesn't; also right-righteous, which show evidence of Palatalization but not Spirantization), Palatalization is absolutely exceptionless; if a suffix triggers Palatalization in some stems, then it triggers Palatalization in all stems. This implies that Palatalization is more productive than Spirantization.

Nevertheless, its level of productivity does not appear to be much greater. As with -y, the suffixes triggering Palatalization are associated with idiosyncratic changes. Unlike the other cases we have discussed so far, however, these idiosyncrasies seem to be confined to allomorphs of the suffixes themselves; the idiosyncratic manipulations are not performed on the stems. Thus, as Aronoff (1976) shows, -ion surfaces in several different unpredictable allomorphs, some examples of which are given below. Very few of these forms show unpredictable changes in the stem, such as destroy-destruction.

\footnote{As with -y, there are problematic cases, such as transmission (as in a car), whose meaning is not transparently related to transmit.}
(4) [examples from Aronoff 1976:100-107]

a. -/æ/tion:
   
   cease - cessation  
   deport - deportation  
   evoke - evocation  
   perturb - perturbation  
   realize - realization  

b. -tion:
   
   define - definition  
   imbibe - imbibition  
   repeat - repetition  
   suppose - supposition  

c. -tion:
   
   absorb - absorption  
   deduce - deduction  
   destroy - destruction  
   receive - reception  
   redeem - redemption  
   resume - resumption  

d. -ion:
   
   commune - communion  
   decide - decision  
   educate - education  
   explode - explosion  
   insert - insertion  

As Rubach (1984) shows, the same is true of -ial and -ious, which may surface with an initial [i], as in dictator-dictatorial, labor-laborious, with initial [y], as in torrent-torrential, rebel-rebellious, or with initial [ə], as in doctor-doctoral, peril-perilous (Rubach
1984:31-32). Interestingly, however, Palatalization is not associated with unpredictable changes in the stem as we saw happen with all the other rules discussed above: s-Voicing (*solve-dissolve*), Velar Softening (*apex-apices*), Spirantization (*lunatic-lunacy*).

Consider the significance of this. If a suffix is associated with arbitrary stem changes, then according to Aronoff's logic, derived forms with these stem changes must be listed in the lexicon. Hence many forms containing this suffix are not derived on-line, and so by my definition are prepatterned. By contrast, if a suffix is associated with arbitrary suffix changes, we do not need to store all the forms derived with this suffix; at most, we need to store all the different allomorphs of the suffix and mark the stems as taking one suffix allomorph or the other. Thus each of the allomorphs of a suffix can potentially remain fully productive. This is not true of suffixes that cause arbitrary changes in the stem.

All the remaining rules seem to apply more generally, so this method of calculating productivity will not work for them. Thus I conclude this section with the results obtained so far for s-Voicing, Velar Softening, Spirantization and Palatalization. Note that ranking the rules according to productivity as in (5) results in a natural progression in all three dimensions. The rules of s-Voicing and Velar Softening are morphologically more restricted than the rules of Spirantization and Palatalization. All morphological operations triggering these rules are associated with semantic anomalies, but s-Voicing is associated by far with the most. Finally, all morphological
operations triggering these rules are associated with unpredictable changes in form, but with Palatalization such changes are restricted to the suffix.

(5)  

<table>
<thead>
<tr>
<th>Domain</th>
<th>Semantics</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>s-Voicing very restricted</td>
<td>no pattern stem changes</td>
<td></td>
</tr>
<tr>
<td>Vel. Soft. restricted</td>
<td>some anomalies stem changes</td>
<td></td>
</tr>
<tr>
<td>Spirant. less restricted</td>
<td>some anomalies stem changes</td>
<td></td>
</tr>
<tr>
<td>Palatal. less restricted</td>
<td>some anomalies suffix changes</td>
<td></td>
</tr>
</tbody>
</table>

3.2.2 Novel forms in linguistics

As discussed in Chapter 1, phonology is different from syntax in a very significant way: while syntactic rules and constraints standardly act on novel utterances, many phonological rules act solely within words. Thus in seeking novel forms to use as direct evidence for the productivity of phonological rules, the researcher is often forced to use material that is atypical of what speakers normally use phonology for. Thus rule productivity is tested by seeing to what extent rules apply in speech errors, in experimental nonce forms and in language games; such evidence will be examined in the following section. There are, however, natural and nonanomalous novel forms as well. Because such forms are considered typical, the study of people's production and use of them is considered linguistics and not psycholinguistics.

One primary example of a linguistic novel form is the foreign loanword. 5 Holden (1976) represents an explicit attempt to use this

---

5 The most obvious cases of linguistic novel forms are invented words and neologisms. As an example, Jeri Jaeger (pc) points to the following example
kind of novel form to measure rule productivity. An examination of nativization rates of foreign words borrowed into Russian shows that native rules spread through borrowings at different rates. Holden (1976:143) hypothesizes that "the different rates of assimilation of certain foreign sequences and segments were a direct measure of the strength (or productivity) of the native constraints which served as targets for the assimilation process."

This source of evidence has yet to be exploited in a study of the relative productivity of the English rules discussed in the last chapter. As far as I know, no such systematic study of borrowed forms has been done with these rules in mind.

There is a nontrivial problem to keep in mind when using foreign borrowings as novel forms in measuring rule productivity. As Janda, Jacobs and Joseph (1992) show, borrowed words often undergo special phonological and/or morphological changes that are not derivable from the native rules of either the borrowing or loan language. Just to take one example, one common English pronunciation of the word *lingerie* as [l̩æʒərɛ] not only violates rules of English (a nasal vowel before a nonalveolar voiced continuant, stress on the last syllable), but also differs from the original French pronunciation as [lɛʒərɪ]. Apparently the

---

noted by John Ohala in the speech of a radio announcer: *witticisms and musicisms*. Interestingly, in spite of the apparently intentional parallelism of the novel form with *witticism*, which ends [sɪzɪm], *musicism* ends [kɪzɪm]. Thus in this novel form Velar Softening does not apply, indicating the relatively low productivity of this rule. My impression, however, is that such neologisms are rather rare in natural speech, certainly far rarer than the speech errors I will be considering below.
pronunciation of the word has been changed to make it sound more "foreign". In other words, if speakers are aware that certain novel forms are foreign borrowings, they may treat them as a special lexical class, subject to special rules, instead of treating them as generic novel forms upon which the regular rules of the language may act.\(^6\)

Another sort of linguistic novel form is the phrase. Phrases count as novel forms in phonology just as they do in syntax; with the exception of stock phrases like *kick the bucket* and frequent word sequences like *in the*, it is safe to hypothesize that most phrases are not stored in memory, and so must be generated on-line. Consequently any phonological rule that applies across words must apply on-line and so by my definition must be very productive. This idea will be exploited extensively in Chapter 4.\(^7\)

Unfortunately, as is well known, most of the rules I discussed in the last chapter do not apply across word boundaries. Two apparent exceptions to this generalization are Palatalization and Nasal Assimilation. As I will show in Chapter 5, however, there is good reason to believe that the versions of these rules that apply exclusively within words are independent of the versions that apply

---

\(^6\)For other examples of this, see Ishihara (1991), Mejia (in preparation), Tsay (1990), and references cited therein.

\(^7\)A caveat must be added here. As Hayes (1990) shows, certain phonological processes that have all the hallmarks of being prepatterned may nevertheless be triggered by information beyond the word boundary, in particular syntactic information. In Chapter 6 I argue that this involves the on-line operation of Lexical Lookup, not on-line rule application. Here and in Chapter 4 I am concerned with rules that apply freely across word boundaries, that is across-the-board.
across word boundaries. We have already seen some of that evidence for Nasal Assimilation in the last chapter, where it was seen that the rules of Obligatory Nasal Assimilation and Optional Nasal Assimilation could not be reduced to a single rule. At this point I will merely note that it is surely no coincidence that of all the rules discussed in the last chapter, it is precisely these two that have parallel rules that apply between words. Unlike Vowel Shift, s-Voicing, Velar Softening, Noncoronal Deletion and the rest, Palatalization and Nasal Assimilation are extremely natural rules (ie, phonetically motivated), and in fact similar rules show up in many of the world's languages. If word-internal rules are relics of formerly much more productive rules, then it is not surprising to find word-internal Palatalization and Nasal Assimilation among them. This just means that an older variety of English had word-external versions of these rules, too, just like Modern English, Russian, Mandarin, and a wide variety of other languages.

In this section I have shown that it is possible to apply traditional linguistic methods in the examination of novel forms to support claims about the relative productivity of rules. It appears, however, that this source of information is not helpful in determining the relative productivity of the rules I discussed in Chapter 2, since these rules apply solely within words. In order to test their productivity, therefore, we must examine the way people deal with novel word-sized units, a methodology beyond the scope of the traditionally defined body of linguistic evidence.
3.2.3 Linguistic evidence for productivity: Summary

The evidence from linguistics, then, thus far seems to support the Productivity Hypothesis. Estimating the relative productivity of s-Voicing, Velar Softening, Spirantization and Palatalization from their associated morphology suggests that these rules are ranked in productivity in precisely the same way as they are ranked in application. In this section I also discussed how novel forms can be used in linguistics to estimate a rule's relative productivity. Although this technique does not seem to be useful for the rules I am currently examining, in Chapter 4 we will see that novel linguistic forms, specifically phrases, are extremely useful in understanding how other rules in English interact.

3.3 Psycholinguistic evidence for productivity

Because the phonological rules discussed in the last chapter apply, in natural speech, primarily in real words, we must turn to psycholinguistic methods to investigate their relative productivity. I will focus on two sorts of psycholinguistic methods: experiments and the study of naturalistic speech errors. We will see that many suggestive findings will come out of an examination of these sources of evidence, although the support for the Productivity Hypothesis is far from conclusive at this stage in the investigation. A detailed survey of the various methodologies available for determining productivity is of course useful in itself; some of the methodologies I will describe, in particular the systematic use of speech errors, have never previously been attempted on this scale.
3.3.1 Experiments

Numerous experiments have been conducted using novel forms to test the productivity of phonological rules. Unfortunately for the purpose of testing the Productivity Hypothesis, however, the goal of virtually all such experiments is not to determine the relative productivity of a set of phonological rules, but rather to prove that a given rule is or is not productive at all. Moreover, with to my knowledge only one exception, no one has ever attempted to test the productivity of different rules using the same methodology. Only if this is done can some conclusions be drawn about the relative productivity of different rules. Consequently the published literature on experimental testing of phonological rules is less than fully satisfying.

Nevertheless, it is worthwhile to summarize some representative results found in this literature, since they do have bearing on the model of Double Lookup as a whole, if not for the relative productivity of the specific rules we are interested in. Interestingly, we will see that the different methods used to test the productivity of rules vary in their success in a way that is understandable within the Double Lookup model.

As we will see, the methodologies that have been used to test the productivity of rules are enormously varied; it is rare to find even two experiments that use precisely the same methodology. Nevertheless, it seems that all the methodologies can be fit into just
five distinct categories: production, judgments, memory, perception and concept formation tasks.

3.3.1.1 The production task

The simplest task is the production task. In this task, the subject is simply told, either directly or through example, what sort of nonword the experimenter wants to hear, and the subject is then supposed to produce it. A classic example of the successful use of a task like this is Berko (1958), where English-speaking children were told to produce the plural form of novel nouns; subjects had no trouble applying the regular English plural rule, implying that it is a very productive rule. Using a very similar test paradigm, Derwing (1976) reports success at testing the productivity of some English derivational rules, such as -ly suffixation. Another successful use of a method like this is found in Anshen and Aronoff (1988), where subjects were asked to produce all words they could think of ending in -iveness, -ivity, -ibility, and -ibleness. It was found that subjects could think of many more words ending in -ness than in -ity; moreover, many more of the -ness words were "nonwords" in that they didn't appear on any published word list. This implies that -ness words are derived on-line as needed, which in turn means that (according to my definition) the rule adding this suffix is more productive than that adding -ity. Other examples of relatively successful usages of this task include the Hsieh (1975) and Wang (1992) studies of Taiwanese tone sandhi.
Often, however, the production task is not successful, at least not in the sense of finding positive evidence for rule productivity. In most versions of the task, the subject is asked to do something rather unnatural: coin a novel word. Given this, as Kiparsky (1975) points out, it is not surprising that such experiments do not show any evidence for the productivity of phonological rules triggered by specific morphemes. For example, if subjects are simply asked to affix -ity on a nonsense word ending in [k], in most cases Velar Softening will not be applied at all; experiments with such null results include Ohala (1974) and Steinberg and Krohn (1975). All this shows is that Velar Softening is not fully productive; it does not rule out the possibility that the rule is still partially productive. Thus it may well be that the production task can only be used to determine if a rule is over a certain threshold of productivity, but not to test how productive the rule is if it is less than fully productive. Consequently such experiments are not as useful a means of determining the relative productivity of the rules discussed in Chapter 2 as are other methods to be discussed below.

Nevertheless, a few interesting results can be reported. Myerson (1976, 1978) (to be examined more fully in 3.3.1.3) found, using a production task, that subjects correctly applied Palatalization in the nonwords prezate+ion and delort+ion over half of the time (55%). By contrast, Ohala (1974), using a similar task, found that subjects correctly applied Velar Softening in the nonwords toxic+ism, public+ism and domestic+ism only about a quarter of the time (26%).
As we saw in Chapter 2, there is evidence that Velar Softening must be ordered before Palatalization, since the former rule feeds the latter in forms such as *Grecian*. On the face of it, then, the results of Myerson (1976, 1978) and Ohala (1974), taken together, constitute support for the Productivity Hypothesis: ordering correlates with differences in productivity. Of course the experiments conducted by the two researchers were not precisely the same (eg, Ohala 1974 used real stems and somewhat older subjects) and further study is needed.

The relevance of any differences between the two studies becomes significant when one compares Ohala's (1974) results for Velar Softening with Myerson's (1976, 1978) results for Vowel Shift. As we saw in Chapter 2, there is evidence that Velar Softening precedes Vowel Shift, and so we would expect that Vowel Shift is more productive than Velar Softening. Yet Myerson (1976, 1978) found no evidence from a production task that Vowel Shift was productive at all (see figure (6) below). The Productivity Hypothesis would predict that if these two rules truly are ordered as claimed, a more carefully controlled experiment would indeed find that Velar Softening is less productive than Vowel Shift.

### 3.3.1.2 The judgment task

The judgment task is a little more sensitive. In the various versions of this task, the subject is given nonwords or pairs of nonwords. The task is to determine how "good" the nonwords sound
as possible words, or, if nonword pairs are given, to determine how "related" the (non)words are.

One question that immediately arises concerns whether such a task really does measure productivity as I have defined it. Making judgments about forms, nonwords or not, does not necessarily require the production of these forms on-line. The task may actually be tapping something quite different, for example, the degree with which the given nonword conforms to generalizations prepatterned across the lexicon. These generalizations aren't necessarily stored separately from the lexical items themselves, but instead may be induced from the lexicon itself during the task, and then immediately forgotten. As I have emphasized, and as will become ever more clear as we continue through later chapters, prepatterning is not a direct measure of productivity. This caveat will in fact be relevant for all of the remaining measures of productivity other than speech error accommodations. It does not reduce results from the judgement task to meaninglessness, but it does require us to be cautious in their interpretation.8

A classic example of the use of this task is found in Zimmer (1969), where native speakers of Turkish were given lists of nonwords that either conformed or failed to conform to vowel harmony generalizations. It was found that some vowel harmony rules were relevant in these judgments, while others were not. Another example is the Ohala and Ohala (1986) study of some

8This was impressed upon me by Jeri Jaeger.
phonotactic constraints in English, where subjects were asked to judge how close various nonwords were to being a possible word of English. Derwing (1976) also used this method to confirm the results of his production study of English morphology.

Other than Myerson's (1976, 1978) study, to be discussed in the next section, the only example I know of where this technique was used on a rule I discussed in Chapter 2 is McCawley's (1986) study of Vowel Shift. In this experiment, subjects were asked to judge the semantic relatedness of pairs of words that varied according to what degree they showed various Vowel Shift alternations. Because of the nature of the task, only real words were used; however, very few cases of what a linguist might call "truly" morphologically related pairs were used. Thus, for example, representative examples showing the proper Vowel Shift alternations include pairs like delight-delicious, creep-decrepit, shade-shadow, mouse-mustard, bloom-blossom, and bone-bonnet. The assumption was that if Vowel Shift alternations are psychologically real, word pairs that show such alternations will be treated the same as word pairs with no alternations at all, such as deceive-deceit, mate-matron, isle-island, yoga-yogurt, and so forth. Word pairs with vowel alternations that are not rule-governed, such as stride-straddle, astound-astonish, clear-clarify, deign-dignity, peace-pacify, should be considered less "related" than either the Vowel Shift or same-vowel pairs.
The results of the experiment, though not unequivocal, seem to give some evidence that Vowel Shift is greater than fully unproductive. Approximately three-fourths of the subjects seemed to treat Vowel Shift and same-vowel pairs as equally related; the other fourth did not. However, McCawley (1986:37) himself calls the results "tentative" and does not even say whether or not they are statistically significant.

3.3.1.3 The memory task

More success has been achieved with memory tasks. The various versions of this task agree in one respect: the subject is given some time before producing a novel form. Presumably, in the intervening time the novel form has had to have been stored in memory in some sort of representation.9

Myerson (1976, 1978) is a classic example of the use of this technique. She taught children of various ages nonsense word pairs in a story context that either did or did not conform to five phonological patterns: Palatalization with -ion, Vowel Shift with -ity, Vowel Shift with -ical, and stress shift with -ity and -ical. Subjects were then tested on the nonsense words one day, one week or six weeks later. The idea was that if the patterns were productive, nonsense words that conformed to the patterns would be easier to recall than those that did not; moreover, nonsense words that violated the patterns would be more likely to be changed to conform

---

9 Again, it must be emphasized that this sort of task does not necessarily measure productivity directly, but at best only the validity of generalizations across forms stored in memory.
to the patterns than the other way around. The nonsense words used in Myerson's studies are listed below.

(6) [after Table 1 in Myerson 1978:383; notation is hers]

<table>
<thead>
<tr>
<th>Stem</th>
<th>Derivation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Palatalization -ion</td>
<td></td>
</tr>
<tr>
<td>prezate</td>
<td>prezəsən</td>
</tr>
<tr>
<td>delort</td>
<td>delorəsən</td>
</tr>
<tr>
<td>Vowel Shift -ity</td>
<td></td>
</tr>
<tr>
<td>verane</td>
<td>verænitı</td>
</tr>
<tr>
<td>trave</td>
<td>trævity</td>
</tr>
<tr>
<td>Vowel Shift -ical</td>
<td></td>
</tr>
<tr>
<td>dəreter</td>
<td>dəretrical</td>
</tr>
<tr>
<td>məgeet</td>
<td>məʒitelı</td>
</tr>
<tr>
<td>Stress shift -ity</td>
<td></td>
</tr>
<tr>
<td>túrel</td>
<td>turálity</td>
</tr>
<tr>
<td>rómmel</td>
<td>rommålity</td>
</tr>
<tr>
<td>Stress shift -ical</td>
<td></td>
</tr>
<tr>
<td>gáthod</td>
<td>gathódical</td>
</tr>
<tr>
<td>néttory</td>
<td>nettórical</td>
</tr>
</tbody>
</table>

Two things make Myerson's study especially interesting. First, she investigated these rules using not only the memory task, but the production task and judgment task as well. Her figures thus show very clearly the relative sensitivities of these tasks. I illustrate this by giving below her results of the various tasks for the different rules obtained with twelfth-graders. The figures for the memory task are from the experiment with the six-week delay; percentages are percent of phonologically expected nonsense forms produced, judged or recalled, respectively. In virtually every case, higher success is achieved with the memory task than with the judgment task, which, in turn, brings better results than the production task.

---

10 This form is notated elsewhere as "məʒeet."
11 In Myerson (1978) these are notated incorrectly as "məʒǐtical" and "dərǐrical."
The second thing of significance about Myerson's (1978) work is that it is the only study that I know of where more than one rule was successfully tested using the same methodology. Of special interest to us is the fact that she got results on the productivity of both Palatalization and Vowel Shift. One thing that is particularly heartening about her results is that the relative productivity of these two rules remains essentially constant across all three methodologies: in all three tasks, Palatalization is more productive than Vowel Shift. In the production task, Palatalization is 55% productive while Vowel Shift is not productive at all; in the judgment task, Palatalization is 67% productive while Vowel Shift is only 44% productive; and in the memory task, Palatalization is fully productive while Vowel Shift is only about 72.5% productive (averaging across the two cases). This is an important finding, since one might object that productivity is purely a task-dependent performance phenomenon of no linguistic significance whatsoever.

Unfortunately, as we have seen, Palatalization and Vowel Shift do not interact and so there is no way to know if the relative productivity revealed by Myerson's studies has any consequences for the ordering of these rules.
The primary significance of the Myerson experiments thus remains as follows. First, the memory task appears to be much more sensitive than either the productivity or the judgment tasks. It would certainly be worthwhile to extend this methodology to other rules in order to examine their relative productivity more systematically. Second, different methodologies seem to rank rules according to productivity in essentially the same way. This implies that productivity is in fact a genuine property of rules, not an epiphenomenon generated by the methodology itself. Hence the evidence obtained in Myerson's experiments provides support for Gradient Productivity: rules differ gradiently in their relative productivities.

3.3.1.4 The perception task

I now move from experiments using memory tasks to those using perception tasks. In such tasks, the subject is presented with a nonsense form that either does or does not conform to the phonological pattern of interest. The subject is then tested (typically by asking her to repeat the nonsense word) to find out how the nonsense word was perceived. Presumably productive rules will show themselves in the ways nonsense words are misperceived.

One study employing this technique is Hochberg's (1988) investigation of the rules of Spanish stress. She had children listen to a list of nonsense words that either conformed to the rules of Spanish stress or did not; the task was for the children simply to repeat the nonwords back immediately. She found that overall, the more
irregularly stressed the nonword, the more children made errors in repeating it. However, she also found that the children had no problems repeating certain stress patterns that are prohibited in the lexicon of real Spanish words.

The only study of this type examining rules discussed in Chapter 2 is Wright's (1975) investigation of Nasal Assimilation in English. The experiment involved teaching one subject, a child, a set of nonwords that violated the Nasal Assimilation pattern, namely [goʊnŋ], [ɒŋd] and [ʧʌŋ]. This subject then had to teach these nonwords to another subject, who then had to pass them on to a third, and so forth for a total of seven subjects, in a sort of controlled version of the children's game "Telephone". The idea was that if Nasal Assimilation is a productive rule of English, misperceptions would eventually cause the nonwords to regularize.

The results certainly do support the hardly surprising conclusion that Nasal Assimilation is productive, but they are complex in some interesting ways. I give the history of the three nonwords in (8) by listing all the various changes they underwent during the course of their transmission.
(8) [after Table 1 in Wright 1975:394]

Original forms  
1st subject  
2nd subject  
3rd subject  
4th subject  
5th subject  
6th subject  
7th subject  

Nasal Assimilation is supported by these data in that every nonword goes through a stage where Nasal Assimilation is not violated, either by changing the place of the nasal to conform with that of the following stop (as in [gump] and [bųŋge]) or by changing the place of the stop to conform with that of the preceding nasal (as in [tŋge]). The fact that the nonwords subsequently underwent other changes is presumably due to other causes.

The pattern in (8) might possibly be used to justify the analysis of [ŋ] as underlyingly /Ng/, though of course the sample size is much too small to be really meaningful. Note that whereas [n] assimilates to the place of [p] and [m] to [g], [ŋ] does not assimilate, but instead causes the following segment to become [ŋ]. This is understandable if we assume that the underlying representation of [bŋd] that is built during the learning of this nonword is something like /bŋNgd/. In this case, the unspecified /N/ cannot assimilate in place to the /d/, which is not adjacent. Instead, it must assimilate to the /g/. The surface violation of Nasal Assimilation must thus be dealt with by
allowing the /ɡ/ to surface; syllable structure then forces the /d/ to be deleted.

This dramatic conclusion is not warranted by the small amount of data, however. There are too many mysterious changes seen in (8) to accord any one of them too much weight. All that is clear is that Nasal Assimilation is more than fully unproductive. It is not even clear if this is Obligatory Nasal Assimilation or Optional Nasal Assimilation, since both apply word-internally.

3.3.1.5 The concept formation task

The final sort of task that has been used in the literature to determine the productivity of rules is the concept formation task. In this task, subjects are taught to put given forms into different categories on the basis of examples whose category is indicated by the experimenter. As a hypothetical illustration, subjects may be told that the words puddle, moon, full belong to the category while water, sun, empty do not; based on their success in categorizing other words the experimenter determines how well the subjects have learned the category (in this case, words containing a double letter). The relative ease with which a category is learned presumably says something about how psychologically "natural" the category is.

The method can be used to test a wide variety of psychologically interesting phenomena. In using it to test rule productivity, the experimenter merely defines the category as consisting of words or nonwords that conform to the rule of interest. If this category is easier to learn than another collection of forms
that is not rule-governed, the rule is assumed to be at least somewhat productive. Presumably the relative productivity of different rules could also be tested in this way; if Velar Softening is truly more productive than s-Voicing, then, we would expect that subjects could learn the category of nonwords conforming to Velar Softening more readily than that of nonwords conforming to s-Voicing.

Examples of this technique include Jaeger (1984, 1986), Wang (1985), and Wang and Derwing (1986), all of whom were investigating the productivity of Vowel Shift. Unfortunately no other rules were tested, so there is no way to know how Vowel Shift compares in productivity with the rules that interact with it (namely the vowel length rules and Velar Softening).

The goal of all of these investigators is to show that Vowel Shift is not necessarily productive at all. They corroborate earlier researchers in finding that subjects do know some of the Vowel Shift alternations, but they argue this knowledge derives entirely from the subjects' knowledge of orthographic rules, that is, rules that determine how written words are to be pronounced.

It is useful to briefly discuss their arguments. Although for the reasons noted above, these experiments do not affect the case for the Productivity Hypothesis directly, they do illustrate one of the pitfalls that can arise in attempts to demonstrate the productivity of a phonological rule. The problem concerns how one can determine if the behavior literate speakers of a language exhibit in an experiment
does indeed result from their knowledge of phonological rules, which are presumably part of their grammar acquired automatically during language development, and not from their knowledge of orthographic rules, which are presumably learned by rote or invented as a mental shorthand through reading.

The argument that the apparent knowledge of Vowel Shift found in experiments such as Myerson (1978) may derive entirely from knowledge of how to pronounce written words is essentially as follows: Literates only appear to apply Vowel Shift productively if it overlaps with orthographic rules; otherwise they use only orthographic rules. Moreover, nonliterates (eg, young children) show no knowledge of the Vowel Shift rules at all.

As we saw in the previous chapter, the vowel alternations that Vowels Shift is intended to account for are those listed below.

(9) [from (2) in Jaeger 1986:78]

<table>
<thead>
<tr>
<th>Underlying vowel</th>
<th>Surface forms</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>/i:/-/i/</td>
<td>[ay]-[i]</td>
<td>divine-divinity</td>
</tr>
<tr>
<td>/e:/-/e/</td>
<td>[iy]-[e]</td>
<td>serene-serenity</td>
</tr>
<tr>
<td>/æ:/-/æ/</td>
<td>[ey]-[æ]</td>
<td>sane-sanity</td>
</tr>
<tr>
<td>/i:/-/i/</td>
<td>[aw]-[a]</td>
<td>profound-profundity</td>
</tr>
<tr>
<td>/ɔː:/-/ɔː/</td>
<td>[uw]-[a]</td>
<td>reduce-reduction</td>
</tr>
<tr>
<td>/ʊː:/-/ʊː/</td>
<td>[ɔy]-[ɔ]</td>
<td>destroy-destruction</td>
</tr>
<tr>
<td>/oː:/-/oː/</td>
<td>[uw]-[a/ɔ]</td>
<td>lose-lost</td>
</tr>
<tr>
<td>/ɔː:/-/ɔː/</td>
<td>[ow]-[a/ɔ]</td>
<td>verbose-verbosity</td>
</tr>
</tbody>
</table>

This pattern closely parallels the alternate pronunciations allowed for vowels and vowel clusters in English orthography. This pattern is illustrated below in (10).
The orthographic pattern illustrated in (10) is often explicitly taught in school: the left-hand pronunciations of these graphemes are called the "long" pronunciations, while the right-hand pronunciations are called the "short" pronunciations. However, one does not have to be explicitly taught the alternate pronunciations of the vowel graphemes to learn them. If one knows both the pronunciation and the spelling of the words, the pattern can be picked up by oneself.

Of course the connection between the orthographic and phonological patterns is not a coincidence. Prior to the diachronic sound changes of the Great Vowel Shift, the vowel graphemes in English had much more consistent pronunciations. But the fact that the alternate pronunciations of these letters almost perfectly mirrors the effects of the hypothesized synchronic rule of Vowel Shift does not mean that when experimental subjects show knowledge of these alternations, as they did in Myerson’s (1978) experiments, they necessarily know this phonological rule. A diachronic sound change has left a systematicity in the orthographic system as well, and it is quite plausible to suppose that people can learn this independently of any phonological rule.

<table>
<thead>
<tr>
<th>Orthography</th>
<th>Pronunciations</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;i&gt;</td>
<td>[ai]-[i]</td>
<td>&lt; site &gt;-&lt; sit &gt;</td>
</tr>
<tr>
<td>&lt;e&gt;</td>
<td>[ie]-[e]</td>
<td>&lt; Pete &gt;-&lt; pet &gt;</td>
</tr>
<tr>
<td>&lt;a&gt;</td>
<td>[ae]-[æ]</td>
<td>&lt; hate &gt;-&lt; hat &gt;</td>
</tr>
<tr>
<td>&lt;u&gt;</td>
<td>[uw]-[ʌ]</td>
<td>&lt; tube &gt;-&lt; tub &gt;</td>
</tr>
<tr>
<td>&lt;o&gt;</td>
<td>[ow]-[a/ɔ]</td>
<td>&lt; note &gt;-&lt; not &gt;</td>
</tr>
</tbody>
</table>
Fortunately for the researcher interested in the productivity of Vowel Shift, however, the orthographic rules do not overlap the phonological ones completely. In particular, there is no systematic orthographic parallel to the Vowel Shift alternations \([\text{aw}]-[\text{A}], [\text{oy}]-[\text{A}], \text{and } [\text{uw}]-[\text{a/ɔ}]\). The alternation \([\text{oy}]-[\text{A}]\) is not reflected in the orthography at all. As noted by McCawley (1986), the alternation \([\text{aw}]-[\text{A}]\) shows up orthographically in only one pair, namely *southern*. The alternation \([\text{uw}]-[\text{a/ɔ}]\) shows up only sporadically as well, as in *lose-lost*.\(^{12}\)

Jaeger (1984, 1986), Wang (1985), and Wang and Derwing (1986) exploit these mismatches between Vowel Shift and the orthographic rules to determine which are causing the extension of alternations like \([\text{ay}]-[\text{t}]\) to novel forms in experiments like those reported in Myerson (1978). All use the concept formation paradigm, where subjects are taught alternations by example; the goal is to see which alternations are easiest to learn. If the alternations are triggered by a unitary rule of Vowel Shift, we should expect the alternations \([\text{aw}]-[\text{A}], [\text{oy}]-[\text{A}], \text{and } [\text{uw}]-[\text{a/ɔ}]\) to be as easy to learn as all the other alternations; if they are instead triggered by orthographic rules, they should not. The results are clear: these alternations are not as easy to learn. The only alternations that were consistently acquired in the concept formation

\(^{12}\)It may be of significance that two of the three Vowel Shift alternations not associated with orthographic alternations require derivations from underlying segments that are fully neutralized on the surface, namely the high back unrounded /\text{i}/ and the high front rounded /\text{u}/.
experiments were the five involved in the orthographic pattern illustrated in (10).

It might be objected that these results are a side-effect of using the concept formation experiment. After all, as Jaeger (1986:90) observes, this method requires the subjects to be relatively conscious of the pattern they are supposed to learn. It may well be that subjects are therefore seeking to complete the task by encoding the problem as an orthographic one rather than as a phonological one. This seems plausible given that the orthographic rules are explicitly taught in school, while the phonological ones are not; someone trying to solve the task consciously might turn to consciously acquired knowledge for help.

However, another experiment using a different task has yielded similar results. Cena (1978) used a version of the memory task of Myerson (1978) to determine how well adults could recall the derived forms of nonsense stems when prompted with the stems. The expectation was that derived forms that showed the proper Vowel Shift alternations would be easier to learn than those that didn't, and that the forms that violated Vowel Shift would tend to be corrected so that they conformed. The results essentially matched those of Myerson (1978): under these conditions, Vowel Shift gave the appearance of being productive. Notably, however, the alternation \[aw\]-[\A] showed no evidence of productivity whatsoever.\(^{13}\) In other words, even using a memory task we find

\(^{13}\)The alternations \[oy\]-[\A] and \[uw\]-[a/o] were not tested.
that what subjects appear to be using are the orthographic rules and not the phonological Vowel Shift rule.\textsuperscript{14}

Such results do not unequivocally support the hypothesis that Vowel Shift is not productive as a purely phonological rule, however. It may be, for example, that the Vowel Shift alternations that subjects show no knowledge of are correlated with a smaller set of relevant forms.\textsuperscript{15} The rule associated with these forms would then be harder to learn and would consequently be less productive. Another possibility is that works such as Chomsky and Halle (1968) and Halle and Mohanan (1985) are simply wrong in ascribing such alternations to a rule of Vowel Shift.

There is a simple way to test the hypothesis, of course, and that is to probe the knowledge of Vowel Shift in nonliterates. The ideal test would involve subjects with fully mature phonological systems (eg, adults who show knowledge of the morphological processes involved in Vowel Shift) who nevertheless cannot read or write. So far all that has been done is to examine the knowledge of Vowel Shift in young children, including preliterate children. Thus Jaeger (1986) found that her three-year-old (nonliterate) daughter did not possess the relevant vocabulary even to extract the Vowel Shift alternations and learn them as rule-governed. Similarly, though Myerson's

\textsuperscript{14}A suggestive datum in the controversy about the role of orthographic knowledge in the use of Vowel Shift alternations is the finding of Armbruster (1978) that subjects' apparent knowledge of this rule is directly correlated with their S.A.T. verbal scores. My thanks to Jeri Jaeger for pointing this out to me.

\textsuperscript{15}This suggestion was made to me by Mike Hammond.
(1978) youngest subjects were not beginning readers (the youngest was eight years old), they showed a significantly weaker knowledge of the Vowel Shift alternations than the oldest subjects. This sort of evidence is not conclusive, since the difference may also be due to their phonological systems being less developed.

Regardless of the resolution of the particular issues involved in Vowel Shift, what remains important to remember is that the possible influence of orthographic knowledge affects virtually all of the English rules discussed in Chapter 2. To take just one example, consider the ordering of s-Voicing before Velar Softening. We saw that the nonvoicing of [ʃ] in words such as Grecian seems to imply that s-Voicing applies before Velar Softening. If speakers' knowledge of this pattern is derived from orthography, however, there is no need to order such rules at all. The original sound change that gave rise to Velar Softening occurred in late Latin, which had used the grapheme <c> to represent [k]. The result is that all segments in English that undergo the [k]~[s] alternation are transcribed with <c>. It so happens, however, that <c> is never pronounced [z]. Thus there is no way for segments that undergo Velar Softening to undergo s-Voicing. Instead, all one needs is the following pair of noninteracting orthographic rules. Note that rule (11b) makes Velar Softening completely superfluous.16

16 As Mike Hammond pointed out to me, independent evidence for an orthographic rule like this for <c> is found in pairs like picnic-picnicking, where the <k> is inserted to prevent the word from being pronounced with [s]. Note that the phonological rule of Velar Softening is not even applicable here,
(11) a. \(< s > \) is pronounced \([ʃ]\) before \(< h >\)
    elsewhere, \([s]\) or \([z]\) in accordance with the rules of
    phonology (s-Voicing, voicing assimilation, etc)

b. \(< c > \) is pronounced \([s]\) before \(< e >, < i >, < y >\)
    as \([ʃ]\) before \(< h >\), elsewhere as \([k]\)

Further discussion of the orthographic problem would take us
far beyond the scope of this dissertation. The key point I wish to
emphasize in this section is just that the concept formation task
seems to be a potentially useful way to measure the relative
productivity of phonological rules. So far, however, the task has not
been used for this purpose.

3.3.1.6 Experiments: Summary

In this section I have considered five classes of tasks that have
been used to investigate the productivity of phonological rules,
namely the production, judgment, memory, perception and concept
formation tasks. The tasks differ in various ways. For example, the
production task is less sensitive than the judgment task, which in
turn is less sensitive than the memory task. Moreover, the concept
formation task seems to allow conscious decision-making to play a
larger role than some of the other tasks. The degree with which all
of the tasks may be treated by subjects as a puzzle to be consciously
solved is a problem for experimental studies of rule productivity in
general, however; this points out just one of the advantages with
examining naturally occurring speech errors instead.

following the common assumption that \(-ing\) is suffixed long after Velar
Softening has ceased to apply.
As far as the Productivity Hypothesis is concerned, results of experiments reported in the literature are only suggestive. Production studies carried out by Ohala (1974) and Myerson (1976, 1978) imply that Velar Softening is less productive than Palatalization, exactly as we would predict. Unfortunately, however, they also imply that Velar Softening is more productive than Vowel Shift, contrary to what we would predict. Part of the problem may be an artifact arising from treating two different experiments as if they were comparable.

Other studies bear even less directly on the Productivity Hypothesis. Myerson (1976, 1978) found that Palatalization is more productive than Vowel Shift. As these two rules do not interact, the significance of this for the Productivity Hypothesis is not clear. Numerous studies have found that Vowel Shift is greater than fully unproductive, but the question of whether this is due to knowledge of a phonological or an orthographic rule has not been settled. Finally, studies have shown that word-internal Nasal Assimilation is productive, but it is unknown whether this is Obligatory or Optional Nasal Assimilation.

### 3.3.2 Speech error accommodations

One major problem with the use of experiments in testing the relative productivity of phonological rules is the narrowness of their focus: each experiment is designed to test a single rule. Only in rare cases has a researcher examined several different rules with the same methodology. Unfortunately, if different rules are tested with
different methodologies, there is no way to compare their relative productivities.

One way out of this dilemma, in addition to performing the obviously necessary task of conducting more and better experiments, is to examine what happens to novel forms generated by speech errors. Unlike the case in experiments, where a different set of nonce forms must be made up and tested for each rule of interest, with speech error forms all one has to do is collect examples; eventually one should have enough examples to say something of statistical significance about any rule one could imagine.

Of course, the use of speech errors as evidence in linguistic theories is not without its pitfalls as well, as Cutler (1982), Stemberger (1983) and Ohala (1986a) emphasize. This is true whether errors are collected from naturally occurring speech or induced experimentally. In the former case, one depends on the reliability of the transcriber to catch the relevant errors. To take an example especially relevant to my interests, suppose one is researching the question of whether people ever accidentally neglect to assimilate nasals in place to a following stop. Such errors may be quite frequent without anyone being aware of them, since nasal-stop sequences are commonly perceived as homorganic, even if they really aren’t.\footnote{This problem was pointed out to me by several people, including Merrill Garrett, Mike Hammond, Stefanie Shattuck-Hufnagel and Joe Stemberger.} It may be, then, that speech errors quite commonly violate productive patterns of a language but are not perceived as
such. Another problem with using naturally occurring speech errors is their relative infrequency; it takes many years to collect a sufficient number of examples.

Yet experimentally induced speech errors are not necessarily much better. Merrill Garrett (pc) has pointed out that that the speech errors induced in such experiments often differ from naturally occurring ones in important ways; for example, the induced errors examined by Davidsen-Nielsen (1975) violated phonotactic constraints that naturally occurring speech errors have never been observed to violate. Moreover, as with all experimental tasks, induced speech errors do not necessarily provide information on all imaginable rules. Studies that use induced speech errors (eg, Stemberger and MacWhinney 1988) typically must devise different tasks depending on what phenomenon they want to study.

My use of naturally occurring speech errors is problematic in another way. Since the number of relevant speech errors that I have been able to cull from published and other sources, as well as on my own, is still rather small, I am not able to make any clear judgments about the relative productivity of different rules.

Part of the problem in using speech error evidence to investigate rule productivity is that in spite of the vast literature on speech errors, very little of it is concerned with the sort of error I am interested in here, namely speech error accommodations. These are speech errors where a novel form has been created by inserting, deleting or moving a segment, segment string or morpheme into or
out of a real word in such a way that a phonological generalization is violated and then corrected ("accommodated" to the new environment). A (constructed) example of such an error would be the mispronunciation of taffy store as staffy tore, where the /t/ of store undergoes Aspiration in the novel form tore, which was created by the leftward shifting of /s/.

Since the errors certainly occurred on-line, these corrections may be thought of as the on-line application of phonological rules. The extent to which a phonological rule tolerates or corrects an error-produced violation can then be taken as a measure of the rule's productivity.

In this section I will therefore investigate the evidence that is currently available in this data source. We will see that, though not conclusive, the findings are suggestive enough to warrant attention. For each rule in Chapter 2 I will discuss what evidence there is for its productivity from speech error accommodations.

3.3.2.1 Vowel length and Vowel Shift

Evidence from speech error accommodations bearing on the question of the productivity of the vowel length and Vowel Shift rules is meager at best. I have found no examples of relevance in all in the error corpora I have examined. Stemberger (1986) notes that out of seventy-seven apparent noncontextual vowel errors in his corpus (ie, errors where a word appears with the incorrect vowel and yet does not appear to have been copied or shifted from another word in the utterance), nineteen might possibly be understood
actually as involving the shift of vowels before the vowel length and Vowel Shift rules have applied. Sternberger (1986) lists the following three examples.\(^{18}\)

(12) [after (7) in Sternberger 1986:6]

a. st86 You subtly \textit{tand} --- tend to change the tune.

b. st86 Oh, he's \textit{snied} -- snowed in.

c. st86 Error: We're gonna \textit{plane} a few things.
   Target: We're gonna plan a few things.

Thus in (12a), it could be that the underlying \(\text{i}:\) of \textit{change} was anticipated, shortened due to its appearing in a closed syllable environment, and then surfaced as [æ]. In (12b) it could be that a hypothetical underlying \(\text{i}/\) of \textit{in} (rather than \(\text{I}/\)) was anticipated and lengthened, thus surfacing as [ay]. And in (12c) it could be that the length of \textit{few} was anticipated, causing the underlying vowel \(\text{æ}/\) to lengthen to \(\text{i}:\) and then change to [ey] by Vowel Shift.

Sternberger (1986) points out, though, that even if such examples did show evidence of Vowel Shift, there are as many errors (sixteen out of the seventy-seven noncontextual vowel errors) that must occur after Vowel Shift. He lists the two errors below.

\(^{18}\)Throughout my discussion of speech errors, I indicate the source of each error with the following code: df = unpublished error corpus of David Fay, f73 = Fromkin (1973), f88 = Fromkin (1988), fc = Fay and Cutler (1977), jm = unpublished error corpus I have compiled, mit = unpublished error corpus compiled at MIT, sh = Shattuck-Hufnagel (1986), ssh = unpublished error corpus of Stefanie Shattuck-Hufnagel, st86 = Sternberger (1986), st83 = Sternberger (1983). It is standard, unfortunately, for errors to be represented in speech error studies with English orthography. I have resisted the temptation to give these in phonetic transcription to reduce biasing the material. Where I give phonetic transcription, it is found in the original sources.
In these examples, vowel length has erroneously been changed, resulting in changes in vowel tenseness; nevertheless, Vowel Shift is not triggered (we would expect \[təmæto\] and \[sayl\], respectively).

Notice that three of the errors cited, namely (12b,c) and (13b), do not even fit the criteria I require of evidence for rule productivity, since the errors have produced real words and not novel forms. It is thus possible that what has occurred is an erroneous form-based word substitution, not a segment shift at all. As we will see later, this is a very different sort of error, one which bears on the question of rule productivity only indirectly.

Sternberger (1986) himself comments on the unsatisfying nature of such errors as follows:

It is difficult to tell if these analyses of "noncontextual" errors are legitimate or if the presence of the appropriate vowel nearby is just chance; note that most (n=41) noncontextual vowel errors cannot be explained in this fashion. (Sternberger 1986:7)

It appears, then, that there is essentially no evidence to be gained from speech error accommodations on the question of the productivity of Vowel Shift. We will see later that the same is true of evidence gained from other sorts of speech errors.
3.3.2.2 s-Voicing and Velar Softening

Evidence from speech error accommodations is just as meager for the productivity of s-Voicing and Velar Softening. I have found no examples of relevance from my own searches. Stemberger (1986), based on an error corpus of some 7220 naturally occurring errors and from error-induction experiments, claims to have found no errors bearing on the question of productivity of any of Halle and Mohanan's (1985) Stratum 1 rules, which include s-Voicing. He cites exactly one error touching on Velar Softening, given below in (14).

(14) [after Stemberger 1986:5; also found in Stemberger 1983:17]

\begin{verbatim}
st86 Error: the easily accessible places
Target: the easily accessible places
\end{verbatim}

In this error, apparently the /p/ from \textit{places} has been anticipated, appearing within the word \textit{accessible}. Note, though, that the /p/ has appeared between what would be an underlying /k/ and the following /e/. The presence of the /p/ might thus be expected to block the application of Velar Softening, but it does not; the /k/ surfaces as [s]. Apparently, then, Velar Softening, if it has applied at all in this word, applied before the speech error anticipation. There is thus no evidence from this example that Velar Softening can apply on-line.

The most we can conclude from this, then, is that there is no evidence from speech error accommodations that s-Voicing and Velar Softening are productive at all. This does not rule out the possibility that other tests will find them to be productive.
3.3.2.3 Spirantization, y-Insertion and Palatalization

The evidence from speech error accommodations for Spirantization and Palatalization is also scarce. I have found no evidence that bears on the productivity of Spirantization. Below I list the only examples of relevance to Palatalization that I have come across.

(15) a. ssh for D, the jester... gesture for G...
b. f73 Error: aspectal
    Target: aspectual

In both errors, a [yu] has erroneously been left out of the word, thus removing the conditioning environment for Palatalization. In both cases, Palatalization is in fact blocked, and the words surface with the surface reflex of the original underlying /t/.

These errors are hardly conclusive. First and most obviously, there are only two of them. Second, the error in (15a) as produced a real word, jester, so this may not be an example of the blocking of Palatalization at all.

Notice, however, that these examples also give evidence on the productivity of y-Insertion, since in both cases [y] and [u] were affected as a unit. Thus the error may be thought of as involving the removal of just /u/, with the subsequent bleeding of y-Insertion.

Due primarily to the work of Shattuck-Hufnagel (1986), there is a relative wealth of published information from speech error accommodations on the productivity of y-Insertion. In addition to the two errors listed above, I have come across the following twenty-three examples. I have arranged them into four categories: (16a)
lists examples where y-Insertion should be triggered by the speech error, and is; (16b) lists examples where y-Insertion should be blocked by the speech error, and is; (16c) lists examples where y-Insertion should be triggered, but is not; and (16d) lists examples where y-Insertion should be blocked, but is not.

(16) Accommodation errors relevant to y-Insertion:

a. should apply + does apply:

ssh Error: supposed to allow for very similar...
Target: supposed to allow for very similar...

ssh Error: these acoustic characteristics...
Target: these acoustic characteristics

sh Error: Lee Yumis
Target: Lee Loomis

st83 Error: chin-skewing
Target: skin-chewing

st83 Well, it's the high risk kyoup --- group.

sh Error: m/yu/sarpial
Target: mars/yu/pial19

19Uttered by a speaker who allows coronal-/y/ onset sequences.
b. should not apply + does not apply:

**ssh** Error: afraid of executing yourself
Target: afraid of executing yourself

**sh** Error: ruising
Target: using

**sh** Error: Luclid
Target: Euclid

**sh** Error: /fluid/
Target: feud

**sh** Error: fluz
Target: fuse blown

**sh** Error: writing rutensil
Target: writing utensil

**sh** Error: Peggy Newing
Target: Peggy Ewing

**sh** Error: completely
Target: completely
c. **should apply + does not apply**

<table>
<thead>
<tr>
<th>ssh</th>
<th>Error:</th>
<th>fuel bools</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>fuel bills</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f73</th>
<th>Error:</th>
<th>spoon feeding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>spoon feeding</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f73</th>
<th>Error:</th>
<th>feet moving</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>feet moving</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f73</th>
<th>Error:</th>
<th>[səpʊwət] the movement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>support the movement</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>f73</th>
<th>Error:</th>
<th>[buwd] and [rʊrm]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>room and board</td>
</tr>
</tbody>
</table>

| f73  | Error:     | deering the foost blur book
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>during the first blue book</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>st</th>
<th>Error:</th>
<th>verbal behaber</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>verbal behavior</td>
</tr>
</tbody>
</table>

d. **should not apply + does apply**

<table>
<thead>
<tr>
<th>sh</th>
<th>Error:</th>
<th>lose aftershave yotion</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>use aftershave lotion</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>sh</th>
<th>Error:</th>
<th>redistrib/yey/...</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target:</td>
<td>redistributed</td>
</tr>
</tbody>
</table>

In the first error in (16a), for example, the [v] in similar is erroneously replaced with [u]. This appears to trigger y-Insertion as it should. In the first error in (16b), an /r/ is erroneously inserted (copied from the /r/ in afraid, or perhaps yourself) into executed.

---

20This form may be [fʊst], not [fuːst], in which case y-Insertion would not be expected to apply.
before the /u/. No English dialects tolerate the insertion of /y/ after /r/; hence y-Insertion is properly blocked here. In the first example in (16c), the /u/ of fuel has been perseverated into bills, but although y-Insertion should be able to apply here, it does not. Finally, in the second example in (16d), the [u] of redistributed has been replaced with [ey], a change that should bleed y-Insertion; nevertheless, the error form surfaces with [y] before the vowel.

Just given this list, as well as the two examples in (15), the ratio of correct applications or nonapplications of y-Insertion to incorrect ones is sixteen to nine, suggesting that y-Insertion may have some degree of productivity, at least better than Vowel Shift, s-Voicing and Velar Softening. The situation is actually a bit more complex, however. First, all but one of the cases in (16b) where y-Insertion properly fails to apply involve the appearance of a liquid, namely [l] or [r], before the relevant vowel. According to Borowsky (1986), there are no dialects where y-Insertion is robustly allowed after liquids. It may be, then, that what the examples in (16b) show is not the productivity of y-Insertion, but the productivity of some Postliquid y-Deletion rule. If we remove these examples, the ratio of support for y-Insertion changes from fifteen-to-nine to five-to-nine.

Yet the cases where y-Insertion apparently incorrectly fails to apply may not be what they seem, either. Notice that five out of the eight examples in (15) and (16a) of correct application of y-Insertion involve clearly non-Germanic words, namely gesture, aspectual, acoustic, similar and marsupial. By contrast, five out of the seven
examples in (16c) of apparently incorrect nonapplication of y-Insertion involve clearly Germanic words, namely *bills, spoon, feed, feet, moving, room, board, blue* and *book*. One of the remaining cases where a non-Germanic word was affected by the error also involved a Germanic word, namely *support the movement*. It may be, then, that the structural description of y-Insertion includes the stipulation that it only applies to non-Germanic forms. Linguistic evidence for this proposition was offered in the last chapter; here I have shown what appears to be psychological evidence as well. Speakers appear to have knowledge about whether a word is Germanic, and this knowledge shows up even in such unconscious behavior as speech errors.

Removing these examples from the list, however, leaves us with only ten relevant examples. Eight show evidence of the productivity of y-Insertion, namely those in (15) and (16a), and two, namely those in (16d), show evidence of the nonproductivity of y-Insertion. The numbers aren't very impressive, but there still appears to be a trend supporting the claim that y-Insertion is more productive than Vowel Shift, s-Voicing and Velar Softening.
3.3.2.4 Nasal Assimilation and g-Deletion

The evidence from speech error accommodations of relevance for Nasal Assimilation and g-Deletion is also relatively good. I have come across twelve errors relevant to Nasal Assimilation, listed below.\(^\text{21}\) All of them imply the productivity of this rule. The errors in (17a) show Nasal Assimilation being properly triggered, and those in (17b) show it being properly blocked.

\(^{21}\)There are additional speech errors that have been considered relevant in the literature as well. Examples include the following:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>inch thi[ŋ]k</td>
<td>inch thick</td>
</tr>
<tr>
<td>[kʌnt] the [strɪŋ]</td>
<td>cut the string</td>
</tr>
</tbody>
</table>

I have not included these errors because they involve the production of real words, namely *think* and *cunt*. 
(17) Accommodation errors relevant to Nasal Assimilation:

a. \textit{should apply} + \textit{does apply}

\begin{itemize}
  \item \textbf{ssh} Error: Intalian ingredients
    Target: Italian ingredients
  \item \textbf{f73} Error: kinchen sink
    Target: kitchen sink
  \item \textbf{f73} Error: cojunction reduction
    Target: conjunction reduction
  \item \textbf{f73} Error: the \textit{[rænd]} orker...
    Target: the rank order of the subjects
  \item \textbf{f73} Error: \textit{[stɪmp]} a \textit{[ʃɪk]}
    Target: sink a ship
  \item \textbf{f73} Error: \textit{[spriɡ]}time for \textit{[hɪntləɹ]}
    Target: springtime for Hitler
  \item \textbf{st83} Error: finger primps
    Target: finger prints
  \item \textbf{st83} Error: albino bungie
    Target: albino budgie
  \item \textbf{f88} Error: po\[ŋ]cutation
    Target: computation
\end{itemize}
b. should not apply + does not apply

- f73 Error: the red tide will [stǐn] up the [siyk]
  Target: the red tide will stink up the sea

- f73 Error: [swīn] and [sweyg]
  Target: swing and sway

- f73 Error: morphemes in the [vARG strīn]
  Target: morphemes in the verb string

Thus in the first example in (17a), a nasal consonant of some sort has been anticipated from *ingredients*, appearing before the /t/ in *Italian* as the appropriately homorganic [n]. In the first example in (17b), the /k/ of *stink* has been shifted to form the coda of *sea*; the underlying placeless nasal consonant has thus surfaced with a default coronal place specification. It appears, then, that there is good reason to suppose that Nasal Assimilation is rather productive.

As Fromkin (1971) first argued, some of these examples also seem to show the productivity of Postnasal g-Deletion. I repeat these below.22

---

22Two other errors cited in this connection are the following. Again, I have not included them because they involve the production of real words, namely *chunk* and *big*.

- f73 Error: Chu[ŋ]k Yug
  Target: Chuck Young

- f88 Error: Big Cronsby
  Target: Bing Crosby
(18) f73 Error: [sprɪg]time for [hɪntlər]
Target: springtime for Hitler

f73 Error: [swɪn] and [sweyg]
Target: swing and sway

f73 Error: morphemes in the [vɑɹg strɪn]
Target: morphemes in the verb string

As evidence for the productivity of g-Deletion, these examples are not very persuasive. They are supposed to show that the [ŋ] of the target words are underlingly /Ng/, since apparently this /g/ has moved independently of the /N/, surfacing in places where g-Deletion does not apply. There are problems with this analysis, however. In the third error in (18), the word that surfaces as [vɑɹg] already has a voiced obstruent stop in coda position in its target form [vɑɹb]. Moreover, as Sternberger (1983) shows, the other errors do not require the assumption that [ŋ] is underlingly /Ng/, either. If [ŋ] is underlingly /ŋ/, we might expect errors to be able to target each of its features separately, as other errors have been shown to do (see, eg, Fromkin 1971, 1988, Sternberger 1983, 1991, among many other sources). Then we can understand the first error in (18) as involving the shifting of just the nasality feature and the second error as involving the shifting of just the place feature. As Sternberger (1983) notes in his discussion of errors involving [ŋ], not enough is known to say whether such feature-level shifts occur more with this segment than with [m] or [n], as Fromkin's (1971) /Ng/ analysis would predict.
Furthermore, as Stemberger (1983) points out, the Fromkin (1971) analysis seems to presuppose that shift, insertion and deletion errors must manipulate only one segment or other constituent at a time. This is not true; there is a class of errors, which Stemberger (1983) calls "complex," that cannot be analyzed as the erroneous placement of just one constituent. Some examples from Fromkin (1973) are listed below.

(19) f73 Error: \[\text{protfæbowkŋ}\]
Target: thought provoking

f73 Error: it's now whittled clad whites
Target: it's now middle class whites

Thus at this point there doesn't seem to be any clear evidence from speech error accommodations for the productivity of g-Deletion.

3.3.2.5 Evidence from speech error accommodations: Summary

In short, it appears that the evidence from speech error accommodations regarding the relative productivity of the rules discussed in Chapter 2 is suggestive but not conclusive. We saw that there appears to be no convincing evidence regarding the productivity of the vowel length rules, Vowel Shift, s-Voicing, Velar Softening, Spirantization and g-Deletion. There was some evidence suggesting that Palatalization and y-Insertion are at least somewhat productive and relatively abundant evidence that Nasal Assimilation is rather productive. Although the information available so far from speech error accommodations is insufficient to allow a direct test of the Productivity Hypothesis, it is clear that this method has promise,
and I intend to continue to pursue it as more large speech error corpora become generally available over the next few years.

3.3.3 Malapropisms

I now turn to another sort of speech error, referred by Fay and Cutler (1977) as *malapropisms*. In this sort of error, a real word is unintentionally substituted for another real word and the basis of this switch does not seem to be similarity in meaning. Some examples are given below. The examples in (20a) are malapropisms, since the error involves substituting a real word with another real word not based on meaning similarity. The examples in (20b) are not, since the word substitution seems to be at least partly conditioned by meaning similarity.

(20) a. f73 Error: white Anglo-Saxon prostitute
Target: white Anglo-Saxon Protestant

f73 Error: the native vowels
Target: the native values

f73 Error: the conquest of Purdue
Target: the conquest of Peru

b. f73 take him to the lab first -- I mean last
f73 I would have gone -- come
f73 When were you last on the west -- east coast

By following Fay and Cutler (1977) in defining malapropisms in this negative way, we provide significance to the following observation: malapropisms involve form similarity, as can be seen

---

23 Their use of this term differs from the common use, which refers not to momentary speech errors but to erroneous word substitutions due to ignorance.
by the examples above in (20a). While this is hardly surprising, it is not a priori necessary. Moreover, we can now investigate the nature of this form similarity without ruling out possibly interesting cases with a biasing conception of what should or should not constitute "similar forms."

Fay and Cutler (1977) point out that the investigation of malapropisms as defined is significant for linguistic theories because they provide a window into the way forms are stored in memory. Because malapropisms involve the incorrect production of real words, the error must occur at the stage of lexical retrieval, the only difference between real words and possible nonwords being the fact that only the former are stored in memory. Moreover, since the errors seem to involve form similarity, they presumably show what is considered "similar" at the level of lexical retrieval. In other words, error forms should be more likely to be retrieved in place of the target forms the more similar the forms are in their representations in memory.

Fay and Cutler (1977), applying this insight to a corpus of one hundred eighty-three malapropisms, find that error and target forms match in a striking respect: they almost always agree in number of syllables and in the placement of main stress. In other words, it appears that prosodic information is present even in forms stored in memory. Independent evidence for this comes from work into what is called the the tip-of-the-tongue phenomenon (eg, Brown and McNeill 1966), where speakers struggle to recall a word but can only
call up selected aspects of it. One of the aspects they are often able to extract is the prosodic structure, specifically the number of syllables and the stress pattern. This sort of evidence also implies that prosodic information is stored along with other phonological information about a word.

Fay and Cutler (1977) suggest that the same technique can be used to determine how the segmental information is stored as well. In their corpus, they find several examples where the erroneously substituted word matches the target word more closely if the effects of y-Insertion are undone; these are listed below in (21). This implies that words may be stored in memory without the /y/ that could be inserted by y-Insertion.

(21) [after (15) in Fay and Cutler 1977:516]

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. musician</td>
<td>magician</td>
</tr>
<tr>
<td>b. museums</td>
<td>machines</td>
</tr>
<tr>
<td>c. emanate</td>
<td>emulate</td>
</tr>
<tr>
<td>d. review</td>
<td>revise</td>
</tr>
<tr>
<td>e. genuine</td>
<td>general</td>
</tr>
<tr>
<td>f. movie</td>
<td>music</td>
</tr>
</tbody>
</table>

I will illustrate the argument with the example in (21f). If these words are compared in their surface forms, there are several mismatches, as indicated in (22).

(22) m u w v i y
    |   |   |
    m y u w z i k
By contrast, if these words are matched in the forms they would have before the application of y-Insertion, the mismatches are fewer, as indicated in (23).

(23)  

| m | u | w | v | i | y |

Given such hints, Fay and Cutler (1977:517) suggest that a comparison of error and target words at the level of Chomsky and Halle's (1968) underlying forms may "result in an overall simplification in the description of our data." According to Anne Cutler (pc), such a step has never been taken. Given the apparent reality of Gradient Productivity, in fact, it might not even be a good idea. Since not all rules are equally productive, we might not expect there to be a single level where malapropisms and target words match closest. Instead, it makes more sense to compare error and target forms by selectively undoing the effects of each rule individually.

The result of such a test would be an indirect measure of the relative productivity of the rules. As with all tests involving real words, some caution is required in the interpretation of the results. Specifically, while evidence that a rule is never prepatterned implies that it must be productive, evidence that it is prepatterned does not necessarily mean that it is not productive. After all, Myerson's (1976, 1978) experiments suggest that at least some English stress rules are rather productive (see last two rows in figure (6) above),
although Fay and Cutler (1977) find that at least some stress patterns are stored on forms in memory. Nevertheless, we might expect some inverse correlation between prepatterning and productivity.

The methodology I used, then, was as follows. First I attempted to build up a large corpus of malapropisms. I did this by combining the examples in Fay and Cutler (1977) with others culled from David Fay's original notes, Fromkin (1973), the MIT corpus and some examples collected on my own; eliminating cases where precisely the same malapropism was made, this results in a corpus of four hundred fifteen different examples. Second, I went through this corpus looking for examples of words that are expected to have undergone the specific rules from Chapter 2 I am interested in. Third, for each rule, I determined what these words should look like if the effects of the rule were undone. Fourth, I eliminated all cases where both error and target form underwent precisely the same rule, so that both were equally the same (or different) at both the surface and more abstract level. Finally, of the pairs that remained I

---

24 I have not done a systematic analysis of stress matches in malapropisms. It is logically possible, for example, that the stresses that match are precisely those that are unpredictable; with predictable stress, it may be that forms match better at a pre-stress level. There are suggestive hints that this is true, such as the error given below.

<table>
<thead>
<tr>
<th>df</th>
<th>Error:</th>
<th>Target:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>original</td>
<td>ordinary</td>
</tr>
</tbody>
</table>

The stressing on these words matches better if one matches the stem of the first (ie, *original*) rather than the derived form, which has undergone stress shift.
determined whether the surface forms or more abstract forms matched better. "Matching better" was determined by a formula based on the number of segments and features in each form. As examples, consider the pairs below.

(24)  

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>jm explosion exposure</td>
<td>optionally optically</td>
</tr>
<tr>
<td>fc</td>
<td></td>
</tr>
</tbody>
</table>

In the surface forms of the first pair, the error and target forms differ in that unlike the target form the error form has [I] and ends in the suffix -ion rather than -ure. In the forms these words have prior to the application of Spirantization, however, they differ not only in these ways but in another way as well: explosion, being derived from explode, has an underlying /d/, while exposure, being derived from expose, has an underlying /z/. Thus for the pair in (24a), the surface forms match better with respect to Spirantization. By contrast, for the pair in (24b), the surface forms have a mismatch that is not found when comparing the pre-Spirantization forms. Namely, in their surface representations, optionally contains [ʃ] where the surface form of optically has [t], while in their pre-Spirantization forms both words contain /t/ (assuming optionally is derived from opt).

Using this methodology, I tested the relative productivity of the rules discussed in Chapter 2. More precisely, for each rule I came up with a list of malapropisms that either showed or did not show evidence of prepatterning according to this rule. Note, however, that finding that a pair of words match better when the
effects of a rule are undone than in their surface forms does not mean the rule must have applied in those forms. As the examples below illustrate, malapropisms can occur even if the match between forms is not particularly strong either on the surface or underlyingly.25

(25) Error  Target
    df  called  came
    df  careful  quite
    fc  do  take
    df  embarrass  abolish
    fc  got  gave

3.3.3.1 Vowel Shift

The evidence from malapropisms for the productivity of Vowel Shift is as equivocal as the other evidence cited above. Below I list all appropriate examples in my collection involving front vowels. In (26a) I list the malapropisms where the error words match the target words better at the pre-Vowel Shift level, and in (26b) I list the ones where the error words match the target words better at the post-Vowel Shift level. As can be seen, the two lists are almost exactly the same length.

25If one defines the term malapropism to cover only error-target pairs which are sufficiently similar in form, some of these pairs are clearly too different to qualify. If we define malapropism in the negative way I have done above, however, following Fay and Cutler (1977), then all of these pairs are by definition malapropisms, since the errors are true lexical substitutions and not derived by blending, segment shifting, copying and so forth, and yet are not semantically conditioned.
(26) a. Match better pre-Vowel Shift:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>df biography</td>
<td>bibliography</td>
</tr>
<tr>
<td>fc emancipated</td>
<td>emaciated</td>
</tr>
<tr>
<td>fc experience</td>
<td>experiment</td>
</tr>
<tr>
<td>df experiences</td>
<td>expenses</td>
</tr>
<tr>
<td>jm feel</td>
<td>fell</td>
</tr>
<tr>
<td>df line</td>
<td>list</td>
</tr>
<tr>
<td>fc map</td>
<td>make</td>
</tr>
<tr>
<td>fc shield</td>
<td>shed</td>
</tr>
<tr>
<td>df standing</td>
<td>staying</td>
</tr>
</tbody>
</table>

b. Match better post-Vowel Shift:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc blankly</td>
<td>blindly</td>
</tr>
<tr>
<td>fc fire</td>
<td>follow</td>
</tr>
<tr>
<td>fc label</td>
<td>level</td>
</tr>
<tr>
<td>df leave</td>
<td>live</td>
</tr>
<tr>
<td>df made</td>
<td>mgt</td>
</tr>
<tr>
<td>df paper clips</td>
<td>pipe cleaners</td>
</tr>
<tr>
<td>df reading</td>
<td>ringing</td>
</tr>
<tr>
<td>fc straddles</td>
<td>sidles</td>
</tr>
<tr>
<td>fc trick</td>
<td>treat</td>
</tr>
</tbody>
</table>

As an example of how the comparisons work, consider the pair \textit{line-list} in (26a). On the surface, these words differ in the vowel, among other things; one has [ay] and the other [l]. Before Vowel Shift, however, the first vowel of the first word is /i:/ and that of the second word is /i/. In fact, the match is even closer if we assume that the underlying vowel in \textit{list} is a long /i:/ which is then shortened by Closed Syllable Shortening. Thus in this case the pre-Vowel Shift forms match better.
By way of contrast, consider the pair label-level in (26b). On the surface, the vowels in the second syllable of these words differ only in length (and consequently tenseness as well); both [ey] and [ε] are mid front vowels. Before Vowel Shift, however, these vowels are /æ:/ and /e/; at this level the vowels differ not only in length but also in height. Thus in this case the post-Vowel Shift forms match better.

At best, then, these malapropisms may show that Vowel Shift is not prepatterned about half the time, implying that the rule is at least somewhat productive. To know whether this is really the case, however, one would have to collect enough examples to determine if the apparent pattern noted above is real or merely due to chance.

Before closing this section, I should briefly mention what evidence there is from malapropisms bearing on the productivity of Prenasal g-Deletion, the rule that deletes the underlying /g/ in pairs like sign-signature, thereby triggering Vowel Shift. As might be expected, there isn't very much. I have found only two relevant cases, namely those listed below. In one case (27a), it appears that the underlying cluster /gm/ has been matched with the cluster /ks/, implying that this /g/ really is present underlyingly. However, in the other case (27b), exactly the opposite situation holds. Here the hypothesized underlying cluster /gm/ is matched with the single consonant [m], implying that in this error the /g/ was not present underlyingly. The small sample makes any firm conclusions premature.
3.3.3.2 s-Voicing

There are not many examples relevant to s-Voicing in my collection, but what there is suggests that this rule is generally prepatterned. I list the examples below; (28a) gives the examples where a pre-s-Voicing match works better, while (28b) shows examples where a surface match works better. In the pre-s-Voicing pairs, the root-initial [s] is matched with a voiceless consonant; in the post-s-Voicing pairs, the root-initial [s] is matched with a voiced consonant. As can be seen, the post-s-Voicing pairs slightly outnumber the pre-s-Voicing pairs, but unfortunately the numbers are too small to know what to make of this.

(28) a. Match better pre-s-Voicing:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
<td>operation</td>
</tr>
<tr>
<td>fc</td>
<td>presented</td>
</tr>
<tr>
<td>df</td>
<td>reticent</td>
</tr>
</tbody>
</table>

b. Match better post-s-Voicing:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>deserved</td>
</tr>
<tr>
<td>df</td>
<td>pregnant</td>
</tr>
<tr>
<td>mit</td>
<td>prevent</td>
</tr>
<tr>
<td>df</td>
<td>providing</td>
</tr>
<tr>
<td>fc</td>
<td>result</td>
</tr>
<tr>
<td></td>
<td>diverged</td>
</tr>
<tr>
<td></td>
<td>present</td>
</tr>
<tr>
<td></td>
<td>present</td>
</tr>
<tr>
<td></td>
<td>presiding</td>
</tr>
<tr>
<td></td>
<td>regard</td>
</tr>
</tbody>
</table>
3.3.3.3 Velar Softening

The number of relevant examples for Velar Softening is not much greater, and again about half of them show better matches post-Velar Softening; that is, this rule seems to be prepatterned at least some of the time. The relevant examples are listed below, organized in the usual way. In (29a), pre-Velar Softening //k, g// are matched with stops; in (29b), post-Velar Softening //k, g//, ie [s, °s], are matched with fricatives.

(29) a. Match better pre-Velar Softening:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>accident</td>
</tr>
<tr>
<td>fc</td>
<td>particle</td>
</tr>
<tr>
<td>fc</td>
<td>photogenic</td>
</tr>
<tr>
<td>cf</td>
<td>problem</td>
</tr>
<tr>
<td>fc</td>
<td>appetite</td>
</tr>
<tr>
<td>fc</td>
<td>participle</td>
</tr>
<tr>
<td>fc</td>
<td>photographic</td>
</tr>
<tr>
<td>cf</td>
<td>process</td>
</tr>
</tbody>
</table>

b. Match better post-Velar Softening:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>analogy</td>
</tr>
<tr>
<td>df</td>
<td>deciding</td>
</tr>
<tr>
<td>df</td>
<td>expect</td>
</tr>
<tr>
<td>df</td>
<td>expected</td>
</tr>
<tr>
<td>fc</td>
<td>subjective</td>
</tr>
<tr>
<td>df</td>
<td>analysis</td>
</tr>
<tr>
<td>df</td>
<td>dividing</td>
</tr>
<tr>
<td>df</td>
<td>accept</td>
</tr>
<tr>
<td>df</td>
<td>suggested</td>
</tr>
<tr>
<td>fc</td>
<td>suggestive</td>
</tr>
</tbody>
</table>

3.3.3.4 Spirantization

I only have two possible examples of relevance to Spirantization when it applies independently of Palatalization. As seen below, in both cases matching is better when surface forms are
considered (ie, fricatives match with fricatives, whether underlying or derived by Spirantization).

(30) Match better post-Spirantization:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>deliverable divisible</td>
</tr>
<tr>
<td>df</td>
<td>leprosy epilepsy [cf epileptic]</td>
</tr>
</tbody>
</table>

There are several other examples of relevant malapropisms that involve both Spirantization and Palatalization. Even in this group, the number of pairs matching better after the application of Spirantization is greater than the number of pairs matching better before application of Spirantization. That is, there are more pairs where a surface fricative or affricate that is not derived is matched with a surface fricative that is ostensibly derived by Spirantization, as in (31b), than there are pairs where a surface /t/ is matched with a surface fricative ostensibly derived from /t/, as in (31a).

(31) a. Match better pre-Spirantization:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>transistor transition</td>
</tr>
<tr>
<td>df</td>
<td>vivisectionist vasectomy</td>
</tr>
</tbody>
</table>
b. Match better post-Spirantization:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>conclusion</td>
</tr>
<tr>
<td>fc</td>
<td>confession</td>
</tr>
<tr>
<td>fc</td>
<td>convergence</td>
</tr>
<tr>
<td>fc</td>
<td>divergence</td>
</tr>
<tr>
<td>jm</td>
<td>explosion</td>
</tr>
<tr>
<td>v73</td>
<td>promotion</td>
</tr>
<tr>
<td>fc</td>
<td>recession</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>accurate</td>
<td>adequate</td>
</tr>
<tr>
<td>article</td>
<td>argument</td>
</tr>
<tr>
<td>emanate</td>
<td>emulate</td>
</tr>
<tr>
<td>emulate</td>
<td>imitate</td>
</tr>
<tr>
<td>flute</td>
<td>fugue</td>
</tr>
<tr>
<td>genuine</td>
<td>general</td>
</tr>
<tr>
<td>juniors</td>
<td>journals</td>
</tr>
<tr>
<td>magicians</td>
<td>musicians</td>
</tr>
<tr>
<td>maid</td>
<td>music</td>
</tr>
<tr>
<td>message</td>
<td>measure</td>
</tr>
<tr>
<td>module</td>
<td>model</td>
</tr>
<tr>
<td>movie</td>
<td>music</td>
</tr>
<tr>
<td>museums</td>
<td>machines</td>
</tr>
</tbody>
</table>

Both of these lists therefore suggest that Spirantization is not very productive, as its effects are usually prepatterned.

3.3.3.5 y-Insertion

At first sight, the evidence I have collected strongly supports Fay and Cutler's (1977) suggestion that forms may match better when considered prior to y-Insertion. All of the examples I have found match better when the /y/ is removed, as seen below.

(32) Evidence from malapropisms for y-Insertion:
Error | Target
---|---
fc ovulations | observations
fc powerful | popular
fc review | revise
df situation | sensation
df situation | solution
df systems | situations
df using | losing
v73 vowels | values

This evidence is clearly in conflict with the evidence from speech error accommodations, however, which imply that y-Insertion is only somewhat productive. That the evidence from malapropisms may be suspect can be seen by considering the following question: What would an example look like that didn’t seem to support the on-line application of y-Insertion? That is, given that by my criteria I only looked at pairs which differed in the relevant way in their surface and more abstract representations, we cannot compare word pairs where in one word there is a prevocalic /y/ inserted by y-Insertion, while in the other there is an underlying prevocalic /y/; all cases of onset clusters with /y/ involve y-Insertion. Because of this constraint on the appearance of prevocalic /y/ in real English words, the best we could hope for is a pair of words where one has a prevocalic /y/ and the other a different glide or consonant in the matched location. Such a case may then be interpreted as showing a better match in the surface representations, where the one word has [y].

As a matter of fact, there are examples of this situation in (32). That is, there are a few cases where [yuw] is matched with [Cuw],
where C represents a consonant that is more sonorant than an obstruent stop. At first sight such cases may seem best characterized as surface matches, since [y] shares more features with such a C than with with the absence of a segment. Unfortunately, in all such cases the C in question is a coronal, after which /y/ cannot surface anyway. The relevant examples are repeated below.

(33)  | Error       | Target       |
      | flute       | fugue        |
      | ovulations  | observations |
      | situation   | sensation    |
      | using       | losing       |

Thus the sort of word pair we would need to show better matching on the surface with y-Insertion must be a case where a prevocalic [y] in one word is matched in the other word with a noncoronal consonant. In order to meet this requirement, though, either the [y] must be word-initial, being matched with a word-initial noncoronal, or the word not containing [y] must contain a consonant cluster that does not end in a coronal. Since virtually all onset consonant clusters in English end in a coronal, both restrictions severely limit the number of appropriate word pairs, and it is not surprising that none have arisen in my relatively small survey of malapropisms.

The evidence from malapropisms regarding the productivity of y-Insertion thus appears to be consistent with that from speech error accommodations: there is no need to consider y-Insertion more than a somewhat productive rule.
3.3.3.6 Palatalization

The next rule I will consider is Palatalization. Although there are not a large number of relevant examples, the evidence available from malapropisms is clearly consistent with the evidence from other sources concerning its relatively great productivity. As seen in (34), there are more examples where forms match better prior to the application of Palatalization than examples where forms match better after the application of Palatalization.

(34) a. Matches better pre-Palatalization

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>df</td>
<td>message</td>
</tr>
<tr>
<td>df</td>
<td>module</td>
</tr>
<tr>
<td>jm</td>
<td>soldier</td>
</tr>
<tr>
<td>df</td>
<td>systems</td>
</tr>
<tr>
<td>fc</td>
<td>transistor</td>
</tr>
<tr>
<td>df</td>
<td>vivisectionist</td>
</tr>
</tbody>
</table>

b. Matches better post-Palatalization

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>conversion</td>
</tr>
<tr>
<td>fc</td>
<td>divergence</td>
</tr>
</tbody>
</table>

However, examples with Palatalization illustrate a different problem: they often match better when one considers the surface forms because the surface forms are derived through the suffixation of different allomorphs of -ion. Examples of this phenomenon are illustrated below. Such examples imply that not even -ion suffixation is necessarily done on-line, which means that
phonological rules that it triggers, such as Palatalization, aren't necessarily done on-line either. Again recall, though, that evidence for prepatterning is not necessarily evidence against productivity.

(35) Error | Target | E -- stem | T -- stem
--- | --- | --- | ---
fc determination | denomination | determine | denominate
fc inclination | intimation | incline | intimate
df information | imitation | inform | imitate
df observation | operation | observe | operate
fc operations | occupations | operate | occupy
fc provocation | indication | provoke | indicate

Although the examples in (34) may possibly indicate that Palatalization is rather productive phonologically, the examples in (35) show that at least some of the morphological processes triggering it are not. We are forced to conclude, therefore, that malapropisms provide at best equivocal evidence for the robust productivity of Palatalization.

3.3.3.7 Nasal Assimilation and Postnasal g-Deletion

The last pair of rules I will discuss are Nasal Assimilation and Postnasal g-Deletion. The evidence from malapropisms suggests, unsurprisingly, that Nasal Assimilation is rather productive. More examples show matches between nasals that disagree in place on the surface (36a) than matches between nasals whose place specification is derived and other segments that have this same place specification underlyingly (36b).
(36) a. Matches better pre-Nasal Assimilation:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>blankly</td>
</tr>
<tr>
<td>df</td>
<td>blintz</td>
</tr>
<tr>
<td>fc</td>
<td>complete</td>
</tr>
<tr>
<td>fc</td>
<td>confidence</td>
</tr>
<tr>
<td>v73</td>
<td>emphasis</td>
</tr>
<tr>
<td>fc</td>
<td>finger</td>
</tr>
<tr>
<td>fc</td>
<td>improve</td>
</tr>
<tr>
<td>fc</td>
<td>punch</td>
</tr>
<tr>
<td>df</td>
<td></td>
</tr>
</tbody>
</table>

b. Matches better post-Nasal Assimilation:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>corrected</td>
</tr>
<tr>
<td>fc</td>
<td>count</td>
</tr>
<tr>
<td>df</td>
<td>simple</td>
</tr>
<tr>
<td>df</td>
<td></td>
</tr>
</tbody>
</table>

Also unsurprisingly, the evidence for the productivity of Postnasal g-Deletion is equivocal at best. I have found only three relevant examples. In one of them, \([\eta]\) is matched with a consonant cluster (37a), while in two others, it is matched with a single consonant (37b). Hence the analysis of \([\eta]\) as being underlyingly the consonant cluster /Ng/ is not strongly supported by these data.

---

26These examples assume that a surface \([n]\) before a coronal consonant is underlyingly placeless, while other coronals such as /t/ are marked for place underlyingly; thus *constructed* matches *corrected* better on the surface, where both are marked for place. If, however, all coronals are underlyingly placeless, as has been argued in works such as Paradis and Prunet (1991), three of these examples are not relevant.
(37) a. Matches better pre-g-Deletion:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc</td>
<td>song</td>
</tr>
<tr>
<td>souM</td>
<td>sound</td>
</tr>
</tbody>
</table>

b. Matches better post-g-Deletion:

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>v73</td>
<td>hanger</td>
</tr>
<tr>
<td>df</td>
<td>reading</td>
</tr>
<tr>
<td>hammer</td>
<td>ringing</td>
</tr>
</tbody>
</table>

I conclude, therefore, that while there is evidence from malapropisms that Nasal Assimilation is productive, there is no evidence that g-Deletion is.

3.3.3.8 Evidence from malapropisms: Summary

As noted above, the evidence from malapropisms concerning the productivity of rules in English is only suggestive. The following is what we appear able to say at this point. First, there is no evidence for the productivity of Vowel Shift, Prenasal g-Deletion, s-Voicing, Velar Softening, Spirantization, or Postnasal g-Deletion. Palatalization appears to be at least somewhat productive, as does Nasal Assimilation. According to the malapropism evidence, y-Insertion appears to be very productive, but this may just be a side effect of the methodology.

3.4 Chapter summary

I summarize the various sorts of evidence regarding the productivity of the rules discussed in Chapter 2 in the chart below. The columns represent four different sources of evidence. The numbers in the columns indicate the approximate ranking according
to productivity that seems to emerge from each source of evidence
("Lx" = "Linguistic evidence", "Exp" = "Experimental evidence",
"Accomm" = "Evidence from speech error accommodations", "Malap" =
"Evidence from malapropisms"; "1" = "very unproductive", "4" = "very
productive"). If a source provides no evidence for a rule's
productivity, the appropriate cell is left blank.

(38) Lx Exp Accomm Malap

<table>
<thead>
<tr>
<th></th>
<th>Lx</th>
<th>Exp</th>
<th>Accomm</th>
<th>Malap</th>
</tr>
</thead>
<tbody>
<tr>
<td>s-Voicing</td>
<td>1</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Vowel Shift</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Velar Softening</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Spirantization</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>y-Insertion</td>
<td></td>
<td>2/3</td>
<td>2/3</td>
<td></td>
</tr>
<tr>
<td>Palatalization</td>
<td>4</td>
<td>4</td>
<td>2/3</td>
<td>2/3</td>
</tr>
<tr>
<td>Nasal Assimilation</td>
<td>4</td>
<td>3/4</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>g-Deletion</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The safest conclusion that can be drawn from this table is that
virtually no conclusion can be safely drawn from this table. Not only
is the evidence from each source rather unsatisfactory for the
reasons discussed throughout the chapter, but there is no way to
know how to relate the evidence from different sources together.

Nevertheless, there does seem to be a general trend consistent
with the Productivity Hypothesis: more productive rules are ordered
later. In particular, s-Voicing appears to be less productive than
Spirantization/Palatalization and Velar Softening appears to be less

---

27This line is left blank because there is no evidence indicating that g-
Deletion is productive, contrary to what works such as Fromkin (1971) have
claimed.
productive than Palatalization, exactly as predicted by the Productivity Hypothesis.

At first sight there appear to be some problems for the Productivity Hypothesis as well. First, while most sources of evidence show both $\mu$-Insertion and Palatalization to be rather productive, the one source where both rules may be compared, namely malapropisms, seems to imply that $\mu$-Insertion is more productive than Palatalization. As noted above, however, this is likely due more to the use of this methodology than to the actual relative productivity of the two rules. Second, if Postnasal $g$-Deletion is a real rule at all, the evidence suggests that its productivity is hardly overwhelming. However, it is preceded by Nasal Assimilation, which is clearly rather productive. There are two possible explanations for this apparent dilemma for the Productivity Hypothesis. One is to suppose that Postnasal $g$-Deletion is not a rule at all. Arguments for supposing this come both from linguistics (Chapter 2) as well as from the psycholinguistic evidence discussed in this chapter. Another way around the problem is to exploit the fact that there are actually two rules of Nasal Assimilation, an obligatory one and and optional one. Since the optional one may apply across word boundaries, it is more productive than the obligatory one (phrases being novel forms). It may be, then, that Obligatory Nasal Assimilation truly is less productive than some rule of Postnasal $g$-Deletion, but that what is seen in the psycholinguistic
evidence discussed above is not this rule, but the much more productive rule of Optional Nasal Assimilation.

Thus in this chapter we have seen that although the evidence for the relative productivity of rules is not yet conclusive, what is known is fully consistent with the Productivity Hypothesis. That is, I have found no unambiguous case in English of a pair of rules A and B, where psycholinguistic evidence suggests that A is more productive than B even though linguistic evidence shows that A must be ordered before B. There is as yet no evidence to reject the claim that Double Lookup can account for both competence and performance data.
CHAPTER 4
FULLY PRODUCTIVE RULES IN ENGLISH

4.0 Introduction

If rules may be arbitrarily ordered only if they differ in productivity, as the Productivity Hypothesis claims, then we predict that two rules that are both fully productive, and therefore also equally productive, cannot be arbitrarily ordered. Consequently any case that we might analyze as involving two apparently fully productive rules that are apparently arbitrarily ordered must be subject to another, better analysis. Specifically, it must either be the case that the rules are really arbitrarily ordered, but the earlier of the rules is actually less productive than the other, or that the rules really are fully productive, but they are not arbitrarily ordered.

In this chapter I test this prediction against a variety of rules in English that have been claimed to be both ordered and fully productive. The primary reason for focusing exclusively on English was that this was the language for which the relevant psycholinguistic evidence was most readily available to me. In addition, the only way to show that the predictions of the Productivity Hypothesis hold is to thoroughly examine all apparent counterexamples. I have chosen to limit this search by attempting to do a thorough survey of the rules of one language.

We will see that such apparent counterexamples divide themselves into a small number of distinct types. Although space does not permit a complete discussion of all claimed examples of

---

1This chapter overlaps in some respects with, and supersedes, J. Myers (1992).
ordered fully productive rules mentioned in the literature, I believe that my typology of examples is exhaustive, so that an examination of a few examples in each type should suffice to make my case. This is not to deny the possibility that the crucial counterexamples may still remain to be found. The most I can say now is that I have attempted to falsify the predictions that the Productivity Hypothesis makes for fully productive rules, but I have not yet succeeded.

What I have found is the following. First, in 4.1 I show that the interaction of two truly fully productive segmental rules, specifically two rules both of which apply across word boundaries, is always in accordance with Automatic Feeding. That is, such rules always interact as if they apply and reapply until they no longer can, resulting in apparent feeding and counterbleeding orders.

Second, in 4.2 I show that the ordering of a fully productive prosodic rule and a fully productive segmental rule is always in accordance with Prosody First. That is, on-line prosodic rules always apply before on-line segmental rules, allowing in some cases for bleeding order (i.e., application of the prosodic rule blocks the application of the segmental rule).

Third, as I show in 4.3, in ordered rule pairs where one of the rules only applies word-internally and the other may also apply across word boundaries, the word-internal rule is always ordered first. This is true even if the earlier rule's inability to apply across word boundaries is not stipulated as part of the rule's description but follows from independent constraints. I argue that this ordering
follows from the Productivity Hypothesis because of the way rules are learned as fully productive. Since phrases are by far the most common nonanomalous novel forms that occur in natural speech, rules that apply fully productively at the phrase level will be learned as fully productive, while rules that apply solely within words (i.e., non-novel forms) need not be learned as fully productive.

Finally, in 4.4 I discuss the claim in Donegan and Stampe (1979) that on-line rules are ordered by the universal principle that fortitions precede lenitions. I argue that this ordering principle is derivable from the claim that fortitions are universally less productive than lenitions.

I summarize these findings in 4.5.

4.1 Automatic Feeding

In Chapter 1 I pointed out that since the Productivity Hypothesis only predicts the ordering of rules differing in productivity, it was not clear what would happen with rules that were equal in productivity, in particular two fully productive rules. I therefore added a new principle, the principle of Automatic Feeding, which claims that on-line segmental rules apply and reapply until they no longer can. This will have the effect of ordering on-line segmental rules in feeding and counterbleeding relationships. In this section I show that given the information currently available, this principle seems accurate. That is, in all cases of pairs of interacting fully productive segmental rules in English known to me, all involve feeding and counterbleeding ordering. I give two examples in this
section. The question of the ordering of fully productive prosodic rules is considered in the next section.

4.1.1 Palatalization and y-Deletion

The word-internal rules of Palatalization and y-Deletion, discussed in Chapter 2, have word-external parallels as well. Significantly for any theory of on-line rule interaction, these rules both apply across word boundaries, and are therefore both fully (and equally) productive, and yet their interaction shows that they cannot be applied simultaneously. Rather, they apply in accordance with Automatic Feeding.

Word-external palatalization causes word-final \([t, d, s, z]\) to surface as \([\text{t}, \text{z}, \text{f}, \text{3}]\), respectively, when the following word begins with \([y]\). Examples are given below.

(1) [after Kaisse 1985:35]

\[
\begin{align*}
\text{could you} & \rightarrow \text{cou[t]} \text{ you} \\
\text{can't you} & \rightarrow \text{can[t]} \text{ you} \\
\text{as you} & \rightarrow \text{a[z]} \text{ you} \\
\text{unless you} & \rightarrow \text{unle[f]} \text{ you}
\end{align*}
\]

Parallel to the formulation of the word-internal rule, we may formalize word-external Palatalization as follows.
As with the word-internal palatalization process discussed in Chapter 2, this one too is associated with the deletion of the triggering [y]. As seen in (3), this process is optional.

(3) could you → [kʊʤyuw] or [kʊʤuw]
can't you → [kæŋʃyuw] or [kæŋʃuw]
as you → [æʒyuw] or [æʒuw]
unless you → [æŋʃyuw] or [æŋʃuw]

Moreover, following Rotenberg (1978), Kaisse (1985) claims that this y-deletion process is independent of palatalization, since, as seen in the examples in (4), [y] may be deleted after non-derived palatals as well.

(4) [after Kaisse 1985:35]

misjudge you → [mɪʃʤuw]
touch you → [tʃuw]
camouflage you → [kæməflæʒuw]
push you → [pʃuw]

Thus she concludes that the variable appearance of [y] in examples like those in (3) is due to the optional application of a rule of y-Deletion, which deletes [y] after palatal segments. This rule may be given as in (5).
Given this analysis, these two rules must apply in the order given in (6a), where Palatalization precedes y-Deletion. If the order were reversed, as in (6b), [y] would be deleted before it could palatalize anything. Note that Palatalization and y-Deletion cannot be applied simultaneously. If they did so, as in (6c), [y] would not be able to delete at all in examples like those in (3).

(6) The ordering of Palatalization and y-Deletion

a. unless you
   UR          Anles yuw
   Palatalization  Anlpsj yuw
   y-Deletion    Anlpsj uw

b. unless you
   UR          Anles yuw
   y-Deletion  {not applicable}
   Palatalization  Anlpsj uw

c. unless you
   UR          Anles yuw
   Palatalization+y-Deletion  Anlpsj yuw
   {y-Deletion not applicable}
As noted above, this ordering is not consistent with a theory that requires on-line rules to apply simultaneously. However, the ordering is fully consistent with Automatic Feeding, since it is a counterbleeding order: the rules apply in such a fashion so that neither bleeds the other. The rules may thus be thought of as applying repeatedly until neither can apply, as illustrated in (7).

(7)  

<table>
<thead>
<tr>
<th></th>
<th>unless you</th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
<td>(\lambda n\in \epsilon yw)</td>
</tr>
<tr>
<td>Pal + y-Del</td>
<td>(\lambda n\in \epsilon yw) {only Palatalization is applicable}</td>
</tr>
<tr>
<td>Pal + y-Del</td>
<td>(\lambda n\in \epsilon uw) {only y-Deletion is applicable}</td>
</tr>
<tr>
<td>Pal + y-Del</td>
<td>---        {neither is applicable}</td>
</tr>
</tbody>
</table>

It may be objected that this case is not truly conclusive support for the claim that on-line rules do not apply simultaneously. It is possible, for example, that the word-external processes of Palatalization and y-Deletion are in fact two aspects of a single rule. In any event, something is not quite right with Kaisse's (1985) analysis, since in some dialects (such as my own), \([y]\) is easily deleted after palatals derived by Palatalization, such as in the examples in (3), but not after underlying palatals, such as in the examples in (4).

The case I discuss next, however, does provide clear evidence for the necessity of some kind of ordering with on-line rules. Again we will see, though, that this ordering is precisely that predicted by Automatic Feeding.
4.1.2 Coronal Deletion and two assimilation rules

Another example of fully productive yet ordered rules is the ordering of what I will call *Coronal Deletion* before the rules of Palatalization and Nasal Assimilation (Selkirk 1972, Guy 1991a,b). As with the ordering of Palatalization before y-Deletion, these orderings will be seen to follow from the Automatic Feeding principle.

Roughly speaking, Coronal Deletion deletes word-final /t, d/, especially when preceded and/or followed by a segment of low sonority such as an obstruent or a nasal; examples are given below. I will not formalize Coronal Deletion now, but save this for the fuller discussion of this rule in Chapter 5.

(8) [after Selkirk 1972]

a. draft-dodger → draf-dodger
   fact-finder → fac-finder
   thefts → thefs
   attempts → attemps
   conflicts → conflics
   last month → las month
   exact sciences → exac sciences

b. band together → ban together
   lend money → len money
   send packages → sen packages
   a sound value → a soun value

This rule feeds the two assimilation rules of Palatalization and Nasal Assimilation, both of which we have discussed before.
Palatalization is the rule formalized above in (2); Nasal Assimilation is the rule referred to as Optional Nasal Assimilation in Chapter 2.

Above I showed how Palatalization could be triggered by [y]. However, nothing in the formalization of the rule in (2) prevents the palatals [ʃ] and [ʒ] from triggering it as well. As the examples in (9) show, Palatalization also assimilates word-final coronals in place to following palatal obstruents, so that in particular /s/ and /z/ become [ʃ] and [ʒ], respectively, before [ʃ] or [ʒ].

(9) [after Selkirk 1972]

this show → thi[ʃʃ]ow
is Sheila here? → i[ʒʃ]eila here?
I gave Chris show tickets → ...Chris[ʃʃ]ow...
Chris shied away → Chr[iʃʃ]ied...
Buzz shrieked → Bu[ʒʃ]rieded

Selkirk (1972) notes that Coronal Deletion must be ordered before both Palatalization and Nasal Assimilation because fricatives and nasals that become word-final through Coronal Deletion can then undergo Palatalization or Nasal Assimilation if the appropriate segment follows in the next word. Examples of the ordering of Coronal Deletion before Palatalization are given below in (10a); examples of the ordering of Coronal Deletion before Nasal Assimilation are given in (10b). Derivations are given in (11).

(10) a. last shot → [læʃʃat]

b. lend beer → [lēm biyr]
send comics → [sɛŋ kəmɪks]
Note that as with Palatalization and y-Deletion, these rules too cannot be assumed to apply simultaneously. If they were to do so, as illustrated in the derivations in (12), the relevant pronunciations would not be derivable.

(12) a. last shot
  UR: læst ñat
  Cor Del + Pal læs ñat {Palatalization not applicable}

b. lend beer
  UR: lend biyr
  Cor Del + Nas Assim len biyr {Nasal Assimilation not applicable}

If we could show that Coronal Deletion is less productive than Palatalization and Nasal Assimilation, it might be argued that this ordering follows from the Productivity Hypothesis. Apparent support for this hypothesis comes from the work of Guy (1991a,b), discussed more fully in Chapter 5. He provides evidence that Coronal Assimilation is sensitive to word-internal morphology in a very specific way. The upshot of this is that although the deletion of word-final /t, d/ depends on information beyond the word
boundary, word-internal information appears to play a large role as well. This could mean that Coronal Deletion is less productive than Palatalization and Nasal Assimilation in the same way s-Voicing is less productive than Velar Softening and word-internal Palatalization.\footnote{This is in fact the gist of the argument I presented in J. Myers (1992).}

Unfortunately, this move will not work. It is easily shown that even when interacting with clearly word-external rules like Palatalization, the application of Coronal Deletion is still triggered by word-external factors in rather specific ways, meaning that it too must apply on-line. The ordering of Coronal Deletion before Palatalization cannot be ascribed to a difference in productivity.

As mentioned above, Coronal Deletion applies more readily the lower on the sonority hierarchy the following segment is. This pattern is illustrated below.

\[(13)\]
\[
\begin{align*}
\text{a.} & \quad \text{exact shortage} \rightarrow \text{exac shortage} \\
\text{b.} & \quad \text{at that exact moment} \rightarrow \text{exac moment} \\
\text{c.} & \quad \text{for that exact reason} \rightarrow \text{?exac reason} \\
\text{d.} & \quad \text{exact year} \rightarrow \text{?exac year} \\
\text{d.} & \quad \text{exact answer} \rightarrow \text{??exac answer}
\end{align*}
\]

Now consider the contrast in (14). Note that when /t/ is followed by a palatal obstruent (low sonority), it is easily deleted, thus allowing Palatalization to affect the previously hidden /s/. However, if /t/ is followed by a palatal glide (high sonority), it is not easily deleted, forcing Palatalization to affect the /t/ and not the
hidden /sl/. Derivations illustrating this phenomenon are given below in (15).

(14) a. last shot → [læʃʃat] or [læʃʃʒat]
bust Jacque → [bʌʃʒak] or [bʌʃʃʒak]

b. last year → [læʃʃjɨyr] { cf [læʃʃjɨyr] }
bust you → [bʌʃʃjɯw] { cf [bʌʃʃjɯw] }

(15) last shot

<table>
<thead>
<tr>
<th>rule</th>
<th>last shot</th>
<th>last year</th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
<td>/læʃʃat/</td>
<td>/læʃʃjɨyr/</td>
</tr>
<tr>
<td>Coronal Deletion</td>
<td>læʃʃat</td>
<td>læʃʃjɨyr</td>
</tr>
<tr>
<td>Palatalization</td>
<td>læʃʃat</td>
<td>læʃʃjɨyr</td>
</tr>
</tbody>
</table>

We cannot ascribe the ordering of Coronal Deletion before Palatalization to a hypothesized lesser productivity because one of the major factors influencing the application of Coronal Deletion in these cases must be determined on-line, coming as it does from across a word boundary. Hence it appears that both Coronal Deletion and Palatalization are applying on-line in the cases discussed above; although there is no direct evidence for it, we have no reason to suppose that this is not also true of Coronal Deletion and Nasal Assimilation.

However, the ordering is in complete accord with the principle of Automatic Feeding. This is because the orderings of Coronal Deletion before Palatalization and Nasal Assimilation are both feeding orders: application of Coronal Deletion allows the other two rules to apply. In other words, we can assume that what happens in these cases is that these fully productive rules apply and reapply until
they no longer can. A derivation illustrating this concept is given below.

(16) | last shot  |
<table>
<thead>
<tr>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>UR</td>
</tr>
<tr>
<td>Cor Del + Pal</td>
</tr>
<tr>
<td>Cor Del + Pal</td>
</tr>
<tr>
<td>Cor Del + Pal</td>
</tr>
</tbody>
</table>

These rules thus provide evidence consistent with the claims made in Chapter 1, in particular with the principle of Automatic Feeding. Specifically, these fully productive segmental rules interact in such a way as to allow all rules a chance to apply.

4.1.3 Automatic Feeding: Summary

In this section I have discussed all cases of interacting unambiguously fully productive segmental rules in English that I am aware of. Certainly further possible counterexamples must be investigated, but at this point it is clear that all orderings seen so far are consistent with the Automatic Feeding principle, as expected.

4.2 Prosody First

In Chapter 1 I argued briefly that due to their differences in function, prosodic rules must always be ordered before segmental rules. More precisely, segmental rules must always apply within a prosodic framework. This is the principle I called Prosody First. As noted in Chapter 1, a principle like this is adopted for apparently independent reasons by researchers as diverse as Donegan and Stampe (1979) and Kaisse (1985), both of whom claim that on-line
segmental rules apply within frameworks laid down by on-line prosodic rules.\(^3\)

We have already seen two sorts of evidence for this principle, linguistic and psycholinguistic. In Chapter 2 we saw linguistic evidence that many partially productive segmental rules make reference to prosodic structure. Interestingly, in many of these cases the prosodic structure must first be added by rule. Thus the application of the Vowel Shift rules presupposes the prior application of rules of syllabification, in particular Resyllabification. Similarly, Palatalization and s-Voicing also seem to make reference to prosodic structures that were built by Resyllabification.

Naturally, although rules like Resyllabification appear to be very productive (the Onset Principle, in fact, as we saw in Chapter 2, applies across word boundaries), the versions of these rules that precede partially productive rules like Vowel Shift cannot apply on-line either. In other words, these prosodic rules are prepatterned.\(^4\) In Chapter 3 we saw psycholinguistic support for this claim: malapropisms almost always match in prosodic structure, implying that at least some of it is stored along with forms in memory.

\(^3\)More precisely, Kaisse (1985) argues that postlexical prosodic rules are universally ordered after what she calls P1 postlexical rules and before what she calls P2 postlexical rules (see discussion in Chapter 6). Given Hayes's (1990) reanalysis of her model with P1 rules as lexical rules, and Mohanan's (1982) and Shattuck-Hufnagel's (1986) suggestion that lexical rules do not apply on-line while postlexical rules do, Kaisse's (1985) claim becomes the claim that on-line prosodic rules are universally ordered before on-line segmental rules.

\(^4\)In Chapter 5 we will see that prepatterned very productive rules always end up appearing to apply cyclically, that is, to apply more than once in the derivation.
Within the framework of Double Lookup, then, the ordering of partially productive prosodic rules before partially productive segmental rules cannot have a synchronic explanation. This is because partially productive rules do not typically apply on-line, and synchronic ordering is caused by the ordering of Lexical Lookup (retrieving prepatterned forms) before Rule Lookup (applying productive rules). Instead, it must be that at some time in the past, when Vowel Shift was fully productive, it was ordered after prosodic rules like Resyllabification according to some principle.

This principle is Prosody First. The necessity of this principle can be seen most clearly when we turn to the application of fully productive prosodic and segmental rules. Given the principle of Prosody First, even on-line rules must be ordered so that prosodic structure is laid down before segmental rules can apply within it.

In this section we will consider examples that must be analyzed with this principle. That is, all of the rules involved appear to be equally fully productive, applying freely across word boundaries; the ordering can therefore not follow from differences in productivity. And yet they also may be ordered contrary to the pattern of interaction allowed by Automatic Feeding. In particular, the earlier rule may bleed the later one, so that its application prevents the application of the later rule. This is not possible if the rules are simply applying and reapplying until they no longer can. Instead, the rules must actually be ordered.
In all the cases of this known to me, however, it is always the prosodic rule that applies first. I consider two cases here. In 4.3.1 I examine the ordering of Resyllabification and the Onset Principle before Flapping. This ordering is a feeding order, and so may also be analyzed as being derived through Automatic Feeding. In 4.3.2, however, I consider the ordering of the Onset Principle before Coronal Deletion. In this case, the ordering is bleeding: the application of the Onset Principle prevents Coronal Deletion from applying. Such an ordering relation cannot arise through repeated free applications of the rules. It appears, then, that something like the principle of Prosody First is right: prosodic rules precede segmental rules.

4.2.1 Resyllabification and Flapping

In this section I consider the ordering of both of the resyllabification rules discussed in Chapter 2 with Flapping. I argue that while the ordering of stress-sensitive Resyllabification before Flapping could conceivably be ascribed to the Productivity Hypothesis the same way as the cases to be discussed below in 4.3, the ordering of the Onset Principle before Flapping should be ascribed to Prosody First.

Flapping is the paradigm example of a fully productive segmental rule in English (Kahn 1976, Withgott 1982, Selkirk 1983, Kaisse 1985, Borowsky 1986, among many others). The generalization this rule captures is roughly this: segments which surface as [d] or [t] in some dialects or social registers surface in
others as the voiced alveolar flap [ɾ] if preceded by a vocalic element and followed by an unstressed vowel. Examples are given below in (17); Flapping only applies in the cases given in (17a,b) where /t/ is preceded by a vowel and followed by an unstressed vowel.

(17)  

<table>
<thead>
<tr>
<th></th>
<th>Dialects/registers without Flapping</th>
<th>Dialects/registers with Flapping</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>BYTES</td>
<td>bÁTER</td>
</tr>
<tr>
<td>b.</td>
<td>BYTES</td>
<td>sÁNITI</td>
</tr>
<tr>
<td>c.</td>
<td>BYTES</td>
<td>butane</td>
</tr>
<tr>
<td>d.</td>
<td>BYTES</td>
<td>potato</td>
</tr>
<tr>
<td>e.</td>
<td>CTV</td>
<td>fÁKTOR</td>
</tr>
<tr>
<td>f.</td>
<td>[TV</td>
<td>tomorrow</td>
</tr>
<tr>
<td>g.</td>
<td>BYTES</td>
<td>butler</td>
</tr>
<tr>
<td>h.</td>
<td>BYTES</td>
<td>cat</td>
</tr>
</tbody>
</table>

The way that stress plays a role in the application of Flapping has been analyzed in two ways, both requiring the prior application of a prosodic rule. Kahn (1976), Hammond (1982), Selkirk (1983),
Kaisse (1985) and Borowsky (1986) argue that Flapping is fed by the rule of Resyllabification discussed in Chapter 2, which makes the onset of an unstressed syllable into the coda of the preceding syllable. By contrast, Kiparsky (1979) and Withgott (1982) argue that Flapping is bled by the rules assigning foot structure. Since feet in English are left-headed (see, eg, Hayes 1980), stressed syllables will appear adjacent to a foot boundary; onset consonants in such syllables will thus not be able to undergo Flapping.

Whether Flapping is preceded by Resyllabification or by foot building, the fact that stress plays a role means that we must assume that Flapping is preceded by a prosodic rule of some sort; to be consistent with my discussion in Chapter 2, I will assume that Flapping is fed by Resyllabification. I formalize the rule in (18). In the formalization of Flapping I use "[+flap]" as a cover symbol to represent whatever feature(s) distinguish(es) [r] from [t, d]. In (19) I give sample derivations showing the ordering of Resyllabification before Flapping.

(18) Flapping

\[ \sigma \rightarrow \sigma \]

\[ C \rightarrow \text{[-son]} \times \text{Cor} \rightarrow \text{ [+flap]} \]

These authors disagree as to whether this new coda also remains as the onset of the original syllable, thereby becoming ambisyllabic.
Flapping must also be preceded by the other resyllabification process discussed in Chapter 2, namely the Onset Principle, whereby an intervocalic consonant becomes the onset of the syllable with the second vowel as nucleus. This rule applies across word boundaries, as illustrated by the fact that it feeds Flapping in cases like those below. As shown by the examples in (20), the Onset Principle first resyllabifies the word-final coda consonant as the onset of the following word; Flapping, which requires the target to be ambisyllabic, may then apply.6

---

6These facts are of course well-known; see eg Hammond (1982).
Like the rule pairs to be discussed below in 4.4, the ordering of Resyllabification before Flapping can be derived from the Productivity Hypothesis because Flapping is fully productive while Resyllabification is not. The productivity of Flapping is demonstrated by the fact that it applies freely across word boundaries as we have just seen. As I pointed out above, phrases count as novel forms; hence rules that apply freely in the phrasal domain apply in novel forms and are therefore fully productive.

By contrast, as pointed out by Kahn (1976) and Kaisse (1985), the rule of Resyllabification does not apply freely across word boundaries. This is seen by the fact that Flapping is not possible in phrases with stress patterns identical to those in (20), but where the /t/ is the onset of the first syllable of the second word and not the coda of the last syllable of the first word; examples are given in (21). In these cases, Resyllabification does not apply; the word-initial /t/
will not be changed into a coda, thus bleeding the Flapping rule as formulated above.

(21) sea  [sīy]  sea tomography  [sīytəməgrəfi]
row  [rōw]  row tamales  [rōwtəməliyz]

In short, it is possible to argue that the ordering of Resyllabification before Flapping follows from the fact that Resyllabification is restricted to apply within words while Flapping is not. This case thus does not constitute evidence for the necessity of the principle of Prosody First, though of course it is fully consistent with it.

On the other hand, the ordering of the Onset Principle before Flapping cannot be so understood. This is because like Flapping, the Onset Principle also applies across word boundaries. If it did not, there would be no way to account for the application of Flapping in cases like those in (20).

It is true that in this case the ordering relationship is feeding. That is, the application of the Onset Principle allows for the application of Flapping. We might then choose to ascribe this ordering to Automatic Feeding, just as we did with the cases in 4.1. However, I will not take this approach, since it does not appear that the interaction which the Onset Principle shows with segmental rules is in general of a sort consistent with Automatic Feeding. In particular, the next case I will discuss suggests that the Onset Principle is ordered before all segmental on-line rules, even if this means they interact in a bleeding relationship. Consequently I
ascribe the ordering of the Onset Principle before Flapping to Prosody First.

I conclude, then, that the ordering of both processes of resyllabification before Flapping is predictable. Although both orderings could be derived by other principles as well, they are also fully consistent with the principle of Prosody First. In any event there is no need to consider these orderings arbitrary. Since both orderings are ascribable to other principles as well, however, evidence for the necessity of Prosody First will have to come from another example. I present just such an example now.

4.2.2 Resyllabification and Coronal Deletion

Earlier we saw how the ordering of Coronal Deletion before both Nasal Assimilation and Palatalization is consistent with the principle of Automatic Feeding. In this section I show that this is not the case of the ordering of Coronal Deletion after the Onset Principle.

Recall that the application of Coronal Deletion depends on the sonority of the following segment (ie, the first segment of the following word). As Guy (1991a) points out, this has a natural explanation if we suppose that the Onset Principle precedes Coronal Deletion.

The Onset Principle is more likely to resyllabify a segment X as the onset of a syllable beginning with a segment Y if the sequence XY makes a good syllable-initial sequence. Among other things, this means that X must be less sonorous than Y, and in fact, the greater the difference in sonority, the better XY will be as a syllable-initial
sequence. Consequently, the more sonorous \( Y \) is with respect to \( X \), the more likely the Onset Principle will take effect, forcing the resyllabification of \( X \) as an onset.

If Coronal Deletion is thought to apply only to codas, then, the prior application of the Onset Principle will result in the effect of sonority discussed above. I repeat the relevant examples below, showing that Coronal Deletion is less likely to apply the more sonorous is the following segment.

(22) a. exact shortage → exac shortage
    b. at that exact moment → exac moment
    c. for that exact reason → ?exac reason
    d. exact year→ ?exac year
    e. exact answer → ??exac answer

Guy (1991a) shows that this can be understood as being caused not by the sensitivity of Coronal Deletion to sonority, but by the sensitivity of the Onset Principle, which applies first. Sample derivations illustrating this are given below.

(23) \[
\begin{array}{c|c}
\text{best pickle} & \text{best answer} \\
\hline
\sigma & \sigma & \sigma \\
/\diagdown & /\slash & /\backslash \\
\text{best pɪ kɛl} & \text{best ɬɛ n sɛr} \\
\hline
\text{Onset Principle} & \sigma & \sigma & \sigma \\
\text{---} & /\backslash & /\slash & /\backslash \\
\text{Coronal Deletion} & \sigma & \sigma & \sigma \\
/\diagdown & /\slash & /\backslash \\
\text{best pɪ kɛl} & \text{---} \\
\hline
\text{Output} & [bɛspɪkɛl] & [bɛstænsɛr] \\
\end{array}
\]
This ordering is not consistent with the principle of Automatic Feeding. If the Onset Principle and Coronal Deletion applied and reapplied until they no longer could, we would expect that both rules would attempt to apply in all cases. Specifically, we would expect Coronal Deletion to apply as readily in forms like best answer as it does in forms like best pickle. It is true that once the Onset Principle applies, Coronal Deletion cannot, but if both rules are applying freely and repeatedly until neither can apply again, we expect this to be irrelevant: Coronal Deletion should always be able to apply when its conditioning environment is met. This is not true, however. Instead it appears that the Onset Principle applies before Coronal Deletion in all cases. Thus Coronal Deletion is only able to apply when the Onset Principle cannot.

This ordering is, of course, follows from Prosody First. It is fully consistent with a picture where the presence of prosodic structure is necessary for the application of segmental rules. If this structure must be added on-line, as it is in the case of word sequences, then the rules adding this structure must apply before all other on-line rules. This is what we see here. The Onset Principle is ordered before Coronal Deletion because the former builds prosodic structure.

4.2.3 Prosody First: Summary

In this section I have shown the necessity of ordering in on-line rules. In these examples, the ordering cannot be ascribed to Automatic Feeding because in some cases the ordering is not feeding
or counterbleeding, but instead bleeding. In all such cases, however, the earlier rule is a rule that builds prosodic structure. The ordering can thus be assumed to follow from the principle of Prosody First.

4.3 Word-internal rules precede word-external rules

All examples proposed in the literature of interacting (apparently) fully productive rule pairs in English (including Selkirk 1972, Bailey 1973, Stampe 1973, Donegan and Stampe 1979, Kaisse 1985, Borowsky 1986) can be divided into two classes: those in which both rules may apply across word boundaries and those in which at least one of the rules applies solely within words. To my knowledge, the only examples of the former sort discussed in the literature are those dealt with above. As I showed in the previous sections, in all such cases the ordering is principled.

This leaves rule pairs where at least one of the rules applies solely within word boundaries. I am not claiming that the restriction to word-internal application is part of the structural description of these rules. Instead, in all cases I discuss in this section this restriction follows from independently required constraints. Nevertheless, as I will show, a rule that is only able to apply within word boundaries is always ordered before a rule that may also apply across word boundaries. The only examples where the ordering of rules seems arbitrary (e.g., different dialects order the rules in different ways) involve rule pairs where both rules are restricted to apply word-internally.
The reason for this ordering pattern, I claim, is the Productivity Hypothesis. This can be seen by considering the way a rule is learned as fully productive. If a rule only applies word-internally, even if this restriction is independently necessary, there will be no evidence that the rule is necessarily fully productive. This is because word-sized units in natural speech are almost always real words; natural phenomena such as speech errors that might demonstrate the productivity of a word-internal rule are relatively rare. Even with speech errors, the learner must know what the speaker intended to say so that she can compare the error and target forms to determine what rules have been triggered. By contrast, rules that apply across word boundaries are transparently productive.

Consequently, word-internal rules, regardless of their productivity in the adult language, are more likely to be learned as less productive than word-external rules. In other words, word-internal patterns are initially not learned as productive rules at all, but rather are learned as prepatterned.

I will discuss three representative cases showing the ordering of word-internal rules before word-external rules.

4.3.1 Raising and Flapping in Canadian English

By far the best known apparent example of an ordered fully productive rule pair is the ordering of Flapping with respect to vowel raising in Canadian English and other North American dialects (Joos 1942, Halle 1962, Chambers 1973, Vance 1987). This example seems to show not only that two fully productive rules can be ordered, but
that the ordering of these rules is arbitrary, varying from dialect to
dialect. I will demonstrate that neither of these claims goes through. Instead, we will see that all dialects with both Flapping and vowel raising order vowel raising first, because vowel raising is restricted to word-internal application, which leads to its lesser productivity. In some cases the lesser productivity of vowel raising will be seen directly by the existence of lexical exceptions.

The rule of Flapping appears to be exactly the same in Canada as in other parts of North America (J. Chambers, pc). The rule of Raising, however, is somewhat more localized, occurring in various forms in many North American dialects, but not all. In Canada the rule can be understood as raising /a/ in diphthongs followed by a voiceless consonant. As Chambers (1973) notes, the specific phonetic values of the raised version of the vowel depends on whether the diphthong contains [y] or [w]; I will follow him in representing both as [ʌ]. Examples illustrating the independence of this rule from Flapping are shown below.

(24) Raising

<table>
<thead>
<tr>
<th>a.</th>
<th>[after (1) in Chambers 1973:115]</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>[tʌyp]</td>
</tr>
<tr>
<td>tight</td>
<td>[tʌyt]</td>
</tr>
<tr>
<td>tyke</td>
<td>[tʌyk]</td>
</tr>
<tr>
<td>rife</td>
<td>[rʌyf]</td>
</tr>
<tr>
<td>rice</td>
<td>[rʌys]</td>
</tr>
<tr>
<td></td>
<td>[tɔwʌt]</td>
</tr>
<tr>
<td></td>
<td>[SAwθ]</td>
</tr>
<tr>
<td></td>
<td>[mʌwʌs]</td>
</tr>
<tr>
<td></td>
<td>[kʌwɔf]</td>
</tr>
</tbody>
</table>
b. [after (2) in Chambers 1973:116]

<table>
<thead>
<tr>
<th>Word</th>
<th>Unraised</th>
<th>Raised</th>
</tr>
</thead>
<tbody>
<tr>
<td>house</td>
<td>[hAWS]</td>
<td>[hawzəz]</td>
</tr>
<tr>
<td>mouth</td>
<td>[mAθ]</td>
<td>[mawðz]</td>
</tr>
<tr>
<td>spouse</td>
<td>[spAWS]</td>
<td>[espawz]</td>
</tr>
<tr>
<td>knife</td>
<td>[nAYf]</td>
<td>[nayvz]</td>
</tr>
<tr>
<td>life</td>
<td>[lAYf]</td>
<td>[layvz]</td>
</tr>
<tr>
<td>wife</td>
<td>[wayf]</td>
<td>[wayvz]</td>
</tr>
<tr>
<td>advice</td>
<td>[ædvAYs]</td>
<td>[ædvayz]</td>
</tr>
<tr>
<td>device</td>
<td>[dəvAYs]</td>
<td>[dəvayz]</td>
</tr>
</tbody>
</table>

The rule is actually somewhat more complicated, however, as Chambers (1973) shows. Like Flapping, the application of Raising also depends on prosodic structure. Note the relevance of stress in the examples below in (25): Raising only affects /a/ if /a/ has a greater degree of stress than the following syllable, or if there is no following syllable in the word. Note also that the syllable following /a/ does not need to be stressless; hence Chambers and the speakers he consulted pronounced *icon* as [ʌykən], not [ʌykən] or [ʌykən].

(25) [after (12) in Chambers 1973:125]
The dependence of Raising on stress suggests an analysis involving Resyllabification. As Paradis (1980) shows, such an analysis appears to be precisely what is needed; the stress patterns illustrated above in (25) are just those where Resyllabification would be expected to apply. Specifically, it applies in the forms in the right-hand column of (25) to make a voiceless consonant the coda of a syllable containing the diphthong /ay/. If we suppose that the application of Raising requires trigger and target to be tautosyllabic, we can understand the examples in (25) where Raising applies as cases where Raising is fed by Resyllabification. 7

Given this insight, the rule of Raising can be formulated as given below.

(26) Raising


The requirement that Raising apply solely within syllables predicts that it cannot apply between words either. This restriction is illustrated by the examples below. The one case where Chambers (1973) observed raising between words, namely the example given in (27b), involves a rather frequent compound that may have been reanalyzed as monomorphemic in some sense (J. Chambers, pc).

7 Mike Hammond points out that if an analysis with Resyllabification is adopted, something must be said about the Raising that occurs in forms such as *fcón. The second vowel of such words is not stressless, meaning Resyllabification as formulated in Chapter 2 cannot apply.
(27) a. [from Chambers 1973:117]

tie clip     [təy klɪp], not *[tAy klɪp]
cow poke    [kɔw pɔwkw], not *[kAw pɔwkw]

b. [after (14) in Chambers 1973:128]

high school  [hAy skɔl] or [hAy skɔl]

The fact that Raising is restricted to apply within words while Flapping is not predicts that Raising should be less productive than Flapping. Thus, in accordance with the Productivity Hypothesis, we expect Raising to be ordered before Flapping. This prediction is supported: in all current dialects that have both rules, Raising is ordered before Flapping (Chambers 1973, Vance 1987).

Of course, it is well known that Halle (1962) claims there are dialects that order the rules the opposite way, with Flapping before Raising, basing this claim on the data in Joos (1942). However, this dialect appears to have disappeared completely by 1973 (Chambers 1973). As Kaye (1990) observes, the utter disappearance of a dialect in just over thirty years suggests that it may never have existed at all, especially since Joos (1942) includes schoolchildren as speakers of this dialect.

Another factor casting doubt on the existence in 1942 of a dialect with Flapping ordered before Raising is the fact that this ordering should result in a complete neutralization between words like rider and writer, because Flapping will change the underlying /t/ and /d/ to the voiced [ɾ], thus removing the environment for Raising. It is not clear, however, if such complete neutralizations
ever really occur in natural speech. Thus Dinnsen (1985) cites a number of studies showing that the underlying distinction between /t/ and /d/ always remains after flapping in the instrumentally measurable spectral or temporal properties of the preceding vowel. In any event, the lack of any currently surviving dialect neutralizing rider and writer, combined with the small amount of data provided in Joos (1942), suggests that Halle's (1962) argument is not as convincing as may first appear.\(^8\)

If we only consider dialects for which we have data, we see that Raising always precedes Flapping. For example, in the Canadian dialect discussed by Chambers (1973), the pronunciation of words like rider and writer are distinct only in the quality of the first vowel: in rider the vowel is an unraised [a], while in writer it is a raised [A]. Both words show evidence of Flapping. Since Flapping changes the voiceless /t/ in writer to a voiced [r], thus destroying the environment for Raising, it must be that in this dialect Raising is ordered before Flapping. Sample derivations are given below.

---

\(^8\) Rudes (1976) also claims that there are dialects where rider and writer are fully neutralized, but like Joos (1942) does not provide much data, all of it impressionistic. Further, Chomsky and Halle (1968) claim that dialects may arbitrarily order Flapping with a rule similar to but distinct from Raising, namely one where voicing affects vowel length rather than height. Again, however, the phonetic data are not as clear-cut as often assumed; see Fox and Terbeek (1977) and Huff (1980) for discussion.
The same ordering is found in other dialects that have a slightly different version of Raising. Vance (1987) observes that some North Americans south of Canada, including at least speakers from Minnesota and northern New York, have a rule which raises /ay/ to [Ay] before voiceless consonants, but which leaves /aw/ unaffected. This rule is stress-dependent in precisely the same way as the Canadian rule and it is also restricted from applying across word boundaries. Significantly, in these dialects, too, the Raising rule is ordered before Flapping, so that rider and writer differ only in the pronunciation of the vowel.

The claim that the ordering of Raising before Flapping is due to a difference in productivity finds support in the fact that the former rule appears to be prepatterned, at least to a much greater degree than the latter. This is seen in the observation that there are lexical exceptions to Raising in all dialects with such a rule. Some exceptions in Canadian English are listed in (29a), culled from examples in Chambers (1973) and my own informal survey of the speech of three Canadian speakers; exceptions in the dialects Vance (1987) discusses are listed in (29b).
(29) Exceptions to Raising
["JC" = Chambers 1973, "JM" = my own data, "TV" = Vance 1987]

a. Canadian English

<table>
<thead>
<tr>
<th>JC</th>
<th>Cyclops [sáyklâps]</th>
<th>cf micron [máykrâns]</th>
</tr>
</thead>
<tbody>
<tr>
<td>JM</td>
<td>outrage [áwtrêyç]</td>
<td>cf tout [táwt]</td>
</tr>
<tr>
<td>JM</td>
<td>mouth [máwθ]</td>
<td>cf south [sÁwθ]</td>
</tr>
<tr>
<td>JM</td>
<td>mouse [máwθs]</td>
<td>cf spouse [spÁwθs]</td>
</tr>
</tbody>
</table>

b. Other dialects

<table>
<thead>
<tr>
<th>TV</th>
<th>nice [náys]</th>
<th>cf like [lÁyk]</th>
</tr>
</thead>
<tbody>
<tr>
<td>TV</td>
<td>icon [áykan]</td>
<td>cf psyche [sÁykly]</td>
</tr>
<tr>
<td>TV</td>
<td>bison [báysen]</td>
<td>cf vital [vÁyrel]</td>
</tr>
</tbody>
</table>

By contrast, Flapping appears to be entirely exceptionless. The only example pointed to in the literature as a possible exception to Flapping is given below in (30). I will deal more fully with what an exception like this really suggests in Chapter 5.

(30) [from Withgott 1982]

militaristic [militɛrˈstɪk] | cf capitalistic [ˈkærɪpəlɪstɪk]

The relative unproductivity of Raising is demonstrated further by the existence in many dialects of raised [AY] or [AW] before voiced consonants, where it could not have been derived by Raising. This means that the distinction between [ay, aw] and [AY, AW] is not always predictable, forcing the learner to posit both sets of diphthongs as underlying in many words. The necessity of this prepatterning makes the learning of Raising as fully productive even
harder. Examples of such unpredictable raised diphthongs are given below.

(31) Unpredictable raised diphthongs

["EK" = Kaisse 1992a, "JM" = my own data, "TV" = Vance 1987]

a. **Canadian English**
   
   **EK**  
   cider  [səːdər]  
   or [səːdər] {cf spider [spəːdər]}
   
   **JM**  
   espouse  [ɛsəːwəz] {cf houses [həːwəz]}

b. **Other dialects**

   **TV**  
   idle  [ədəl] {cf idol [ədəl]}
   
   **TV**  
   tiger  [təːɡər] {cf Niger [nəːɡər]}

The lesser productivity of Raising is not to be taken to imply that this rule is fully unproductive; novel words such as Nike and plout will be expected to have raised [ʌ] (J. Chambers, pc). The fact that Raising can apply in novel forms does not mean that it is fully productive; as we saw in Chapter 3, even rather unproductive rules are capable of applying in novel forms. I predict, therefore, that if experimental or speech error studies were to be done on Raising, they would show this rule to be only somewhat productive, corroborating the linguistic evidence cited above. I claim that this lesser degree of productivity follows from the fact that it is restricted to apply word-internally. This lesser productivity, in turn, results in Raising being ordered before Flapping in all known dialects.
4.3.2 Regressive Nasalization and Flapping

The interaction of English Regressive Nasalization and Flapping also seems to follow properly from the Productivity Hypothesis. Specifically, since Regressive Nasalization is restricted to apply word-internally while Flapping is not, it is less productive than Flapping and therefore applies earlier. Before I discuss its interaction with Flapping, a word or two about Regressive Nasalization itself is in order.

Malecot (1960) found that in the perception of nasal consonants before voiceless stops, the presence of actual nasal consonants is not required; all that seems to be necessary is nasalization on the preceding vowel. With speech production, it was found that speakers seem to reduce the duration of nasal consonants before voiceless stops. The two most important factors influencing the extent of the reduction are dialect (New Yorkers do not seem to reduce nasal consonants at all) and vowel quality. Nasal consonants are shortened most after [æ], less so after [ʌ], and the least after [ʊ]. This same dependence of the effect on vowel quality was also found with speech perception; [æ] followed by a voiceless stop was perceived as a vowel-nasal consonant sequence more readily than [ʌ] or [ʊ] followed by a voiceless stop.

Subsequent discussions of these phenomena (eg, Kahn 1976, Donegan and Stampe 1979, Kaisse 1985, among many other places) have made a significant assumption that is not derived from what is found in Malecot (1960). As Bourgeois (1990, 1991) points out, what
Malecot unambiguously describes as a reduction in the duration of /n/ before voiceless stops, with the amount of reduction depending on vowel quality, is assumed by all later commentators to be a categorical deletion of /n/.

In other words, it is claimed that two independent processes are interacting here. In Regressive Nasalization, a vowel preceding a nasal consonant becomes nasalized; in Nasal Elision, a nasal consonant is reduced or deleted before a voiceless stop. Based on these later summaries of Malecot (1960), pronunciations of the words in (32) should be as given below.

(32) a. Regressive Nasalization

\[
\begin{array}{ll}
\text{pun} & [p\tilde{n}] \\
\text{pin} & [p\tilde{n}] \\
\text{pan} & [p\tilde{\alpha}n] \\
\text{fund} & [f\tilde{\alpha}nd] \\
\text{pinned} & [p\tilde{\alpha}nd] \\
\text{panned} & [p\tilde{\alpha}nd] \\
\end{array}
\]

b. Nasal Elision

\[
\begin{array}{ll}
\text{punt} & [p\alpha t] \\
\text{flint} & [f\alpha t] \\
\text{pant} & [p\alpha t] \\
\end{array}
\]

If these rules had to be ordered, it must be that Regressive Nasalization applies before Nasal Elision. This is seen by comparing the derivations in (33), where only the ordering in (33a) derives the correct form.
Given the description in Malécot (1960), however, this analysis is entirely wrong. Nasal stops are reduced, not elided, before voiceless stops, and the amount of reduction depends on the quality of the preceding vowel. In the cases where the nasal is not entirely eliminated, there is therefore no evidence that Regressive Nasalization and "Nasal Elision" have to apply in a particular order with respect to each other; the rules do not in fact interact at all. Consequently, in the remainder of this discussion, I will refer to Regressive Nasalization and concomitant nasal reduction as aspects of a single process.

The restriction of Regressive Nasalization to word-internal domains follows from the fact that it only applies within the syllable (Kahn 1976, Kaisse 1985). That is, a vowel is nasalized by a following nasal consonant only if both are tautosyllabic. Likewise, a nasal consonant is reduced before a voiceless stop only if both are tautosyllabic. These patterns are illustrated with examples in (34); those in (34a) show vowel nasalization occurring only if the vowel and nasal consonant are tautosyllabic, either underlyingly or through
application of Resyllabification, while those in (34b) show nasal 
reduction occurring only if the nasal consonant and voiceless stop are 
tautosyllabic. The contrast between the last two examples in (34b) 
provides further evidence that Resyllabification cannot apply across 
word boundaries, because otherwise the [n] in span would be 
reduced.

(34) ['." marks syllable boundaries]

a. pán [pæn]  
pánic [pæn.ɪk]  
panáche [pæn.æʃ]

b. pánt [pæt]  
bánter [bæt.ər]  
spán togethér [spæn.tə.ge.θeər]

Given this restriction to word-internal application, the 
Productivity Hypothesis predicts that Regressive Nasalization should 
apply before Flapping, which easily applies across word boundaries. 
As Donegan and Stampe (1979) show, this is in fact true.

The evidence comes from examples where both rules can 
apply, such as the phrase plant it. This phrase may be pronounced in 
a number of different ways, including with the nasalized flap [ɾ]; 
these are listed in (35a). Notably, the pronunciation in (35b) with 
both a nasal consonant and a flap is not possible.9

9Selkirk (1972) presents a different analysis of these facts predicated on the 
assumption that a flap may appear after a nasal stop. As this does not occur in 
any of the dialects I have heard or found in the literature, I will assume that 
her analysis is mistaken.
(35) Pronunciations of plant it.

   a.  [plæntɪt]
       [plætɪt]
       [plærɪt]
       [plæʃɪt]

   b.  *[plæŋtɪt]

This pattern can be understood if we suppose that Regressive Nasalization applies before Flapping. If Nasalization applies with insufficient reduction of the nasal consonant, Flapping cannot apply, giving the first pronunciation in (35a). If nasal reduction is sufficient, Flapping may or may not apply, giving the next two pronunciations in (35a). Finally, given that the nasal reduction associated with Regressive Nasalization tends not to be complete, the feature [+nasal] may also surface on the adjacent flap.

The incomplete nature of nasal reduction explains two facts that Donegan and Stampe (1979) note but do not account for. First, there are speakers who have a fully productive Flapping rule, as evidenced by examples like pat it [pærɪt], but who do not flap with /nt/ in examples like plant it. Second, even speakers who can flap in plant it do not apply the rule as readily there as in examples like pat it; in the former case, Flapping is optional in formal styles, while in the latter Flapping is essentially obligatory. Both facts can be understood if we recall that Regressive Nasalization is only associated with reduction of the nasal consonant, not necessarily with its complete elision. The amount of nasal reduction that occurs is variable, depending on phonetic factors such as speaking rate. If
Flapping follows Regressive Nasalization, it may then be thought of as essentially obligatory; its apparently optional application in cases like *plant it* actually follows from the optional nature of nasal reduction. Derivations illustrating this analysis are given below in (36).

(36)

<table>
<thead>
<tr>
<th>UR</th>
<th>plant it</th>
<th>plant it</th>
<th>plant it</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regressive Nasalization</td>
<td>plant it</td>
<td>plant it</td>
<td>plant it</td>
</tr>
<tr>
<td>Flapping</td>
<td>plant it</td>
<td>plant it</td>
<td>---</td>
</tr>
</tbody>
</table>

By claiming that the ordering of Regressive Nasalization before Flapping follows from the Productivity Hypothesis, I am claiming, then, that Regressive Nasalization is prepatterned and not in fact fully productive as often assumed. Since it cannot apply across word boundaries, the language learner is given no evidence that it must be fully productive. Its ordering before Flapping is thus expected.¹⁰

4.3.3 Prevelar Raising and Nasal Assimilation

My final example of the ordering of a word-internal rule before a word-external rule concerns Prevelar Raising, discussed in Chapter 2, and Optional Nasal Assimilation.

In contrast with Optional Nasal Assimilation, Prevelar Raising only applies word-internally. This is especially clear in dialects where /æ/ and /ʌ/ are only raised before [ŋ], since [ŋ] never appears word-initially. Thus if it is to act as a trigger for Prevelar Raising, the process must occur entirely within the word. The fact

¹⁰Thanks to Samuel Wang for suggesting to me in another context that Regressive Nasalization may not be fully productive in English. See Read (1972:98-101) for further (external) evidence for the prepatterning of Regressive Nasalization in English.
that this process is restricted to apply solely within words, while Optional Nasal Assimilation can apply across word boundaries, predicts that it is less productive than Nasal Assimilation and is therefore ordered before it.

As we saw in Chapter 2, this prediction is correct: when /ŋ/ is derived by assimilating across word boundaries, Prevelar Raising does not apply. Recall that Obligatory Nasal Assimilation does apply before Prevelar Raising. This is consistent with the Productivity Hypothesis because like Prevelar Raising, Obligatory Nasal Assimilation only applies word-internally. The word-internal rule of Prevelar Raising and Optional Nasal Assimilation thus provide yet another example in support of the ordering predicted by the Productivity Hypothesis.

4.3.4 Word-external rules last: Summary

I have hoped to show in this section that the ordering of word-internal and word-external rules is predictable, with the former always preceding the latter. I claimed that this ordering follows from the way rules are learned as fully productive; word-internal rules tend to be learned as less productive than word-external rules, so that they end up ordered through application of the Productivity Hypothesis. In at least one of the cases I discussed, namely Raising, there is further evidence of the lesser productivity of the earlier rule, in that it tolerates exceptions.
4.4 Word-internal rules and the Productivity Hypothesis

To the best of my knowledge, all other rule pairs in English that have been suggested as showing the ordering of fully productive rules do not involve word-external rules at all. Without further information on the rules' relative productivity, it is therefore difficult to know what the Productivity Hypothesis would predict about such rule pairs. Examples of these include Vowel Reduction and Sonorant Syllabification (Selkirk 1972), ə-Deletion and h-Deletion (Kaisse 1985), and a large number of rules discussed in Bailey (1973). All of these examples involve the ordering of word-internal rules only.

There is some evidence that such rule pairs may be ordered (apparently) arbitrarily, as indicated by different orderings in otherwise (apparently) identical dialects. Thus Bailey (1973) claims that many of his rules may be applied in different orders by different speakers, and the same may be true of the apparently very productive, but nevertheless word-internal rules in German dialects discussed by Hall (1992). Of course, Double Lookup claims that if the relative productivities of these rules are determined, it will be seen that the rules are not ordered arbitrarily at all, but instead are arranged in order by differences in productivity.

Support for this claim comes from those cases where there does seem to be external evidence for the word-internal rules' relative productivity. In this section I discuss a general category of such cases. Specifically, I examine the claim made by Donegan and Stampe (1979) that on-line fortition rules are universally ordered
before on-line lenition rules. I show that this principle, which appears to hold even of pairs of word-internal rules, actually derives from the Productivity Hypothesis because fortitions are universally more prepatterned and thus likely to be less productive than lenitions.

4.4.1 Fortition precedes lenition

In the last section we saw evidence for the universal principle, derivable from the Productivity Hypothesis, that word-internal rules are ordered before word-external rules. Here I consider another universal ordering principle that was proposed by Donegan and Stampe (1979) and show that it too appears to be derivable from the Productivity Hypothesis.11

Donegan and Stampe (1979) claim that fortition processes always precede lenition processes. Fortitions and lenitions are defined in the following two passages:

Fortition processes... intensify the salient features of individual segments and/or their contrast with adjacent segments. They invariably have a perceptual teleology.... Dissimilations, diphthongizations, syllabifications and epentheses are fortition processes.

[Donegan and Stampe 1979:142]

Lenition processes... have an exclusively articulatory teleology, making segments and sequences of segments easier to pronounce by decreasing the articulatory "distance" between features of the segment itself or its adjacent segments. Assimilations, monophthongizations, desyllabifications, and deletions are lenition processes.

[Donegan and Stampe 1979:142]

11 Thanks to Greg Iverson for calling my attention to this.
Donegan and Stampe (1979) give many examples supporting their ordering principle. I will simply take it as given and consider the question of why this ordering principle might be expected to hold. I suggest that their principle, instead of operating on the ordering of application of on-line rules as they claim, is actually derivable from the Productivity Hypothesis. That is, I suggest that fortitions are universally ordered before lenitions because fortitions are universally less productive than lenitions.

The first clue that this may be true is found in Donegan and Stampe's (1979) own description of their model. In their discussion of the difference between what they call rules (i.e., rules that do not tend to apply on-line) and processes (i.e., rules that do apply on-line), they note that "[r]ules... are formed through the observation of linguistic differences of which the speaker is or was necessarily conscious", while "[p]rocesses apply involuntarily and unconsciously" (Donegan and Stampe 1979:144). This distinction between on-line rules and prepatterned rules in terms of conscious learning and application is paralleled in their description of fortitions and lenitions. While lenitions "apply most widely in styles and situations which do not demand clarity" (p. 143), fortitions "apply in situations and styles where perceptibility is highly valued" (p. 142). In order for a speaker to apply a rule with a "perceptual teleology" (i.e., a fortition), the speaker must be aware enough to know what is and what is not perceptually salient for the listener. By contrast, the application of a lenition process need not be conscious at all, as its
function is merely to aid in articulation. Thus fortitions are like prepatterned rules in applying more consciously the times they do apply; lenitions are like on-line rules in their unconscious application. If consciousness correlates with productivity, as Donegan and Stampe (1979) imply, this means that fortitions are less productive than lenitions.

More specific evidence for this is found in Ohala (1986b). There the argument is made, based on experimental studies and studies of diachronic sound change, that dissimilations (a type of fortition) occur because speakers are overcorrecting for perceived articulatory assimilations. For example, through an automatic, purely articulatory process, high back rounded vowels tend to be fronted before coronals, so that a speaker's intended /ut/ may surface as [üt]. Speakers of languages without a phonemic /ü/ tend to mentally undo this distortion, perceiving [üt] as /ut/. If the language does have a phonemic /ü/, however, so that a speaker says [üt] because she intends to say /üt/, a surface ambiguity arises. Ohala (1986b) suggests that in such a situation, listeners may mentally "correct" the vowel here as well, undoing what was thought to be a purely articulatory process and thereby ending up perceiving /ut/. If this mistake becomes grammaticalized, the language will show evidence of a dissimilation, either as a diachronic sound change or perhaps as a synchronic rule.

In addition to the experimental and sound change evidence, Ohala (1986b) observes that this analysis makes another interesting
prediction, namely that the features that are able to dissimilate should be precisely those that can assimilate through mechanical mistiming of articulations. This seems to be true. For example, as McCarthy (1988) notes, the features [consonantal] and [sonorant] are not subject to the Obligatory Contour Principle.\(^\text{12}\) This remains a mystery in his paper, but it finds a natural explanation within Ohala's model. Namely, these features are not associated with particular articulatory gestures and so cannot be spread through mistiming; listeners thus have no motivation to "undo" examples involving these features.

In his comment on Ohala (1986b), Kiparsky (1986) observes that since in this view dissimilations involve the reanalysis of lexical forms (eg, /üt/ for one speaker becomes /ut/ for another), dissimilation rules must be lexical rules. He points out that this explains why dissimilations are structure-preserving and may have lexical exceptions (see Chapter 6 for a more complete discussion of Lexical Phonology). In terms of Double Lookup, this observation means that dissimilations are prepatterned, and therefore less productive than other rules, all else being equal.

If Donegan and Stampe's (1979) ordering principle can be empirically verified, then, it may provide more support for the validity of the Productivity Hypothesis. Note that Donegan and Stampe's typology of fortitions and lenitions should apply equally well with within-word rules. Hence if I am correct in supposing that

\(^{12}\text{But compare the response of Kaisse (1992b).}\)
fortitions are universally less productive than lenitions, we may have another source of evidence for the relative productivity of word-internal rules.

4.4.2 Word-internal rules: Summary

In this section I argued first that evidence of arbitrary ordering among apparently fully productive word-internal rules does not contradict the Productivity Hypothesis. Since such rules cannot apply in the clearly novel forms represented by phrases, there is no way to know if the rules really do apply on-line as claimed. I argued further that when something about the relative productivities of such rules is known, the ordering is in fact consistent with the Productivity Hypothesis. Of particular interest is the finding that the ordering of fortitions before lenitions can be derived from the Productivity Hypothesis; there is no need to stipulate this as a special property of on-line rule pairs.

4.5 Chapter summary

In this chapter I have argued that there are no cases of arbitrarily ordered fully productive rule pairs in English. All apparent counterexamples fall into three classes. First, the two rules truly are fully productive and so apply in accordance with Automatic Feeding. Second, the two rules truly are fully productive, and the first is a prosodic rule, so that they are ordered in accordance with Prosody First. Finally, the earlier of the rules is actually less than fully productive. This difference in productivity can be seen in the fact that the earlier rule is always a word-internal rule, while the
later one may not be. Rule pairs where both rules apply word-internally thus appear to be ordered arbitrarily without further information about their relative productivity. In cases where such information is known, the ordering is consistent with the Productivity Hypothesis. Thus I argued that Donegan and Stampe's (1979) principle that fortitions precede lenitions follows from the Productivity Hypothesis since fortitions are less productive than lenitions.

I summarize the rules discussed in this chapter by listing them below. I indicate the universal nature of these orderings by dividing the rules into four blocks: word-internal prosodic rules, word-internal segmental rules, word-external prosodic rules, and word-external segmental rules. Both within words and across words, prosodic rules are ordered first, in accordance with Prosody First. Word-internal rules are ordered before word-external rules, since the former are learned as less productive, which by the Productivity Hypothesis results in their being ordered earlier. Word-external segmental rules are ordered last, and interact among themselves in accordance with Automatic Feeding.

(37) **Word-internal prosodic rules**

a. Resyllabification  [before d, e, h]

b. Onset Principle

**Word-internal segmental rules**

c. Prevelar Raising  [before i]

d. Canadian Raising  [before h]

e. Regressive Nasalization  [before h]
**Word-external prosodic rules**

| f. | Onset Principle [before g, h] |

**Word-external segmental rules**

| g. | Coronal Deletion [before i, j] |
| h. | Flapping |
| i. | Nasal Assimilation |
| j. | Palatalization [before k] |
| k. | y-Deletion |

In this chapter I have only examined rules in English. Although I see no reason for English to be atypical in this respect, it may be possible that other languages will show the counterexamples needed to falsify my predictions concerning (apparently) fully productive rules. Hence further work is needed. At this stage, however, I think I have made a good case for the claim that fully productive rules cannot be ordered except by Automatic Feeding and Prosody First. At the very least, I have shown that many cases that have been analyzed in the literature as involving ordered fully productive rules actually involve nothing of the sort.
CHAPTER 5
PREPATTERNING, PRODUCTIVITY
AND CYCLICITY

5.0 Introduction

Beginning with Chomsky and Halle (1968), it has been claimed that in certain cases a phonological rule can apply more than once in a derivation. Such multiple applications are dependent on morphology, so that different applications of the same rule occur under distinct morphologically defined conditions. I will call this cyclicity.

My use of this term conflates two different but very similar concepts in the current literature, both defined in terms of the Lexical Phonological concept of ordered strata or levels, mentioned in Chapter 1. The first is the idea that an entire level may reapply; specifically, the phonological rules in the level may reapply every time a morphological operation associated with that level applies. Such a level is called a cyclic level. The second is the idea that a phonological rule may apply in more than one level; that is, such a phonological rule may reapply every time the derivation reaches a new level. Such a rule should not technically be referred to as "cyclic," but in terms of the relevant property of my definition of cyclicity it is the same: it may reapply in a different morphological domain. Indeed, theories without the concept of ordered levels (e.g., Chomsky and Halle 1968, Halle and Vergnaud 1987) treat these two forms of what I call cyclicity the same way as well. This conflation will actually turn out to be an argument in favor of Double Lookup,
since I will show that certain cases handled in Lexical Phonology with very different mechanisms actually arise from the same underlying cause. Hence throughout the remainder of the discussion I use the term "cyclicity" to refer simply to the reapplication of rules in new morphologically defined domain.

The reason why this discussion of cyclicity is necessary at all is because as I have defined it, multiple applications of a rule are entirely incompatible with Double Lookup. This is because according to the Productivity Hypothesis, the ordering of rules is determined solely by their relative productivity. A rule cannot appear more than once in a derivation because it cannot be less productive or more productive than itself.

The concept of rule cyclicity must be distinguished from two others that are compatible with Double Lookup. The first is rule iterativity. Iterativity also involves the repeated application of a rule, but unlike cyclicity, these repeated applications must occur adjacent in the derivation; that is, no other rule application can intervene. An example of this is the iterative spreading of some feature across a word, such as the spread of roundness to all the segments to the left of [u] in the pronunciation of *strew* as *[s\textit{p}tw\textit{r}wuw]*. It is commonly assumed (see, eg, Archangeli and Pulleyblank forthcoming) that iterativity is a characteristic of the form of individual rules, and is not a property of rule derivations. Because of this, iterative rule application is fully compatible with
Double Lookup; iterative rules may only apply once, though their subparts may be thought of as repeating.

A second concept that must be distinguished from cyclicity is that of Automatic Feeding, discussed in Chapter 4. I argued that two interacting fully productive rules automatically feed each other, the application of one rule triggering the application of the other, and vice versa, until neither rule can apply again. This concept clearly does allow for the repeated application of a rule in a derivation. Note, however, that Automatic Feeding can only occur under very specific circumstances. First, all the rules involved must be fully productive. By contrast, the model of Lexical Phonology explicitly claims that the postlexical stratum, which contains rules that are almost always fully productive (see Chapter 6), is not cyclic (Kiparsky 1982, 1985; Mohanan 1982). Thus all cases of cyclicity in the literature involve partially productive rules (more accurately, prepatterned rules), a fact that will be significant in my account of them. Second, the only evidence that fully productive rules really do apply repeatedly is that they always surface with feeding and/or counterbleeding order. It is conceivable that the avoidance of bleeding and counterfeeding interactions among fully productive rules is really due to some mechanism other than ordered rule application. The evidence for the repeated application of cyclic rules, on the other hand, is of a very different sort, as we will see.

1 This idea will be pursued in later work.
Double Lookup predicts, then, that rule cyclicity *per se* does not exist. I will show that the various effects that have been ascribed to cyclic rule application actually do not have a unary explanation at all.

As Cole (1990) notes, there are three types of arguments in the literature for cyclicity, namely arguments from derived environment effects, from rule ordering paradoxes, and from evidence that directly shows cyclic application of a rule. I follow this division in the discussion that follows.

I begin in 5.1 by pointing out that derived environment effects have nothing to do with cyclicity and hence do not provide evidence for it.

In 5.2 I discuss rule ordering paradoxes, showing that the diversity of such cases actually requires two distinct explanations, neither of which involves cyclicity. The first such mechanism is what I call *fake cyclicity*, the application of two or more rules in a derivation that may be very similar but are nevertheless distinct; like real cyclicity, the applications of the distinct rules occur in distinct morphologically defined domains. Since claimed cases of real cyclicity can always be reanalyzed as cases of fake cyclicity, where the different rules happen to be identical and not just similar, it suffices to show that fake cyclicity is independently needed; parsimony considerations will then prefer the theory allowing only fake cyclicity. In 5.2.1 I do just this, giving several cases that can only be analyzed as fake cyclicity.
In 5.2.2 I discuss the other mechanism needed to understand rule ordering paradoxes, namely *idiosyncratic ordering*, which allows certain lexical items to be marked as undergoing rules in an order different from other lexical items. Cases of idiosyncratic ordering do exist and real cyclicity cannot handle them; however, Double Lookup can. Since phonological rules can be prepatterned, it is possible for certain lexical items to preserve patterns originally derived through a given ordering of rules even after these rules have later switched order. Thus only the dominant ordering pattern results from the Productivity Principle; the items with idiosyncratic ordering may be thought of as a special kind of lexical exception. Like all lexical exceptions in Double Lookup, such lexical items are not marked with a diacritic, but are simply stored in memory with the older ordering pattern and do not undergo either of the rules on-line. In 5.2.2 I examine some clear cases of this, one of which, significantly enough, mimics real cyclicity to a striking degree, in particular in its apparent sensitivity to morphology.

In 5.3 I consider the source of the best evidence for cyclicity, namely sets of words that directly reflect multiple applications of a given rule. Unlike the fake cyclicity cases discussed above, it appears that in these cases the different rule applications are always exactly alike, rather than merely being similar. The common feature shared by the two cases I discuss here is that they both involve the apparent repeated application of fully productive rules, ordered both before and after partially productive morphological operations. I
argue that these cases involve fake cyclicity just like those discussed in 5.2.1. The fake cyclicity arises from the ordering of an on-line rule after a prepatterned version of itself; this prepatterned version is identical to the on-line version because of the rule's high level of productivity. This argument is supported with psycholinguistic evidence that fully productive rules can nevertheless be prepatterned.

In 5.4 I summarize the evidence for cyclicity. I conclude that real cyclicity does not exist. Instead, apparent cases of it actually involve fake cyclicity or idiosyncratic rule ordering, both consistent with Double Lookup and the Productivity Hypothesis. Even the best apparent cases of real cyclicity show a constellation of properties indicating that they too should be understood as fake cyclicity, specifically fake cyclicity involving prepatterned and on-line versions of fully productive rules.

5.1 Derived environment effects

Following Mascaro (1976), phonologists have assumed the Strict Cycle Condition, which claims that structure-changing cyclic rules and only structure-changing cyclic rules are obliged to apply in a derived environment. A structure-changing rule is one that removes structure from the input and replaces it with new structure; by contrast, a structure-building rule is one that merely adds extra structure to the input, leaving the rest alone.

It has since been shown that the constraint preventing certain rules from applying in non-derived environments is dissociable from
constraints on cyclic rule application. Hualde (1989) and Cole (1990) have found examples of structure-changing noncyclic rules in Basque, and in Icelandic and Arabic, respectively, that also fail to apply in nonderived environments. In other words, whether or not a rule only applies in derived environments has nothing to do with whether it is cyclic.

Hualde's (1989) example from Basque shows this clearly. Many dialects of this language have a rule of Vowel Assimilation, whereby /a/ is raised to [e] when the nearest vowel to the left is high, as illustrated in (1a) below. This rule is restricted to derived environments, as can be seen from the fact that in nonderived forms [a] often follows nonhigh vowels, as illustrated in (1b). And yet the rule is noncyclic. This is seen by the fact, illustrated in (1c), that /a/ is only affected if adjacent to a word boundary, whether or not /a/ is adjacent to a word-internal morpheme boundary. Thus the rule can only apply once in the derivation (ie, at the word level).

(1)  a. [from (1) and (3) in Hualde 1989:675-6]

| /gišon-a/   | [gišona] | 'the man (abs. sg.)' |
| /lagun-a/   | [layune] | 'the friend (abs. sg.)' |
| /mutil-a/   | [mutiæe] | 'the boy (abs. sg.)' |

b. [from (2) in Hualde 1989:676]

| /muga/      | [muya]   | 'limit' |
| /eliša/     | [eliša]  | 'church' |
c. [from (3) and (5) in Hualde 1989:676-7]

/ɡiʃon-ak/ [ɡiʃonak] 'the men (abs. pl.)'
/lagun-ak/ [layunak] 'the friends (abs. pl.)'
/mutil-ak/ [mutiʌak] 'the boys (abs. pl.)'
/mutil-a/ [mutiʌe] 'the boy (abs. sg.)'
/mutil-a-k/ [mutiʌak] 'the boy (erg.)'

It remains true that some rules require a constraint preventing them from applying in nonderived environments. However, there are constraints available for doing this that do not make reference to cyclicity. For example, there is the Alternation Condition, originally proposed in Kiparsky (1973b), which states that obligatory neutralization rules cannot apply to all occurrences of a morpheme. Kiparsky (1982) removes some of the problematic features of this constraint by revising it to state that obligatory neutralization rules apply only in derived environments. In either case, the Alternation Condition must be thought of as a constraint on rule learning, not on rule application. Thus Kiparsky (1982) notes:

The only sense that can be made out of it is as a strategy of language acquisition which says that a learner analyzes a form "at face value" unless he has encountered variants of it which justify a more remote underlying representation. [Kiparsky 1982:36]

Hammond (1992) adopts a different interpretation of the derived environment effect, though like the Alternation Condition

---

2 As Iverson and Wheeler (1988:326-7) note, the changes in the revised version of the Alternation Conditions are actually not that great, and not all positive; the revised version in fact lets certain problematic cases get through that had been blocked in the earlier version.
one that does not make reference to cyclicity. According to his Morphology Hypothesis, a phonological rule that is morphologically restricted is in fact a form of morphological rule; being morphology, the rule may then only apply as part of a morphological process, so that monomorphemic forms may not be affected.

In summary, then, it appears we must reject the Strict Cycle Condition as an explanation for the derived environment effect. Rules that are unambiguously noncyclic may nevertheless show derived environment effects; thus if we choose to keep the Strict Cycle Condition, we must supplement it by another principle, such as the Alternation Condition or the Morphology Hypothesis, that does essentially the same thing more generally.

5.2 Rule ordering paradoxes

In a model of rule ordering that does not allow cyclicity but does require all rules to be fully ordered, at least two properties must hold of derivations. First, transitivity violations should be impossible: given any three rules A, B and C, the ordering of A before B and B before C means that A must be ordered before C. Second, idiosyncratic ordering should be impossible: for any two rules A and B and any two forms P and R that may potentially undergo both A and B, it must be the case that the ordering of A and B as displayed in P must be the same as their ordering as displayed in R.

Anderson's (1974) model of Local Ordering denies both of these properties. Contrary to the model sketched above, Local Ordering
supposes that rules are ordered solely in a pairwise fashion, allowing for the possibility of transitivity violations. Moreover, lexical items or affixes may be marked to undergo rules in an order different from the order that occurs in the rest of the lexicon. Ordering of either of these sorts, that is transitivity violations or idiosyncratic ordering, are known as rule ordering paradoxes. Apparent examples of rule ordering paradoxes have been pointed to in Finnish and Icelandic (Newton 1971, Kiparsky 1973b, Anderson 1974, Bley-Vroman 1975), modern Greek (Kaisse 1975), the Athapaskan language Hare (Rice 1980), and various dialects of Swiss German (Robinson 1976, 1977), among other places.

Rule ordering paradoxes are not discussed much any more because there is a general feeling that all such cases can be accounted for through the single mechanism of rule cyclicity, in particular the version adopted in Lexical Phonology. Apparent transitivity violations, where A precedes B, B precedes C but A follows C, can be handled by supposing that A, B and C appear in a cyclic level, so that the first application of C may precede the second application of A. Some cases of apparent idiosyncratic ordering can also be handled by cyclicity, specifically cases where a certain set of affixes triggers one ordering of a rule pair, while another set of affixes triggers another. This is handled by supposing that different affixes may be associated with different morphologically defined levels, which are themselves ordered.
By contrast, Double Lookup has to claim that apparent transitivity violations and idiosyncratic ordering result from different mechanisms. Thus according to the Double Lookup model, transitivity violations should be impossible. The Productivity Hypothesis claims that rules are ordered according to their productivity values; since the relation "less productive than" is transitive (i.e., if rule A is less productive than rule B, and B is less productive than rule C, it must be the case that A is less productive than C), rule ordering must be transitive as well. All apparent cases of transitivity violations, therefore, must involve fake cyclicity. That is, if C appears to follow B but also precede A, this is because the version of C preceding A is in fact a distinct rule from the version following B.

Fake cyclicity may also be used in the cases where real cyclicity has been argued to account for apparent idiosyncratic orderings, namely those involving the association of certain orderings with affixes. This is because, as noted in the introduction to this chapter, fake cyclicity can account for all cases real cyclicity can. In addition, as we will see, it can also account for cases where real cyclicity is not useful.

In some cases of idiosyncratic ordering, though, neither fake cyclicity nor real cyclicity explain the facts. These are cases where the marked ordering is associated not with affixes but with individual lexical items. As we will see, such cases do exist. This sort of idiosyncratic ordering, however, is expected within the Double
Lookup model. Just as certain lexical items, especially frequent ones, can preserve the effects of rules that have since become very unproductive, they can also preserve the record of older orderings, even after the formerly later rule has since become less productive than the formerly earlier one, and hence has been reordered before it. Within Double Lookup such cases are thus theoretically no different from any other sort of prepatterning, from lexical exceptions (really cases of "pre-nonpatterning") to the prepatterning involved in standard rule ordering. The difference is only that here what has been prepatterned is the output of two ordered rules, one of which has since been reordered as a later rule. Such a situation does not seem to be very common. Still, it is worthwhile to examine some clear cases of this to see exactly how they manage to mimic real cyclicity as closely as they do.

The organization of this section is as follows. In 5.2.1 I consider cases from English and other languages that require fake cyclicity. I show that real cyclicity will not work here because the rules in question are not in fact identical, contrary to what has been claimed for some of these cases. I argue that both the similarity of the rules as well as their interaction with morphology should be understood in diachronic, not synchronic terms. The existence of such cases, as well as the fact that fake cyclicity can account for all apparent cases of real cyclicity, makes fake cyclicity a more parsimonious hypothesis than real cyclicity. In 5.2.2 I consider cases involving true idiosyncratic rule ordering. Their existences show
that there are rule ordering paradoxes that cannot be accounted for
with real cyclicity. Finally, in 5.2.3 I summarize the arguments made
in this section.

5.2.1 Real cyclicity and fake cyclicity

As noted above, all reanalyses of rule ordering paradoxes
involving real cyclicity can also be handled by fake cyclicity. That is,
every apparent case of the repeated application of a single rule can
be analyzed as single applications of distinct though formally
identical rules. Since, as I will show, fake cyclicity is a genuine
phenomenon, real cyclicity therefore becomes redundant.

I am hardly the first to comment on the necessity of fake
cyclicity; several other researchers have noticed it as well, though
they apparently were not aware of each other's work. Following an
unpublished paper by Foley on Spanish, Bley-Vroman (1975) calls
the separate applications of the distinct but very similar rules he
discusses in Icelandic an interrupted rule schemata. Working
independently, Robinson (1976) calls a case of fake cyclicity in a
dialect of Swiss German a "scattered" rule. Hooper (1976) shows that
patterns in Spanish dialects argued by Harris (1969, 1973) and Reyes
(1972) to show evidence of cyclicity actually involve what I would
call prepatterning and fake cyclicity. Rice (1980), also apparently
unaware of earlier work, discusses a case in the Athapaskan
language of Hare. Finally, again working independently, Cole (1990)
document several more such cases, including stress in Chamorro and
syncope and elision rules in dialects of Bedouin Arabic.
My claim that very similar rules can nevertheless be distinct finds a parallel in the necessity in Lexical Phonology of postulating the existence of distinct affixes, associated with distinct levels, whose meaning and phonological form are nevertheless very similar. We saw this earlier in the discussions of -ism in Chapter 1 and prefixes like re- in Chapter 2. Because Lexical Phonology assumes that different levels are associated with a different set of affixes, so that different affixes may trigger a different set of phonological rules, cases where an affix appears to trigger different phonological rules in different words (eg, in resume the prefix re- triggers s-Voicing while in reseek it does not) must be analyzed instead as involving distinct affixes that only appear to be the same (ie, there really are two prefixes, re-1 and re-2). If such redundancy is allowed in morphological rules, there is no a priori reason to expect it to be disallowed in phonological rules as well.

The fact that the rules in unambiguous cases of fake cyclicity are not phonologically identical shows that they cannot be reduced to a single rule. The question of why they nevertheless appear so similar is thus, I claim, a matter of concern to theories of diachronic sound change and not theories of synchronic phonology. I will suggest two different diachronic mechanisms that may give rise to fake cyclicity. Interestingly, both predict the interaction with morphology that is characteristic of cyclicity.

I begin in 5.2.1.1 by considering the two versions of Nasal Assimilation in English which in Chapter 2 I called the obligatory and
optional versions. I first review the evidence showing we cannot collapse them together as involving either a single rule application or the repeated applications allowed by Automatic Feeding. I then show that the different properties of the rules are not in fact predictable from properties of the levels where they are thought to apply, contrary to what Borowsky (1986) has claimed. Thus these two rules are in fact distinct. I argue that the rules are historically related in that the earlier rule is a prepatterned reflex of the ancestor of the later rule. Similar arguments will be sketched for the obligatory and optional versions of Palatalization.

In 5.2.1.2 I examine the interaction of the rules of Syncope and Umlaut in Icelandic, which Kiparsky (1984) argues involves cyclic rule application. Following Bley-Vroman (1975), however, I show that the case actually involves fake cyclicity in that the version of Umlaut applying before Syncope is distinct from that applying afterwards. Bley-Vroman (1975) argues that this situation arose through a different diachronic mechanism than the one I postulate for Nasal Assimilation and Palatalization. In this case, the version of Umlaut that follows Syncope is a historically later addition to the language, added when surface patterns produced by the earlier rule were reanalyzed and extended.

In 5.2.1.3 I examine a similar case, namely the ordering of two versions of a rule of vowel lowering, one before and one after a rule of umlaut, in Schaffhausen, a dialect of Swiss German discussed by Robinson (1976). As Robinson (1976) shows, the historical origin of
this situation is essentially the same as that independently argued by Bley-Vroman (1975) for the Icelandic case.

Finally, in 5.2.1.4 I summarize the arguments made in this section, focusing especially on the fact that Double Lookup predicts that simple diachronic mechanisms automatically lead to fake cyclicity. The stipulations that other theories like Lexical Phonology must make to account for apparent cyclic effects (e.g., the division into cyclic and noncyclic strata) are therefore unnecessary.

5.2.1.1 Nasal Assimilation and Palatalization in English

In Chapter 2 I argued that in English there are two rules assimilating nasal consonants to the place of the following consonant, which I called Obligatory and Optional Nasal Assimilation. In this section I show that these two rules are indeed separate rules, not two applications of the same rule, as has been argued. I then briefly argue that the obligatory and optional rules of Palatalization, discussed in Chapters 2 and 4, respectively, are also two separate rules for essentially the same reasons.

Borowsky (1986) claims that Obligatory and Optional Nasal Assimilation are actually two applications of a single rule, one applying within the lexical strata and the other applying postlexically. The differences in their properties are argued to follow from independently motivated differences between lexical and
postlexical rules. I will show that in fact not all of the properties are predictable, implying that they are indeed two separate rules.

The differing properties that the cyclic analysis would expect to be predictable are as follows. First, Obligatory Nasal Assimilation is obligatory, while Optional Nasal Assimilation is not. Second, the optional rule is affected by temporal closeness, while the obligatory one is not. Third, the optional rule can apply across word boundaries while the obligatory one cannot. Fourth, the optional rule is feature changing while the obligatory one is not. And finally, the obligatory rule, being a lexical rule and therefore only able to derive phonemes, allows the creation of [n] and [m] but not [ŋ], while the optional rule, being postlexical, creates all equally easily.

Before considering whether Lexical Phonology can predict all of these differences, recall what Double Lookup says about these rules. In Chapter 4 I argued that the ordering of Prevelar Raising before Optional Nasal Assimilation was due to the earlier rule being less productive, since it is restricted to apply within word boundaries. What this means, then, is that Prevelar Raising is prepatterned, while Optional Nasal Assimilation applies on-line. Since Obligatory Nasal Assimilation precedes Prevelar Raising, Obligatory Nasal Assimilation must also be prepatterned. Hence according to Double Lookup, the psychological status of the obligatory and optional versions of Nasal Assimilation are very different.

3Nasal assimilation in Catalan is analyzed in a similar way in Kiparsky (1985). See Goldsmith (1990) and Iverson (1991) for critical reanalyses of this case.
This different psychological status predicts several but not all of the differences between Obligatory and Optional Nasal Assimilation. The fact that Optional Nasal Assimilation applies on-line means that we expect it to be influenced by factors relevant to articulation such as temporal closeness. By contrast, since Obligatory Nasal Assimilation does not apply on-line, we expect factors like temporal closeness to be irrelevant, as they are. The ability of the optional rule to apply across word boundaries is also expected, since word sequences are novel forms; the obligatory rule, not being fully productive, should not be able to do this.

But not all the properties distinguishing the obligatory and optional rules are predictable within Double Lookup. First, Double Lookup cannot predict whether or not a prepatterned rule will be feature-changing. This is because claims about the ways in which prepatterned rules, which do not apply on-line, may affect representations are really claims about how forms are stored in memory. Without a theory of what forms look like in memory, there is no reason to expect patterns across forms in memory to be restricted to feature fill-in patterns as opposed to feature-changing patterns.

Second, for similar reasons, Double Lookup cannot predict the fact that Obligatory Nasal Assimilation cannot derive [ŋ] in prefixed forms. This is because such a constraint must again be understood not as a constraint on rule application, but on how forms are stored in memory, which strictly speaking is beyond the domain of Double
Lookup. As we saw in Chapter 2, of course, Obligatory Nasal Assimilation is perfectly free to derive [t] in monomorphemic forms, meaning the explanation for the nonobligatory appearance of prefixes like [kaŋ] must be due to something about the prefixes, not the rule. Specifically, an independently motivated assumption concerning affix allomorphy allows us to explain the nonobligatory appearance of [kaŋ] as being due directly to the marginal status of the /t/ phoneme. In this view, there is actually no difference at all between Obligatory and Optional Nasal Assimilation with regard to the creation of this segment.

Finally, and most interestingly, it is not true that all on-line rules are optional while all prepatterned rules are obligatory. The point that on-line rules can be obligatory is made by Donegan and Stampe (1979). One example both they and Selkirk (1983) cite as obligatory is Flapping, which in the dialects that have it always applies when it can (unless it is bled by the nonapplication of Regressive Nasalization; see Chapter 4). Moreover, rules that Double Lookup would claim are prepatterned need not be obligatory. One striking example of an optional rule that it is at least partially prepatterned, Coronal Deletion, will be discussed later in this chapter. Other examples mentioned in this dissertation include Resyllabification (Chapter 2), Vowel Reduction in English (Chapter 6) and i-Spirantization in Finnish (this chapter). In Chapter 6 I claim that the existence of optional prepatterned rules in fact argues for the superiority of Double Lookup over Lexical Phonology.
I now return to the question of whether Lexical Phonology can predict all of the differences between Obligatory and Optional Nasal Assimilation. The answer is no, and for the same reasons that Double Lookup cannot predict them either. Like Double Lookup, Lexical Phonology can predict the fact that the obligatory rule is word-internal while the optional rule is not (lexical rules apply word-internally, postlexical rules apply across word-boundaries), the fact that the optional rule is sensitive to temporal closeness (assuming with Mohanan 1982 and Shattuck-Hufnagel 1986 that postlexical rules apply on-line; see Chapter 6), and the fact that the optional rule is feature-changing (all postlexical rules are). Like Double Lookup, however, Lexical Phonology cannot predict the difference in optionality (Flapping is an obligatory postlexical rule, Coronal Deletion is an optional lexical rule) and the fact that the obligatory rule is a feature fill-in rule (Velar Softening, as formulated in Chapter 2, is a feature-changing lexical rule).

To summarize, the properties distinguishing Obligatory and Optional Nasal Assimilation, unpredictable for both Double Lookup and Lexical Phonology, are listed below.

(2) **Obligatory Nasal Assimilation**
- obligatory
- feature fill-in

**Optional Nasal Assimilation**
- optional
- feature-changing

Thus we are forced to conclude that Obligatory Nasal Assimilation and Optional Nasal Assimilation are not reducible to a single rule that applies twice in the derivation. The question then
arises where two such similar rules come from if they are not reducible to one rule in the synchronic phonology.

I suggest that the rules are in fact reducible to one, but only diachronically. Rules assimilating nasals to the place of following consonants are among the most common fully productive rules in the world's languages. Given this near universal, it is expected that even dead languages should show evidence of it. In particular, we expect that older varieties of English, in particular varieties of English before the introduction of Noncoronal Deletion and Prevelar Raising, had a rule essentially identical to the modern rule of Optional Nasal Assimilation. As it became prepatterned over time, presumably due to language learners taking forms derived on-line in the adult grammar as underlying representations, Double Lookup would predict that it would appear to become embedded more deeply in the phonology, now being ordered before both the more recently added rule of Prevelar Raising and the currently still very productive rule of Noncoronal Deletion. Being prepatterned, this less productive version would be restricted to word-internal application. Meanwhile, however, a version of the rule remained fully productive, due to the phonetic naturalness and universality of nasal assimilation rules. It is therefore not surprising that we should find both a prepatterned and on-line version of the same process.

4Recall the discussion concerning the origin of prepatterning in Chapter 1, as well as the comment of Kiparsky (1982) quoted above in 5.1.
The same can be said of the obligatory and optional rules of Palatalization in English, discussed in Chapters 2 and 4, respectively. Rubach (1984), Halle and Mohanan (1985), Kaisse (1985) and Borowsky (1986) all assume that these rules are different applications of the same rule, with the differences in their lexical and postlexical applications being predictable from independently motivated principles. However, as with Nasal Assimilation, this does not seem to be true. Among other things, the two rules differ in optionality, which is not predictable, as I showed above. Thus we must conclude that these two rules cannot be reduced to a single rule synchronically. Diachronically, however, both rules may be thought of as having a common source, because like nasal place assimilation, palatalization is an extremely natural and ubiquitous rule.

I conclude, therefore, that neither Nasal Assimilation nor Palatalization in English present evidence that a single rule can apply in two different places in a derivation. Instead we must assume that there are two rules of Nasal Assimilation and two of Palatalization that happen to be historically related. In this case the separation of the rules resulted from a very natural, virtually universal rule being added to the end of a derivation that already contained a less productive phonologization of the same natural process, added earlier. The common source for both rules, then, is phonetics. In the following sections I examine cases that cannot be explained this way, as the rules in question are relatively phonetically unnatural, or at
least far from ubiquitous. In these cases, however, we will see that there is evidence for an alternative diachronic explanation.

5.2.1.2 Syncope and Umlaut in Icelandic

Kiparsky (1984) uses an analysis involving cyclicity to solve a rule ordering paradox in Icelandic discussed in Anderson (1974). The rule of Syncope, which deletes certain lax unstressed vowels before a syllable beginning with certain consonants, and the rule of u-Umlaut, which changes /a/ to [ö] before a following /u/, appear to be ordered paradoxically, since with some suffixes Syncope precedes u-Umlaut, while with others the order is reversed. Anderson (1974) takes this as support for Local Ordering, which allows just this sort of apparently lexically marked ordering of rules. I illustrate these rules, and the rule ordering paradox associated with them, in the figures below.

(3) Two rules in Icelandic [from (4) - (7) in Anderson 1974:141-142]:

a. **u-Umlaut:** \( a \rightarrow \ddot{o} / \_ C_0 u \)

\[
\begin{align*}
\text{barn} & \quad '\text{child}' \\
(\text{ég}) & \quad \text{ka}l\text{la} \quad '\text{I call}' \\
\text{börnum} & \quad '\text{dat. pl.}' \\
(\text{viô}) & \quad \text{köllum} \quad '\text{we call}'
\end{align*}
\]

b. **Syncope:** \( V \rightarrow \emptyset / C \_ C + V \)

\[
\begin{align*}
\text{hamr} & \quad '\text{hammer}' \\
\text{morgun} & \quad '\text{morning}' \\
\text{hamri} & \quad '\text{dat. sg.}' \\
\text{morgni} & \quad '\text{dat. sg.}'
\end{align*}
\]
(4) Syncope precedes u-Umlaut:

a. [from (8) in Anderson 1974:142-143]

<table>
<thead>
<tr>
<th>word</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>alin</td>
<td>'ell of cloth'</td>
</tr>
<tr>
<td>regin</td>
<td>'(the) gods'</td>
</tr>
<tr>
<td>ölnum</td>
<td>'(dat. pl.)'</td>
</tr>
<tr>
<td>rögnun</td>
<td>'(dat. pl.)'</td>
</tr>
</tbody>
</table>

b. Underlying: alin + um
   Syncope: alnum
   u-Umlaut: ölnum

(5) u-Umlaut precedes Syncope:

a. [from (9) in Anderson 1974:143]

<table>
<thead>
<tr>
<th>word</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>baggi</td>
<td>'pack, bundle'</td>
</tr>
<tr>
<td>böggull</td>
<td>'parcel, package'</td>
</tr>
<tr>
<td>böggli</td>
<td>'parcel, package (dat. sg.)'</td>
</tr>
<tr>
<td>jaki</td>
<td>'piece of ice'</td>
</tr>
<tr>
<td>jökull</td>
<td>'glacier'</td>
</tr>
<tr>
<td>jökli</td>
<td>'glacier (dat. sg.)'</td>
</tr>
</tbody>
</table>

b. Underlying: bagg + ul + i
   u-Umlaut: bögguli
   Syncope: böggli

The paradox is resolved, however, if it is supposed that affixes like -um and affixes like -ur apply at different levels, with the -ur level ordered before the -um level; each level can be associated with both Syncope and u-Umlaut, ordered with Syncope before u-Umlaut. This analysis is given credence by the fact that -ur is a derivational suffix and -um is an inflectional suffix; in Icelandic, inflectional affixes are added outside derivational affixes, implying that their affixation must be ordered this way in any case. The derivations of the forms ölnum and böggli can therefore be handled as in (6).
The foremost advantage of this analysis over that provided by Local Ordering is its exploitation of the independently motivated ordering of derivation before inflection. Note, however, that the ordering of derivation before inflection is also predicted by the Productivity Hypothesis, assuming Icelandic is like English and its derivational morphology is less productive than its inflectional morphology.\(^5\) The phonological rules associated with these different morphological processes must therefore also differ in their productivity. This implies that the different applications of u-Umlaut are actually distinct rules, so that this is a case of fake and not real cyclicity.\(^6\)

\(^5\)Evidence that this is the case in Icelandic is seen in the semantics of the derivational and inflectional suffixes illustrated above. While the inflectional suffixes all mark clear syntactic features, the meaning of the derivational suffix \(-ul\) is rather obscure: it changes "pack" to "package" but also "piece of ice" to "glacier". Recall from the discussion of Aronoff (1976) in Chapter 3 that semantic coherence can be used to determine the productivity of morphological rules.

\(^6\)Jennifer Cole (personal communication) has suggested that this rule ordering paradox can be resolved without recourse to rule ordering at all if it supposed that Syncope deletes not the entire vowel, but just its prosodic slot, leaving the features, particularly [+round], intact. Rounding would thus be able to spread in both \(\text{ölnum}\) and \(\text{böggli}\) even after Syncope has ostensibly removed the conditioning environment. I reject this analysis for two reasons. First, it ignores the fact that the ordering of derivation and inflection in

\begin{align*}
\text{cycle 1:} & \quad \text{Syncope} & \text{aln} & \text{bagg} + \text{ul} \\
& \quad \text{u-Umlaut} & - & \text{-} \\
\text{cycle 2:} & \quad \text{morphology} & \text{aln} + \text{um} & \text{böggul} + \text{i} \\
& \quad \text{Syncope} & \text{aln} + \text{um} & \text{böggul} + \text{i} \\
& \quad \text{u-Umlaut} & \text{öln} + \text{um} & \text{-}
\end{align*}
Striking independent support for this view comes from the work of Bley-Vroman (1975). He shows that the sources of the two versions of Umlaut in Modern Icelandic discussed by Kiparsky (1984), namely their ancestors in Old Icelandic, were actually two separate rules with distinct phonological effects.\(^7\)

Bley-Vroman (1975) criticizes the Local Ordering analysis of Icelandic for very different reasons from those offered by Kiparsky (1984). The problem with Anderson's (1974) analysis, Bley-Vroman (1975) claims, is that he only considered the ordering of Syncope with u-Umlaut. Although the version of u-Umlaut preceding Syncope does indeed appear to be formally the same as the version following, there is another aspect of the umlauting process, namely \(i\)-Umlaut, that is ordered paradoxically with Syncope as well, in precisely the same way. When \(i\)-Umlaut is considered, it is clearly seen that the version preceding Syncope is phonologically distinct from that following Syncope. The similarity of the two versions of u-Umlaut must thus be left as a synchronically uninteresting historical relic. This is especially true if, as Bley-Vroman (1975) assumes, \(i\)-Umlaut and u-Umlaut can easily be collapsed into a single rule of Umlaut.

---

Icelandic is independently motivated; this ordering should show up as productivity differences between the two versions of u-Umlaut, which in Cole's analysis are treated as if they applied simultaneously. Second, as we shall shortly see, the Umlaut rule before Syncope and the Umlaut rule after Syncope are not phonologically identical. In other words, Kiparsky's (1984) example is fake cyclicity, not real cyclicity.

\(^7\)Although Bley-Vroman (1975) uses exclusively Old Icelandic examples in his discussion, he explicitly intends his arguments to apply to Modern Icelandic as well, as his discussion of Anderson (1974) shows.
The apparently paradoxical ordering of Syncope with i-Umlaut, which fronts a stressed vowel if followed by /i/, is illustrated in (7).

(7) [after Bley-Vroman 1975:74]:

<table>
<thead>
<tr>
<th></th>
<th>Underlying:</th>
<th>Syncope:</th>
<th>i-Umlaut:</th>
</tr>
</thead>
<tbody>
<tr>
<td>a.</td>
<td>vake + ëi</td>
<td>vak_ + ëi</td>
<td>vek + ëi</td>
</tr>
<tr>
<td></td>
<td>'wake (3rd past obl.)'</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b.</td>
<td>do:mi + ëa</td>
<td>do:mi + ëa</td>
<td>dö:mi + ëa</td>
</tr>
<tr>
<td>i-Umlaut:</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Syncope:</td>
<td>tal_ + ëa</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The figure in (7) is misleading, however, since the rules labelled "i-Umlaut" are actually phonologically distinct. The rule following Syncope, call it i-Umlaut-2, umlauts a vowel regardless of its length, as seen in (7a), where it umlauts a short vowel. The rule preceding Syncope, call it i-Umlaut-1, however, only umlauts long vowels. This can be seen by contrasting the form in (7b) with that given below in (8), both of which involve exactly the same suffix.

(8) [after Bley-Vroman 1975:74]

| Underlying: | tali + ëa  | 'tell (1st sg past)' |
| i-Umlaut:   |            |                    |
| Syncope:    | tal_ + ëa  |                    |

Bley-Vroman (1975) concludes, therefore, with the following synchronic description of rule ordering in Old Icelandic.

---

As far as I can tell, the vowels at the end of the verb stems in these examples from Old Icelandic can be considered theme vowels, similar to those found in modern Spanish. They are thus a form of derivational affix.
1. Umlaut with the $i$ case limited to long cluster forms.
2. Medial syncope.
3. Umlaut generally -- no special conditions.

The similarity of the two rules in such interrupted rule schemata can be understood if one considers the historical changes that lead to such a situation in the first place. Bley-Vroman (1975) gives evidence that the rules of $i$-Umlaut-1, Syncope and $i$-Umlaut-2 were added to Old Icelandic in precisely the same order as synchronically represented in (9). What can account for this ordering?

Bley-Vroman (1975) notes that at the stage where Syncope has been added, the earlier rule of $i$-Umlaut-1 became more opaque. This is because in some words Syncope removes the conditioning environment for $i$-Umlaut-1 while in others it creates the conditioning environment, leaving surface "exceptions" like [dö:ma], which should not undergo Umlaut but do, and [vakði], which should undergo Umlaut but don't. The rule of $i$-Umlaut-1 thus becomes harder to learn. If we assume that in other cases its effects remained clear, so that learners at some point simplified it to apply to short vowels as well as long, the addition of a new general $i$-Umlaut-2 rule is not surprising. The reason the restricted $i$-Umlaut

9Suggestive evidence for this is also found in Anderson (1975:143, figure (8)). There it is noted that álnum, where u-Umlaut-2 has applied, is in fact merely an alternate pronunciation of álnum. By contrast, all forms showing evidence of u-Umlaut-1 surface only with umlaut. This implies that u-Umlaut-2 has not fully spread through the lexicon, and is therefore a more recent addition to the language than u-Umlaut-1.
rule did not simplify *in situ*, Bley-Vroman (1975) argues, is because doing so would have actually had the effect of making it more opaque, since it would have increased the number of forms that disobey Umlaut on the surface, through the addition of forms like [telbɑ].

Note that according to Double Lookup, this should be true in general: non-surface rules cannot be simplified *in situ*, but can only simplify through the addition of a simpler version of themselves at the end of the derivation. But the significance of this historical observation goes beyond this. As Bley-Vroman (1975:78) emphasizes, "[t]his view of phonological change makes sense only in a framework which views diachronic 'generalization' not in terms of the form of a rule, but in terms of its phonetic effects...." This means that a non-surface but productive rule can act to simplify and generalize its pattern on the surface not only through the addition of a simplified copy of itself, but also through the addition at the end of the derivation of an *exact* copy of itself. Thus cases that seem to involve repeated applications of precisely the same rule can arise historically in exactly the same way as interrupted rule schemata. The only empirical evidence that such cases involve fake cyclicity and not real cyclicity would come from their historical ordering and differences in the rules' productivity.

---

10 See Chapter 6 for further discussion of the generalization that new rules are added at the end of the derivation.
The difference in productivity between derivational and inflectional affixes is also clearly relevant for the Double Lookup analysis. At the historical stage where Umlaut-2 was added to the language, Umlaut-1 had already been prepatterned. The fact that Umlaut-1 is associated with the less productive derivational morphology is therefore not surprising. The fact that it was associated with less productive morphological processes may even have been a factor in its becoming prepatterned in the first place. The close relation between apparent cyclicity and the relative productivity of the associated morphological rules will be seen again in the discussion of the Swiss German dialect of Kesswil in 5.2.2.2.

I have shown in this section that an apparently clear example of real cyclicity, that of Umlaut and Syncope in Icelandic, is actually a clear example of fake cyclicity. Moreover, I argued that the most powerful argument offered in favor of the cyclic analysis, namely the independent necessity of ordering derivation before inflection, is in fact an argument in support of the Double Lookup analysis. Further support for this analysis comes from the historical order with which the different versions of Umlaut were added to the language.

5.2.1.3 o-Lowering and Umlaut in Schaffhausen

The same is true of the rules in Swiss German discussed by Robinson (1976). In some Schaffhausen dialects, there was a rule of Lowering, whereby /o/ becomes /ɔ/ before all coronal obstruents and also /r/. This rule seems to be ordered paradoxically with respect to the morphological rule of Umlaut, which fronts the vowel
of the stem in a variety of morphological environments. The problem is that while Umlaut usually changes the /ɔ/ created by Lowering into [ɔ], implying that Umlaut is ordered before Lowering, the /ɔ/ created by a following /r/ is changed by Umlaut into [œ], implying that this subpart of Lowering is actually ordered before Umlaut. The contrast is presented below in (10a) and (10b).

(10) a. [after (7) in Robinson 1976:149]

<table>
<thead>
<tr>
<th>English</th>
<th>Norwegian</th>
</tr>
</thead>
<tbody>
<tr>
<td>bode</td>
<td>bøde</td>
</tr>
<tr>
<td>trotte</td>
<td>trøtli</td>
</tr>
<tr>
<td>götte</td>
<td>götti</td>
</tr>
<tr>
<td>'floor'</td>
<td>'floors'</td>
</tr>
<tr>
<td>'sidewalk'</td>
<td>'little sidewalk'</td>
</tr>
<tr>
<td>'godmother'</td>
<td>'godfather'</td>
</tr>
</tbody>
</table>

b. [after (8) in Robinson 1976:149]

<table>
<thead>
<tr>
<th>English</th>
<th>Norwegian</th>
</tr>
</thead>
<tbody>
<tr>
<td>torn</td>
<td>torn</td>
</tr>
<tr>
<td>xorrb</td>
<td>xorblî</td>
</tr>
<tr>
<td>fortø</td>
<td>fôris</td>
</tr>
<tr>
<td>'thorn'</td>
<td>'thorns'</td>
</tr>
<tr>
<td>'basket'</td>
<td>'little basket'</td>
</tr>
<tr>
<td>'Scots pine'</td>
<td>'wood from Scots pine'</td>
</tr>
</tbody>
</table>

As the examples in (11) show, it is not the case that an underlying /œr/ surfaces as [œr]. Rather, only when umlauting is triggered morphologically does /or/ become [œr]. Thus the pattern in (10b) cannot be explained by adding a stipulation allowing the Lowering rule to affect /œ/ as well as /o/, when followed by /r/.

(11) [after Robinson 1976:149]

<table>
<thead>
<tr>
<th>English</th>
<th>Norwegian</th>
</tr>
</thead>
<tbody>
<tr>
<td>töre</td>
<td>'dare'</td>
</tr>
<tr>
<td>törffø</td>
<td>'be allowed to'</td>
</tr>
<tr>
<td>dort</td>
<td>'there'</td>
</tr>
<tr>
<td>jwörsedel</td>
<td>'plant name'</td>
</tr>
</tbody>
</table>
The ordering of the rules, then, must be as in (12), with a special rule of r-Lowering preceding Umlaut, while general Lowering follows.

(12) /bodə/ /bodə/ + PL /torn/ /torn/ + PL
r-Lowering -- -- torn torn + PL
Umlaut -- böde -- torn
Lowering böde -- --

This case is thus synchronically exactly the same as the Old Icelandic case discussed by Bley-Vroman (1975), in that the later Lowering rule is a more general version of the earlier one. Robinson (1976) is also able to show that, as with the Old Icelandic case, the more general version of the rule is a historically later development. One difference between the Swiss German and Icelandic cases is the fact that in German, the Umlaut rule is demonstrably far older than either Lowering rule. Robinson (1976) presents a chronology where r-Lowering is first added after Umlaut and then is reordered before it. The more general Lowering rule is then added after Umlaut, parallel to the Old Icelandic case.

5.2.1.4 Fake cyclicity: Summary

We have seen that fake cyclicity can arise in at least two different ways. First, fake cyclicity can arise as I argued was the case with the two rules of Nasal Assimilation, namely a very productive rule can become prepatterned and yet remain very productive, resulting in an on-line version of the rule being ordered after a prepatterned version of the same rule. In this section I have discussed cases where the prepatterned rule is slightly different
from the on-line rule; in 5.3 I discuss this phenomenon more thoroughly, showing that cases exist where the prepatterned and on-line versions appear to be identical. The other way fake cyclicity can arise was illustrated with cases from Icelandic and Swiss German, where it appears that a less productive rule, ordered before a more productive one, has left a pattern on the surface that has led a later generation to induce a new version of the rule, adding it, as with all productive rules, at the end of the derivation.

In both situations, however, the end result is the same: the ordering of a more productive rule after a less productive version of itself. Fake cyclicity can thus be thought of as an expected and natural consequence of the principles we have been discussing up to this point. In Chapters 2 and 3 I discussed rules that tend not to apply on-line; all of these rules are very prepatterned, and their ordering is therefore primarily a consequence of the relative productivity they had in the past. In Chapter 4 I discussed rules that are fully productive; these rules apply on-line and are thus ordered in accordance with the principles of Prosody First and Automatic Feeding. Now in this chapter I discuss cases of rules which are (or, in cases like Icelandic or Swiss German, were) both prepatterned and fully productive.

5.2.2 Idiosyncratic ordering

However, there still remain rule ordering paradoxes which cyclicity, whether fake or real, cannot handle. These are cases of
idiosyncratic ordering, where individual lexical items, not affixes or affix classes, choose a particular ordering.

In Chapter 2 we already saw a possible example of this in English. In the word *equation*, the underlying /t/ becomes /s/ by Spirantization; this /s/ then appears to be voiced by s-Voicing, whose triggering environment is precisely met. The problem is that in all other relevant words, such as *relation*, s-Voicing does not apply, a fact that is standardly taken to mean that s-Voicing is ordered before Spirantization. The word *equation* therefore appears to be a case of a lexical item marked to undergo rules in a special order.\(^\text{11}\)

I discuss two further examples of this, namely Contraction and i-Spirantization in Finnish and a-Rounding and Umlaut in the Swiss German dialect of Kesswil. In both cases, the idiosyncratic ordering appears to have arisen as a side effect of rule reordering. That is, the lexical items that show the more unusual ordering are in fact preserving, through prepatterning, the historically older ordering pattern. The Kesswil case is especially interesting in that the conservative forms all involve less productive morphology than the forms with the more common rule ordering pattern. This is

\(^{11}\) Another possible example of this in English is seen in the pair *pharmacological-pharmaceutical*. The surface [u] of *pharmaceutical* should not trigger Velar Softening. The appearance of [s] in this word can be explained, however, if we suppose that first /y/ is inserted before /u/ by y-Insertion, and that it is this /y/ that triggers Velar Softening. The /y/ is then deleted by general processes disallowing /y/ after coronals. It is standardly assumed, though, that Velar Softening precedes y-Insertion, on the assumption that Velar Softening precedes Vowel Shift, which in turn is thought to precede y-Insertion.
completely consistent with Double Lookup, in that it is the forms with less productive morphology that show prepatternning.

5.2.2.1 Contraction and i-Spirantization in Finnish

A typical example of idiosyncratic ordering is presented in Finnish by Kiparsky (1973b:92-101), although he does not analyze it as such. Finnish has a rule of Contraction, whereby the vowel of a verb stem is deleted before the past tense suffix -i, so that /a + i/, /ä + i/ and /e + i/ all become [i]. This rule interacts with a rule of i-Spirantization, whereby /t/ becomes /s/ before the past tense suffix -i; other suffixes beginning with /i/, such as the past impersonal, the conditional, and plural markers, do not have this effect.

The strange thing about this interaction is the following. If the verb stem ends in /t/, the past tense suffix always triggers Spirantization, as illustrated in (13). However, if the verb stem ends in a vowel that is subject to Contraction, whether or not the Spirantization rule applies depends on the individual verb. In verbs such as the one in (14a), Spirantization always applies. In verbs such as the one in (14b), Spirantization never applies. Finally, in verbs such as the one in (14c), Spirantization is optional.

(13) [after Kiparsky 1973b:95]

/\textit{halut+i}/ $\rightarrow$ halusi "wanted"
/\textit{hakkat+i}/ $\rightarrow$ hakkasi "hewed"
/\textit{turpot+i}/ $\rightarrow$ turposi "swelled"
a. /piirtä+i/ → piirsi "drew"
b. /pità+i/ → piti "held"
c. /kiitä+i/ → kiiti "sped" kiisi

Since Spirantization is morphologically conditioned, it is, in the terminology of Lexical Phonology, a lexical rule; it thus represents yet another example of an optional lexical rule. But what is especially interesting about this case is the implications it has for rule ordering.

As Kiparsky (1973b) notes, there are two ways the interaction between Contraction and i-Spirantization could be handled synchronically. One would be to order Contraction before i-Spirantization. The exceptional forms in (14b,c) would have to be marked as lexical exceptions to Spirantization. The other option would be to order i-Spirantization before Contraction. In this case, the form in (14b) would be the expected pattern, and the apparent application of i-Spirantization in (14a,c) would have to be accounted for by another rule ordered after Contraction.

The strange thing is that both of these analyses are supported by the historical evidence. In support of the ordering of Contraction before i-Spirantization is the observation that it is /s/ forms like those in (14a,c) that are more recent, implying that i-Spirantization has begun to affect forms that have already undergone Contraction. In other words, i-Spirantization has become more productive than Contraction. At the same time, Kiparsky (1973b) gives numerous
arguments for supposing that i-Spirantization is historically older than Contraction, so that the form in (14b) which fails to undergo i-Spirantization was, at a time before the introduction of Contraction, a fully regular form.

What appears to have happened is this. After the rule of i-Spirantization was added to the language, Contraction was added at the end of the derivation, in accordance with King's (1973) generalization about rule additions (see Chapter 6). At that stage, then, the synchronic ordering of the rules was with i-Spirantization before Contraction. Since that time, however, the rules have begun to reorder themselves. In terms of Double Lookup, this can be understood as i-Spirantization somehow becoming more productive than Contraction, perhaps because it is less opaque.

This rule reordering need not occur all at once all across the lexicon; in fact, Double Lookup does not allow rule reorderings to occur all at once. Rules can only be reordered instantaneously if their existence is fully independent of the lexicon, that is, if the patterns they represent are stored entirely as rules and not as prepatterned forms in memory. However, according to Double Lookup, such rules cannot be ordered in the first place, since they would then both be fully, hence equally, productive. Thus Double Lookup predicts that rule reorderings can only occur through lexical diffusion. This means that any intermediate stage between the two ordering patterns, there will be some lexical items, typically the most
frequent (being the most resistant to on-line rule application), that continue to preserve the older order.

This appears to be precisely what is occurring in Finnish. In verbs such as that in (14a), the synchronic order of rule application is with Contraction before i-Spirantization. This is the most typical pattern since i-Spirantization is synchronically more productive than Contraction. In verbs such as that in (14b), the synchronic order of rule application is with i-Spirantization before Contraction. This does not mean that this ordering is part of a speaker's knowledge about this word, only that this ordering is preserved in its prepatterned form stored in memory. This is not a special situation within the model of Double Lookup; the ordering of rules that typically do not apply on-line is not part of one's knowledge in general, even if all relevant words order the rules this way. Finally, in verbs such as that in (14c), there are at least two forms stored, one with the ordering as in (14a) and one with the ordering as in (14b).

Derivations illustrating this analysis are given below in (15).

(15) a.  
Contraction /piirtä+i/  
i-Spirantization piirti  
           piirsi

b.  
i-Spirantization /pitä+i/  
           --  
Contraction piti
c. /kiitä+i/
   Contraction: kiiti
   i-Spirantization: kiisi

/iitä+i/
   i-Spirantization: --
   Contraction: kiiti

In explaining the synchronic pattern, this analysis is effectively the same as an analysis where Contraction is ordered before i-Spirantization for all words, and words like *pita* are simply marked as lexical exceptions. The Double Lookup analysis has an advantage, however, when considering the diachronic pattern, since it accounts for the source of such exceptions in the first place.

5.2.2.2 a-Rounding and Umlaut in Kesswil

Another case showing the same thing is found in Swiss German dialects like Kesswil (Robinson 1977).12 This case is worth discussing, however, since unlike the Finnish case, at first glance it appears to involve real cyclicity, since the forms that preserved the older ordering tend to fall into different morphological categories, with words containing derivational suffixes being more conservative than words containing inflectional suffixes. The existence of minimal pairs involving precisely the same suffix, however, shows that neither a real cyclicity nor a fake cyclicity analysis will work.

The two relevant interacting rules in Kesswil are a-Rounding, whereby an underlying long /a:/ becomes /o:/, and Umlaut, the

---

12 Robinson (1977) was the work that first led me to the reconception of cyclicity outlined in this chapter. I have Jane Tsay to thank for bringing this paper to my attention.
morphological rule discussed above, whereby vowels are fronted in certain morphologically defined environments.

In most cases, the umlauted form of \[\text{o}:] is \[\text{æ}:] as illustrated below in (16).

(16) [after Robinson 1977:73, 78]

<table>
<thead>
<tr>
<th>Word</th>
<th>Original Form</th>
<th>Umlauted Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>o:bad</td>
<td>&quot;evening&quot;</td>
<td>o:bad</td>
</tr>
<tr>
<td>rø:t</td>
<td>&quot;counsel&quot;</td>
<td>rø:t</td>
</tr>
<tr>
<td>o:ðæl</td>
<td>&quot;vein&quot;</td>
<td>o:ðælL</td>
</tr>
<tr>
<td>jɔ:ff</td>
<td>&quot;sheep&quot;</td>
<td>jɔ:ffL</td>
</tr>
<tr>
<td>nɔ:x</td>
<td>&quot;near&quot;</td>
<td>nɔ:xr</td>
</tr>
<tr>
<td>hɔ:kkə</td>
<td>&quot;hook&quot;</td>
<td>hɔ:kkəL</td>
</tr>
</tbody>
</table>

This implies that a-Rounding is ordered before Umlaut, as the derivation in (17) shows. This ordering is hardly surprising, since it appears that long [a:] is neutralized with [o:] in virtually all environments; the evidence that [a:] exists at all comes solely from distributional patterns in morphemes. This naturally makes the a-Rounding rule very difficult to learn, hence very unproductive, as a learning principle like the Alternation Condition would lead us to expect.

(17) a-Rounding: /ra:t/  Umlaut: /ra:t/ + PL

In many words, however, the umlaut of [o:] is not [æ:] but [e:], which is the historically older umlaut form of [a:]. Examples are given below in (18).
These words seem to undergo a-Rounding after Umlaut, as illustrated by the derivation in (19). As Robinson (1977) shows through a comparison with other Swiss German dialects, this is the historically older ordering pattern.

(19) /raːt/ /raːt + ɪɡ/
Umlaut -- ræːttɪg
a-Rounding rot --

This is a rule ordering paradox that seems to call for a cyclic analysis, since there is an obvious connection between morphology and the ordering of a-Rounding and Umlaut. It appears that a-Rounding precedes Umlaut for all cases involving inflection, and some cases involving derivation, particularly apparently productive derivation such as comparative and diminutive formation. By contrast, Umlaut precedes Rounding only for forms involving derivation, particularly the derivational suffix -ɪɡ. This suffix
appears to be less productive than the comparative and diminutive suffixes since it is less regular: it affixes to both nouns ("needle") and verbs ("to cook"), and does not seem to carry any specific semantic force other than "adjective vaguely related to the stem." On independent grounds, then, it is reasonable to suppose that -īg is ordered before the comparative and diminutive suffixes, as well as before inflectional processes such as pluralization. An analysis involving real cyclicity would thus appear as in (20).

(20)  

<table>
<thead>
<tr>
<th>Cycle 1:</th>
<th>/raːt/ + PL</th>
<th>/raːt + ıg/</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ıg added</td>
<td>--</td>
<td>raːt + ıg</td>
</tr>
<tr>
<td>Umlaut</td>
<td>--</td>
<td>ræːttıg</td>
</tr>
<tr>
<td>a-Lowering</td>
<td>rəːt</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cycle 2:</th>
<th>raːt + PL</th>
<th>--</th>
</tr>
</thead>
<tbody>
<tr>
<td>PL added</td>
<td>rəːt</td>
<td>--</td>
</tr>
<tr>
<td>Umlaut</td>
<td>rəːt</td>
<td>--</td>
</tr>
<tr>
<td>a-Lowering</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Unfortunately this analysis will not work for all cases, since the division between -ıg affixation and other unproductive derivational processes on the one hand and inflection and productive derivational processes on the other is not absolute. Some unproductive derivation, including -ıg , triggers the ordering of a-Lowering before Umlaut, too, just like inflection and productive derivation. Examples are given in (21).
The ordering of the rules is illustrated in (20) with the minimal pair of "hairy" and "active".

(22) a. /haːrɪg/
   a-Lowering  hɔːrɪɡ
   Umlaut      hɔːrɪɡ

   b. /taːtɪɡ/
   Umlaut      tæːtɪɡ
   a-Lowering  --

While it is true that in Kesswil all inflection triggers the ordering of a-Lowering before Umlaut, Robinson (1977:78) points out that in other dialects of Swiss German, inflection shows the ordering typical of unproductive derivation, with Umlaut first, as seen in the example in (23) from the dialect of the Bündner Herrschaft.

(23) "counsel" "counsels"\(^{13}\)
   /raːt/   /raːt/ + PL
   Umlaut   --       rɔːt
   a-Lowering rɔːt  --

It becomes clear, then, that this is not a case of cyclicity at all, but rather of idiosyncratic ordering, just like the Finnish case. Originally a-Lowering was ordered after Umlaut, being more productive; now it is apparently ordered before Umlaut, having lost in productivity. While most forms were regularized to the new

\(^{13}\)This form can also optionally appear to undergo the rules in the opposite order, coming out as [rɔːt].
pattern, some forms were conservative and retained the old ordering.

If this is not a case of cyclicity, then, what accounts for the influence of morphology, as nondeterministic as it is? My answer, unsurprisingly perhaps, involves rule productivity. Recall that the dominant ordering of a-Lowering before Umlaut tends to be associated with relatively productive morphology, such as inflection and productive derivation, while the conservative ordering of Umlaut before a-Lowering tends to be associated with relatively unproductive morphology. The phonological rule associated with this morphology, namely Umlaut, therefore varies in productivity, too: Umlaut associated with -ίς is less productive than the Umlaut rules associated with inflection, comparative and diminutive formation. This in turn means that forms containing -ίς are less likely to undergo Umlaut on-line, and will therefore tend to resist regularization to a new pattern longer. Words with -ίς thus end up being conservative, but not because they are necessarily more frequent, since what makes a word conservative is not its frequency per se but the likelihood that it undergoes a rule on-line. If words with -ίς tend not to undergo Umlaut on-line, they will end up being more conservative than other words, all else being equal, and will thus tend to preserve the older, hence prepatterned, rule ordering pattern.
5.2.2.3 Idiosyncratic ordering: summary

We see, then, that cases of idiosyncratic ordering, far from being a problem for Double Lookup or an unpleasant stipulation, are actually an expected consequence of rule reordering. In fact, as noted above, if rules come to change in relative productivity, the subsequent reordering must occur with different lexical items at different times, thus producing idiosyncratic ordering. This is because all ordering is argued to involve the earlier rule's being more prepatterned than the later. If the later rule now comes to be more prepatterned, the prepatterning on the other items will not immediately disappear until they are regularized according to the new pattern.

My analysis of such cases thus differs in a major way from those of both Local Ordering (Anderson 1974) and Lexical Phonology (eg, Kiparsky 1984). In both of these other models, rule ordering is always considered part of the speaker's knowledge. Within Double Lookup, however, forms that are describable using idiosyncratic ordering are simply stored in memory with the relevant pattern already present. They may be thought of, then, merely as lexical exceptions to the regular pattern. One advantage of the approach taken here, however, is that it accounts for where such exceptions come from in the first place.

5.2.3 Rule ordering paradoxes: summary

In this section I have shown how fake cyclicity can handle all the cases of rule ordering paradoxes analyzable with real cyclicity.
Moreover, fake cyclicity was seen to be independently necessary in other cases that real cyclicity could not handle. This includes cases, such as Nasal Assimilation in English and Umlaut and Syncope in Icelandic, that have been argued in the literature to involve real cyclicity. Finally, I discussed some rule ordering paradoxes that cannot be handled with cyclicity at all, showing that they instead involve idiosyncratic ordering, which is expected in Double Lookup given Redundancy of Patterning. Idiosyncratic ordering is especially interesting since within Double Lookup lexical items marked as if they undergo rules in an atypical order are treated theoretically no different from lexical exceptions, and yet cases such as that of Kesswil nevertheless show that this simple phenomenon may mimic real cyclicity quite closely. Such parallels get at the heart of Double Lookup: the ordering of partially productive rules, whether it is standard, "cyclic" or idiosyncratic, always depends on prepatterning.

5.3 Direct evidence for cyclic rule application

The arguments for cyclicity we have examined so far have been indirect. First was the argument that if a rule obeys the Strict Cycle Condition, it must be cyclic; this was shown to be false. Second was the argument that if rules seem to be ordered paradoxically, they must be ordered cyclically; this too was shown to be false.

Now I turn to the final sort of evidence that has been used to support claims of cyclicity, namely cases where the result of multiple applications of a single rule appear directly in individual forms. That
is, something about the surface forms shows directly that a rule has applied more than once to it.

Cole (1990) mentions several apparent examples of this type in a variety of languages, noting that for at least some of them noncyclic analyses are available. I will not discuss these here. Double Lookup would force me to predict that to the extent a cyclic-type analysis is plausible in these cases, they involve fake cyclicity and not real cyclicity.

Instead, in this section I will discuss two clear cases of apparently direct evidence for cyclicity in English. Interestingly, in both cases the different applications appear to be the same, not merely similar. I suggest that the earlier rule or rules forms part of the prepatterning retrieved in Lexical Lookup, while the later rule applies on-line in Rule Lookup. In other words, I suggest that these apparent examples of cyclicity are very similar to the case of Obligatory and Optional Nasal Assimilation discussed earlier. The difference is that in this case the prepatterned rule is not different from the fully productive rule it is derived from.

In 5.3.1 I motivate this hypothesis by showing that it is possible, perhaps even common, for fully productive rules to become prepatterned. I then show how this fact can lead to fake cyclicity where the different rules appear very similar, perhaps even identical. This predicts that the more similar the rules in a case of fake cyclicity are, the more productive the later rule should be. I
test this prediction on the cases discussed in this chapter and show that it is correct.

In 5.3.2 I discuss the case that, as Cole (1990) notes, is the one that originally raised the hypothesis of cyclicity in the first place, namely the cyclic derivation of secondary stress in English (Chomsky and Halle 1968 and later works). In 5.3.3 I discuss the apparently cyclic application of the optional rule of Coronal Deletion in English (Guy 1991a,b). I show that both cases can be reanalyzed in terms of the hypothesis that the earlier applications of the rules are prepatterned.

5.3.1 Prepatterened fully productive rules

As I have repeatedly emphasized, showing that a rule is fully productive does not mean that the rule cannot be prepatterned. Many rules that are clearly fully productive are nevertheless prepatterned as well.

Consider the rule of Flapping. This rule is clearly fully productive, applying readily across word boundaries; it also applies readily in novel forms produced by speech errors and presumably also in experimental nonce forms as well. Evidence from speech errors suggests that it is nevertheless also prepatterned.

The following is a list of all malapropisms in my collection that seem to involve Flapping. As with the malapropisms examined in Chapter 3, I included only those cases where the surface forms of the error and target were different in either the representation before Flapping or in the representation after Flapping. Since the forms
were not transcribed phonetically, I cannot be sure if all of the
speakers used by my sources really had this rule or not. Based on
the pattern illustrated below, however, I believe they did: in every
single case the /t/ in one word (error or target) that would be
changed into the voiced flap [ɾ] is matched with an underlyingly
voiced segment in the other word. Thus the /t/ in meeting is
matched with the /v/ in movie, the /t/ in hospital is matched with
the /d/ of holiday, and the /nt/ in tentative is matched with the /n/
of tenable. There are no examples where /t/ matches underlying
voiceless consonants, as in the hypothetical examples in (25). In
other words, in all of the malapropisms involving Flapping in my
collection, the surface forms match better than the underlying forms.
This implies that the misselections are occurring at a level after
Flapping has applied. This level, however, also happens to be the
level at which these words are stored. The evidence thus suggests
that Flapping may be prepatterned.\footnote{There is some internal evidence for the same thing. The word buttnsk contains a flap in an environment where Resyllabification and therefore Flapping should not be able to occur, namely before a stressed syllable. The source of this flap is clearly the phrase butt in, where the flap is allowed because the underlying /t/ becomes ambisyllabic by the stress-insensitive Onset Principle. A cyclic analysis is not possible here unless we suppose that affixation to phrases occurs within the lexicon and that Flapping is a cyclic
rule, both of which undermine basic assumptions of Lexical Phonology. By contrast, if the flap in buttnsk is underlying, having only a historical
relation to that in the phrase butt in, there is no dilemma. I am indebted to
Oliver Myers for bringing this example to my attention.}
(24) Malapropisms involving Flapping match better *afterwards*: [df = unpublished error corpus of David Fay, fc = Fay and Cutler (1977)]

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>fc article</td>
<td>argument</td>
</tr>
<tr>
<td>fc bothered</td>
<td>buttered</td>
</tr>
<tr>
<td>df college</td>
<td>cottage</td>
</tr>
<tr>
<td>fc confidence</td>
<td>competence</td>
</tr>
<tr>
<td>fc hospital</td>
<td>holiday</td>
</tr>
<tr>
<td>df letter</td>
<td>lever</td>
</tr>
<tr>
<td>df movie</td>
<td>meeting</td>
</tr>
<tr>
<td>fc tentative</td>
<td>tenable</td>
</tr>
<tr>
<td>df undevoted</td>
<td>undivided</td>
</tr>
<tr>
<td>fc water</td>
<td>weather</td>
</tr>
</tbody>
</table>

(25) Hypothetical malapropisms that do not happen

<table>
<thead>
<tr>
<th>Error</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>article</td>
<td>archetype</td>
</tr>
<tr>
<td>bustled</td>
<td>buttered</td>
</tr>
<tr>
<td>muffin</td>
<td>meeting</td>
</tr>
</tbody>
</table>

How does a fully productive rule become prepatterned?

Although this question is beyond the domain of Double Lookup, which is concerned with how rules apply and not how forms come to be represented in memory the way they are, I will suggest a possibility.

Recall that in Chapter 1 I noted that one of the primary factors determining the ease of a form's retrievability is its frequency. This means, among other things, that forms that do not conform to a rule tend to be very frequent; if they were less frequent, they would be harder to access and it would be more likely that they would have to
be derived on-line by rule, thus regularizing them (Marcus et al. 1990).

Moreover, it has been found that automatic phonetic changes tend to affect frequent words first (Phillips 1984, Johnson 1983). This suggests that surface forms are stored in memory more easily if frequent; changes due to automatic phonetic effects are thus stored first as part of the representation of frequent forms.

I suggest the hypothesis, then, that surface forms, even those derived by predictable rule, tend to be stored in memory as is. The likelihood of this happening depends on factors such as word frequency, with more frequent surface forms being stored before less frequent surface forms.

The prepatterning of fully productive rules, particularly with very frequent words, is thus expected. This prepatterning results naturally through the storage of surface forms as underlying. Doing this need not have any effect on the rule's productivity.

The frequency effect can be seen with Flapping very clearly. As noted in Chapter 4, Resyllabification, the rule that feeds Flapping, does not apply across word boundaries, as shown by its nonapplication of the /t/ in a phrase like that given below in (26a). When one considers very frequent words like today or tomorrow in (26b), however, Flapping may nevertheless apply. This suggests that
the flap in these words is not in fact derived on-line, but instead is prepatterned.\(^{15}\)

\begin{align*}
(26) & \quad \text{tomatoes} \quad [\text{t}\text{m}\text{é}\text{yrowz}] & \text{see tomatoes} & \quad [\text{s}\text{y t}\text{m}\text{é}\text{yrowz}] \\
& \quad \text{today} \quad [\text{t}\text{d}\text{é}\text{y}] & \text{see today} & \quad [\text{s}\text{y r}\text{d}\text{é}\text{y}] \\
& \quad \text{tomorrow} \quad [\text{t}\text{m}\text{á}\text{row}] & \text{see tomorrow} & \quad [\text{s}\text{y r}\text{m}\text{á}\text{row}] \\
\end{align*}

In any case, the effect of this is that rules like Flapping can be both prepatterned and fully productive. I now turn to the question of how this can lead to fake cyclicity.

Since Flapping is both prepatterned and fully productive, the surface pattern motivating this rule may arise in two different ways: through the retrieval of prepatterned forms in Lexical Lookup, or through the on-line application of the rule in Rule Lookup. The possibility thus exists that both things will happen during the same derivation, so that a form prepatterned according to Flapping retrieved from memory will then undergo Flapping on-line. In most cases this will not be apparent, since the on-line rule will not be able to affect a form that already conforms to the rule.

\(^{15}\)Compare Hammond's (1982) explanation for this pattern cited in Chapter 2, which supposes instead that \textit{today} and \textit{tomorrow} contain the word \textit{to}, which is inherently unstressed. This analysis is also compatible with a frequency-based analysis, since \textit{to} is of course extremely frequent, and the effects of the word-initial destressing rule Hammond (1982) postulates would be prepatterned in this word. The frequency effect in prepatternning will be seen again in the discussion of Vowel Reduction in Chapter 6.
The situation is different, however, when the application of morphological rules is considered as well. If one allows that such rules can build new words from forms stored in memory, and if at least some of these stored forms are prepatterned according to some phonological rule R, then the possibility arises that the addition of a new affix will create a form that may then undergo the on-line version of R. The result will appear to be cyclicity. Since in this case there may be no linguistic evidence that the prepatterned and on-line versions are distinct, this term is descriptively adequate. Psychologically, however, the two applications of the rule will be very different, one being prepatterned and the other being on-line; the cyclicity is thus fake.

This analysis predicts that almost any fully productive rule can potentially appear to be cyclic, given the right circumstances. Surprisingly, this seems to be true; in this section I will discuss just three cases of this in English. From the viewpoint of a theory like Lexical Phonology, however, each new case of cyclicity is unexpected. In that theory most very productive rules are placed in the postlexical stratum (stress is the only exception); since this stratum is claimed not to be cyclic, there is not only no reason to expect fully productive rules to be cyclic, but it is in fact not expected at all.16

16Lexical Phonology does allow for rules to appear in both the lexical and postlexical strata, as we saw earlier. Thus Borowsky (1986:36) explicitly claims that Flapping appears three times in the derivation, once for each of her three morphologically defined strata; its application in the lexical strata is blocked by Structure Preservation. The possibility of such cases is merely stipulated, however, as in the Strong Domain Hypothesis of Kiparsky (1984) (see below), and does not derive from general properties of the system.
Although I have found no evidence suggesting that Flapping can appear to apply cyclically, there is evidence that another very productive rule, namely Aspiration, can. Speech error accommodations show this rule to be very productive, if not fully productive (e.g., Fromkin 1971, Stemberger 1983). I have no psycholinguistic evidence that it is prepatterned, but given the arguments made in Chapter 4 we may expect that it is, for the following reason. As we saw in my introduction to the Resyllabification rule in Chapter 2, Aspiration only applies within word boundaries; a word-final coda resyllabified as the onset of the following word does not undergo Aspiration. In Chapter 4 we saw that rules that only apply within word boundaries tend to be less productive than those that apply across word-boundaries; in any event, they cannot be more productive than such rules. This suggests that Aspiration is not as productive as Flapping and should therefore be more prepatterned. This prepatterning, combined with its high level of productivity, allows for the possibility of it appearing to apply in a cyclic fashion.

The pair below suggests that Aspiration may in fact apply cyclically. That is, the derived forms on the right show exactly the same stress pattern. The word militaristic, however, appears to have carried over the aspiration on /t/ from the form it is derived from.

17 Note in this regard that Aspiration is a fortition rule while Flapping is a lenition rule; that is, Aspiration increases a segment's perceptual salience while Flapping decreases it. This example thus provides more evidence for both Donegan and Stampe's (1979) ordering principle as well as my own reinterpretation of this principle in terms of productivity.
(27) [from Withgott 1982].

military [mɪlɪtəri] militaristic [mɪlɪtərɪstɪk]
capital [kæpɪrəl] capitalistic [kæpɪrəlɪstɪk]

The apparent cyclicity of the application of Aspiration is illustrated in the derivations below.

(28) Cycle 1:  
<table>
<thead>
<tr>
<th>Stress</th>
<th>militaristic</th>
<th>capitalistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspiration</td>
<td>mɪlɪtəri</td>
<td>kæpɪtəl</td>
</tr>
</tbody>
</table>

Cycle 2:  
- *-istic* added  
| Stress         | mɪlɪtəri+ɪstɪk | kæpɪtəl+ɪstɪk |
| Aspiration     | mɪlɪtərɪstɪk  | kæpɪtəlɪstɪk |

Postcyclic level:  
| Aspiration     | ---           | ---           |
| Flapping       | ---           | kæpɪrəlɪstɪk |

In terms of Double Lookup, however, what has happened is this: Aspiration has become prepatterned in the form *military*. Thus when a new word is derived from this form, aspiration is carried over into the derived word in exactly the same way other stored aspects of the word are carried over, such as the initial /m/. The form *militaristic* thus goes through four historically ordered stages, as illustrated below. At each stage, the form that is stored in memory is the form that was derived at an earlier stage.

---

18The question of the application of the stress rules is left to the following section.
Another prediction of the hypothesis that fully productive rules naturally lead to fake cyclicity is that the more similar the rules are in a case of fake cyclicity, the more productive the later rule should be. This is because if the prepatterning occurs through the storage of surface forms derived by a very productive rule, the prepatterning is expected to mirror the on-line rule very closely. By contrast, if the case of fake cyclicity arises through the addition of a rule at the end of a derivation that mirrors an older (and thus less productive) rule, these rules need not be particularly similar at all.

Both of the above predictions are supported by the examples discussed in this chapter. The cases of fake cyclicity where the different rule applications appear most similar all involve very productive rules, namely Aspiration, Nasal Assimilation and Palatalization, discussed above, and English stress rules and Coronal Deletion, discussed below. The cases where the different rules
clearly appear different involve Icelandic Umlaut and Swiss German o-Lowering, both relatively less productive rules. This correlation is utterly ignored in a model where cyclicity is a primitive.

In other words, I claim that it is no coincidence that the best apparent examples of cyclicity involve precisely the most productive rules. Instead, given the fact that the prepatterning of productive rules is likely to be highly systematic, we should expect virtually all fully productive rules to appear to apply in a cyclic fashion: the earlier application will be prepatterned while the later one will be on-line.

5.3.2 Secondary stress in English

In this section I examine the evidence for the cyclic assignment of stress in English. In contrast to other opponents of the Lexical Phonological notion of cyclicity, such as Halle and Vergnaud (1987) and Cole (1990), I conclude that English stress does in fact appear to apply cyclically. The fact that stress patterns are both demonstrably prepatterned and very productive, however, suggests that this is actually a case of fake cyclicity, exactly parallel to that sketched above for Aspiration.

The Chomsky and Halle (1968) argument for the cyclic assignment of stress in English rests on examples such as those listed in (30) below. The observation is that secondary stress falls on a heavy syllable preceding a main stress in English only in words that are derived from stems with main stress on this same heavy syllable. Hence the forms in (30a) show secondary stress before the main
stress because they are derived from stems with main stress in the appropriate place. The forms in (30b) do not have secondary stress before the main stress because in the stem there is no stress on this syllable; the form in (30c) does not have secondary stress before the main stress because it is nonderived.19

(30) [after (5) in Halle and Kenstowicz 1991:460 and Hammond 1989:139]

a. còndênsâtìon (from còndênsé)  
   instrumêntâlity (from instrumêntal)  
   àttêstâtìon (from attêst)  
   ôbjectîvîty (from ôbjectîve)  
   êlâstîcîty (from êlâstîc)  

b. còntêmplâtìon (from còntêmplâtê)  
   còmpênsâtìon (from còmpênsâtê)  
   dêmônstrâtìon (from dêmônstrâtê)  
   ànêcdôtal (from ànêcdôte)  

c. sèrendîpîty (underived)

The cyclic analysis of these facts is as illustrated in (31). The stress assignment rules in each cycle are the same. Stress assigned on a previous cycle surfaces as secondary stress.

(31) condense + ation  compensate + ion

Cycle 1:  
Stress condénsé  cómpênsâtê

Cycle 2:  
suffixation condénsé + ation  cómpênsâtê + ion  
Stress còndênsâtìon  còmpênsâtìon

19Hammond (1989) follows Chomsky and Halle (1968) in indicating three degrees of stress in his examples; I follow Halle and Kenstowicz (1991) and transcribe secondary and tertiary stress the same way.
As first pointed out by Halle and Vergnaud (1987), however, this analysis does not appear to hold up as given. There are many derived words, like those in (32a), which do not pass on the secondary stress from an earlier cycle. Moreover, there are many nonderived words, like those in (32b), which contain a secondary stress directly before the main stress.

(32) [after Halle and Kenstowicz 1991:460, following stress in Kenyon and Knott 1944; stress on stems in (a) from Kenyon and Knott 1953 added by JM]

a. affirmation (from affirm)
cônfimation (from confirm)
cônservation (from conserve)
cônsultation (from consult)
cônversation (from converse)
inîrmation (from inform)
làmentation (from lament)
préervation (from preserve)
trànsportation (from transport)
ûsurpation (from usurp)

b. Haliçànássus
incântation
incârnation
ðstêntation

Based on this, Halle and Vergnaud (1987), Halle and Kenstowicz (1991) and Cole (1990) conclude that there is no evidence whatsoever for the cyclic assignment of secondary stress in English. Instead, whether or not secondary stress shows up on heavy

---

20 As usual, I use "cyclic" here in the sense of applying a rule more than once in a derivation. Halle and Vergnaud (1987) use this term in a different sense, so that technically for them English stress is in fact cyclic; they just argue that there is no copying of stress from earlier cycles.
syllables preceding main stress is determined on a word-by-word basis; thus it shows up in words like condensation and ostentation but not in words like compensation and lamentation.

However, there is another explanation for the problematic forms in (32). To see this, compare the derivation of lamentation below in (33) with the cyclic derivation of condensation in (31).

Although it derives an incorrect form, namely *lamentation, this can easily be corrected by assuming that this word, unlike condensation, is subject to a special rule of Destressing, whereby secondary stresses are removed from syllables adjacent to a main stress.

(33) lament + ation

<table>
<thead>
<tr>
<th>Cycle 1:</th>
<th>Stress</th>
<th>Destressing</th>
</tr>
</thead>
<tbody>
<tr>
<td>lament</td>
<td>lamént</td>
<td>--</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Cycle 2:</th>
<th>suffixation</th>
<th>Stress</th>
<th>Destressing</th>
</tr>
</thead>
<tbody>
<tr>
<td>lament + ation</td>
<td>lamént + ation</td>
<td>lamentation</td>
<td>lamentation</td>
</tr>
</tbody>
</table>

Strong motivation for this analysis comes from the observation that with the exception of lamentation, all of the syllables before the main stress in the words in (32a) end in the liquids /r/ or /l/. By contrast, none of the syllables before the main stress in the words in (30a) end in liquids; rather, they end in /n/, /s/, or /k/. Moreover, lamentation itself contains a sonorant after the relevant vowel. The fact that /n/ is the only segment that appears in the relevant syllables in the two lists is also significant, since /n/ is midway between obstruents and liquids in sonority, and hence might be
expected to be involved in the destressing process in a less regular fashion.

Thus it appears that the destressing rule illustrated in (33) is parallel to the rule of *Sonorant Destressing*, discussed in Chomsky and Halle (1968), Kiparsky (1979), Hayes (1980), Hammond (1984), Halle and Vergnaud (1987) and other works.\(^\text{21}\) This rule removes a stress from a syllable ending in a sonorant if next to another stress. Among other things, this rule is used to account for contrasts such as those illustrated below, where the heavy syllable in the words in (34a) is stressed, while the heavy syllable in the words in (34b) is not, due to its ending in a sonorant and being next to a stressed syllable.

(34) [from Halle and Vergnaud 1987:256-259, including examples from Kiparsky 1979 and Hayes 1980]

a. perfunctory reféctory projéctile
   smarágdine staláctite

b. répertöry Gîlbertîte Álgernón
   Wáshington Lîvingstôn Pálmerton
   Bîrmingham Rûtherford

This approach has an advantage over an analysis where all secondary stresses are idiosyncratic in that it accounts for the fact

\(^{21}\)Thanks to Mike Hammond for pointing out this parallel to me. I warn the too gullible reader that he also points out numerous problems with my specific use of Sonorant Destressing here. See also Kager (1989) for fuller discussion of the uses of Sonorant Destressing in English. Whatever the correct solution is, however, the fact remains that in all of the problematic cases cited in the literature, i.e. those collected in (30), the problematic syllable always ends in a sonorant.
that a stem stress may be absent in derived forms only on a syllable preceding the main stress. This rules out possible pairs like *ànticipàte~ *ànticipátiōn, where a stress in the stem is lost in the derived form even though it doesn't precede the main stress of the derived form. Examples of nonderived forms like those in (32b) are not a serious problem to this analysis. All they show is that some secondary stresses are indeed marked underlyingly. In derived forms, however, this does not seem to be necessary.

Moreover, there are other apparent examples of stem stress being preserved in derived forms. Hammond (1989) points out the examples below, where it appears that if a stem receives secondary stress on the last syllable this may show up as secondary stress in the derived form, as illustrated in (35a), while if there is no such stress, the derived form has no such stress either, as in (35b).

(35) [after (3) in Hammond 1989:140]

a. mānifēstātiōn (from mānifēst)
   ōriēntātiōn (from ōriēnt)
   expērīmēntātiōn (from expērīmēnt [verb])
   dīphthōngīze (from dīphthōng)
   frāgmēntātiōn (from frāgmēnt [verb])

b. fōrestātiōn (from fōrest)
   stāndardīze (from stāndard)
   bāstardīze (from bāstard)
   fēcundīze (from fēcund)

Thus there does seem to be a generalization here that would be lost if all secondary stresses were considered idiosyncratic. Halle and Vergnaud's (1987) analysis of this pattern, which treats it as an
epiphenomenon of other stress rules that are noncyclic, is effectively
demolished in Hammond (1989).

I argue that this situation is in fact expected given the
assumptions of Double Lookup. Specifically, I claim that the pattern
discussed above is fake cyclicity which arises through the English
stress rules being both prepatterned and very productive, exactly
parallel to the fake cyclicity of Nasal Assimilation, Palatalization and
Aspiration.

Evidence for the prepatterning of stress rules comes from some
of the speech errors discussed in Chapter 3. Recall that in
malapropisms, where a real word is erroneously substituted for the
intended word, several properties are consistently found to be
shared by the error and the target word (see, eg, Fay and Cutler
1977). These similarities tell us something about how forms are
stored in memory. This is because malapropisms involve
substitutions between real words, and real words are distinct from
nonreal but possible words only in that they are stored in memory.

First, the error and target word in malapropisms are
overwhelmingly the same part of speech, presumably because part
of speech plays a role in how words are organized in memory.
Second, the error and target are similar at a segmental level,
implying segmental information is used in organizing words in
memory as well. And third, the error and target are almost always
prosodically identical. That is, they almost always agree in the
number of syllables and the placement of stress.
What this last point means is that prosodic structure, specifically syllable count and stress placement, forms part of the representation of English words stored in memory. Why might this be so? Given Double Lookup, this might mean that the syllabification and stress rules of English are synchronically extremely unproductive; patterns seen in surface forms are not necessarily associated with on-line rules at all. The available evidence seems to argue strongly against this, however. For example, in the experiments by Myerson (1978) discussed in Chapter 3, it was found that children of all ages had no trouble knowing how to stress novel words, even in using the straight production task; Hochberg (1988) found the same thing was true of stress in Spanish. Thus it appears that the stress rules associated with the stress patterns stored in memory are very productive.

In other words, the stress rules in English appear to be both very productive and prepatterned. It may be in this case that the prepatternning is due to something like Prosody First, but in any event the important thing is to observe that the stress rules are in precisely the same situation as we saw for Aspiration in the previous section. Being both prepatternned and very productive, we expect the stress rules to show evidence of fake cyclicity. And, as we have seen, they do.

Below I give a series of derivations showing how we may view the apparently cyclic application of the stress rules within Double
Lookup. As with Aspiration, in each stage the derivation begins where it left off in the diachronically earlier stage.

(36) a. Stage 1:  
   Lexical Lookup: condense  
   Rule Lookup: condense  
   Stress rules condénsé  

b. Stage 2:  
   Lexical Lookup: condénsé  
   Rule Lookup:  
   Stress rules ---  

c. Stage 3:  
   Lexical Lookup: condénsé  
   Rule Lookup:  
   -ation added condénsation  
   Stress rules condénsation  

d. Stage 4:  
   Lexical Lookup: condénsation  
   Rule Lookup:  
   Stress rules ---  

Note that showing that this case is analyzable as exactly parallel to cases involving "postlexical" rules, such as Nasal Assimilation, provides further *prima facie* support for my claim that fake cyclicity is a superior explanation to real cyclicity. By contrast with Double Lookup, Lexical Phonology is forced to ascribe the apparently repeated applications of Nasal Assimilation and the stress rules to two very different causes. In the former case, Nasal Assimilation is called into application at different levels. In the latter case, however, the stress rules appear in only one level, but this level is marked to apply repeatedly. By accounting for these two
cases in such different ways, the fact that Nasal Assimilation and the stress rules are both repatterned and both very productive is left as a coincidence.

According to my analysis, then, the carrying over of stress from *condéns*e to *condénsá*tion, created with a "cyclic" affix, is no different in principle from the carrying over of stress from *org*inal to *org*inal*ness*, created with a "noncyclic" affix. In both cases the stress is carried over diachronically, not synchronically, since the form that undergoes the morphological process is a word where stress is already repatterned. The difference between affixes like these is not a matter of cyclicity but instead that noted in Chapter 1. There I mentioned the observation of Goldsmith (1990) that affixes like *-ity* and *-al* give rise to stress patterns typical of monomorphemic words (eg, antepenult stress in *or*igin, *org*ina/ and *org*iná/ity), while suffixes like *-ness* give rise to stress patterns that may not appear in monomorphemic forms (eg, stress on the fourth-to-last syllable in *org*inal*ness*). This implies that English stress rules in some sense apply only within "monomorphemic" forms, assuming these to include forms with relatively unproductive suffixes like *-ity* and *-al*; very productive suffixes like *-ness* fall outside of this domain and behave prosodically no different from clitics. One can still imagine that word-level stress rules apply appropriately in a form like *org*inal*ness*; the application of these rules is vacuous, however, since

---

22 Various other researchers have suggested approaches like this to the interaction of morphology and word-internal prosody, among them Cole (1990), Inkelas (1991), Ishihara (1991), and Lemus (in preparation).
they only apply within the the domain defined by original, which is retrieved from memory with the stress already in place.

5.3.3 Coronal Deletion in English

In this section I review the evidence for Coronal Deletion being a cyclic rule. Though at first it will appear to be a case completely different from that of the English stress rules, in the end we will see that this is simply another example of fake cyclicity, due to the rule being both prepatterned and very (in this case fully) productive.

Coronal Deletion presents two unique problems for such an analysis, however. The first is that it is optional, but unlike Nasal Assimilation and Palatalization, optional in all of its versions, not just the one nearest the surface. My explanation for this is that Redundancy of Patterning, which allows multiple allomorphs of words, therefore allows for the optionality of prepatterned rules. This arises from Lexical Lookup being free to select between multiple allomorphs; its choice may then be influenced by the same factors that influence the optional application of on-line rules. This claim is further supported in Chapter 6 with clear examples of Lexical Lookup choosing on-line between different allomorphs.

The second problem is that unlike all the other rules I have discussed in this chapter, Coronal Deletion appears to apply not twice but three times. According to Double Lookup, only one of these applications, namely the last one, can occur on-line; I therefore conclude that both of the first two applications are prepatterned.

---

23This section overlaps in some ways with parts of J. Myers (1992).
Their interaction can be shown to arise in essentially the same way as the single prepatterned version of Aspiration discussed above.

The evidence for the cyclic application of Coronal Deletion comes from work by Guy (1991a,b, 1992). As a reminder, examples of Coronal Deletion are given below in (37).

(37) **Coronal Deletion** [after Selkirk 1972:193-4]

<table>
<thead>
<tr>
<th>Original Form</th>
<th>Resulting Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>draft-dodger</td>
<td>draf-dodger</td>
</tr>
<tr>
<td>exact sciences</td>
<td>exac sciences</td>
</tr>
<tr>
<td>thefts</td>
<td>thefs</td>
</tr>
<tr>
<td>lend money</td>
<td>len money</td>
</tr>
</tbody>
</table>

Not only is this rule optional, but it also appears to be noncategorical, in that the likelihood of its application varies gradiently with the relative sonority of adjacent segments; as we saw in Chapter 4, /t,d/ is more likely to drop next to a less sonorant segment than next to a more sonorant segment. This is illustrated in the varying degrees of acceptability of the forms in (38).

(38) a. exact sciences → exac sciences  
    b. at that exact moment → exac moment  
    c. for that exact reason → ?exac reason  
    d. exact answer → ??exac answer  

Guy (1991a) accounts for the gradient effect of sonority by proposing that Coronal Deletion is actually a special case of what he calls Consonant Cluster Simplification, formalized as in (39). Since this rule makes reference to syllable structure, syllabification must precede it and feed it. It is the gradient effect of sonority on syllabification that produces the effect seen in (38).
Consonant Cluster Simplification [after (3) in Guy 1991a]

\[
\begin{array}{c}
\sigma \\
\rightarrow
\end{array}
\]

Since Coronal Deletion delinks the second consonant in a consonant cluster (i.e., the /t/ or /d/) only if both consonants in the cluster are linked to the same syllable, it follows that this consonant cannot delink if it has been resyllabified as part of the following syllable. The likelihood of this resyllabification, in turn, depends on the sonority of the following segment. If the following segment is a vowel, it's quite likely the /t/ or /d/ will be resyllabified as the onset of the following syllable, thereby bleeding Coronal Deletion. If the following segment is a liquid, resyllabification is less likely, making Coronal Deletion more likely, and so forth.

Coronal Deletion thus has all the hallmarks of a prototypical postlexical rule. It is optional, is phonetically natural, is fed by the gradient rules of syllabification, and applies across word boundaries.

The problem is, as sociolinguists have known for a long time, Coronal Deletion is also sensitive to morphology, just like a lexical rule. Specifically, the likelihood of the deletion of a word-final /t/ or /d/ depends on whether this segment is used as an inflectional suffix or not. Guy (1991a,b) goes further, showing that Coronal Deletion not only appears before the postlexical stratum, but in fact appears twice

---

24 The significance of the right bracket in Guy's formulation of the rule will be explained later.
in the lexicon, once at level 1 and then again at level 2; it then appears a third time as a postlexical rule.

There are two independent arguments for this claim. First, by examining the recorded speech of several English speakers, Guy (1991a) shows that the cube root of the rate of retention of /t/ and /d/ in monomorphemic (M) forms like lift is very close to the rate of retention $p_r$ in the past tense forms of regular verbs like laughed (P forms); for many speakers, the square root of the rate of retention in semiweak verbs like left (S forms) is very close to $p_r$ as well. Thus, as was already known, Coronal Deletion is more likely to apply in M forms than in S forms, which in turn drop /t/ and /d/ more often than P forms. What is new is the finding of a specific exponential relation in the likelihoods of application in the three morphological types. Guy (1991a) argues that this relation can only be understood if it is supposed that Coronal Deletion has three chances to apply in M forms, two times in S forms and only one time in P forms, each time with a retention rate of $p_r$. This makes sense if one recalls that [t,d] is present for a level 1 application of Coronal Deletion only in M forms, since neither the regular suffix -ed nor the semiweak suffix -dlt has been added yet. For S forms [t,d] only appears in time for an application of this rule in level 2, namely after the semiweak suffix -dlt has been added in level 1. P forms only get the opportunity to lose [t,d] postlexically, after the regular suffix -ed has been added in level 2. This conception is illustrated in (40).
The second argument in support of this view concerns the sensitivity of Coronal Deletion to adjacent segments. As we saw above, the likelihood of /t,d/ dropping depends on the sonority of adjacent segments. If the picture in (40) is correct, we would expect that the relevance for word-final deletion of [t,d] of a preceding, ie word-internal, segment should be greater in M forms than in P forms. By contrast, the relevance of a following segment, ie the first segment of the following word, should be equal for M and P forms. This is because Coronal Deletion would only get a chance to apply in P forms at a level where the word-external environment is relevant, while in M forms Coronal Deletion can apply before this level as well. As Guy (1991b) shows, this is just what is in fact found to be the case.
Guy (1991a,b) therefore concludes that Coronal Deletion is a cyclic rule. More specifically, it applies after syllabification, and prior to morphology and bracket erasure, at all three morphological levels in English.25

The gist of the Double Lookup reanalysis of these facts should come as no surprise: the earlier applications of Coronal Deletion are prepatterned, while the later one applies on-line. What makes this case interesting are the two peculiarities noted at the beginning of this section. First, Coronal Deletion is optional, and second, it appears to be prepatterned twice, once at the pre-irregular-inflection stage, and again afterwards. Neither presents a serious problem for Double Lookup, however, if additional independently motivated assumptions are taken into account.

Consider first the optionality of Coronal Deletion. At first sight this appears to show that Coronal Deletion applies on-line three times in a row, with the on-line application of past tense suffixation in

25One interesting consequence of the cyclic analysis of Coronal Deletion is that it demonstrates the inferiority of the Strict Cycle Condition compared to an acquisition-based constraint. As John Goldsmith pointed out at the UWM Symposium in Milwaukee, April 1992, where Guy presented some of his work, it is not clear, if Coronal Deletion is cyclic, how it is able to apply in non-derived forms like /ift/ in the first place. Dropping the /t/ in /ift/ is easy to understand if Coronal Deletion is merely postlexical, since postlexical rules need not obey the Strict Cycle Condition, but if Coronal Deletion applies in level 1, it should not be able to affect words like /ift/. Guy's (1991a) numbers, however, imply that it does nevertheless. There's no problem, however, if the Strict Cycle Condition is replaced by something like the Alternation Condition or the Morphology Hypothesis. Since Coronal Deletion is not an obligatory rule, forms like /ift/ will occasionally surface with a surface [t], making the rule learnable for these forms. That is, Coronal Deletion is not an obligatory neutralization rule, and so may apply with all forms of a morpheme.
interleaved throughout. Further consideration shows that this cannot be the case.

The major problem is the abundance of evidence that other rules that apply within the lexicon do not typically apply on-line. Whether or not the Productivity Hypothesis is true, theorists from all perspectives would consider it paradoxical to have an on-line rule applying before a prepatterned rule. And yet by claiming that Coronal Deletion applies in level 1, Guy (1991a,b) is claiming that Coronal Deletion applies before s-Voicing, Velar Softening, Vowel Shift, y-Insertion, and a wide range of other partially productive rules. If Coronal Deletion applies before these rules, these rules must apply on-line as well, and yet the evidence adduced in Chapter 3 shows that they typically do not.

I suggest therefore that the versions of Coronal Deletion that precede prepatterned rules are themselves prepatterned. Because of Redundancy of Patterning and the on-line nature of Lexical Lookup, the fact that even these prepatterned rules can be optional is not a problem. That is, with the prepatterned versions of Coronal Deletion there is no on-line application that is optional; rather, what is optional is the selection of allomorphs in Lexical Lookup. Freedom in choosing allomorphs can be seen with speakers who may pronounce the noun *address* with stress either on the first or second syllable, or with those who pronounce *economic* with either [iy] or [e] in the first syllable. Recall also the apparently optional application of i-Spirantization in Finnish. Further examples of this phenomenon are
given in Chapter 6. I suggest that the same thing happens with the prepatterned versions of Coronal Deletion; the selection of allomorphs either prepatterned or not prepatterned according to Coronal Deletion occurs on-line and may then be influenced by the same factors influencing the optionality of on-line rules.

Given Double Lookup, we would expect that the likelihood of retrieving lif over lift should closely mirror the likelihood of on-line Coronal Deletion pretty closely, since the source of the two stored forms is the pair of surface forms generated in the past by the on-line application of Coronal Deletion. If on-line Coronal Deletion applies X% of the time, lif will be the allomorph of lift that appears X% of the time, all else being equal. The allomorphs in memory will thereby be reinforced in accordance with this same proportion, so that X% of the time, lif will be the allomorph of lift that is retrieved from memory.

All else is of course not equal. If there is a productive morphological rule that adds [t] or [d] at the end of a word retrieved from memory, the surface percentage of nondeleted /t,d/ will be different from that given in the above scenario. Thus if laugh is retrieved from memory in Lexical Lookup and laughed is then generated by an on-line rule, the likelihood of this final [t] being deleted is not X%, but is much lower. This is because a monomorphemic form like lift has two chances to conform to the Coronal Deletion pattern: once, when it is retrieved from memory as lift or as lif, and twice, through the on-line application of Coronal
Deletion. By contrast, a form like *laughed*, whose [t] is derived by the application of an on-line morphological rule, can only undergo Coronal Deletion once, namely on-line.

This scenario is represented schematically below. As with my analyses of the apparently cyclic behavior of Aspiration and stress assignment, this situation must arise in separate diachronic stages, even if these stages all occur during an individual speaker's acquisition of the language. As Marcus et al. (1990) and Stermerger and MacWhinney (1988) have shown, *-ed* suffixation does not appear to be prepatterned except perhaps for very frequent words; hence I assume that even at the second stage *-ed* suffixation occurs on-line. What this analysis means, then, is that it will appear that Coronal Deletion gets an extra chance to apply in monomorphemic forms, while even at the later stage it will in fact only have one chance to apply in regularly inflected forms.

(41) a. Stage 1:  

<table>
<thead>
<tr>
<th>Lexical Lookup:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule Lookup:</td>
</tr>
<tr>
<td>-ed added</td>
</tr>
<tr>
<td>/ \</td>
</tr>
<tr>
<td>Cor Del</td>
</tr>
</tbody>
</table>

b. Stage 2:  

<table>
<thead>
<tr>
<th>Lexical Lookup:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rule Lookup:</td>
</tr>
<tr>
<td>-ed added</td>
</tr>
<tr>
<td>/ \</td>
</tr>
<tr>
<td>Cor Del</td>
</tr>
</tbody>
</table>
This analysis can be extended quite naturally when the semiweak suffix \(-d/t\) suffix is taken into account as well. As with my discussion of the optionality of Coronal Deletion, I begin my discussion by showing that the cyclic analysis of the behavior of this suffix cannot be right.

One major problem with the idea that Coronal Deletion applies on-line three times, one at each morphologically defined level, is that it crucially depends on the assumption that English does in fact only have three morphologically defined levels. However, it is not clear how many levels English really has; Kiparsky (1982) and Borowsky (1986) argue that it has three levels, while Halle and Mohanan (1985) argue that it has five.

What is certain is that there are only three morphologically defined classes of words in English ending in [t] or [d], namely Guy's M, S and P classes. It is also clear that the \(-d/t\) past tense rule of semiweak verbs is far less productive than the \(-ed\) past tense rule of regular verbs. Double Lookup would thus order these rules exactly as is done in Lexical Phonology, with the semiweak rule before the regular rule. If and when these rules apply on-line, the input they receive is from the set of M-class words, namely verb stems. Thus Double Lookup, like Lexical Phonology, places M class words "prior" to S class words, which in turn are "prior" to P class words. Guy's (1991a) argument thus has nothing whatsoever to do with level ordering in general, but only with the ordering of these specific morphological classes.
I suggest therefore that the versions of Coronal Deletion ordered before prepatterned morphological rules are themselves prepatterned. Since -ed suffixation is a fully productive morphological rule (Marcus et al. 1990), it always applies on-line; thus the version of Coronal Deletion that applies after it must apply on-line as well. However, since -d/t suffixation is only partially productive, it is not affixed on-line in every case. When it is affixed on-line, it is affected by on-line Coronal Deletion the same way as the regular past tense verbs with -ed. When the -d/t suffix is prepatterned, the Coronal Deletion that affects it can either be prepatterned or apply on-line. Finally, since the word-final /t, d/ of monomorphemic forms is never added on-line at all but is always "prepatterned", the Coronal Deletion that affects it can be the on-line rule or a prepatterned rule.

The idea is that in monomorphemic forms, the /t/ that may be deleted by Coronal Deletion is always stored with the form in memory; in regular forms, the /t/ is (virtually) never stored with the form in memory; while in semiweak forms, the /t/ is sometimes prepatterned and sometimes not. A series of stages illustrating the effect of this is given below.

(42) a. Stage 1:

<table>
<thead>
<tr>
<th></th>
<th>lift</th>
<th>left</th>
<th>laughed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lex Lk:</td>
<td>lift</td>
<td>left</td>
<td>læf</td>
</tr>
<tr>
<td>Rule Lk:</td>
<td>--</td>
<td>--</td>
<td>læft</td>
</tr>
<tr>
<td>-ed</td>
<td>/ \</td>
<td>/ \</td>
<td>/ \</td>
</tr>
<tr>
<td>Cor Del</td>
<td>lift left left læf læft</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
It is here that Double Lookup makes some predictions crucially different from those made by the Lexical Phonology model argued for in Guy (1991a,b, 1992). Although they have unfortunately not yet been tested, there are some suggestive signs in what data we do have that imply that the Double Lookup approach is better.

First of all, consider the fact that speakers apparently do not recognize the S class as distinct from the M class until adulthood, if even then (Guy and Boyd 1990). Guy (1991a,b) studied the speech of people who had already shown signs of having an S class. What about people who do not yet have this class? The Lexical Phonology prediction, according to Guy (personal communication) would be as follows. From other morphological evidence or through innate knowledge (see Gordon 1989), such people would already have
acquired the three levels. The only difference between them and people with the S class would be that they have not yet learned the specific rule of semiweak past tense formation. Thus for them we would still expect the likelihood of retention of \([t,d]\) in M forms to be the cube root of that in P forms, just as with speakers who have the S class. This follows from the Strong Domain Hypothesis (Kiparsky 1984:142), which says that "the grammar may stipulate merely where a rule ceases to apply"; thus if Coronal Deletion applies lexically and also postlexically, as it would for speakers without the S class as well as those with it, it must also apply at both level 1 and level 2.

By contrast, Double Lookup would predict that speakers without the S class would treat M class and P class words as if Coronal Deletion had two chances to apply in the former and only one in the latter. These predictions have unfortunately not yet been tested. However, Guy (1991a) notes that he never obtains as good numbers for the S class as for the M and P classes: the retention rate is always higher or lower than what would be expected given the cyclic analysis.

Another prediction made by the Double Lookup analysis is that token frequency should play a role. Just as M class forms should conform to Coronal Deletion more often than S forms, because of the possible analysis of S forms as derived on-line, frequent words of all sorts should conform to Coronal Deletion more often as well, and for precisely the same reason: such forms will be more likely to be
stored in memory without /t/ or /d/. Guy (personal communication) grants this possibility, noting that extremely frequent words like *and* conform to Coronal Deletion so often that it is plausible to ascribe this to allomorphy. However, frequency has not yet been separated out from the morphological information involved in the M, S and P class distinctions. This seems crucial, since as noted above, irregular forms tend to be more frequent than regular forms. The effect cannot reduce entirely to frequency, since Guy (1991a,b) finds that irregulars ending in /t/ or /d/, like *sent*, seem to drop the coronal at the same rate as monomorphemic forms; morphological information is necessary. However, it would be interesting to see what role frequency does play.

In short, the finding that Coronal Deletion appears to apply cyclically is actually an argument for Double Lookup. The conclusive tests remain to be done, but so far it appears that the facts are more in accord with a fake cyclicity analysis than a real cyclicity analysis. First, the real cyclicity analysis requires Coronal Deletion to apply online before prepatterned rules, a clear impossibility. Second, it crucially requires there to be exactly three morphologically defined levels in English, something that is far from clear. The facts Guy discusses can be accommodated within Double Lookup quite simply if we allow for the possibility that Lexical Lookup can choose between allomorphs in a variable fashion, and if we recognize that *-t/d* suffixation is a partially productive, not fully unproductive, morphological rule.
5.3.4 Direct evidence: Summary

The direct evidence for cyclicity presented by English stress and Coronal Deletion, though at first impressive, thus turns out to be fully consistent with Double Lookup. In particular, both cases involve rules that are prepatterned as well as very productive, precisely the situation that Double Lookup predicts should give rise to the sort of fake cyclicity that mimics real cyclicity most closely (such as with Nasal Assimilation and Aspiration). Within Lexical Phonology, by contrast, this parallelism in prepatterning and productivity is irrelevant, since the multiple applications of the stress rules and of Coronal Deletion must be handled by two completely different mechanisms: the stress rules apply in a single stratum marked "cyclic," while Coronal Deletion applies in all strata.  

5.4 Chapter summary

In this chapter I have argued that cyclicity per se does not exist. This is as predicted by the Productivity Hypothesis, which does not allow a single partially productive rule to apply more than once in a derivation. Instead, apparent cases of cyclicity involve fake cyclicity or idiosyncratic ordering. Fake cyclicity was shown to be necessary anyway because it accounts for phenomena real cyclicity cannot. Moreover, fake cyclicity was shown to arise through natural historical processes that can be couched in terms of Double Lookup, even predicting significant properties such as the dependence on ordered morphological processes. Idiosyncratic ordering, which for similar reasons also mimics real cyclicity and yet is not reducible to
it, also finds a simple explanation within Double Lookup. Given that such cases must be handled anyway, the device of real cyclicity becomes redundant. Finally, I showed that Double Lookup correctly predicts that any prepatterned very productive rule can show signs of fake cyclicity of a particularly striking sort, namely where all versions of the rule appear to be formally identical. I illustrated this with two cases of prepatterned very productive rules which in Lexical Phonology must be handled with rather different mechanisms.
CHAPTER 6
THEORIES OF RULE ORDERING

6.0 Introduction

Two fundamental questions about phonological rules, namely what they are and what the evidence for their ordering really means, have been addressed in a variety of different ways in the literature. Part of the novelty of the Double Lookup model lies in its claim that answering the first question directly leads to an answer for the second. In other words, if one accepts the claims of Double Lookup about what rules are, that is, Redundancy of Patterning, Gradient Productivity, and Gradient Retrievability, then a claim about rule ordering, that is, the Productivity Hypothesis, automatically follows. The fact that the Productivity Hypothesis need not be added as an independent stipulation is quite a point in Double Lookup's favor, since in most theories of rule ordering, the ordering of rules is not explained, in the sense of being reducible to other independently motivated claims.

In this, the final chapter of the dissertation, I summarize the advantages of Double Lookup over these other models of rule ordering.

I begin in 6.1 with the foremost theory of rule ordering, which I will term the Standard Model, which claims that rules apply sequentially in an order specific to a speaker's idiolect; the historically most prominent proponent of this model is Chomsky and Halle (1968). The remaining sections concern theories of rule ordering that attempt to make up for the conceptual and empirical
failings of the Standard Model but which bring conceptual and empirical problems of their own.

In 6.2 I discuss theories which claim that rules are never ordered at all, but instead apply simultaneously.

In 6.3 I discuss theories which claim that rules may apply sequentially, as in the Standard Model, but where the ordering is always predictable; theories like this include that of Koutsoudas, Sanders and Noll (1974).

In 6.4 I discuss the strength and weaknesses of theories which claim that rules come in fundamentally different types whose ordering is predictable, focusing attention on Lexical Phonology. Lexical Phonology will be seen to overcome many of the problems inherent in the other models. Double Lookup, however, will turn out to share its good points but not its bad points.

I conclude the chapter with a chapter summary in 6.5.

6.1 The Standard Model

Chomsky and Halle (1968) present what I will call the Standard Model of rule ordering in the following convention.

(1) [= (29) in Chomsky and Halle 1968:341]

Rules are applied in linear order, each rule operating on the string as modified in all earlier applicable rules.¹

¹Chomsky and Halle (1968:344) modify this later somewhat to allow for the simultaneous application of the subparts of an infinite rule schemata. Since the need for such conventions has been eliminated by the formalism of autosegmental phonology, I will not consider this here. Other modifications to this simple convention allowed by Chomsky and Halle (1968), in particular cyclicity, were discussed in Chapter 4.
In support of this model, Chomsky and Halle (1968) note two empirical generalizations concerning phonological patterns, repeated below.

(2) a. [= (13) in Chomsky and Halle 1968:18]

It is always possible to order the rules in a sequence and to adhere strictly to this ordering in constructing derivations without any loss of generality as compared to an unordered set of rules or a set ordered on a different principle.

b. [= (14) in Chomsky and Halle 1968:18]

Such linear ordering makes it possible to formulate grammatical processes that would otherwise not be expressible with comparable generality.

As many researchers have noted, the Standard Model faces several empirical and conceptual problems. I will single out three, which I call the combinatorial problem, the real-time problem and the rule insertion problem.

The combinatorial problem is as follows. Notice that the statement in (1) implies that all rules in a phonological system are fully ordered; that is, the rules can be put into a one-to-one correspondence with an ordered subset of the natural numbers. In practice, however, phonological patterns are only capable of showing evidence for partial ordering, since many rules do not interact at all. The result is that if a speaker's knowledge of phonology includes a full ordering of rules, there is no way for her to learn the specific ordering relations of many of the rules.
The severity of the problem can be seen if we look at an actual example. In their argument against theories of predictable rule ordering (to be discussed in 6.3 below), Cathey and Demers (1976) give the following list of rules of Old Icelandic. Arcs indicate ordering relations between rules for which there is claimed to be evidence.

(3) [after Cathey and Demers 1976:622]

As densely packed as they are, however, these arcs only represent a partial ordering of these rules. The full ordering represented by the listing of all the rules on the page is only one out of many possible alternative full orderings consistent with the indicated partial ordering. The number of these full orderings is difficult to calculate, since it depends on the specific ordering relations of each of the rules. However, it is not difficult to appreciate just how large the number is, even in a case such as this with such a relatively complete partial ordering.
Consider the rule of æ-Raising. The only restriction on the ordering of this rule is that it must be preceded by i-Mutation. Assuming all the other rules remain where they are, æ-Raising can be ordered in seven possible ways: directly before u-Mutation, directly after u-Mutation, directly after Terminal Syncope, directly after UV-Raising, and so forth. This means that there are seven possible derivations that preserve all the orderings illustrated above:

<p>| | | | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(4)</td>
<td>a.</td>
<td>Syllabification</td>
<td>Stress Shift</td>
<td>Internal Syncope</td>
<td>i-Mutation</td>
<td>æ-Raising</td>
<td>u-Mutation</td>
<td>Terminal Syncope</td>
<td>UV-Raising</td>
</tr>
<tr>
<td></td>
<td>b.</td>
<td>Syllabification</td>
<td>Stress Shift</td>
<td>Internal Syncope</td>
<td>i-Mutation</td>
<td>u-Mutation</td>
<td>æ-Raising</td>
<td>Terminal Syncope</td>
<td>UV-Raising</td>
</tr>
<tr>
<td></td>
<td>c.</td>
<td>Syllabification</td>
<td>Stress Shift</td>
<td>Internal Syncope</td>
<td>i-Mutation</td>
<td>u-Mutation</td>
<td>Terminal Syncope</td>
<td>æ-Raising</td>
<td>UV-Raising</td>
</tr>
<tr>
<td></td>
<td>d.</td>
<td>Syllabification</td>
<td>Stress Shift</td>
<td>Internal Syncope</td>
<td>i-Mutation</td>
<td>u-Mutation</td>
<td>Terminal Syncope</td>
<td>UV-Raising</td>
<td>æ-Raising</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glide-drop</td>
<td>Vowel Elision</td>
<td>Glide Formation</td>
<td></td>
<td>Glide-drop</td>
<td>Vowel Elision</td>
<td>Glide Formation</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Glide-drop</td>
<td>Vowel Elision</td>
<td>Glide Formation</td>
<td></td>
<td>Glide-drop</td>
<td>Vowel Elision</td>
<td>Glide Formation</td>
<td></td>
</tr>
</tbody>
</table>
Syllabification
Stress Shift
Internal Syncope
i-Mutation
u-Mutation
Terminal Syncope
UV-Raising
Glide-drop
æ-Raising
Vowel Elision
Glide Formation

Syllabification
Stress Shift
Internal Syncope
i-Mutation
u-Mutation
Terminal Syncope
UV-Raising
Glide-drop
æ-Raising
Vowel Elision
Glide Formation

i-Mutation, on the other hand, has no ordering relation with Stress Shift. Thus for each of the seven orderings listed above, i-Mutation could be listed either before or after Stress Shift, resulting in a total of 2x7=14 different orderings.

Stress Shift, in turn, has no ordering relationship with Syllabification. If Stress Shift were ordered before Syllabification and all other rules left where they are, there would again be seven ways æ-Raising could be ordered, bringing the total up to twenty-one

2Syllabification, in Cathey and Demers's (1976) terms, is a rule that changes /j/ to /l/ in a specific environment, and thus does not necessarily relate to syllable-building, which of course must precede stress assignment and hence stress shift.
full orderings. This number must be quadrupled, since for any of these twenty-one orderings, Terminal Syncope can be ordered immediately before UV-Raising, immediately before Glide-Drop, immediately before Vowel Elision, or immediately before Glide Formation (before inserting æ-Raising in one of its seven places), bringing the total up to eighty-four possible full orderings.

This game can be extended for quite a while. Clearly if one considers the full set of phonological rules of a language (Chomsky and Halle 1968:239-245, for example, list at least thirty-four rules in English, to which Bailey 1973, Halle and Mohanan 1985 and others add many more), the number of possible full orderings consistent with the evidence can become quite enormous.

There are at least three ways the combinatorial problem can be faced. The simplest is to reject the claim that all rules need to be fully ordered; noninteracting rules need not be considered ordered at all. This is the view adopted by many theories of rule ordering, such as Local Ordering (Anderson 1974) and the theory of Koutsoudas, Sanders and Noll (1974). Such a view, however, allows for transitivity violations. If rule A is ordered before rule B and B is ordered before rule C, but rules A and C do not interact, this view would have A and C unordered although according to transitivity, it must be the case that A is ordered before C. As observed in Chapter 5, transitivity violations are incompatible with the Productivity Hypothesis; apparent cases of it always involve fake cyclicity or idiosyncratic ordering. Hence I will not adopt this solution here.
Another resolution to the combinatorial problem is to suppose that there are universal principles that can determine the ordering of rules even if the rules don't interact; this idea is discussed in 6.3 and 6.4 below, where aspects are shown to be fully compatible with Double Lookup.

The third way to deal with this problem is to suppose that in acquiring phonology, the ordering of noninteracting rules can be settled on arbitrarily, perhaps due to the order of acquisition of the rules, or the order in which crucial data is encountered. This concept is also compatible with Double Lookup, as I will now demonstrate.

An interesting argument for the claim that the ordering of noninteracting rules can be settled on arbitrarily is found in Smith (1973). Smith records a large number of phonological rules accounting for mispronunciations of English words, called realization rules, produced by his young son Amahl at various stages of development. At early stages (ie, around two years three months, and thereafter), Amahl pronounced the alveolar stops /n/ and /t, d/ as the velars [ŋ] and [g], respectively, before a syllabic [l], as in the examples in (5a). The examples in (5b) show that the rule does not apply to fricatives. The notation throughout is Smith's.

\[(5)\]

a. handle → ɛŋu
   pedal → ɡɛɡu
   beetle → ɡiːɡu
   bottle → ɡιɡu
   puddle → pʌɡeɪl \{later developmental stage after loss of /l/ rounding\}
b. whistle → wibu.  puzzle → padəl        {later stage}  muscle → matəl        {later stage}

At this same stage, Amahl also had a rule deleting /s/ preconsonantally, as in the examples in (6).

(6) biscuit → bɪɡɪk  escape → ĝeɪp  skin → ĝin  Smith → mit  spoon → ɬuːn  scream → ĝiːm  swing → wɪŋ

At this early stage, Smith (1973) found no words in Amahl's speech where these two rules had to interact. Later, however, around the age of two years eight months, Amahl added his version of the word pistol to his active vocabulary, pronouncing it [pɨtəl]. Assuming the velarization rule does not apply if the alveolar stop is preceded by /s/, this pronunciation can be derived by an ordered application of the velarization rule followed by the s-deletion rule, as illustrated in (7a). The opposite order of application does not work, as illustrated in (7b).

(7)  a. UR  pistol  
      Velarization {not applicable}  
      s-Deletion pɨtəl

b. UR  pistol  
      s-Deletion pɨtəl  
      Velarization *pɨkəl
Smith (1973:160) rejects the possibility that Velarization and s-Deletion were unordered until words such as pistol came along for which the ordering was crucial, for three reasons. First, upon being confronted by words such as pistol, Amahl produced a consistent pronunciation immediately, indicating a consistent order of application of the two rules. Second, both Velarization and s-Deletion were already ordered in the earlier stage with respect to other rules. Third, if rule orderings were settled on only when faced with specific examples, one would expect rules to be reordered constantly as new and more complex data came to light; instead, Smith (1973) only found one apparent case of rule reordering during the course of Amahl's phonological development. Smith (1973) concludes that at every stage all of Amahl's rules, even noninteracting ones, were fully ordered.

This solution to the combinatorial problem is fully consistent with the Productivity Hypothesis. Except for rules that are fully productive, it is unlikely that any two rules will agree precisely in their productivity. The relation "less productive than" imposes a full ordering on a set of rules where no two rules agree in productivity. Thus it may be that Velarization was already less productive than s-Deletion at the earlier stage, so that their ordered application in forms like pistol was automatic.\(^3\)

\(^3\)Naturally, we would still expect that fully productive rules, even in children's speech, should interact solely in accordance with Automatic Feeding and Prosody First.
The strength of the Double Lookup model goes beyond this, however, since the combinatorial problem is only one part of the more general problem faced by the language learner in acquiring knowledge of rule ordering. According to the Productivity Hypothesis, rule ordering follows automatically from productivity. This means the language learner doesn't need to acquire knowledge of rule ordering as part of her phonological competence, only knowledge of each rule's productivity. Since speakers do demonstrate knowledge of gradient productivity of different rules (see Chapter 3), they must have acquired this knowledge somehow. Thus even defenders of the Standard Model, who require language learners to acquire knowledge of rule ordering, must also require them to learn rule productivity. Considering merely what it formally states, therefore, the Productivity Hypothesis cuts the learner's work in half. Actually the saving is even more than this. It is likely that at least a first guess at the productivity of a rule is automatically acquired during the learning of the rule, since the degree to which a rule is surface-true should have an effect on how strongly recorded the rule becomes in memory. Thus within the Double Lookup model, knowledge of a rule's productivity, and therefore also its ordering, is acquired automatically with the learning of the rule itself.

I turn now to a second major problem with the Standard Model, the real-time problem. This essentially is as follows. It may be that linearly ordered phonological rules as described in the Standard Model never apply on-line, or it may be that they always
apply on-line. If they never apply on-line, it is not clear why the rules would be ordered at all, since ordering is an inherently asymmetric relationship, while the relationship between forms stored in memory need not be. As Bromberger and Halle (1989) point out, the asymmetry allowed by ordering makes more sense if rules are instead thought of as relating forms in memory with forms produced during speech:

Since underlying phonological representations of words are stored in speakers' permanent memory, whereas phonetic surface representations are generated only when a word figures in an actual utterance, there is a clear and theoretically significant sense in which underlying representations are prior to surface representations, a sense that justifies thinking of the surface form as "derived" from the underlying form. [Bromberger and Halle 1989:53]

And yet the contrary assumption, that rules always apply on-line, is even more problematic. Specifically, if all rules apply in real-time, each rule must apply extremely rapidly, since retrieval of phonological forms occurs in a very short period of time. For example, Schriefers, Meyer and Levelt (1990) give evidence that the phonological form of a word may be retrieved within 150 milliseconds of seeing a picture illustrating that word. At this rate, the thirty-four ordered rules of English phonology postulated by Chomsky and Halle (1968) would have to be checked (and applied if appropriate) in less than 4.5 milliseconds each; the rate of course would be greater if retrieval of full phonological forms is actually faster and if there are more than thirty-four ordered rules in English.
Since the assumption that rules never apply on-line and the assumption that rules always apply on-line both lead to problems, one clear solution is to suppose that rules sometimes apply on-line and sometimes don't. This solution can be interpreted either as Double Lookup does, where partially productive rules are thought of as sometimes applying on-line and sometimes not, or as the models invoking rule typology do, where one rule type never applies on-line while the other always does. In 6.4 I compare these two approaches and show that the Double Lookup solution is better.

One final weakness of the Standard Model of the several more that could be mentioned is the rule insertion problem. It has been shown (e.g., King 1973) that diachronically new rules are added to a phonological system at the end of the derivation. In other words, the insertion of a new rule between two older rules does not seem to occur. Halle (1962:66) notes this fact, and accounts for it as follows:

All other things being equal, a rule will affect intelligibility less if it is added at a lower point in the order than if it is added higher up.

---

4 In this regard, King (1973:553) observes that "rule insertion does not correspond systematically to traditional categories of change. Rule loss and types of rule re-ordering correlate... with analogical leveling and extension." Rule addition at the end of a derivation, of course, corresponds with the concept of a sound law.

5 Halle (1962) actually holds that adding rules at the end is a strong tendency, but not a universal, giving a putative example of true rule insertion in Middle English. Halle (1962:68) notes, however, that the example may be "hypothetical" since the facts are not fully clear.
In other words, rule addition behaves like the addition of a new biological trait in the evolution of a species. As Darwin (1859:341-2) theorized, ontogeny tends to recapitulate phylogeny because new traits that manifest themselves in early stages in development tend to disrupt the later stages, leading to a reduction in fitness. The species that survive do so because by chance they have added their new traits later in development. Likewise, Halle (1962) seems to be arguing, Universal Grammar contains no constraint against inserting rules early, but an idiolect that chooses to do so will decrease its intelligibility and hence the functionality of its phonological system, reducing its fitness in the struggle against competing idiolects.

Although this explanation has a superficial plausibility, upon reflection it becomes clear that it only does so if one assumes essentially the opposite of what Halle (1962) wants to argue. This can be seen if one considers the question of why the addition of a rule at the end of a derivation is expected to disrupt intelligibility less than the addition of a rule nearer the beginning. Halle (1962) seems to assume that a rule added at the end of a derivation will have a more regular effect and therefore be more obvious and easier to undo by a listener whose idiolect lacks the rule; a rule added near the beginning will have effects that appear more irregular. The total number of affected forms in both cases may be roughly the same; the key difference is how regular the change seems to be. However, the very foundation of the Standard Model rests on the premise that
apparent "irregularities" caused by the interaction of rules can be
counted for in a perfectly regular fashion through rule ordering.
The people who have not added the new rule and who are trying to
understand the innovators presumably know all the other rules of
the language and their ordering. Thus it should be as easy for them
to deduce where the new rule was inserted as it was for the
innovators to insert it there.

The real solution to the rule insertion problem seems to lie
with the innovators, not their listeners. It does not seem equally
easy to insert a rule in every place in the derivation. If it is true that
later rules are easier to perceive and therefore learn than earlier
rules, as Halle's (1962) argument seems to require, then earlier rules
should take longer to learn than later rules. As discussed in Chapter
3, this seems to be true; Jaeger (1986), for example, finds that pre-
literate children continue to show no signs of knowing the rule of
Vowel Shift long after having mastered later rules. It seems
reasonable to suppose that the ease of learning a rule is correlated
with the rule's productivity. Given this, it is hardly surprising that
innovators (primarily children) will not be able to insert new rules

6The same is true of morphological rules. Thus while it has long been known
that children display knowledge of inflectional rules from a very young age
(see, eg, Berko 1958, Marcus et al. 1990), Tyler and Nagy (1989) find that the
knowledge of the rules of English derivational morphology, ordered before
inflectional rules, continues to increase gradually between fourth and eighth
grades. Likewise, within the class of inflectional rules, Guy and Boyd (1990)
find that recognition of the morphological status of early-applied suffixes /t,
d/ in semi weak verbs such as kept and told only comes, if it comes at all, in
adulthood, that is, over the age of twenty.
before earlier rules, simply because they have not yet been learned as rules.

This is essentially the solution to the rule insertion problem offered by Double Lookup; typological models of rule ordering, to be discussed in 6.4, also exploit this sort of explanation.

In this subsection I have discussed the weaknesses of the Standard Model of rule ordering. Three of these are the combinatorial problem, the real-time problem and the rule insertion problem. The models described in the remainder of this chapter owe their existence, in many cases, to attempts to overcome some or all of these problems.

6.2 Simultaneous rule application

The Standard Model is not without its strengths as well, though. Primary among these is the claim that rules can be ordered at all. That is, it claims that in many cases the interaction of rules can best be understood by thinking of them as ordered. In this section I show that this is true. Models that disallow rule ordering entirely, such as Lamb (1966), thus face both empirical and conceptual challenges.

The empirical force of the claim that rules can be ordered can be seen by constructing artificial examples of rule interactions that could not be handled by nonlinear theories of rule ordering, such as obligatory simultaneous ordering. A phonological rewrite rule has the basic form $A \rightarrow B / C$, which represents the concept "$A$ is replaced by $B$ in the context of $C$," where $A$, $B$ and $C$ are phonological representations (features, segments, or something more complex).
Thus there are thirty-four possible ways the rules I: $A_1 \rightarrow B_1 / C_1$ and II: $A_2 \rightarrow B_2 / C_2$ can interact (assuming $X_1 \neq Y_1$ and $X_2 \neq Y_2$ for all $X_1, Y_1 \in \{A_1, B_1, C_1\}$ and all $X_2, Y_2 \in \{A_2, B_2, C_2\}$), depending on whether $A_1 = A_2$, and/or $B_1 = C_2$, and so on. In twenty-five out of these thirty-four possibilities, applying rules I and II simultaneously will have the same effect as ordering rule I or rule II first. In five cases, simultaneous application is simply incoherent, since the two rules affect the same form in the same environment, but replace it by two different things. For example, consider the rule pair I: $A \rightarrow B / C$ and II: $A \rightarrow D / E$. Thus in a language with these rules, CA will become CB and AE will become DE. Now consider the form CAE. If rule I applies first, CAE will become CBE, and rule II will be bled. If rule II applies first, CAE will become CDE, and rule I will be bled. The two rules clearly cannot apply simultaneously, since one rule has A becoming B and the other has A becoming D. A theory that requires all rules to apply simultaneously would have to claim therefore that such rule pairs (or lexical items such as CAE) cannot exist.

The remaining four types of rule interactions are cases where simultaneous application of the two rules results in output that is distinct from the output when the rules are applied in either order. Chomsky and Halle (1968:19,fn5) discuss a case of this sort. Another familiar example of this is the class of exchange rules, discussed by Chomsky (1967), Anderson and Browne (1973), Janda (1987) and others, that is, rule pairs of the form I: $A \rightarrow B / C$ and II: $B \rightarrow A / C$. If rule I applies first, the form ABC will first become BBC, and then
AAC by rule II (assuming A, B and C are all local to each other). If rule II applies first, ABC will first become AAC, and then BBC by rule II. If the two rules apply simultaneously, however, ABC will become BAC.

It is significant that apparent cases of this sort always turn out to be explainable in other ways. For example, the simultaneous application of exchange rules plays a large role in folk descriptions of accents. John Fowles has his Cockney character in *The French Lieutenant's Woman* switch /h/ and /w/, pronouncing "horrible" as "orrible" and "awful" as "hawful"; speakers of standard German in Erlangen believe that the local dialect, Franconian, switches /p,f,t,s,k/ and /b,v,d,z,g/; Americans imitating a Japanese accent systematically switch /r/ and /l/, as in "velly sirry" for "very silly". Of course, in cases like this, no pair of exchange rules actually exists. At most one could say there is a pair of rules that neutralize two distinct segments as a sound "intermediate" between the originals; nonnative listeners of the nonneutralizing dialect then perceive this intermediate form, which is clearly distinct from both originals, as the opposite of whatever the appropriate original segment should be.

The hypothesis that all rules apply simultaneously can be saved if one allows "later" rules to repeat some of the information contained in "earlier" rules. For example, consider a rule pair where the first rule accounts for the alternation between /f/ and /p/ in forms such as *father*~*paternal*, and the second changes /p/ into [ph]. Suppose the aspiration rule only applies to stops, so that these rules
cannot apply simultaneously as given, or else *paternal would surface as *[p paternal, not *[p paternal. This obstacle could be overcome by modifying the aspiration rule so that it applies in precisely the same environment as the /f/-to-/p/ rule, in addition to applying in other, more general environments. This would work, but it would be formally unsatisfying for at least three reasons. First, it would require redundancy unnecessary under the rule ordering analysis. Second, the intuitive sense of aspiration as a single process would be lost. And third, as Chomsky and Halle (1968) point out for similar cases, if there were evidence of further rules ordered after aspiration or before the /f/-to-/p/ rule, the redundancy of the later rules would simply compound indefinitely. This is the idea captured in generalization (2b).

I conclude, therefore, that standard linguistic evidence supports the claim of the Standard Model that rules must at least sometimes be ordered. Support for this claim comes from psycholinguistic considerations as well. Primary among these is the observation that motivates the Double Lookup model in the first place: rules differ in productivity. It seems implausible to suppose that a rule as unproductive as the /f/-to-/p/ rule must necessarily apply at the same time as a very productive rule like aspiration.

Arguments such as these lead me to suspect that interacting rules are rarely if ever applied simultaneously. This is of course
fully consistent with the Productivity Hypothesis and the principle of Automatic Feeding.  

6.3 Predictable rule ordering

As observed above, the combinatorial problem is just one aspect of a general problem faced by the Standard Model, namely the problem of learning the ordering of rules. If a solution like Smith's (1973) is adopted, where language learners are assumed to arbitrarily order each rule relative to all the others as it is acquired, then Smith cannot be right when he claims that rule reordering is a rare occurrence. If there are no constraints on rule ordering, a child acquiring rules A and B may arbitrarily choose to order A before B, forcing her to reorder them when evidence makes it clear that there is a rule C that precedes A and follows B. In order to prevent this sort of constant reordering, and to account for what happens with rules that don't interact, it seems plausible to hypothesize that all interacting rules are ordered intrinsically, that is, by universal principle.

This is, of course, exactly what the Double Lookup model does, the relevant principle being the Productivity Hypothesis. It must be kept in mind, however, that the fact that rule ordering is principled in Double Lookup does not mean it is predictable. On the contrary,

---

7By contrast with phonological rules, Anderson and Browne (1973) and Janda (1987) observe that morphological exchange rules, which must be thought of as applying simultaneously, do exist. This seems to be a result of the fact that the rules they discuss are not very productive. If such rules are primarily used to relate forms in memory, there is no need for them to be ordered, as noted in the previous section; the fact that some of them interact in a way inconsistent with an ordered application is thus expected.
learning the productivity of a rule depends at least partly on what percentage of forms conforms to the appropriate pattern, and this percentage depends on the order in which the learner is exposed to adult forms and other arbitrary factors.

The most prominent example of a theory claiming all rule ordering is predictable is the Universally Determined Rule Application hypothesis (henceforth UDRA), espoused in works such as Koutsoudas, Sanders and Noll (1974) and Koutsoudas (1980). The UDRA essentially consists of the following claims:

(8) Universally Determined Rule Application

a. Rules not subject to special principles apply wherever their structural descriptions are met.

b. Noninteracting rules apply simultaneously.

c. Most problematic cases can be handled by means of special rule ordering principles.

d. Remaining problematic rule pairs can be handled by changing the form of one rule to include information handled by the other rule.

The fact that rules apply wherever their structural descriptions are met means that all interacting rules (except those for which special principles are relevant) appear to apply in a feeding order.


8 It should be noted, however, that in these works the motive given for predictable rule ordering is often not the combinatorial problem, but the idea that rule ordering is too powerful a device, blinding the phonologist to other possible explanations for phonological patterns. In this regard, see also Kenstowicz and Kisseberth (1970), Churma (1980) and Hayes (1986a).
1977, Churma 1980, Coates 1982, Kanai 1982, Noske, Schinkel and Smith 1982). The debate hinged on two issues. First, is rule ordering really always predictable? The evidence seems to suggest that it is not, since virtually every one of the proposed universal ordering principles has its counterexamples (see the above works). Second, is the UDRA hypothesis truly an empirical hypothesis? For example, the Standard Model explicitly rejects analyses such as those advocated in (8d) because they do not allow rules of maximal generality, even though such analyses handle the facts just as well. Moreover, finding cases of rule ordering that disprove particular universal principles can never show that there are not as-yet undiscovered principles that do predict the ordering.

Koutsoudas (1980:22) argues that the UDRA hypothesis has empirical content, noting that it "can be falsified only by showing that there are two well-motivated grammars of natural language which must be identical in every respect except in the way their rules interact in derivations." This is an extremely tall order, but cases of precisely this sort have been claimed to exist by Halle (1962), Bailey (1973), Robinson (1977), Noske, Schinkel and Smith (1982), Halle and Mohanan (1985), and Hall (1992), among others.

Since some of these cases are problematic (especially that discussed by Halle 1962; see Chapter 4), it may be that closer examination will vindicate the UDRA hypothesis. I see no reason for expecting this to happen, however. If it is true that rules are added to phonological systems at the end of the derivation, and if over
historical time rules may be added for any number of arbitrary reasons, then it would be extremely surprising to find that there are absolutely no cases of extrinsic rule ordering in synchronic phonological systems.

Moreover, even if these problems with the UDRA hypothesis could be dealt with, it would still only solve the combinatorial problem. The real-time problem and rule insertion problem would remain unsolved.

Consider for example the real-time problem. Even if rules are ordered by universal principle, the question of whether they apply in real time or not must be addressed. If all rules apply in real time, the fact that they are ordered by principle and not by stipulation surely cannot speed up their application. If, on the other hand, no rules apply in real time, then the existence of universal ordering principles becomes more mysterious than the problems it seeks to resolve.

In short, theories claiming that all rule ordering is predictable suffer from several shortcomings. First and foremost, they are virtually impossible to falsify. Second, they offer no explanation for how universally ordered synchronic rules can arise out of the arbitrary diachronic ordering in which rules are added to a language. Finally, these theories do not address the real-time and rule insertion problems. I therefore conclude that Double Lookup, which does not face these difficulties, is a superior model of rule ordering.
6.4 Rule typology

In the discussion of the Standard Model, it was noted that one solution to the combinatorial problem is to suppose that rules apply in intrinsically ordered blocks of rules of well-defined types; the ordering of a rule could thus be approximated by a language-learner, even if the rule did not interact with other rules, if the rule had characteristics of some particular type. This solution does not solve the combinatorial problem entirely, of course, since rule ordering within each block may not be predictable, but it does reduce its severity.

In this section I consider the currently most influential theory of rule typology, namely Lexical Phonology (see, eg, Kiparsky 1982, 1984, 1985, 1988; Mohanan 1982, Mohanan and Mohanan 1984, Halle and Mohanan 1985; Borowsky 1986; Pulleyblank 1986; Kaisse and Shaw 1985; Kaisse 1985, 1990; Hayes 1990; among many others). I show that although it shares some good points with Double Lookup, it is ultimately undone because it insists, unlike Double Lookup, that rules come in fundamentally distinct types.

The fundamental division in rule types posited by Lexical Phonology is the distinction between *lexical rules* and *postlexical rules*. All lexical rules are supposed to apply before all postlexical rules. The reason typically given for this ordering is that lexical

---

9 As noted in Chapter 1, this typology matches rather closely the traditional division of phonology into morphophonemics and allophonics (see eg Sapir 1925) and the distinction in Natural Phonology between rules and processes (see eg Stampe 1973, Donegan and Stampe 1979).
rules apply before the insertion of lexical items into sentences, while postlexical rules apply after lexical insertion.

Below I list the characteristics that have been given to distinguish these two rule types.

(9) Lexical rules and postlexical rules [after (8) in Pulleyblank 1986:7]

<table>
<thead>
<tr>
<th>Lexical rules</th>
<th>Postlexical rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. may refer to word-internal structure</td>
<td>cannot refer to word-internal structure</td>
</tr>
<tr>
<td>b. may not apply across words</td>
<td>may apply across words</td>
</tr>
<tr>
<td>c. may be cyclic</td>
<td>cannot be cyclic</td>
</tr>
<tr>
<td>d. if cyclic, then subject to strict cycle</td>
<td>non-cyclic, hence across-the-board</td>
</tr>
<tr>
<td>e. can only produce phonemes</td>
<td>need not produce phonemes</td>
</tr>
<tr>
<td>f. may have lexical exceptions</td>
<td>cannot have lexical exceptions</td>
</tr>
<tr>
<td>g. must precede all post-lexical rule applications</td>
<td>must follow all lexical rule applications</td>
</tr>
</tbody>
</table>

In addition to this rule typology, one of the primary ways Lexical Phonology differs from the Standard Model is its claim that phonological lexical rules are interleaved with morphological rules. Specifically, lexical rules are subdivided into ordered language-particular blocks that are morphologically defined.10

Although Lexical Phonology is not meant to be interpreted as a theory of phonological performance, processing claims based on it

---

10The specific way this is done varies among different versions of Lexical Phonology. In the version advocated by Kiparsky (1984) and Borowsky (1986), for example, there is only one ordered set of rules, and each morphological level is associated with an ordered subset of this. See Chapter 5 for implications of such a model.
have been made. For example, Kiparsky (1988) suggests that during sound change, new lexical rules entrench themselves in the language through lexical diffusion, while postlexical rules are added in an exceptionless "NeogrammARIANT" fashion. Postlexical rules behave this way because they are psychologizations of automatic processes of articulation and perception. This neatly derives the Natural-Phonology-like property, often attributed to postlexical rules, of "phonetic naturalness."

In addition, Shattuck-Hufnagel (1986), following Mohanan's (1982) claim that speech errors do not trigger lexical rules, suggests that perhaps lexical rules do not apply on-line, while postlexical rules do. Finally, several studies (e.g., Guy 1991a and Gordon 1989) have claimed experimental evidence for the psychological reality of morphological level ordering.

With this as background, Lexical Phonology has several strengths and weaknesses. First, the rule insertion problem is solved, since new rules begin life as psychologizations of mechanical phonetic effects, which because of their origin will be exceptionless, across-the-board, and so forth, and so will be postlexical rules, which apply last in the derivation. Second, by dividing the derivation into blocks, it reduces but does not solve the combinatorial problem. Third, if it is true that lexical rules do not apply on-line, the real-time problem is reduced but not eliminated as well, since postlexical rules are still assumed to be ordered as in the Standard Model. Finally, Lexical Phonology suffers from the problem faced by all rule
typology theories: the diagnostics distinguishing between its rule types do not work, as I demonstrate below.\textsuperscript{11}

By contrast, Double Lookup has all the strengths and none of the weaknesses of Lexical Phonology. It too solves the rule insertion problem by having new rules first show up as fully productive, hence ordered last in the derivation. It too reduces the combinatorial problem by having the ordering of rules follow from other differences in the rules; in this case, however, the ordering is exhaustive because productivity distinguishes between any two rules, not just between rule blocks. Third, Double Lookup also claims that most rules do not apply on-line; the difference is that the rules that do apply on-line are not really ordered either. Finally, Double Lookup does not claim that there is a fundamental distinction between on-line and prepatterned rules. The evidence presented in Chapters 3, 4 and 5 should have made it clear that some rules can be both partially productive and yet sometimes apply on-line, while others can be both fully productive and yet also be prepatterned.

The similarities and differences between Double Lookup and Lexical Phonology can be thought of this way. Lexical Lookup, with its retrieval of prepatterned allomorphs, parallels the application of lexical rules (as they are ideally conceived), and Rule Lookup, with its retrieval and on-line application of rules, parallels postlexical rules (again, as they are ideally conceived). Thus something like the

\textsuperscript{11}In addition, Lexical Phonology faces special problems of its own, in particular in its claims about cyclicity, discussed in Chapter 5.
lexical/postlexical distinction is real, but indeed only something like it; what is truly real is Double Lookup. The major difference between Lexical Phonology and Double Lookup is that the latter claims the "lexical/postlexical" border can shift with every derivation; sometimes the allomorph input to the rules will be rather abstract, sometimes it will be rather concrete.

I now spend the rest of this section showing that the rule typology posited by Lexical Phonology does not work. In pursuing some of the relevant evidence, we will find additional support for the view of rule application presented in Double Lookup.

For the rule typology of Lexical Phonology to make sense, none of the properties exclusively associated with lexical rules can cooccur with any of the properties exclusively associated with postlexical rules. It is well known, however, that this is not true. Some rules in Tiberian Hebrew both apply across word boundaries and are (apparently) cyclic (Dresher 1983; for arguments that these rules are in fact lexical, see Churchyard 1992). Some rules in English and other languages both refer to word-internal structure and yet are not structure-preserving (Sproat 1985, Borowsky 1986, Harris 1989). Rules have been claimed to exist that follow all lexical rules and yet are phonetically unnatural (Halle and Mohanan 1985).

Let me mention just two more significant mismatches. First, as the examples discussed in Chapter 3 make clear, Mohanan (1982) and Shattuck-Hufnagel (1986) are simply wrong in claiming that speech errors never trigger lexical rules. More significantly, the sorts
of speech errors triggering these rules often make reference to information beyond the word boundary, as in segment shifts where the origin of the segment is in another word in the utterance (as in Fromkins's 1971 *Sprigtime for Hintler* example). Thus lexical rules can follow speech errors, and they also apparently can follow lexical insertion, casting doubt on what is often assumed to be the primary difference between lexical and postlexical rules.

A second significant mismatch is that rules in many languages both apply across word boundaries and also have lexical properties such as lexical exceptions, cyclicity, and/or phonetic unnaturalness (Kaisse 1985, 1990). I will discuss this mismatch more fully because the search for an explanation for it not only leads us out of Lexical Phonology but also directly into Double Lookup.

Kaisse (1985, 1990) offers an explanation for the existence of such cases, which, while it works, is not in fact explanatory at all. Because the primary distinction between lexical and postlexical rules is standardly held to be when they apply relative to lexical insertion, she takes such rules to be postlexical rules that happen to have some lexical properties. Kaisse (1985, 1990) calls such rules P1 rules, contrasting them with P2 rules which follow all P1 rules and truly do have all of the properties of postlexical rules listed in (9).

This solution is not explanatory for two reasons. First, it merely stipulates a distinction between the two types of postlexical rules. For instance, no reason is offered for the fact, left as a

---

12Thanks to Merrill Garrett for pointing out this discrepancy.
coincidence here, that the lexical-like P1 rules are ordered close to the lexical stratum. Second, it reduces the meaningfulness of the lexical/postlexical distinction by letting some postlexical rules have lexical properties.

Hayes (1990), however, offers a solution that can accommodate P1 rules within the lexical level, even though they are sensitive to the environment beyond the word boundary. Hayes (1990) proposes that apparent cases of P1 rules should be handled by means of Precompilation Theory. P1 rules, he claims, are lexical rules which generate different allomorphs of forms for all possible word-external environments. The set of such environments is always very small, especially since syntactic information often plays a role in limiting the application of the rule. Since these are lexical rules, they may well have lexical exceptions, be structure-preserving, and have all the other properties lexical rules are believed to have. Then, after the allomorphs have been "precompiled" in this fashion (which Hayes 1990:94 is careful to emphasize is not intended to imply anything about the on-line nature, or lack thereof, of this precompilation process), postlexical allomorph selection occurs, so that the appropriate allomorph appears in the appropriate phrasal environment.\(^\text{13}\)

One of Hayes's (1990) strongest arguments for this position and against the P1 hypothesis is the existence of postlexical rules with

---

\(^{13}\)In Myers (1992) I argue independently for a virtually identical solution to the P1 problem.
lexical properties in Hausa, Mende and Kimatuumbi that actually precede unambiguously lexical rules. On the basis of this, and other considerations discussed in Myers (1992), I conclude that Hayes (1990) is right: P1 effects are simply the result of allomorph-generating lexical rules with postlexical allomorph selection.

Unfortunately for Lexical Phonology, however, Hayes's (1990) solution to the P1 problem brings up an entirely different set of problems, in that it allows lexical items to have predictably different allomorphs that are themselves lexical. That is, these allomorphs become predictable through allomorph selection in the postlexical level, but within the lexicon they are essentially unpredictable and so must appear separately at the output of the lexical phonology.

This seems to violate the spirit, if not the letter, of Lexical Phonology. The forms found at the level of output of the lexical rules are typically assumed to comprise what a native speaker of a language considers the set of lexical entries in the language, ie the "real words" (see, eg, Mohanan 1982). Allomorphs of the same word cannot appear separately at this level, since then they would then be considered different words. Double Lookup, which happily accommodates different allomorphs of single lexical items, does not face this problem.

But the situation for Lexical Phonology is even worse, since it turns out that not only can lexical rules result in distinct allomorphs that are then chosen postlexically due to syntactic and phonological considerations, as Hayes (1990) has shown, but they can in fact be
entirely optional. In other words, lexical rules can fail to apply for no reason at all. In Precompilation Theory, this would mean that such rules produce allomorphs at the output of the lexical phonology that are then chosen between by lexical insertion either at random, or in consideration of all the unruly factors playing a role in the application of optional rules.

One example of an optional lexical rule is of course Coronal Deletion in English, discussed in Chapter 5. Another, perhaps more familiar, example in English is that of vowel reduction, whereby certain vowels become schwa when unstressed. For many words, this rule is optional; some examples are given in (10).

(10) abstain  [æbstéyn]  [əbstéyn]  
paprika  [pɛprɪkə]  [pəprɪkə]  
acknowledge  [æknálədʒ]  [əknálədʒ]  

As Fidelholtz (1975) documents, however, whether or not a word is able to undergo this process often depends on a purely lexical property, namely its relative frequency. This can be seen in the near minimal pairs listed in (11), where the bolded and underlined vowels in the relatively frequent word undergo reduction, while the similar bolded and underlined vowels in the relatively infrequent word do not.
Not only is frequency not a phonological property, it is not even a grammatical property. Hence the pattern in (11) cannot be indicated systematically in the rules of English, but rather must be marked individually on each word. Thus the forms on the right that do not undergo vowel reduction must be marked as lexical exceptions, and vowel reduction must be a lexical rule.

Halle and Mohanan (1985) give further evidence of the lexical status of vowel reduction, arguing that it appears to be ordered before the lexical rule of Vowel Tensing, which refers to word-internal structure (specifically, word boundaries).

The relevant evidence for Vowel Tensing is as follows. There are claimed to be dialects in English which show a pattern in the tenseness of the indicated vowel in examples like those in (12). Halle and Mohanan (1985) account for this with a rule that tenses short vowels that are either followed by a word boundary or by the feature [-consonantal].

(11) [after (1) in Fidelholtz 1975:200]

<table>
<thead>
<tr>
<th>RELATIVELY FREQUENT</th>
<th>RELATIVELY INFREQUENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>[undergo reduction]</td>
<td>[do not undergo reduction]</td>
</tr>
<tr>
<td>astronomy</td>
<td>gastronomy</td>
</tr>
<tr>
<td>mistake</td>
<td>mistake</td>
</tr>
<tr>
<td>mosquito</td>
<td>Muskegon</td>
</tr>
</tbody>
</table>

(12) [from (67) in Halle and Mohanan 1985:81]

\[ \text{vary} \ [\text{veəri}] \]
\[ \text{variation} \ [\text{veərjən}] \]
In these dialects, this rule is bled by vowel reduction, since as seen in (13), short vowels do not tense if followed by schwa. This is because vowel reduction is assumed to operate by deleting all vowel features, leaving a bare timing slot; the short vowel is followed neither by a word-boundary nor the feature [-consonantal], and the rule cannot apply. As (13b) shows, in these dialects schwa blocks the tensing of a preceding short vowel even if the schwa is unambiguously derived, since it alternates with a nonreduced vowel in another word.

(13) [from (67) in Halle and Mohanan 1985:81]

a. variation [vərˈeiʃən]
   various [vərˈeərəs]

b. theatric [θiˈætrɪk]
   theater [θiˈætər]

Since vowel reduction bleeds a lexical rule, it must precede it and is therefore itself a lexical rule. Thus we have another example of an optional lexical rule. We have seen why such cases are problematic for Lexical Phonology. How might Double Lookup handle them?

Of course Double Lookup does more than handle such cases; it is premised on the notion that cases like this are in fact typical. The existence of allomorphs of the same word that act as input to on-line rules, required both with P1-type rules discussed by Hayes (1990) as well as with optional lexical rules, is exactly what one would expect given Redundancy of Patterning. The very role of Lexical Lookup is
to choose between memorized allomorphs and then pass one of them on to Rule Lookup. Normally the choice between allomorphs is based solely on their relative frequency, but other factors may be able to play a role, too. As Hayes (1990) observes, allomorph selection in more uncontroversial cases (as in the choice of an affix) can depend on factors as diverse as phonological environment, inflectional features, and the identity of individual morphemes. Moreover, allomorph selection in such cases, as expected if allomorph selection in general is done by Lexical Lookup, is also influenced by word frequency, as noted above with respect to the findings of Marcus et al. (1990).

In other words, Lexical Lookup as an on-line process is not parallel so much to the application of lexical rules as it is to Hayes's (1990) concept of allomorph selection. Since these allomorphs are stored in memory and not derived on-line, it is not surprising to find that patterns associated with them may have lexical exceptions.

To summarize, then, I have argued that the boundary between lexical and postlexical phonology is fuzzy because it corresponds to the boundary between Lexical Lookup and Rule Lookup. While in any particular on-line derivation, the boundary between Lexical Lookup and Rule Lookup is sharply delineated, where this boundary lies during the production of utterances can vary.

In short, the problem with theories of rule typology like Lexical Phonology is that they ignore the reality of Gradient Retrievability and Gradient Productivity. It is not the case the lexical rules never
apply on-line and postlexical rules always do, as Mohanan's (1982) comments about speech errors imply. Instead, the likelihood of a rule's on-line application varies gradiently from rule to rule. Patterns that are not derived on-line and those that are do differ from each other in systematic ways (the latter tend to regularize exceptions, and so on), but because of Redundancy of Patterning, there is no typological distinction between these classes of patterns. Any given pattern can potentially fall into either class at different times.

6.5 Chapter summary

In this chapter I have compared and contrasted Double Lookup with other models of rule ordering. None of these was seen to handle the combinatorial, real-time, and rule insertion problems as well as Double Lookup.

Some models brought new problems as well. Thus theories allowing for simultaneous application incorrectly predict the existence of phonological exchange rules. Theories allowing noninteracting rules to remain unordered incorrectly predict the existence of transitivity violations. Theories claiming that all rule ordering is predictable suffer from the difficulty of their falsifiability. Theories claiming that rules fall into clearly defined ordered strata suffer from the fact that the strata cannot be clearly distinguished.

Most interestingly, exposing the weaknesses of these other models often leads directly to further positive support for Double
Lookup. As we saw in section 6.4, the answer to the PI problem and related puzzles involves the assumption that surface phonological patterns can arise through allomorph selection, exactly as postulated in Lexical Lookup. More generally, resolution of the real-time problem seems to require the assumption that some rules apply in real time while others do not. Either we suppose that rules may be divided into a class that always acts in real time and a class that never does, or else we suppose that each rule may sometimes act in real time, and sometimes not. With the rejection of rule typology noted in the previous paragraph, the only option remaining is the latter, which is, of course, the fundamental assumption underlying Double Lookup.

In this chapter I hope to have shown the superiority of Double Lookup over other models of rule ordering that have appeared in the literature. This is only the final step in the long argument I have been developing through the entire dissertation. In previous chapters I showed how the ordering of partially productive rules in English is consistent with what is known about their relative productivity, in accordance with the Productivity Hypothesis (Chapters 2 and 3), how the interaction of fully productive rules supports the principles of Automatic Feeding and Prosody First (Chapter 4), and how the phenomenon referred to in the literature as cyclicity is explained and in fact expected by Double Lookup (Chapter 5).
It goes without saying that further work remains to be done. Thus it seems prudent to collect more reliable psycholinguistic information on the relative productivity of rules. In addition to continuing to subject English to intense scrutiny, other languages, particularly those that differ markedly from English, should be investigated. Finally, and most importantly, to my mind, more work is needed to determine precisely how forms come to be prepatterened in the first place. In many ways I feel that this is the key to understanding not only how phonological rules are learned and used, but also what role they play in theories of the phonological component of grammar.
REFERENCES


Cena, R. M. (1978) "When is a phonological generalization psychologically real?" Indiana University Linguistics Club: Bloomington, Indiana.


Wang, H. S. (1992) "An experimental study on the productivity of Taiwanese tone sandhi," National Tsing Hua University, Taiwan, ROC ms.


