The Identification of Bases in Morphological Paradigms

Adam C. Albright

University of California, Los Angeles

June, 2002

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Linguistics

UCLA Linguistics Department

Thesis committee:
Bruce Hayes, co-chair
Donca Steriade, co-chair
Carson Schütze
Brent Vine

Contents

List of Figures

List of Tables

ABSTRACT OF THE DISSERTATION

The Identification of Bases in Morphological Paradigms

by

Adam C. Albright
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 2002
Professor Bruce Hayes, Co-chair
Professor Donca Steriade, Co-chair

Many theories, in many domains of linguistics, assume that some members of morphological paradigms are more basic than others. Bases of paradigms are privileged in various ways: they may determine phonological properties of other forms, they may determine the direction of analogical changes, and so on. In this thesis, I propose that such effects are a result of the procedure by which learners seek to develop a grammar that allows them to project inflected forms as accurately and confidently as possible. I present a computationally implemented model of paradigm acquisition that attempts to use one form in the paradigm as the base to project the remaining forms, using stochastic morphological rules. I pursue two hypotheses about how this is done. The first is that learners are limited to selecting a single form as the base, and that the base must be a surface form from somewhere within the paradigm. Furthermore, the choice of base is global, meaning that the same slot must serve as the base for all lexical items. The second hypothesis is that learners select the base form that is maximally informative, in the sense that it preserves the most contrasts, and permits accurate productive generation of as many forms of as many words as possible.

As evidence for this approach, I analyze three cases in which an typologically marked form served as the base of a historical analogical change: Yiddish present tense paradigms (in which all forms were remodeled on the 1st sg), Latin noun paradigms (in which nominatives were remodeled on oblique forms), and Lakhota verbs (in which unsuffixed forms are being remodeled on suffixed forms). In each case, I show how the model correctly selects the base form, and also correctly predicts asymmetries in the direction of subsequent paradigmatic changes. I show that these asymmetries are not predicted by a more traditional model of underlying forms, in which learners compare all of the parts of the paradigm to construct abstract underlying representations that combine unpredictable information from multiple forms. Finally, I discuss possible extensions of this model to accommodate larger paradigms with multiple, local bases.

Acknowledgments

I would like to thank my committee members (Bruce Hayes, Donca Steriade, Carson Schütze and Brent Vine) for all of their patience, encouragement, and friendly criticism, and especially for agreeing to show up for a defense at 9am on a Sunday morning. First and foremost, I am thankful to Bruce Hayes, who has been a patient and wise teacher, an attentive advisor, and an enthusiastic collaborator during my time at UCLA. He has helped me more in my research and growth than many pages of acknowledgments could do justice to. I am also greatly indebted to Donca Steriade, for her wisdom and guidance, and for her knack for making hard problems seem so clear, at least until the time when I left her office. I am equally grateful to Carson Schütze for our many stimulating and challenging discussions in which he pushed me to take on hard issues, clarify my analysis, and see things from another angle. Finally, I am much obliged to Brent Vine for his thoughtful and insightful questions, for offering intriguing parallels and additional data, and for clearing up at least a few of my many misconceptions about Latin and Russian.

I would also like to thank all of the other members of the UCLA Linguistics Department who have made it such a vibrant and educational environment. I owe particular thanks to Colin Wilson for his discussions and advice, and to Pamela Munro for introducing me to Lakhota, and for generously providing me with her insights and the use of her Lakhota verb list. I am also especially grateful to Victoria Anderson, Marco Baroni, Rebecca Brown, Leston Buell, Ivano Caponigro, Taehong Cho, Christina Foreman, John Foreman, Matt Gordon, Amanda Jones, Sun-Ah Jun, Pat Keating, Sahyang Kim, Peter Ladefoged, Harold Torrence, Motoko Ueyama, Jie Zhang, Kie Zuraw, and all of the other members of the UCLA Phonetics Lab and Linguistics department (too numerous to name) who have made my years here stimulating and entertaining. In addition, I would like to thank my teachers during my undergraduate years at Cornell who instilled me with a love of phonology and historical linguistics—especially Abby Cohn, Wayne Harbert, and Jay Jasanoff.

Much of the data presented here would have been difficult or impossible for me to collect or make sense of on my own; I am particularly indebted to my Lakhota teacher Mary Rose Iron Teeth, for many hours patiently and cheerfully teaching me about her language. I am also extremely grateful to Angelo Mercado for his generous Latin help, and likewise to Jongho Jun, Taehong Cho, and Sahyang Kim for Korean, and to Katya Pertsova for Russian.

Various portions of this dissertation have benefited from questions and comments by audiences at UCLA, MIT, BU, the LSA (Chicago, January 2001), the ZAS Conference on Paradigm Uniformity (Berlin, March 2001), BLS 28 (Berkeley, February 2002), WCCFL 21 (Santa Cruz, April 2002), and the GLOW Phonological Acquisition Workshop (Utrecht, April 2002).

This dissertation work was supported in part by an NSF Graduate Fellowship, NSF grant

BCS-9910686, and a UCLA Dissertation Year Fellowship, all of which I gratefully acknowledge.

Writing a dissertation can be an ordeal, and it is made tolerable only by the unswerving support and encouragement of friends—even when you appear to have momentarily abandoned them. For this, I am eternally grateful to Moto and Marco, to Jie, Ivano, and Harold, to Scott, and most of all to Angelo. Last but not least, I am thankful to my parents, for their years of support and for giving me my first introduction to language.

Chapter 1

Introduction

The problem of bases in inflectional paradigms is neither new nor forgotten in linguistic theory. Consider, for example, the full set of forms which a noun could take in Sanskrit (Whitney 1924), a language with a moderately large set of distinct case forms:[1]

(1) Full paradigm for Skt. *pā́d* 'foot'

	Sg.	Dual	Pl.
Nom.	*pā́t*	*pā́dāu*	*pā́das*
Acc.	*pā́dam*	*pā́dāu*	*padás*
Instr.	*padā́*	*padbhyā́m*	*padbhís*
Dat.	*padé*	*padbhyā́m*	*padbhyás*
Abl.	*padás*	*padbhyā́m*	*padbhyás*
Gen.	*padás*	*padós*	*padā́m*
Loc.	*padí*	*padós*	*patsú*
Voc.	*pā́t*	*pā́dāu*	*pā́das*

It is commonly observed, starting at least as far back as Paul (1920, chap. 5), that for languages like Sanskrit, it would be impractical to memorize every form of every word, since there are so many forms and so many words. Even more dramatic examples include Hungarian, which has 924 possible forms for each noun (Tihany 1996), Archi, which is claimed to have up to 1,502,839 forms for each verb, if deverbal and commentative forms are included (Kibrik 1998), and Shona, which may have up to 16,000,000,000,000 verbal forms (Odden 1981).

Fortunately, it is intuitively clear that individual inflected forms are not unrelated, isolated words; rather, if we compare them (and consider a variety of other nouns), we can see that they stand in definite relations to one another. For example, we might observe that for this particular Sanskrit noun, the genitive and locative duals are identical, and the locative plural can be obtained by concatenating the nominative singular with a suffix *-su*, shortening the [aː], and shifting the accent to the second syllable; or we might even observe that all of the forms share a common root (*pā́d-*), with a certain set of suffixes used to mark the cases. Statements of this sort, which tell the speaker how to create forms based on other forms, can greatly reduce the

[1] I will use the following abbreviations here: *sg.* = singular, *pl.* = plural, *nom.* = nominative, *acc.* = accusative, *instr.* = instrumental, *dat.* = dative, *abl.* = ablative, *gen.* = genitive, *loc.* = locative, *voc.* = vocative; for features, *cons.* = consonantal, *cont.* = continuant, *cor.* = coronal, *lab.* = labial, *dors.* = dorsal

amount of information that must be learned. Rather than memorizing the entire paradigm of every word, the speaker must simply memorize a single form for each word—either a privileged surface form, or an abstract underlying form. In addition, the speaker must learn a grammar of morphological and phonological rules that derive the remainder of the paradigm from the memorized base form. The base therefore serves as the *input* to the grammar; it is what the grammar operates on to produce the remaining forms of the paradigm. With a grammar in place, the speaker is ideally able to produce any form of any word without having memorized it—or, at least, of most words, with a small residue of exceptions that must be memorized as such. The goal of this thesis is to explore how bases can be identified algorithmically, and to look for evidence that human language learners employ a similar approach.

Bases are more than just a computational convenience. Memorization is not only impractical in highly inflected languages, but it is also inadequate: speakers need to be able to produce and comprehend forms that they have never encountered before, and have thus had no chance to memorize. There are many sources of evidence showing that speakers are able to construct some forms in the paradigm based on information from other forms. First, there is the simple and easily observed fact that speakers can utter — often without hesitation — forms of words that they almost certainly have never heard before. Spanish speakers, for example, have no trouble producing the first person singular indicative forms of 'to decaffeinate' (*descafeíno*) or 'to Italianize' (*italianizo*), even though few, if any, have ever encountered or had occasion to use these forms before. Thus, even if we were to accept the idea that speakers memorize every inflected form that they have encountered (the *full listing hypothesis*, Butterworth 1983), we would still need some explanation for how speakers can produce novel forms based on other, previously listed forms.

In addition, there is abundant historical evidence that speakers construct relations between parts of the paradigm, since forms are often rebuilt on the basis of other forms within the paradigm. Consider, for example, the change in Yiddish verb paradigms shown in (2), in which the 2nd and 3rd singular forms of 'to dig' and the plural forms of 'to know' have been rebuilt to match the 1sg form. (The notation $*A \Rightarrow B$ indicates that the expected form A has been replaced by the analogical form B.)

(2) Leveling in Yiddish verbal paradigms

a. *grɔbn* 'to dig'

	sg.	pl.
1st	*gr**ɔ**b*	*gr**ɔ**bn*
2nd	**gr**e**bst* ⇒ *gr**ɔ**bst*	*gr**ɔ**bt*
3rd	**gr**e**bt* ⇒ *gr**ɔ**bt*	*gr**ɔ**bn*

b. *visn* 'to know'

	sg.	pl.
1st	*v**ey**s*	**v**i**sn* ⇒ *v**ey**sn*
2nd	*v**ey**st*	**v**i**st* ⇒ *v**ey**st*
3rd	*v**ey**s(t)*	**v**i**sn* ⇒ *v**ey**sn*

Changes of this sort, in which alternations are eliminated by replacing some members of the paradigm, are known as *paradigm leveling* or *analogical leveling*. Such changes poses a well-known problem in historical linguistics, and there are two basic approaches to explaining them. On the one hand, they occur frequently, and are natural and unsurprising. For this reason, it is often assumed that the drive towards nonalternating paradigms is simply a primitive of language, sometimes referred to as "Humboldt's Universal" (one form for one meaning). The analysis that this implies is that paradigm leveling is an output-output (OO) effect between related surface forms. Even in a language with paradigmatic alternations (such as *grɔb* ∼ *grebst*), there is some force that compels speakers to consider the possibility of uttering non-alternating forms (e.g., *grɔbst*), and furthermore, when speakers are faced with this possibility, they find the innovative forms appealing because of their resemblance to other forms within the paradigm. The idea that paradigms like to be uniform was never formalized in rule-based generative phonology (though on the need for it, see Kenstowicz and Kisseberth 1977, p. 74, as well as Hock 1991, p. 260). It has been formalized in recent years in Optimality Theory (OT) (Prince and Smolensky 1993), however, as UNIFORM EXPONENCE, LEVEL, or PARADIGM UNIFORMITY (Kenstowicz 1997b; Kenstowicz 1997a; Steriade 2000; Kager 2000; Raffelsiefen 2000; Kenstowicz 2002). Under these approaches, the force that suggests non-alternating forms is GEN, and the force that prefers them is the set of Uniform Exponence or Paradigm Uniformity constraints. Thus, OO constraints provide us with a formalism to describe the way that surface forms might influence one another to favor paradigm leveling.

A blanket preference for nonalternating paradigms can only go so far in explaining paradigm levelings, however. It is often noted that a paradigm uniformity preference can tell us that an alternation is likely to be leveled, but it cannot necessarily tell us when, or in which direction. For example, why was the desire for uniform paradigms stronger in Yiddish than in other German dialects, most of which have retained alternating paradigms (cf. Modern German *grabe, gräbst, gräbt*)? Why was it the 1sg form that was extended, and not some other form, such as the 3sg, yielding paradigms like **greb*, **grebst*, **grebt*? In this and other cases, the challenge is to explain why the change went in this direction, and not in other, logically possible directions, such as in (3):

(3) Other, logically possible (but unattested) changes in Yiddish

a. *grɔbn* 'to dig'

	sg.	pl.
1st	*gr**ɔ**b* ⇏ **gr**e**b*	*gr**ɔ**ben* ⇏ **gr**e**ben*
2nd	*gr**e**bst*	*gr**ɔ**bt* ⇏ **gr**e**bt*
3rd	*gr**e**bt*	*gr**ɔ**ben* ⇏ **gr**e**ben*

b. *visn* 'to know'

	sg.	pl.
1st	*v**ey**s* ⇏ **v**i**s*	*v**i**sn*
2nd	*v**ey**st* ⇏ **v**i**st*	*v**i**st*
3rd	*v**ey**s(t)* ⇏ **v**i**s(t)*	*v**i**sn*

Many proposals over the years have attempted to explain the direction of analogical change. The usual approach, pioneered by Kuryłowicz (1947) and Mańczak (1958) and continued by Bybee (1985) and others, has been to focus on tendencies, or groups of factors that may compete in making one form the *base*, or *pivot* of a change. Some of the factors that seem to play a role include: (1) presence or absence of suffixes, such that leveling is often to forms with no affixes or with shorter affixes (Mańczak 1958; Hayes 1995, Bybee 1985, pp. 50-52), (2) token frequency, with leveling often extending the members of the paradigm with the highest token frequency (Mańczak 1980, pp. 284-285), and (3) some sense of morphosyntactic markedness, with leveling often extending the "unmarked" member of the paradigm (Jakobson 1939; Greenberg 1966; Bybee and Brewer, 1980; Tiersma 1982).[2]

Often these factors all coincide to favor a single form, so in language after language, it is the 3sg present form of verbs, or the nominative singular of nouns, that is extended. In Polish, for example, the paradigms of many diminutives have been rebuilt on the basis on the nominative singular. In the language in general, there is a regular [ɔ] ∼ [u] alternation before underlyingly voiced obstruents, with /ɔ/ occurring as [ɔ] in open syllables (especially those followed by a lax vowel), and [u] in closed syllables (Gussmann 1980, chap. 4; Kraska-Szlenk 1995, pp. 108-114; Kenstowicz 1997), shown in (4) for the words *dół* [duw] 'ditch' and *krowa* [krova] 'cow':[3]

(4) [ɔ] ∼ [u] alternations in Polish nouns

a. [duw] 'ditch' (masc.)

	sg.	*pl.*
nom.	[d**u**w]	[dɔwɨ]
gen.	[dɔwu]	[dɔwuf]
dat.	[dɔwovi]	[dɔwom]
acc.	[d**u**w]	[dɔwɨ]
instr.	[dɔwem]	[dɔwami]
loc.	[dɔle]	[dɔwax]

b. [krɔva] 'cow' (fem.)

	sg.	*pl.*
nom.	[krɔva]	[krɔvɨ]
gen.	[krɔvɨ]	[kr**u**f]
dat.	[krɔvje]	[krɔvom]
acc.	[krɔvẽ]	[krɔvɨ]
instr.	[krɔvã]	[krɔvami]
loc.	[krɔvje]	[krɔvax]

The Polish diminutive suffix -*(e)k* (masc.)/-*(e)ka* (fem.) contains an initial deletable vowel (a so-called "yer") that disappears when the following syllable contains a full vowel. The result is that sometimes this suffix is vowel-inital, and sometimes it is consonant-initial. We would expect, therefore, that it should condition [ɔ] ∼ [u] alternations in the final syllable of the noun, just as in (4); the expected forms for *dołek* [dɔwek] 'little ditch' and *krówka* [krufka] 'little cow' are shown in (5):

[2]The definition of "unmarked" is often problematic, though various authors have attempted to find a non-circular basis for deciding that one part of the paradigm is "less marked" than another.

[3]This alternation was originally a lengthening of [ɔ] to [ɔː] conditioned by a following coda voiced obstruent; this lengthening was subsequently made opaque by raising/tensing of [ɔː] to [u], and the conditioning environment of open and closed syllables has been made opaque in some cases by the loss of a vowel (the "yer") in some forms of some suffixes, including the diminutive suffix -*ek* (Gussman 1980, p. 30). Furthermore, loans and paradigm leveling have introduced many exceptions to the raising alternation seen in (4), and it has been argued that this alternation is no longer synchronically productive (Buckley 2001, Sanders 2001).

(5) Expected paradigms of Polish diminutives

a. [dɔwek] ‘little ditch’ (masc.)

	sg.	*pl.*
nom.	[dɔwek]	[d**u**wki]
gen.	[d**u**wka]	[d**u**wkuf]
dat.	[d**u**wkovi]	[d**u**wkom]
acc.	[dɔwek]	[d**u**wki]
instr.	[d**u**wkjem]	[d**u**wkami]
loc.	[d**u**wku]	[d**u**wkax]

b. [krufka] ‘little cow’ (fem.)

	sg.	*pl.*
nom.	[kr**u**fka]	[kr**u**fki]
gen.	[kr**u**fki]	[krɔvek]
dat.	[kr**u**ftse]	[kr**u**fkom]
acc.	[kr**u**fkẽ]	[kr**u**fki]
instr.	[kr**u**fkã]	[kr**u**fkami]
loc.	[kr**u**ftse]	[kr**u**fkax]

In fact, this alternation has been leveled in many diminutives; in all cases, it is the form that is expected in the nominative that has been extended to the remainder of the paradigm (6):

(6) Actual paradigms of Polish diminutives

a. [dɔwek] ‘little ditch’ (masc.)

	sg.	*pl.*
nom.	[dɔwek]	[dɔwki]
gen.	[dɔwka]	[dɔwkuf]
dat.	[dɔwkovi]	[dɔwkom]
acc.	[dɔwek]	[dɔwki]
instr.	[dɔwkjem]	[dɔwkami]
loc.	[dɔwku]	[dɔwkax]

b. [krufka] ‘little cow’ (fem.)

	sg.	*pl.*
nom.	[kr**u**fka]	[kr**u**fki]
gen.	[kr**u**fki]	[kr**u**vek]
dat.	[kr**u**ftse]	[kr**u**fkom]
acc.	[kr**u**fkẽ]	[kr**u**fki]
instr.	[kr**u**fkã]	[kr**u**fkami]
loc.	[kr**u**ftse]	[kr**u**fkax]

The fact that we get an [ɔ] in [dɔwki] but an [u] in [kr**u**fki] is not conditioned by any phonological difference in the suffixes of these forms; rather, the oblique and plural forms seem to be influenced by the (phonologically expected) difference in the nominative singular forms. In Polish, as in many other languages, the nominative singular is also the unsuffixed (or, at least, the least suffixed) form, the morphosyntactically unmarked form, and possibly also the most frequent form, so its privileged status in the paradigm is unsurprising.

However, it is not always true that paradigm leveling extends the nominative singular. A famous counterexample, discussed by Hock (1991, pp. 179-180), Kenstowicz (1997b), and many others, is a change that occurred in the history of Latin. In pre-classical Latin, rhotacism of /s/ to [r] intervocalically created [s] $\sim$ [r] alternations within noun paradigms, as in (7a). This alternation was leveled out in the late pre-classical period, extending the [r] of the oblique and plural forms to the nominative singular.

(7) Change of Latin *honōs* to *honor*

a. Pre-leveling

	sg.	*pl.*
nom.	[honoː**s**]	[honoː**re**ːs]
gen.	[honoː**ri**s]	[honoː**ru**m]
dat.	[honoː**ri**ː]	[honoː**ri**bus]
acc.	[honoː**re**m]	[honoː**re**ːs]
abl.	[honoː**re**]	[honoː**ri**bus]

b. Post-leveling

	sg.	*pl.*
nom.	[hono**r**]	[honoː**re**ːs]
gen.	[honoː**ri**s]	[honoː**ru**m]
dat.	[honoː**ri**ː]	[honoː**ri**bus]
acc.	[honoː**re**m]	[honoː**re**ːs]
abl.	[honoː**re**]	[honoː**ri**bus]

The comparison of Yiddish, Polish and Latin in Table 1.1 illustrates the basic conundrum. In Yiddish and Polish, it was the unsuffixed member of the paradigm (the 1sg and nominative singular, respectively) that was extended, while in Latin, it was a suffixed form. In Polish, it was the universally unmarked member of the paradigm (the nominative singular) that was extended, while in Yiddish and Latin, marked forms (the 1sg and a non-nominative form) were extended. In Latin, it was perhaps the member of the paradigm with the highest token frequency that was extended (see section 4.4.2), but in Yiddish, the form that was extended was most likely not the most frequent member of the paradigm, or even the most frequent alternant.[4] In Yiddish and Latin, the form that was extended was the form that occurred in the majority of slots in the paradigm, but in Polish masculines, it was the minority form. The conclusion that is generally drawn from such facts is that no single factor guarantees that a particular form will be extended in paradigm leveling.

This is a problem for theories that try to explain basehood using "static" factors like frequency or markedness. Every language has differences in the frequency of forms, differences in the "degree of suffixation" of forms, differences in markedness, and so on, but it appears that speakers weight these factors differently in deciding which form should get extended in leveling. Thus, proposals that use such factors as an explanation allow us to derive typological predictions, but not to make predictions about a given language at a given time, because we do

[4] I do not have token frequency counts for the various members of the Yiddish verb paradigm. However, Bybee (1985, p. 71) gives some equivalent counts for Spanish, showing that the 3sg is almost twice as frequent as the 1sg. Furthermore, between 65% and 67% of the tokens are either a 2sg or 3sg, meaning that the umlaut alternant would have been the most frequent alternant, but it was not extended.

Table 1.1: Comparison of factors encouraging the extension of a form

Language	Form	Unsuffixed	Unmarked	Highest Freq. Form	Highest Freq. Alternant	Majority Form
Yiddish	1sg	*yes*	*no*	*no*	*no*	*yes*
Polish	nom.sg.	*yes*	*yes*	??	*yes*	*no*
Latin	oblique	*no*	*no*	*no* (?)	*yes*	*yes*

not know which factors will win in that particular case.[5] As Bybee and Brewer (1980, p. 215) state:

> A hypothesis formulated in such a way makes predictions of statistical tendencies in diachronic change, language acquisition and psycholinguistic experimentation. It cannot, nor is it intended to, generate a unique grammar for a body of linguistic data.

In this thesis, I will take a different approach, in the spirit of Paul (1920) and Kiparsky (1965), that focuses on the role of language learners in historical change. I will treat paradigm leveling not as an output-output effect, but rather as an effect of the way that speakers use their grammar to project unknown forms — that is, as an input-output (IO) effect. In particular, I will pursue the hypothesis that learners impose structure on paradigms, as part of an effort to construct phonological and morphological grammars that generate unknown forms as accurately or as confidently as possible. The way that they do this, I will claim, is by seeking a base form within the paradigm that is "maximally informative" — that is, that suffers the least serious phonological and morphological neutralizations — and then deriving the remaining forms in the paradigm from the base form. Under this approach, we can use the direction of the grammar (base form → derived forms) to predict the direction of possible analogical change.

Before beginning with the task of identifying bases in paradigms, it is useful to recognize from the outset that different models of morphology operate on radically different types of inputs. In some, termed "Item and Process" (IP) models by Hockett (1958), morphological rules create words from other words — for example, "add *-su* to the nominative to create the locative plural." In such a model, morphological and phonological rules are assumed to operate on a free-standing surface form from somewhere within the paradigm; examples of such models include Aronoff's word-based model (Aronoff 1976), Anderson's Extended Word and Paradigm (EWP) model (Anderson 1992), Bochner's Lexical Relatedness model (Bochner 1993), and Ford and Singh's whole word morphology (Ford and Singh, 1996; Neuvel, to appear; Neuvel and Fulop, to appear). For models that operate on words, a base selection procedure must be able to choose a surface form that will serve as the base. For example, in the case of Sanskrit, we might choose the accusative, and then formulate a set of rules that changes the suffix, moves the accent, and performs various phonological adjustments to derive the remainder of the paradigm.

[5] In fact, this lack of predictiveness is considered appropriate by many, since leveling is classified as an analogical change, and as such is held not to be rule-governed. I will take the opposite approach here, of trying to pursue a hypothesis that makes strong and falsifiable predictions about possible changes (Gvozdanović 1985).

In other models, termed "Item and Arrangement" (IA) by Hockett, morphological rules combine sub-parts of words (stems and affixes) to create surface forms — for example, "the locative plural morpheme is *-su*, and it occurs after the nominal root." In this type of model, the grammar is assumed to operate on a set of underlying forms, which are combined and readjusted to yield surface forms. This is in fact the type of model usually assumed by phonologists. The underlying forms of morphologically complex words are typically represented in phonological analyses as something like /*root+affixes*/, where the /*root*/ and /*affix*/ are independent entities. Phonologists do not, on the whole, devote much attention to the question of how the morphemes came to be in that particular configuration, except to suppose that there is a separate morphological module that takes them out of the lexicon and arranges them somehow. A few recent formalizations of how this is actually done using an IA approach include Distributed Morphology (Halle and Marantz 1994), DATR (Evans and Gazdar 1996), and Lieber's syntactic approach to morphology (Lieber 1992). For models that operate on (possibly bound) stems, a base selection procedure must be able to discover the underlying forms of word roots, and whatever morphological information is necessary to determine which affixes they should combine with. For Sanskrit, this would involve learning that the underlying form for 'foot' is something like /*pā́d*/ (in spite of surface variations), and that it is in the class of nouns that take *-am* in the accusative, *-as* in the plural, etc.

On the face of it, it would seem that the different requirements of word-based and stem-based approaches would demand fundamentally different base identification procedures: for an IP approach, we need to isolate a surface form in the paradigm that will act as the base, while for an IA approach, we need to be able to compare the surface forms to arrive at a (possibly abstract) underlying form for the stem of each word. What I will argue in the course of this thesis, however, is that although we could imagine very different strategies for selecting whole-word bases vs. stems, they are both compatible with a range of possible strategies, and furthermore, in many cases the empirical evidence drives us to parallel conclusions for both.

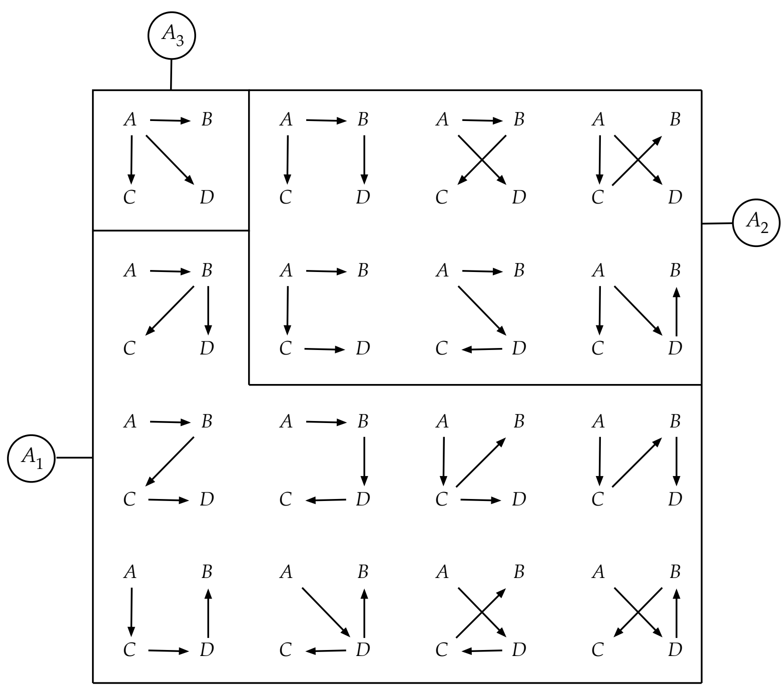
1.1 The problem of bases in word-based morphology

As noted above, selecting a base has the potential to greatly reduce the amount of information that must be learned; rather than having to memorize 30 forms for each word, in most cases learners can simply memorize one form, and use a set of rules to derive the rest of the paradigm. Unfortunately, the rules listed above were only two of the many possible rules that could be formulated to relate parts of the Sanskrit noun paradigm. In addition to relating the genitive dual to the locative dual, and the nominative singular to the locative plural, we could equally well have observed that the instrumental singular is formed by taking the locative dual and replacing the *-os* suffix with *-ā*, the dative dual is equal to the ablative plural minus the final *-s* and adding a final *-am*, and that the nominative singular is the accusative plural minus the *-as* suffix, possibly shifting the accent to the first syllable, and changing the [c] to a [k]. If we were to say simply that a grammar contains rules relating forms in the paradigm to one another and leave it at that (as, for example, Bochner 1993, Barr 1994, and Neuvel and Singh 2001 do), we would be left with an enormous number of pairwise relations to include in the grammar. Intuitively, the answer to this problem is that speakers probably do *not* learn rules for every pairwise relation in the paradigm, but rather for only a subset of the possible relations. The question, then, is which relations are part of the grammar, and which relations are not. Stated differently, what are the possible structures of the paradigm?

A common way to restrict the set of possible paradigm structures in word-based morphology is to place restrictions on bases. Take, for example, the four-member paradigm *walk*, *walks*, *walked*, and *walking*. In a completely unrestricted system, we could say that *walked* is derived from *walk* by the rule $[\,\text{X}\,]_{\text{pres.}} \rightarrow [\,\text{X}\ \textit{ed}\,]_{\text{past}}$, from *walks* by the rule $[\,\text{X}\ \textit{s}\,]_{\text{3sg.pres.}} \rightarrow [\,\text{X}\ \textit{ed}\,]_{\text{past}}$, or from *walking* by the rule $[\,\text{X}\ \textit{ing}\,]_{\text{pres.participle}} \rightarrow [\,\text{X}\ \textit{ed}\,]_{\text{past}}$. Suppose, however, that we restrict grammars such that they may only operate on certain forms within the paradigm. These privileged forms, which serve as the input to the grammar, we will call *bases*. One possible restriction would be to say that each slot in the paradigm must be derived from at most one unique base, but different slots may be derived using different bases . Now there can only be one way to derive a given form, and we would have to pick just one of the rules deriving *walked*; intuitively, we would probably want to derive *walked* from *walk* by the rule $[\,\text{X}\,]_{\text{pres.}} \rightarrow [\,\text{X}\ \textit{ed}\,]_{\text{past}}$.

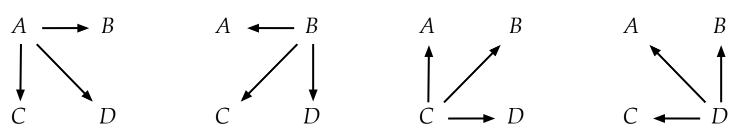
Requiring that each slot in the paradigm have at most one base greatly reduces the number of pairwise relations that are included in the grammar. If we assume a particular version of this restriction in which one form in the paradigm is an underived base form and all other forms are derived by a single relation, then we reduce the number of pairwise relations in the grammar to $n - 1$. The resulting set of possible paradigm structures is much more constrained, but still perhaps larger than we might like. When we consider a hypothetical paradigm with four members (*A*, *B*, *C*, *D*), even if we assume that A is always a base and that there is only one way to derive each form, there are still sixteen possible paradigm structures, as in (8). In particular, there is one structure using A to derive all three of the other forms (A_3), there are six structures using it to derive just two of the other forms (A_2), and there are nine structures using it to derive just one of the other forms (A_1).

(8) Possible paradigms using just one mapping per form, *A* as base



The paradigm structures in (8) are more manageable than a completely unrestricted system would be, and in chapter 6 I will present some evidence that structures like these (in particular, the A_2 structures) may sometimes be needed. However, for now we may note that this restriction still leaves us with quite a large number of hypotheses to explore. In practice, many analysts seem to assume an even more restrictive hypothesis, which is that, in the usual case, the entire inflectional paradigm is derived from a single base form. If we adopt this assumption, then for a paradigm of n forms, we need to consider only n possible candidates for base status. Furthermore, once we have selected a base form, we will have only one possible paradigm structure, consisting of statements about only $n-1$ pairwise relations. For example, in the case of Sanskrit, if we choose the nominative singular as the base, then we are left with a maximum of 23 statements about how to relate the other forms to the nominative. Furthermore, a learner operating under this restriction must consider only n possible paradigm structures of $n-1$ relations each. This is shown schematically in (9).

(9) The single base hypothesis: only n possible paradigms



The single base hypothesis is appealing not only from a learning point of view, but also because of its restrictiveness as a linguistic analysis. Once we have limited the possible paradigm structures to those in (9), all that remains is to construct a grammar that derives each of the non-basic forms in the paradigm.

For purposes of concreteness, I will make the following assumptions about the mechanisms by which speakers can produce forms: first, forms may be produced by retrieving them from the lexicon ready-made, already bearing whatever features are required for the syntactic context. I will refer to this mode of production as retrieval of listed forms, or resorting to memorized word-specific knowledge. For obvious reasons, this option is available only if the speaker has memorized the relevant form ahead of time. Second, forms may be produced by synthesizing them with the rules of the grammar. This option is available only if the speaker has one or more rules to derive the desired part of the paradigm, and if the speaker knows the base form (the input) for the word in question. For example, if you need form B of word w (w_B), you need two things: a rule to derive form B from another form (e.g., $A \rightarrow B$), and you need to know word w in form A, so you can apply the rule to derive $w_A \rightarrow w_B$. I will refer to this mode of production as synthesis, or derivation by the grammar.

In some linguistic theories, limitations are placed on what forms may or may not be stored in the lexicon, requiring a theory of how speakers decide which forms to memorize and which to exclude. In this thesis, I will make what I take to be the rather simpler assumption that speakers are potentially able to memorize any form that they have had sufficient exposure to, whether it is a base form or an inflected form. Indeed, there are several situations in which this capability is necessary. First, when learners are just starting out and have not yet constructed any grammatical rules, they will need to memorize whatever forms they hear, both as a production mechanism until rules for synthesis are in place, and also to use as input data for morphological learning.

Second, even after speakers have constructed a grammar of rules, there may be some forms that the grammar cannot productively derive the correct output for (such as lexical exceptions). In such cases, the correct form can be produced only by using some form of word-specific knowledge—either by retrieving it as a listed exception, or by some sort of lexically-specified rule. We may assume, following Aronoff (1976), that forms produced using word-specific knowledge take precedence over or *block* productively synthesized forms. For inflected forms that the grammar *can* derive productively, memorizing the inflected form is not necessary, but it also not harmful. In the discussion that follows, I will assume that speakers may memorize regular, grammatically derivable forms at least some of the time — following, for example, Baayen, Dijkstra and Schreuder (1997) and Gordon and Alegre (1999) — but this is not crucial.

The hypothesis to be tested here, then, is that although there are no a priori restrictions on what forms can be memorized, there *are* a priori restrictions on the structure of grammar. In particular, the single base hypothesis means that for one form in the paradigm (the base), there are no rules that can be used to synthesize it, and memorization is the only option. Other forms in the paradigm may be memorized or may be synthesized, but synthesis must be done via operations on the base form. Since we are assuming here a word-based model of morphology, the base is a fully formed surface member of the paradigm, and for this reason, I will call this the *single surface base* hypothesis.

The single surface base hypothesis makes strong predictions about the types of errors that a speaker may make. There is only one way to produce base forms (retrieving them from the lexicon), so if lexical access fails for some reason, the speaker will have no way to synthesize a base form. In other words, base forms will be produced correctly, or not at all. In contrast, non-base forms may either be retrieved from memory as listed inflected forms, or they may be synthesized, using the base form together with the relevant rules. In the case of regular, grammatically derivable forms, either method will yield the same result. For exceptional non-base forms, however, only the stored inflected form will yield the right result; the grammatically synthesized form will be an overregularization (Marcus, Pinker, Ullman, Hollander, Rosen, and Xu 1992). This is shown in Figure 1.1.

This model predicts several asymmetries: first, it is possible to produce incorrect non-base forms, but there is no way to derive an incorrect base form. Second, it is possible to overregularize non-base forms by uttering the grammatically expected form instead of a lexically listed exceptional form, but there is no way to "overirregularize" by creating new listed exceptions without any positive evidence for them. For any given language, if you know which form is the base form and what the grammar for deriving the rest of the paradigm looks like, there is only one class of items that should be open to change: exceptional forms in non-basic parts of the paradigm. The goal of this thesis is to develop an algorithm to determine what the base form is and what the grammar looks like.

For the first several chapters of this thesis, I will consider relatively small, "local" paradigms involving just one tense or a handful of noun cases, and I will these small examples to explore the procedure by which one might identify a single, privileged base form. It is important to keep in mind, however, that more complicated structures may also be necessary, especially when we consider larger paradigms or multiple tenses, moods, etc. Language descriptions frequently refer to multiple stems or bases for a single lexical item—for example, Latin nouns are listed in the dictionary in two forms (nominative and genitive), and Latin verbs are listed in four principal parts. Descriptions involving multiple stems, or multiple listed root allomorphs all

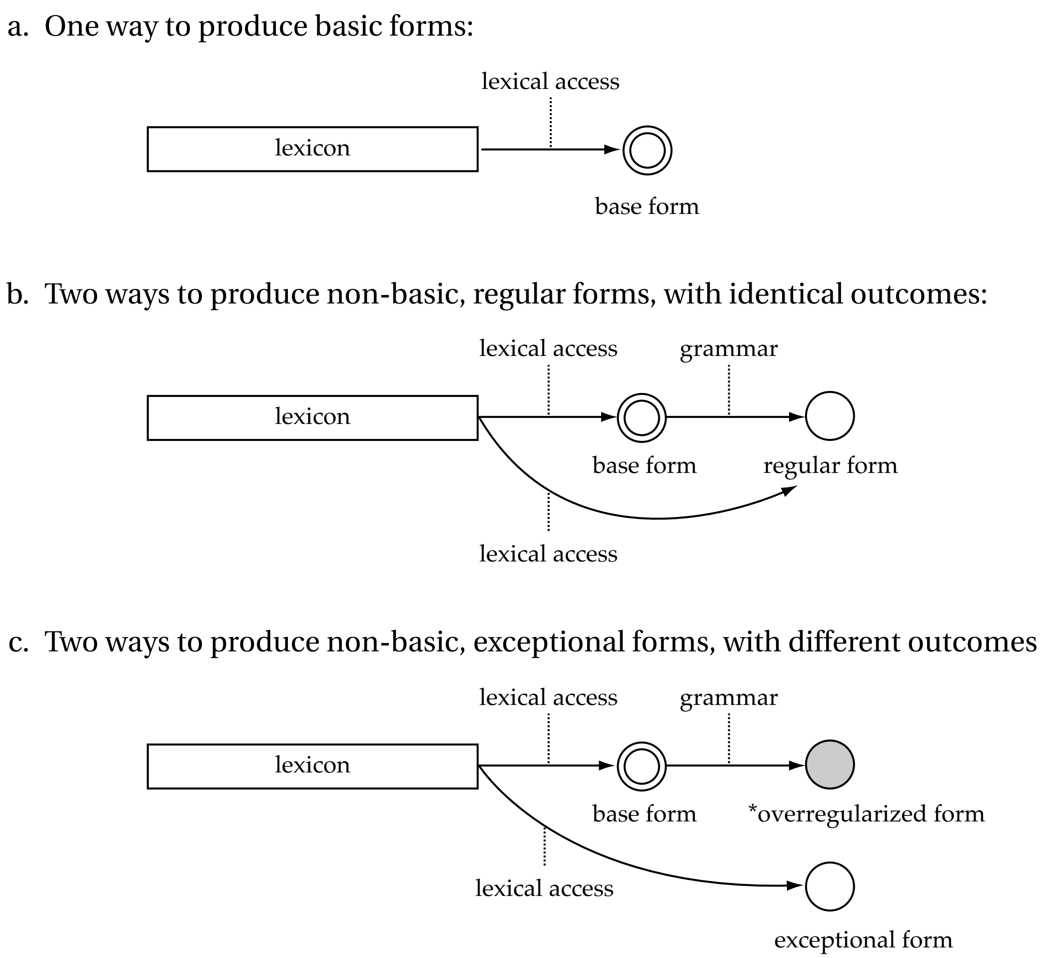


Figure 1.1: Routes for deriving different types of words

seem to require more than just a single base form within the paradigm. In chapter 6, I will discuss a case from Spanish that appears to require multiple, local bases, and I will propose an extension of the base identification algorithm that could be used to identify when a more complicated paradigm structure might be helpful.

1.2 The problem of bases in stem-based morphology

The base identification algorithm proposed in chapter 3 is couched in a word-based model of morphology, and much of the discussion here will focus on the task of selecting a whole word base. It is important to remember, however, that many models of morphology, including those usually assumed (implicitly or explicitly) by phonologists, do not convert surface forms to other surface forms. Rather, they combine independent morphemes (typically stems and affixes) to construct the surface forms of the paradigm. Under such an approach, the input to the morphology is a set of underlying forms, which may or may not match any of the surface forms. What would base identification consist of in a stem-based model?

Within generative phonology, the most serious attempt to develop an all-purpose model of input discovery came in the late 1970s, as part of a quest to constrain the abstractness of underlying representations (URs). As in word-based morphology, there are many possible theories of what should be allowed to serve as bases (or URs) in a stem-based model, and many possible restrictions that could be placed on them. Kenstowicz and Kisseberth (1977, chap. 1) review a series of intuitively reasonable and appealing hypotheses about how to restrict underlying forms, including:

(1) requiring that URs match unsuffixed forms

(2) allowing URs to come from suffixed forms, but requiring that they come from the same place in the paradigm for all words

(3) allowing URs to come from anywhere in the paradigm, but requiring that the form chosen as the UR always occur in the most slots in the paradigm

(4) allowing the UR to come from different parts of the paradigm for different words, but requiring each UR to match a surface form somewhere in the paradigm

(5) allowing URs to combine information from different parts of the paradigm, but requiring that each segment in the UR must surface as such somewhere in the paradigm.

In each case, Kenstowicz and Kisseberth present well motivated arguments showing that if we were to adopt the restriction, we would be unable to construct phonological rules to predict certain alternations. The end conclusion is that we must allow URs to contain abstract structure that is found nowhere in the paradigm, but is needed to derive alternations.

Unfortunately, this conclusion also makes the job of the UR discovery mechanism much more difficult. It is perhaps telling that the field has yet to produce a general purpose model of UR selection that can compare all the forms in the paradigm, find the necessary segments from each one, and posit abstract segments where needed. This is not to say that it could not be done. However, in the course of this thesis, I will present some arguments in favor of adopting one of the more restrictive hypotheses above. In particular, an algorithm that selects a single surface base form within the paradigm is essentially the same as restriction (2) above. In chapters 2 and 4, I will show that in some cases (Yiddish and Latin), this hypothesis makes more specific predictions than the more conventional, unrestricted use of URs, while in chapter 5 I will show that in other cases (such as Lakhota) it makes completely different predictions; moreover, in all cases, these predictions appear to be correct.

1.3 The problem of bases in correspondence theory

I have been using the term "base" here to refer to the entity from which complex forms are derived — that is, as the input to the morphological grammar. It is worth noting, however, that the motivation behind an algorithm to identify bases is not just morphological; in fact, many phonological theories also have a vested interest in identifying bases. Within the framework of OT, various output-based versions of Correspondence Theory (McCarthy and Prince 1995, Steriade 2000, Benua 1997, (Kenstowicz 1997a, 1997b), McCarthy 1998) have made explicit reference to bases. This usually takes the form of faithfulness constraints which demand that derived forms preserve properties of the base in their paradigm. In the strongest versions of output-output correspondence, which seek to eliminate URs altogether and rely completely on surface relations (e.g., Burzio 1996, Cole and Hualde 1998), it is especially vital that bases be correctly identified. The problem of identifying bases has largely been ignored or deferred in the OT literature. In the case of reduplication, it is only a minor problem, since there are only two entities involved (the base and the reduplicant), we generally know that one was used to create the other, and to a certain extent, it doesn't matter which one is which.

In the larger context of transderivational output-output constraints, there seem to be two opposing camps in the literature. The first argues that output-output constraints are inherently symmetrical, and that any form may potentially influence any other form, in order to achieve a globally more harmonic paradigm (However, it should be noted that change towards a more "basic" form may be enforced in a roundabout way by the existence of a UR, and IO-Faithfulness constraints demanding surface forms to match the UR.) An alternative approach, advocated by Benua (1997), Kenstowicz (1998), and others, is to treat output-output correspondence as an asymmetrical relation, where one form is given priveleged *base* status, and allowed to influence other forms of the paradigm. (See Kiparsky, in prep., for a discussion of the pros and cons of these two approaches.)

If we rely on asymmetrical output-output faithfulness constraints, then we need a procedure for determining which form should be considered the base. This is not a trivial problem; Noyer (1998) and Buckley (1999) discuss several cases in which derived forms are apparently faithful to things other than their smallest or most immediate consitutents. Note also that the base identification problem is difficult to solve using an OT-internal learning strategy, because we would need to evaluate three things that are changing simultaneously: the set of hypothesized URs, the hypothesized base, and the ranking of OO faithfulness with respect to markedness and IO faithfulness. Finding the right base for the purposes of evaluating OO correspondence constraints will not be the focus of my discussion here, but it is hoped that the same considerations may hold in both tasks, and that a procedure like the one outlined here could provide an independent means of identifying such bases.

1.4 Plan of the thesis

The puzzle that must be solved, then, is as follows: bases appear to play a role in several areas of linguistics, including in historical changes like paradigm leveling, in psycholinguistic models of lexical organization, in evaluating output-output faithfulness between surface forms. However, the form that serves as the base seems to vary somewhat from language to language. Therefore, learners must be equipped to learn somehow what forms to use as the bases in their language. The factors which would lead learners to use different bases in different languages, and a procedure for exploring hypotheses about basehood as part of the acquisition process, remain largely unexplored and unformalized; I am pursuing here the strong hypothesis that the choice of base is always determined by a single, universal principle. The crucial observation is that previous attempts have met with only partial success because they have focused solely on inherent properties of the base forms themselves — their frequency, their morphosyntactic markedness, etc. The hypothesis which I will explore in this thesis is that bases are identified in the process of learning the relations between forms:

- Learners begin by exploring all relations that are available to them (i.e., all relations between forms that they have actually encountered)
- The goal of morphological acquisition is to find the relations which make the morphological projection problem "easier," in a way which can (and will) be quantified. In other words, they are looking for what would make the best bases.
- Once a global decision has been made about the best all-purpose base, learners concentrate on relations from that form to the rest of the paradigm

I will start with a schematic example in chapter 2, showing for one language (the older, preleveling stage of Yiddish, shown in (2)) how one might go about comparing the informativeness of different parts of the paradigm and selecting the form that is globally most informative. It turns out that this form is the 1sg for Yiddish verb paradigms, and furthermore, Yiddish has subsequently undergone widespread paradigm leveling to precisely this form. This example is meant to show conceptually how one could derive predictions about paradigm leveling using such a restrictive model of base identification.

In order to make testable predictions about bases in different languages, it is useful to have a computational model of morphological acquisition. Therefore, in chapter 3, I will propose a formal system for modeling the acquisition of basehood computationally, building on the system for learning morphological rules developed by Albright and Hayes (1999a). I will show how this system can be used to identify bases in several small artificial languages. Then, in chapter 4, I will discuss its application to the more difficult and realistic problem of the Latin *honor* analogy, a change which violates numerous typological generalizations. I will show that the model is able to select an oblique form as the base for Latin noun paradigms, and that the resulting grammar predicts leveling for exactly the right set of words, in the right direction. The Latin example will serve as a demonstration of how the model is able to select a form other than the nominative singular as the base for this change, and how it can predict paradigm leveling without relying on a formal notion of paradigm uniformity or uniform exponence.

In chapter 5, I will turn to a rather different type of paradigmatic change: the creation of new paradigm types. Using data from Lakhota, I will show how the single surface base hypothesis makes predictions not only about paradigm leveling, but also about other types of analogical change as well. The changes that have taken place in Lakhota are especially interesting because they are completely unexpected under a less restrictive model of UR or stem discovery, for reasons that will be discussed.

Finally, in chapter 6, I will return to some comparisons with other traditional explanations of paradigm leveling. I will contrast the proposed model with explanations that rely on factors like markedness, token frequency, and frequency of occurrence within the paradigm, showing that the current model makes stronger predictions, which appear to be correct. I will also discuss a class of cases, in which it appears that analogical changes have been based on *less* informative members of the paradigm; this includes Korean (Hayes 1995; Kenstowicz 1997b), Maori (Hale 1973), and others. I will suggest a way in which the proposed model could be extended to these cases as well, by making use of the idea that not all forms in the paradigm are actually available in equal numbers to learners. I will show how this extension of the system could be used to handle many of the typological tendencies observed by Kuryłowicz, Mańczak, Bybee, Hock, and others. Finally, I will discuss some possible ways in which the single base hypothesis could be relaxed to allow local bases, in order to handle larger inflectional paradigms with multiple tenses, moods, and aspects.

Chapter 2

Paradigm leveling in Yiddish

A notable difference between Yiddish and German verb paradigms is that Yiddish has no vowel alternations in the present tense.[1] Whereas Middle High German (MHG) and Modern German (NHG) verbs often have vowel alternations among the singular forms (10a), or between singular and plural forms (10b), Yiddish never does (11).[2]

(10) MHG present tense vowel alternations

	‘dig’	sg.	pl.		‘know’	sg.	pl.
a.	1st	*grabe*	*graben*	b.	1st	*w**ei**ʒ*	*wiʒʒen*
	2nd	*gr**e**best*	*grabet*		2nd	*w**ei**st*	*wiʒʒet*
	3rd	*gr**e**bet*	*graben*		3rd	*w**ei**ʒ*	*wiʒʒen*

(11) Yiddish paradigms have no vowel alternations

	‘dig’	sg.	pl.		‘know’	sg.	pl.
a.	1st	*grɔb*	*grɔbən*	b.	1st	***vey**s*	***vey**sən*
	2nd	*grɔbst*	*grɔbt*		2nd	***vey**st*	***vey**st*
	3rd	*grɔbt*	*grɔbən*		3rd	***vey**st*	***vey**sən*

As I will show in Section 2.1, the form that has been extended in Yiddish is always the expected 1sg form. This is a puzzle, because Bybee (1985, chap. 3) argues that it is the 3sg, not the 1sg, that is most often the ‘basic’ form in paradigms—why would Yiddish have chosen the 1sg instead? Furthermore, although this change is across the board in Yiddish, it is apparently unattested in any other German dialect, leading us to wonder: what made Yiddish so different from other, closely related languages? In this chapter, I will show that we can gain insight

[1] With the term ‘Yiddish’, I am referring here exclusively to the eastern dialects of Central and Eastern Europe; I do not know if the same holds true of the western dialects of Austria, Germany, and points west, or not.

[2] For MHG forms, I will use the standardized orthography of Paul, Wiehl, and Grosse (1989, §§18–20), in which ˆ marks long vowels, *ë* represents a short open [e], and *ʒ* represents a coronal sibilant fricative, possibly fortis or possibly postalveolar (Paul et al, §151). For all Yiddish examples, I will use the YIVO transliteration system (http://www.yivoinstitute.org/yiddish/alefbeys.htm), with two minor modifications: I will use the IPA symbol ɔ instead of YIVO *o* for *komets-aleph*, and I will use *-ən* instead of YIVO *-en*/*-n* for the infinitive and 1pl/3pl present tense suffix. The change of MHG short [a] > Yiddish [ɔ] reflects a regular sound change; the correspondences between MHG orthographic ⟨w⟩ and Yiddish ⟨v⟩, MHG ⟨ei⟩ and Yiddish ⟨ey⟩, MHG ⟨s⟩ and Yiddish ⟨z⟩, and MHG ⟨ʒ⟩,⟨ʒʒ⟩ and Yiddish ⟨s⟩ are also completely regular.

into both of these questions by comparing different forms in the paradigm to see which is the "most informative"—that is, which form contains the most information about how to project the remainder of the paradigm. In section 2.2, I will show that it is in fact the 1sg form that preserves the most contrasts in Yiddish, and thus would be selected as the base form in the proposed model. Finally, in section 2.3, I will argue that the advantages of the 1sg form are unique to Yiddish, due to small but crucial differences between Yiddish and other German dialects. Thus, considering the informativeness of forms gives us insight not only into the question of why verbs were leveled to the 1sg in Yiddish, but also into why this did not occur elsewhere.

2.1 Leveling to the 1sg in Yiddish verb paradigms

As illustrated in (11) above, Modern Yiddish has no root vowel alternations in present tense paradigms (Rockowitz 1979; Katz 1987). The goal of this section is to show that in virtually all cases, it is the etymologically expected vowel of the 1sg form that has been extended to the remainder of the paradigm. In order to show this, we will consider the candidates in (12) as sources for the modern present tense stem, successively eliminating all forms except the 1sg.[3]

(12) Candidates for the source of the modern present tense stem:

1sg	1pl	infinitive
2sg	2pl	
3sg	3pl	/UR/

I will start with the fairly traditional assumption that the origin of Yiddish was some form of Middle High German, so it is useful to begin by considering the possible types of present tense paradigms that occurred in MHG.

2.1.1 MHG present tense patterns

Most MHG verbs had the same vowel throughout the entire present tense, with no alternations, as in (13); verbs of this type included the "strong" classes I, IIIa, and some of VII, as well as all of the "weak" verbs.[4]

(13) No alternations (Strong I, Stong IIIa, some Strong VII, all weak)

a.	'live'	*lëbe*	*lëben*	b.	'say'	*sage*	*sagen*
		lëbest	*lëbet*			*sagest*	*saget*
		lëbet	*lëben*			*saget*	*sagen*

[3] The list in (12) is a nearly comprehensive list of all of the verb forms that occur in Yiddish; the only other forms are the present participle, the stem ("shtam"), the past participle, and the imperative. The present participle and stem are always based on the infinitive, and the past participle is demonstrably not the source of the modern present tense forms. The singular imperative form is in fact always identical with the 1sg form, and could equally well have served as the base for the leveling discussed here. For expository ease, I will refer throughout this chapter to the 1sg form, but I cannot preclude the possibility that it was the singular imperative form instead.

[4] For a description of the strong and weak verb classes of MHG, see Paul, Wiehl, and Grosse 1989, chap. 7.

In another set of verbs, an *a* in the root surfaced as an *e* in the 2sg and 3sg, due to a process known as *umlaut* (14), originally conditioned by an [i] in the suffixes of these forms. This occurred in strong class VI and the remainder of strong class VII. The umlaut alternation is illustrated in (14) for the verb *graben* 'dig'; some other common verbs with umlaut alternations included *varn* 'travel', *halten* 'hold', *lâʒen* 'let', *laden* 'invite', *slahen* 'beat', and so on.

(14) 2sg, 3sg different due to Umlaut ($a \sim e$): Strong VI, some Strong VII

'dig'	*grabe*	*graben*
	*gr**e**best*	*grabet*
	*gr**e**bet*	*graben*

Finally, a third set of verbs showed vowel alternations between the entire singular (1,2,3sg) and the entire plural (1,2,3pl). This pattern occurred in two types of verbs. In some, the present tense derived from a Proto-Indo-European perfect, and the singular/plural alternation reflected a PIE alternation in the perfect tense (*ablaut*). These verbs are known as *preterite presents*, and are exemplified by *wiʒʒen* 'to know' in (15a); other preterite present verbs included *kunnen* 'can', *durfen* 'need', and *suln* 'should'. In a second set of verbs, the alternation was due to a phonological process in Old High German that raised mid vowels (*ë*, *ie*) to high vowels (*i*, *iu*) before a following high vowel, causing the singular to diverge from the plural (15b); this pattern is sometimes referred to as *Wechselflexion* ("alternating inflection"), and occurred in strong class II (*ie* ∼ *iu*), as well as IIIb, IV, and V (*ë* ∼ *i*). This pattern is shown in (15b) for the verb *gëben* 'give'; other *Wechselflexion* verbs included *nëmen* 'take', *ëʒʒen* 'eat', and *gieʒen* 'pour'.

(15) Singular ∼ plural alternations

a. Preterite presents

'know'	*w**ei**ʒ*	*wiʒʒen*
	*w**ei**st*	*wiʒʒet*
	*w**ei**ʒ*	*wiʒʒen*

b. *Wechselflexion*

'give'	*g**i**be*	*gëben*
	*g**i**best*	*gëbet*
	*g**i**bet*	*gëben*

2.1.2 Yiddish present tense patterns

Let us now consider the fate of each of these patterns in Yiddish. Unsurprisingly, verbs with no alternations in MHG continue to have no alternations in Yiddish, as seen in (16).

(16) Non-alternating verbs remain non-alternating in Yiddish

a.

'live'	*leb*	*lebən*
	lebst	*lebt*
	lebt	*lebən*

b.

'say'	*zɔg*	*zɔgən*
	zɔgst	*zɔgt*
	zɔgt	*zɔgən*

Umlaut alternations (1sg *grabe* vs. 2sg *grebst*) were leveled to the non-umlaut (*a*) alternant, as in (17). Thus, it appears that the base of the leveling was not the 2nd or 3rd singular, or else the modern Yiddish paradigm would have *e* throughout (*greb*, *grebst*, *grebt*, etc.). This is shown schematically in Table 2.1.

Table 2.1: Generalized form is not the 2,3sg

1sg	1pl	infinitive
~~2sg~~	2pl	
~~3sg~~	3pl	/UR/

(17) Umlaut verbs leveled to non-umlaut (*a*) alternant

'dig'	*gr**ɔ**b*	*gr**ɔ**bən*
	gre**bst* ⇒ *gr**ɔ**bst*	*gr**ɔ**bt*
	gre**bt* ⇒ *gr**ɔ**bt*	*gr**ɔ**bən*

This leaves a number of possible candidates for the source of the Modern Yiddish present tense vocalism: the 1sg, a plural form, the infinitive, or some abstract underlying form. Can we say anything more specific?

Considering next the preterite present verbs, we find that for these, the Yiddish present tense forms come from MHG singular forms. This is shown in (18) for the verbs *darfən* 'need' and *visən* 'know', whose present tense forms are derived from the MHG singular forms *darf-* and *weiʒ-*, and not the plural forms *dürf-*/*durf-* and *wiʒʒ-*. Other examples include *muzən* 'must' (< MHG sg. *muoʒ*, not pl. *müeʒʒen*), *tərən* 'must' (< MHG sg. *tar*, not pl. *türren*), and *zolən* (< MHG sg. *sol*, not pl. *süln*). (In some cases, the infinitive has also been rebuilt, while in others, the etymologically expected infinitive has been retained.)

(18) Preterite Present verbs leveled to singular

a.

'need'	sg.	pl.
1st	*d**a**rf*	**d**ü**rfən* ⇒ *d**a**rfən*
2nd	*d**a**rfst*	**d**ü**rft* ⇒ *d**a**rft*
3rd	*d**a**rf*	**d**ü**rfən* ⇒ *d**a**rfən*
infin.	**d**ü**rfən* ⇒ *d**a**rfən*	
UR	/*d**ü**rf-*/, /*d**a**rf-*/ ⇒ /*d**a**rf-*/	

b.

'know'	sg.	pl.
1st	*v**ey**s*	**v**i**sən* ⇒ *v**ey**sən*
2nd	*v**ey**st*	**v**i**st* ⇒ *v**ey**st*
3rd	*v**ey**s(t)*	**v**i**sən* ⇒ *v**ey**sən*
infin.	*v**i**sən*	
UR	/*v**i**s*/, /*v**ey**s*/	

We can conclude from the fate of the preterite presents that the generalized form was not a plural form or the infinitive—in fact, most infinitives of preterite presents were also rebuilt on the basis of singular forms. Furthermore, the two MHG stem alternants (*darf-*, *dürf-*) cannot easily be reduced to a single UR, since they involve an idiosyncratic vowel alternation that is attested in only one other verb, and it is not clear how to derive *ü* from *a* or vice versa. The most promising analysis seems to be to list two alternants for these verbs (e.g., /darf/, /dürf/), in which case the form that was generalized in Yiddish does match one of the available MHG UR's (/darf/). However, simply saying the UR has been generalized does not explain why one

UR was chosen and not the other. Putting this conclusion together with the conclusion from the umlaut verbs, we have now eliminated the 2sg, 3sg, all of the plural, the infinitive, and the UR as sources of the Yiddish present tense paradigm. Thus, it appears that the 1sg is only remaining possibility.

The data up to this point converge neatly on a single form as the source for Yiddish present tense paradigms. Unfortunately, when we turn to the *Wechselflexion* verbs (15b), the situation appears to be more complicated. From what we have seen thus far, we would expect that these verbs should generalize the vocalism of the singular (*i*), and indeed this is what we find with the verb *gebən* 'give':

(19) Generalized *i* throughout the paradigm: *gebən*

'give'	sg.	pl.
1st	***gi**b*	**__ge__bən* ⇒ ***gi**bən*
2nd	***gi**bst*	**__ge__bt* ⇒ ***gi**bt*
3rd	***gi**bt*	**__ge__bən* ⇒ ***gi**bən*
infin.	***ge**bən*	

For most MHG *Wechselflexion* verbs, however, Yiddish seems to have generalized the *e* of the plural/infinitive, as in *nemən* 'take':

(20) Generalized *e* throughout the paradigm: *nemən*

'take'	sg.	pl.
1st	**__ni__m* ⇒ ***ne**m*	***ne**mən*
2nd	**__ni__mst* ⇒ ***ne**mst*	***ne**mt*
3rd	**__ni__mt* ⇒ ***ne**mt*	***ne**mən*
infin.	***ne**mən*	

The pattern of generalized *e* is found not only in *nemən*, but also in verbs like *esən* 'eat' (1sg *es*, not **is*), *fargesən* 'forget' (1sg *farges*, **fargis*), *zeyn* 'see' (1sg *zey*, **zi*), *vern* 'become' (1sg *ver*, **vir*), *helfən* 'help' (1sg *helf*, **hilf*), and so on. Why do these verbs show a different pattern from all other verbs in the language? Is this an exception to generalization of the 1sg form?

I would like to argue that verbs like *nemən* and *esən* are not exceptions, but rather that these verbs already contained an *e* in the 1sg at the time that Yiddish "diverged" from other German dialects.[5] I began this section with the assumption that Yiddish began as some form of Middle High German, as exemplified by the standard literary MHG forms in (13)-(15). However, it turns out that the history of the *Wechselflexion* in German is somewhat complicated and controversial, and it is not at all clear that the paradigm in (15) is the correct starting point for Yiddish. According to the standard account (Sonderegger 1987, pp. 146-147; Paul et al. 1989 §§31-35), *Wechselflexion* was due to a phonological process in Old High German that raised /e/ to [i] when there was a high vowel (*u*, *i*) in the following syllable. Since the singular suffixes had high vowels and the plural suffixes had mid vowels, this created an alternation between

[5] I am not making any particular commitment here as to where or when Yiddish ceased to be a sociolect of German and became a separate language, except to suppose that the two probably continued to co-evolve at least until the beginning of the Middle Yiddish period (c. 16th century), when Yiddish literature began to flourish in the east, eastward migrations trickled off, and significant east-west dialect differences emerged (Weinreich 1980, p.724-726).

raised *i* in the singular and non-raised *e* in the plural, shown in the first column of (21). In Middle High German, all suffix vowels were then reduced to schwa, eliminating the conditioning environment for *e*-raising, and making the *e*∼*i* alternation a purely morphological difference between the singular and the plural. This pattern is found in all MHG texts until the mid-15th C (Dammers, Hoffmann, and Solms 1988, §148.4). Finally, at some point during late MHG or early NHG times, the vowel of the 1sg lowered back to *e*, probably under the influence of the umlaut pattern (1sg vs. 2,3sg, as in (14) above).[6] First person singular forms with *e* began to occur regularly in "middle German" (Fränkisch, Thüringisch, Böhmisch, Schlesisch) some time during the fifteenth century (Paul et al. 1989, §242, note 1; Philipp 1980, p. 66), taking hold earlier in the west than in the east (Dammers, et al. 1988, §148.4). The change apparently proceeded verb-by-verb, with considerable variation between verbs and even between occurrences of the same verb in the same text (Kern 1903, pp. 47-60; Geyer 1912, §31-§32) , but eventually affected all *Wechselflexion* verbs of German. This chronology is summarized in (21).

(21) The standard history of *Wechselflexion*

essen 'eat'	OHG 700-1050	MHG 1050-1400	Early NHG 1400-1650	NHG 1650-present
1sg.	*issu*	*isse*	{ *isse, esse* }	*esse*
2sg.	*issis(t)*	*issest*	*isst*	*isst*
3sg.	*issit*	*isset*	*isst*	*isst*
1pl.	*ëssêm*	*ëssen*	*essen*	*essen*
2pl.	*ësset*	*ësset*	*esst*	*esst*
3pl.	*ëssent*	*ëssent*	*essen*	*essen*

What we see, then, is that the use of 1sg forms with *e* in German—either as relic forms (see fn. 6) or as 15th century innovations—predates the rise of Yiddish literature in the 16th century and the last large-scale migrations from west to east during the Thirty Years war (1618-1648). Thus, I hypothesize that Yiddish already had *e* in the 1sg of *Wechselflexion* verbs before wholesale, across-the-board paradigm leveling occurred. If this is the case, then the *e* of verbs like *nemen* and *esn* is not an exception to the generalization that leveling was to the form found in the 1sg.

I have found only three other exceptions to generalization of the 1sg in Yiddish: (1) the verb *zayn(ən)* 'to be' retains a suppletive paradigm, and has not undergone leveling, (2) the future auxiliary verb *velən* is derived from a conditional form, not the 1sg. present indicative, and (3) the verb *gefelən* 'be pleasing' is used predominantly in the 3rd person, and derives from a 3sg form (*gefelt* 'it is pleasing', *gefelən* 'they are pleasing').[7] These exceptions are not particularly

[6]Various scholars have even suggested that 1sg forms with *e* were not pure innovations, but may actually have been relics of a much older stage of the language that survived in the spoken language or in certain dialects; for summary, see Dammers et al 1988, §148. Joesten (1931) argues persuasively that a following *u* (as in the 1sg suffix) never conditioned $e > i$ raising at all, but that the *i* of *ich nime* and *ich isse* was actually an analogical extension of *i* from the 2,3sg. If this is the case, then it seems possible that some dialects may not have undergone this analogical change, and retained the expected *e* in the 1sg all along.

[7]This effect, in which the semantics of a word influence the direction of leveling, is discussed by Tiersma (1982) and others under the rubric of *local markedness.* It should be noted, though, that this is the only such case in Yiddish, so it may be extravagant to invoke local markedness to explain just one case. It is also possible that *gefelən* may derive from a MHG variant of *gefallen*; another example is Yiddish *fregən* 'ask', which derives from MHG *vrëgen*, a variant of *vragen* (Paul et al. 1989, §30). I will return to the issue of local markedness in section 4.4.3, and again in section 6.2.2.

surprising—two are extremely high frequency verbs, and the third has a restricted occurrence for semantic reasons.

In sum, for every type of MHG verb, it appears that the form originally found in the 1sg has been extended to the remainder of the paradigm in Yiddish. This leveling has been remarkably complete, affecting virtually all verbs in the language. In the next section, I will consider the question of why Yiddish paradigms were rebuilt on the basis of this, and not some other form.

2.2 The 1sg as the "optimal base" in Yiddish

2.2.1 Identifying the optimal base

Why did the 1sg have a privileged status among the forms of the present paradigm in Yiddish? In this section, I will argue that the 1sg is "maximally informative" in Yiddish, suffering from the fewest phonological neutralizations, and maintaining distinctions between as many lexical items as possible. The strategy for showing this is to examine a version of Yiddish before any paradigm leveling took place, considering which parts of the verbal paradigm would have been affected by neutralizations, and how many lexical items would have been affected in each case.

Yiddish, like German, English, and many other languages, disallows sequences of tauto-syllabic obstruents with voicing disagreement (*bs]$_\sigma$, *pd]$_\sigma$, etc.). When a suffix consisting of voiceless obstruents (such as 2sg *-st*, 3sg/2pl *-t*) is added to a root ending in a voiced obstruent, the root-final obstruent is devoiced to create voicing agreement. The result is that in the 2sg, 3sg, and 2pl, the contrast between root-final voiced and voiceless obstruents is neutralized.[8] (Shading is used here to indicate a neutralization.)

(22) Neutralization in the 2sg/3sg/2pl: voicing assimilation to suffix

	libən 'to love'	*zipən* 'to sift'
1sg	*lib*	*zip*
2sg	***lipst***	***zipst***
3sg	***lipt***	***zipt***
1pl	*libən*	*zipən*
2pl	***lipt***	***zipt***
3pl	*libən*	*zipən*
infin.	*libən*	*zipən*

This neutralization affects all obstruent pairs with a voicing contrast, of which there are seven in Yiddish (*p/b, t/d, k/g, f/v, s/z, ʃ/ʒ, tʃ/dʒ*). A hypothesis of the current approach is that the seriousness of a neutralization depends not on the number of phonemes involved, but rather on the number of lexical items whose underlying form cannot be recovered because of

[8]A number of studies in recent years have shown that voicing neutralizations of this type may not always be complete, and that the contrast may potentially be preserved through secondary cues, such as preceding vowel length, in some languages (German, Port and O'Dell (1986, Port and Crawford (1989); Catalan, Dinnsen and Charles-Luce (1984); Russian, Chen (1970, pp. 135-137)), but not in others (Turkish, Kopkalli (1993); Italian, Baroni (1998)). I am assuming here that the neutralization caused by devoicing is complete in Yiddish; nevertheless, even if it turned out to be only a partial neutralization, it would be possible to argue that a form with no devoicing at all is still a better source of information about the underlying voicing status of final obstruents than a form with partially neutralizing devoicing.

the neutralization. In order to get an estimate of the number of verbs whose final segment would be ambiguous because of voicing assimilation, I counted the number of verbs ending in these fourteen obstruents in the German portion of CELEX (Burnage 1991).[9] For CELEX counts, I considered only verb lemmas that had a token frequency of 1 or greater (i.e., verbs that actually occurred in the corpus), and that were not "compound", in the sense of having a separable initial element (separable prefix, incorporated object, adverb); this left a total of 4877 verbs. As it turns out, 1988 of these end in an obstruent with a voicing contrast, meaning that approximately 41% of all verbs have an ambiguous final segment in the 2sg, 3sg, and 2pl.

Another set of neutralizations in Yiddish verbal paradigms comes from a ban on geminate consonants within a word. For example, adding the 2sg suffix *-st* to a verb ending in *s* or *z* should yield the sequence *-sst* (with devoicing of *z* to satisfy the voicing agreement requirement). However, this sequence actually surfaces as degeminated *-st*: /veys-st/ → [veyst], not *[veysst]. The result is that *s*- and *z*-final verbs are neutralized with vowel-final verbs in the 2sg, as seen in (23a). For the 3sg and 2pl forms, the suffix is *-t*, and an equivalent degemination of *tt* (fed by /d/ → [t] devoicing) applies in these forms as well (23b).[10]

(23) Neutralizations caused by degemination

a. Neutralization in the 2sg: devoicing of *z*, degemination of *ss*

	geyn 'to go'	*visən* 'to know'	*vayzn* 'to show'
1sg	*gey*	*veys*	*vayz*
2sg	***geyst***	***veyst***	***vayst***
3sg	*geyt*	*veys(t)*	*vayst*
1pl	*geyən*	*veysən*	*vayzən*
2pl	*geyt*	*veyst*	*vayst*
3pl	*geyən*	*veysən*	*vayzən*
infin.	*geyən*	*visən*	*vayzən*

b. Neutralization in the 3sg/2pl: devoicing of *d*, degemination of *tt*

	faltən 'to fold'	*falən* 'to fall'	*redən* 'talk'
1sg	*falt*	*fal*	*red*
2sg	*fal(t)st*	*falst*	*retst*
3sg	***falt***	***falt***	***ret***
1pl	*faltən*	*falən*	*redən*
2pl	***falt***	***falt***	***ret***
3pl	*faltən*	*falən*	*redən*
infin.	*faltən*	*falən*	*redən*

[9]Ideally, we would really like to make these counts on a lexicon of Middle Yiddish, but this does not exist in searchable form, and counts from Modern German form a reasonable approximation. There are certainly numerous lexical differences between the verbal vocabularies of Yiddish and German, and even some phonological ones—for example, Yiddish has some verb roots ending in [dʒ], which is absent in German, and it has a fair number of roots ending in [v], which is quite rare in German. However, the bulk of common Yiddish verbs are shared with German, and there is no reason to believe that the lexical differences would significantly alter the proportion of major classes like obstruent-final verbs, strident-final verbs, etc.

[10]This degemination was common already in MHG (Paul et al., §53d): *valt* or *valtet* 'fold-3sg'. Modern Yiddish has obligatory degemination, while Modern German has obligatory epenthesis.

How many lexical items would be affected by these neutralizations? The voicing neutralization of *s*/*z* and *t*/*d* was already included in the count for voicing assimilation above, but degemination means that vowel-final roots are also ambiguous in these forms—an additional 227 words in CELEX, or 5% of the verbal vocabulary.

So far, we have examined neutralizations in forms with obstruent suffixes—the 2sg, 3sg, and 2pl. Turning to the 1pl, 3pl, and infinitive forms, the suffix for all of these forms is *-ən*. Since this suffix is vowel-initial, and Yiddish allows vowels to occur in hiatus, it does not give rise to illegal sequences that can trigger assimilation or deletion, with one exception: if the verb root ends in a schwa (e.g., *pɔrə-* 'fiddle with') then the 1pl/3pl/infinitive form ends simply in *-ən*, and not **-əən*. This reduction of /əə/ to [ə], motivated by a ban on long schwa (*[əː]), means that in these forms, schwa-final verbs are neutralized with non-schwa-final verbs.[11] This is shown in (24) for the minimal pair *pɔr-ən* 'to match' vs. *pɔrə-n* 'to fiddle with'.

(24) Neutralizations in the 1pl/3pl/infinitive: stem-final /ə/

	pɔrən 'to match'	*pɔrən* 'to fiddle with'
1sg	*pɔr*	*pɔrə*
2sg	*pɔrst*	*pɔrəst*
3sg	*pɔrt*	*pɔrət*
1pl	***pɔrən***	***pɔrən***
2pl	*pɔrt*	*pɔrət*
3pl	***pɔrən***	***pɔrən***
infin.	***pɔrən***	***pɔrən***

How serious is this neutralization? German does not have schwa-final verbs, so it is impossible to use CELEX to estimate the number of lexical items that would be affected by it. Instead, I took a sample from Weinreich (1990), counting all of the verbs beginning with [l]. (This segment was chosen to avoid skewing the sample by including uniquely Slavic onsets like *shtsh-* or *tl-*, or characterically Hebrew onsets like *mə-*; words beginning with [l] seem to come from Germanic, Slavic, and Hebrew in representative proportions.) Of the 90 verbs beginning with [l], 9 of them (10%) have stem-final ə. Thus, a contrast that is seen in a significant portion of the Yiddish verbal vocabulary is neutralized in the 1pl/3pl/infinitive forms.

As with other neutralizations, it is worth considering whether the presence of stem-final schwa is truly neutralized in the 1pl/3pl/infinitive forms, or whether it could be predicted using secondary cues. In casual speech in many dialects, the schwa of the *-ən* suffix may be lost, resulting in a syllabic nasal agreeing in place with a preceeding consonant: [libən] ∼ [libm̩] 'love-1pl/3pl/inf.'. This process affects suffix schwas, but not stem-final schwas—meaning that verbs with stem-final schwa could possibly be distinguished by lack of a schwa-less variant (*pɔrən*/*pɔrn̩* 'match' vs. *pɔrən*/**pɔrn̩* 'fiddle with'). This difference would be rather poor evidence about the status of final schwas, however. First, it requires distinguishing a syllabic nasal from a schwa-nasal sequence, which is not always easy to do, particularly after consonants other than stops. Furthermore, this form is only informative if it is determined to end in a syllabic nasal; if it ends in *-ən*, no conclusion can be drawn. Finally, reduction of *-ən* to syllabic

[11] The same holds true if we assume that the [ə] of the infinitive suffix is epenthetic, and not part of the suffix itself. In this case, the problem is that we do not know if a surface [ən] sequence is the result of adding *-n* to a root that ends in schwa, or of adding *-n* to a root that ends in a consonant and inserting a schwa by epenthesis.

-n does not seem to occur in all environments. This is reflected in some fashion in the YIVO orthography, which uses *-en* after *m*, *n*, *ng*, *nk*, and syllabic *l*, but *-n* elsewhere. In practice, the reduction is probably not as categorically restricted as the orthography implies, but occurs most often after stops, least often after vowels, and so on. Therefore, we would be able to use the 1pl/3pl/infinitive form to infer a lack of final schwa for at best only a subset of verbs in the language.

Another potentially relevant fact is that virtually all schwa-final verbs come from Slavic or Hebrew (e.g., *lyubə-n* 'caress', from Russian/Ukranian *lyubit'* 'love'; *taynə-n* 'claim', from Hebrew *ta'anâ* 'claim').[12] If a speaker could identify a verb as non-Germanic, perhaps by recognizing that it contains a sequence that is illegal in German (such as ***lyub**ən* 'caress', ***pyeshtsh**ən* 'caress', ***tli**ən* 'smolder', ***str**ashən* 'threaten'), there is a much higher likelihood that it will have a stem-final schwa.[13] In addition, there are two derivational suffixes with final schwa: the verbal suffix *-eve* (e.g., *ratevə-n* 'rescue', *zhalevə-n* 'use sparingly', *bushevə-n* 'rage'), and the mimetic suffix *-ke* (e.g., *shushkə-n* 'whisper', *hafkə-n* 'bark', *kvakə-n* 'quack', *bekə-n* 'bleat', *khryukə-n* 'grunt'). These suffixes contribute a large number of schwa-final verbs, and if a verb ends in *-kən* or *-evən*, it is extremely likely to have final schwa. These two facts make it somewhat easier to guess whether a new word should behave like *pɔr-ən* 'match' or *pɔrə-n* 'fiddle with', but it is still far from predictable. Indeed, in addition to *pɔr-ən* vs. *pɔrə-n*, there are a number of other minimal or near-minimal pairs with and without schwa, including *bray-ən* 'brew' vs. *brayə-n* 'talk endlessly', *blank-ən* 'gleam' vs. *blɔnkə-n* 'stray', *kvetsh-ən* 'squeeze' vs. *kvitshə-n* 'squeak', and so on. The upshot is that although it may be possible to guess about the status of a final schwa in some cases, it would still be easier and more accurate to choose a form that shows it unambiguously, such as a singular form or the 2pl.

The neutralizations discussed so far would have affected forms with overt suffixes—that is, all forms except the 1sg. The 1sg form would not have been subject to such severe neutralizations, because Yiddish had no phonological processes affecting segments in stem-final position.[14] This is not to say that the 1sg would have been completely free from neutralizations, however; in fact, there are two properties of verbs that could not have been predicted from the 1sg form alone. Umlaut verbs like *fɔrən* would have had the same vowel (ɔ) as non-umlaut verbs like *pɔrən* in the 1sg, and preterite present and *Wechselflexion* verbs would likewise have

[12]Weinreich does list some Germanic words with stem-final schwa, such as *veyə-n* 'blow' and *knurə-n* 'snarl'; cf: German *wehen, knurren*. Lass (1980) cites Mieses (1924) in identifying also *zeyə-n* 'sow' and *kreyə-n* among this group, although Weinreich lists both as having free variation (e.g., *zeyə-n*/*zeyə-n*). One might imagine that these are isolated relics of a time when MHG endings had *e* (*wehet, knurret*), but curiously, they seem to involve stems ending in *r* and *h*, which were among the first environments for syncope of *e* in final syllables (Paul et al. 1989, §53, §240). I have no explanation for how these verbs came to have stem-final schwas in Yiddish.

[13]The problem of how speakers identify members of separate lexical strata is a general one in phonology; see Itô and Mester (1995, Itô and Mester (2002) for a discussion of the different phonotactics for different lexical strata in Japanese. Lass (1980) claims that identifying non-Germanic words in Yiddish based on general phonological properties is a "non-starter" (p.263), based on the fact that there are no systematic differences in their stress patterns. Presumably, he does not consider the possibility of using certain phonemes or phoneme clusters as indicators of non-Germanic status because they do not work 100% of the time; there are plenty of non-Germanic words that by chance happen to be composed of elements that are legal in the Germanic part of the lexicon. We might, however, for the sake of argument suppose that speakers could identify at least a subset of the verbs of Slavic origin, and use this to help predict the occurrence of final schwa.

[14]It appears that an earlier stage of Yiddish did have final devoicing, but this was lost early on in most dialects; see Sapir (1915), Kiparsky (1968), Sadock (1973), and King (1980) for discussion.

been indistinguishable from non-alternating verbs in this form. A crucial difference from the neutralizations discussed above, though, is that umlaut, preterite present, and *Wechselflexion* verbs would have been ambiguous with some non-alternating verb not only in the 1sg, but in *every* part of the paradigm. (Recall that we are considering here a version of Yiddish prior to paradigm leveling, in order to try to predict the base of the subsequent leveling; in actual Modern Yiddish, all of these verbs have uniform paradigms (17-20).) The shading in (25) shows that in some parts of the paradigm, these verbs resembled the non-alternating verbs in the middle column, while in other parts of the paradigm, they resembled those in the final column.

(25) Neutralizations that include the 1sg (pre-leveling forms)

a. Umlaut verbs neutralized with non-umlaut verbs

	fɔrən 'to travel'	*pɔrən* 'to match'	*hern* 'to hear'
1sg	***fɔr***	***pɔr***	*her*
2sg	***ferst***	*pɔrst*	***herst***
3sg	***fert***	*pɔrt*	***hert***
1pl	***fɔrən***	***pɔrən***	*hern*
2pl	***fɔrt***	***pɔrt***	*hert*
3pl	***fɔrən***	***pɔrən***	*hern*
infin.	***fɔrən***	***pɔrən***	*hern*

b. Preterite presents and *Wechselflexion* neutralized with non-alternating verbs

	visən 'to know'	*heysən* 'to order'	*vishən* 'to wipe'
1sg	***veys***	***heys***	*vish*
2sg	***veyst***	***heyst***	*vishst*
3sg	***veys(t)***	***heyst***	*visht*
1pl	***visən***	*heysən*	***vishən***
2pl	***vist***	*heyst*	***visht***
3pl	***visən***	*heysən*	***vishən***
infin.	***visən***	*heysən*	***vishən***

Since these neutralizations affect all parts of the paradigm equally and do not favor any particular choice of base, it is perhaps unnecessary to count the number of lexical items involved. It may be noted, however, that compared with the neutralizations discussed above, these would have affected only a very small number of words. In MHG, umlaut occurred in only in a handful of verbs, mostly in the strong classes VI and VII; it is difficult to get a comprehensive list of all MHG umlaut verbs, but there were perhaps less than two dozen altogether (Paul et al. 1989, §§ 251-253). Added to these were about a dozen preterite present verbs (§§ 269-275) and around 70 verbs in the *Wechselflexion* classes (IIIa, IV, and V; §§ 247-250), for a total of around 2% of the verbal vocabulary.

The combined effect of these neutralizations is summarized in Table 2.2, which shows the proportion of lexical items whose underlying form could not be unambiguously recovered from each part of the paradigm. The conclusion from all of these counts is that the 1sg form in Yiddish preserves the greatest number of phonemic distinctions, including the voicing of stem-final obstruents, the presence of stem-final *t*, *d*, *s*, and *z*, and the presence of stem-final ə. Thus, given a 1sg form, it would be possible to predict virtually every form of every word with absolute

Table 2.2: Summary of neutralizations in Yiddish

	singular	*plural*
1st	umlaut ($<$ 1%??)	final -ə umlaut } (10% ?)
2nd	obstruent voicing final -s/-z/-V } (46%)	obstruent voicing final -t/-d/-V//umlaut } (46%)
3rd	obstruent voicing final -t/-d/-V } (46%)	final -ə umlaut } (10% ?)
infin.	final -ə umlaut } (10% ?)	

certainty, with the exception of the 2sg/3sg of umlaut verbs and the plurals of preterite present and *Wechselflexion* verbs.

2.2.2 Using the 1sg as the base to derive Yiddish verb paradigms

Suppose that you are a language learner, trying to acquire Yiddish at a time before the paradigm leveling has occurred. Your goal is to be able to produce and comprehend all forms of all words, and in order to do this, you need to learn the distinctive phonological properties of each verb—the number of phonemes, their voicing, and so on. I have shown that the 1sg form on its own could provide almost all of this information, and would thus be the optimal choice of base form to predict other forms. A hypothesis of the current approach is that once the learner has identified this fact, she goes on to develop a grammar that derives the rest of the paradigm from the base form. For Yiddish, this grammar would include:

(26) Rules for deriving Yiddish paradigms

a. Morphological rules:
- 2sg: $\varnothing \rightarrow$ *-st*
- 3sg: $\varnothing \rightarrow$ *-t*
- 1pl: $\varnothing \rightarrow$ *-ən*
- 2pl: $\varnothing \rightarrow$ *-t*
- 3pl: $\varnothing \rightarrow$ *-ən*
- infinitive: $\varnothing \rightarrow$ *-ən*

b. Phonological rules (or their constraint-based equivalents):
- Obstruent devoicing: $[\text{-son}] \rightarrow [\text{-voice}]\ /\ \_\_ \begin{bmatrix}\text{-son}\\ \text{-voice}\end{bmatrix}\#$
- *t* Degemination: t $\rightarrow \varnothing$ / ___t#
- *s* Degemination: s $\rightarrow \varnothing$ / ___s#
- ə Degemination: ə $\rightarrow \varnothing$ / ___ə#

If we use 1sg forms as the input for these rules, they will unambiguously yield the correct result for almost all forms of all words. The only exceptions are the 2,3sg of umlaut verbs, for which we would predict incorrect forms like **fɔrst* and **fɔrt*, and the plurals of preterite present and *Wechselflexion* verbs, for which we would predict incorrect forms like **veysən* and **gibən*. Under this approach, then, forms with umlaut (*ferst*) and with sg.∼pl. alternations (*visən*) would have to be learned as exceptions.

It does not matter for present purposes whether umlaut forms are stored individually as whole-word exceptions, or whether we posit an umlaut rule that applies only to words that are lexically specified for it.[15] As with all analyses that make recourse to lexical exceptions, there is an issue of how lexical exceptions can be constrained to avoid arbitrarily different exceptions for each word. In this case, the problem is to ensure that exceptional 2,3sg forms are by and large exceptional in having vowel fronting, and not in other ways. A traditional approach within generative grammar is to set up a diacritic marking exceptional words for a different grammar—either one with or without an additional rule, or one with a different constraint ranking (Itô and Mester 1995, 2002). There are some cases in which "co-phonologies" with different constraint rankings seems well motivated (for example, as it is used by Itô and Mester to distinguish the strata of the Japanese lexicon), particularly if one could formulate meta-constraints on ways in which co-phonologies are allowed to deviate from one another (e.g., Anttila 2002). However, in the cases discussed in this thesis (Yiddish, Latin, and Lakhota), it may be less appealing to consider exceptional words as essentially belonging to a separate stratum. Furthermore, this approach is problematic from a learnability point of view, because the learner must somehow be able to distinguish between cases where the grammar simply needs to be refined, and cases that require a separate co-phonology/diacritic (i.e., when do you keep learning, and when do you give up and use a diacritic?). Another possibility is to list exceptions as full surface forms, but to constrain the possible alternations within the paradigm using output-output surface constraints (Burzio 1996). The latter approach is more consistent with the general model proposed here, and something along these lines is what I will assume.[16] All that is crucial, however, is that forms like *ferst* and *visən* could not be derived productively by the grammar, and require some overriding word-specific mechanism that may fail in certain conditions. If a speaker forgets or is unable to access the correct exceptional form, she will use the grammar to produce an "overregularized" form (*fɔrst*, *veysən*). Furthermore, if these mistakes are accepted and adopted by the speech community, they will eventually replace the old, exceptional forms.

There are clearly many factors at play in determining how willing a community is to adopt new forms; I conjecture that the thoroughness of the change in Yiddish may have been facilitated by the lack of a standard language or widespread literacy, and perhaps even by a conscious desire to differentiate Yiddish from German. The model that I am presenting here is simply an attempt to predict which forms would have been available as *potential* regularizations for Yiddish speakers, and which would not.

---

[15]For a selection of proposals on the handling of exceptions in phonology, see Kenstowicz and Kisseberth (1977, pp. 114-130), Zonneveld (1978), and Zuraw (2000, pp. 67-71); for more recent OT proposals, see (Ellison 1994), Tranel (1996), Zuraw (2000), and Green (2001).

[16]Another common observation about exceptions is that they tend to fall into phonologically similar "gangs", which can help support each other in their irregularity. It is possible that the set of exceptions is constrained not just by statements about what alternations are *possible* in the language, but also by what alternations are *likely*, and in what environment. The model proposed in chapter 3 may be able to capture this by making use of the less reliable, unproductive rules somehow; if an exception is not the grammatically preferred form, at least it should get some support as a second choice losing form.

2.3 Comparison with other German dialects

The Yiddish paradigm leveling discussed above seems like quite a natural change, even if the completeness of the leveling is perhaps a bit striking. If this change is really so natural, however, we would expect that it might have occurred in some other related dialects as well. An informal survey of dialect descriptions revealed several candidates for dialects that bear a superficial resemblance to Yiddish in their present tense forms, but all turned out to have different explanations. In Dutch and some northern German dialects, for example, the present singular paradigm is always uniform (*graaf*, *graaft*, *graaft* 'dig'), as it is in Yiddish. However, these languages never had umlaut to begin with, so their failure to alternate is not due to leveling. Some southern German dialects (Schwabian, Frankish, Bavarian, etc.), which did historically have an umlaut rule, also have uniform present tense paradigms (e.g., Bavarian *grab*, *grabsd*, *grabd*) (Schirmunski 1962; Zehetner 1989). However, these dialects show leveling only of the singular present paradigm, while maintaining alternations between the singular and plural in preterite present verbs. It appears that these dialects have lost the umlaut rule for the 2,3sg, rather than undergoing paradigm leveling in the same way that Yiddish has. Finally, some Early NHG texts occasionally have *e* in the 2,3sg of *Wechselflexion* verbs, such as Fischart's use of *schmelzt*[17] 'melts-3sg' (Standard German *schmilzt*), reminiscent of the generalization of *e* seen in Yiddish (20). This seems to happen mainly with verbs that are also sometimes given regular (weak) pasts, however. Weak verbs never have *Wechselflexion*, and the loss of it in the singular of these verbs was probably part of a larger trend to create weak counterparts of strong verbs in Early NHG. In none of these cases do we find compelling evidence of paradigm leveling of the sort seen in Yiddish.

In contrast, there have been numerous changes in German that have either introduced new alternations, or have leveled to a form other than the 1sg. The change from *i* to *e* in the 1sg of *Wechselflexion* verbs, for example, is usually analyzed as an extension of the umlaut pattern (in which the 1sg had a lower vowel than the 2,3sg) to new verbs (though an alternative possibility was discussed above), and umlaut has been extended to other verbs as well. In addition, many verbs in MHG and NHG have been rebuilt on the basis of 3sg forms, such as MHG weak *schricken* 'scare' from Strong V *schrecken*, NHG *ziemen* from Strong IV *zëmen*, and *wiegen* 'rock' from Strong V *wëgen* 'move'.

It would be extremely difficult, of course, to prove that the leveling found in Yiddish has never occurred in any other form of German, but my tentative conclusion from this comparison is that the prevailing tendencies in German have been in different directions. So why would Yiddish have departed so radically in this respect? When we consider the differences between Yiddish and German, we find that two of the neutralizations discussed above do not occur in German. First, the degemination of /dt/ and /tt/ to [t] ((23b) above) is found in only a few dialects (Schirmunski 1962), meaning the 3sg and 2pl forms preserve the voicing contrast between stem-final *t* and *d*, and keep both distinct from stem-final vowels. This actually creates quite a significant difference between Yiddish and German, because 562, or 12% of the verbs in the CELEX corpus end in coronal stops. Furthermore, German has no stem-final schwas, eliminating a major source of ambiguity in the 1pl/3pl/inf. forms. The neutralizations of Standard NHG are summarized in Figure 2.1. Comparison with the equivalent graph for Yiddish

[17] Ehzuchtbüchlin (1578), S. 228: "vnd schmeltzet inn der liebe vnd … wie inn eynem Ofen zusammen," cited by Geyer (1912, §23.6).

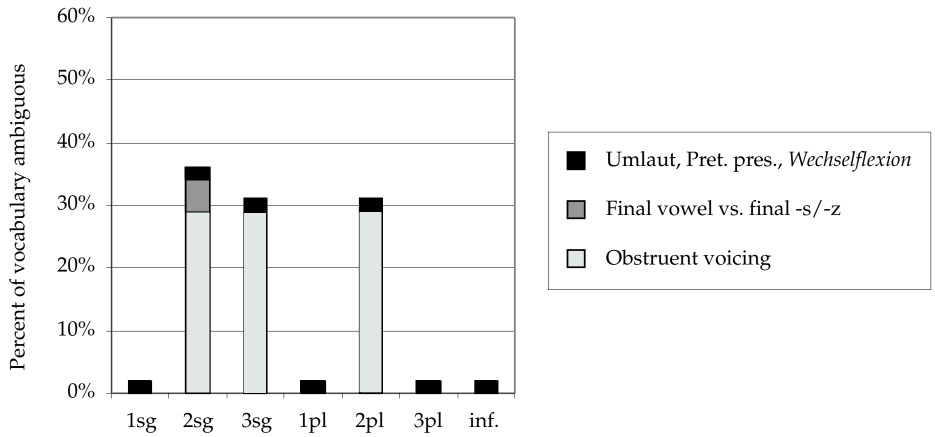


Figure 2.1: Summary of neutralizations affecting German verb forms

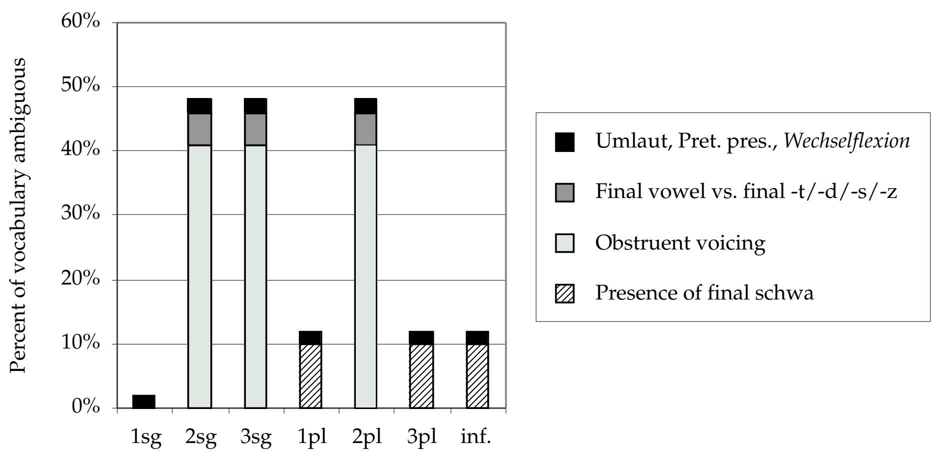


Figure 2.2: Summary of neutralizations affecting Yiddish verb forms

(Figure 2.2) shows that in German, the 1sg form is not uniquely informative; the plural and infinitive forms are just as good. Furthermore, even the 2sg/3sg/2pl forms are not nearly as ambiguous in German as in Yiddish. Thus, a plausible explanation of the difference between Yiddish and German is that in German, some other form (such as the infinitive) is acting as the base.

2.4 Local summary

In this chapter, I have presented evidence from Modern Yiddish that a cross-linguistically marked form (the 1sg) has served as the base of paradigm leveling, affecting almost every verb of the language. A comparison of neutralizations showed that even before the leveling, the 1sg form would also have had the unique property of unambiguously preserving as many properties of the root as any one form could preserve. The only properties that were neutralized in this form were umlaut and *Wechselflexion* alternations, which in fact could never be preserved in a single form, since observing an alternation necessarily requires at least two forms. Comparing the neutralizations of Yiddish with those of German also provided some insight into why the 1sg form may not be so privileged in other, closely related dialects.

The strategy of comparing neutralizations in various parts of the paradigm is based on the premise that language learners need to be able to produce and understand forms that they have never heard before, and they do this by concentrating on the part of the paradigm that reveals properties of the word as unambiguously as possible. Phonological properties (such as obstruent voicing), morphophonological properties (such as the difference between umlaut *a* vs. non-umlaut *a*), and morphological properties (like verb class) are all taken into account when considering which part of the paradigm is most informative. The calculations of how many lexical items would be affected by each neutralization that I have been using here may seem a bit informal, but they are meant to serve as a conceptual example for a more rigorously defined, computationally implemented algorithm that will be described in more detail in chapter 3. The purpose of this Yiddish example is to demonstrate the basic strategy of base identification: we start by considering each member of the paradigm as a potential base, and then we construct grammars of morphological and phonological rules to derive the remainder of the paradigm. Finally, we compare how effective these grammars are by calculating how many forms are derived correctly by the grammar, how many exceptions must be stored, how reliable the stochastic rules are, how confident (probable) the rules generating the correct outputs are, and so on.

2.5 Significance of the single surface base hypothesis

I have assumed in this chapter that the task of the learner is to select a single surface form as the base—but is this assumption really necessary? In order to assess this, it is useful to consider the predictions of a more traditional model, without the single surface base restriction. Under such a model, a learner could notice that some contrasts (like obstruent voicing and final schwas) are seen in some parts of the paradigm, while other contrasts (like umlaut) are seen in others. The learner could then combine information from multiple parts of the paradigm to create an underlying form that captures *all* unpredictable information.

In the case of Yiddish, comparing different forms to synthesize abstract URs might yield the following results: for non-alternating words like *pɔrən* 'match' (13b) or *herən* (25a), the UR of the verb root would be identical to its SR in all forms: /pɔr-/, /her-/. For umlaut verbs like *fɔrən* 'travel', comparing the two surface alternants (*fɔr-* and *fer-*) would lead the learner to posit some sort of underlying distinction between alternating *fɔrən* and non-alternating *pɔrən*. Many different mechanisms for doing this have been proposed in the literature, including:

- Positing a lexically restricted umlaut rule (/ɔ/ → [e]) and then marking *fɔrən* to take this rule ($/\text{fɔr-}/_{[+\text{umlaut}]}$)
- Listing both *fɔr-* and *fer-* as possible allomorphs (Kager 1999, pp. 413-420; Burzio 1996), or using a disjunctive representation with two choices of vowel: /f{ɔ,e}r-/ (Hooper 1976).

In this case, the choice of [ɔ] or [e] is determined by surface constraints, which in this case may admittedly be rather ad-hoc (e.g., "3SG-[e]: The 3sg form must have the vowel [e]")

- Using an abstract, underspecified archiphoneme /O/ (or, a fully specified phoneme that never actually surfaces in Yiddish), along with rules or constraints favoring the addition of front/unrounded features in some contexts, and back/rounded features in others

Any of these approaches could explain why learners might sometimes incorrectly fail to apply umlaut, yielding the Yiddish change (*fɔr, *fɔrst, *fɔrt*). For example, they may simply not yet have learned the umlaut rule, or the constraints favoring [e] in the 2,3sg. Moreover, even once they have learned to produce the umlaut alternation, they may have heard a particular verb in only non-umlaut forms, and in the absence of positive evidence that the verb should undergo umlaut, they may not yet have marked it as [+umlaut], or set up an umlaut allomorph for it. Failure to apply umlaut is not the only possible error, however; there are also possible scenarios that should lead to extension of the umlaut vowel to the entire paradigm (**fer, ferst, fert*). Just as learners may have encountered a verb in only non-umlaut forms, it should sometimes happen that they have heard a particular verb in only umlaut forms, yielding a provisional UR with /e/ instead of $/ɔ/_{[+umlaut]}$, or a UR with just one listed allomorph.

The problem here is that in cases like Yiddish, where there are three possible surface patterns (non-alternating [ɔ], non-alternating [e], and alternating [ɔ] ∼ [e]), it takes evidence from multiple surface forms to be sure that the correct UR has been established. If learners are able to posit URs on the basis of evidence from anywhere in the paradigm, however, then we predict that symmetrical errors should be possible. Suppose, for example, that there is a word with alternating [ɔ] ∼ [e] (such as *fɔr* ∼ *fer*), but the learner has heard it in only one of these forms. If that form was an [ɔ] form (*fɔr*), then the learner may incorrectly conclude that the word is a non-alternating /ɔ/ word (*fɔr, *fɔrst, *fɔrt*). If, on the other hand, that form happened to be an [e] form (*fert*), then the learner may incorrectly conclude that the word is a non-alternating /e/ word (**fer, ferst, fert*). Thus, we are unable to predict the observed asymmetry, that all alternating verbs became invariant /ɔ/, and none became /e/.

What if we allowed learners to set up underlying forms for alternating [ɔ] ∼ [e] without actually hearing both [ɔ] and [e]? For example, what if the learner, upon hearing 3sg *fer*, could infer that the *e* must correspond to an ɔ in the 1sg, and posit a UR like $/fɔr\text{-}/_{[+umlaut]}$?[18] This would help to eliminate mistakes like (**fer, ferst, fert*), because the learner would not have to wait around to hear both *fɔr* and *fer* before positing an alternating UR; she could predict *fɔr* on the basis of *fer*. This strategy would have the unfortunate consequence of introducing other, unattested mistakes, however. In particular, it could also lead the learner to posit alternating URs for words that should not alternate. For example, on hearing the 3sg form *hert* 'hear-3sg', the learner might posit the UR $/hɔr\text{-}/_{[+umlaut]}$, predicting the incorrect paradigm **hɔr, herst, hert* instead of the correct invariant *her, herst, hert.*

The challenge, then, is to explain why nonalternating [ɔ] and [e] were consistently learned as such, while alternating [ɔ] ∼ [e] was sometimes learned as [ɔ], but never as [e]. What we need is a principle that tells the learner to ignore 2,3sg forms with [e].For example, they may notice that [e] is notoriously ambiguous in these forms, and learn not to draw any conclusions from it, setting up neither alternating URs nor invariant [e] URs.

[18] A model that does just this has been proposed by Harrison and Kaun (2000); I will discuss this idea at greater length in chapter 5.

The problem is even harder when we turn to the preterite present and *Wechselflexion* verbs. As with umlaut, there are various possible approaches to handling the alternations seen in these verbs: we may set up lexically restricted rules (though in this case most would apply to just one or two verbs), we may list separate allomorphs, etc. Once again, the challenge is to explain why speakers extended one vowel (that of the singular), rather than another (that of the plural), and rather than extending the alternation to new lexical items. In the case of umlaut, we saw that perhaps the learner could identify [e] in the 2,3sg as a particularly ambiguous phoneme; in this case there is no such red flag. There are a number of different alternations involved: *veys* ~ *vis* 'know', *darf* ~ *durf* 'need', *muz* ~ *müz*/ *miz* 'must', *tɔr* ~ *tür*/ *tir* 'must', *gib* ~ *geb* and so on. What property do these words have in common that would tell learners to ignore their plural forms?It is difficult to see what considerations could have privileged the vowel of the singular for this particular set of words.

The upshot is that in order to make asymmetrical predictions about possible errors, we need to find a way to restrict the set of forms that learners are considering as possible URs. (I will return to this issue in greater detail in chapter 5.) For this reason, the single surface base hypothesis is relevant for all models of morphology and phonology, and not just those that limit themselves to statements about relations between surface forms. It makes strong predictions about which errors should occur and which should not, by limiting speakers to using information from just one place in the paradigm to project new forms, even if it does not reveal all of the information necessary to project all forms of all words correctly.[19] The learning procedure is designed to mitigate this problem by selecting the base form that preserves the most contrasts, in order to minimize the number of exceptions that must be represented separately. The prediction is that contrasts that are preserved in the base will be maintained, while contrasts that are neutralized in the base will be open for leveling—which, in the case of Yiddish, appears to be correct.

[19] I will return in chapter 6 to the issue of local bases for subsets of forms within the paradigm.

Chapter 3

Identifying bases algorithmically

In chapter 2, I informally demonstrated the approach of comparing different forms of the paradigm to determine their “informativeness,” and showed that the most informative form was also the one that acted as the base of a paradigm leveling that took place in the history of Yiddish. In this chapter, I propose an algorithm for comparing informativeness in a more systematic fashion. The algorithm starts by considering each member of the paradigm as a candidate for base status, and constructs grammars that use each as an input to derive the remainder of the paradigm, as seen in Figure 3.1.

More formally, the model starts by taking paradigms of related forms (A, B, C, D, E, ...), and considering all of the pairwise relations between them (A→B, A→C, ..., B→A, B→C, B→D, ...). I will call such pairwise relations *mappings*. For each mapping between two slots in the paradigm (X→Y), the model constructs a set of rules transforming X’s into Y’s. The result is what I will call a *sub-grammar*, consisting of all of the rules that describe a particular morphological mapping between just two slots in the paradigm (X→Y). A complete morphological grammar is a set of such sub-grammars, including at least one sub-grammar to derive each slot in the paradigm.

Given this terminology, it is possible to give a more precise formulation of the learner’s task:

- The learner must be able to produce and comprehend forms from all parts of the paradigm (A, B, C, D, E, ...) accurately and confidently
- Thus, the learner must identify which member of the paradigm contains the most information about how to derive the remainder of the paradigm

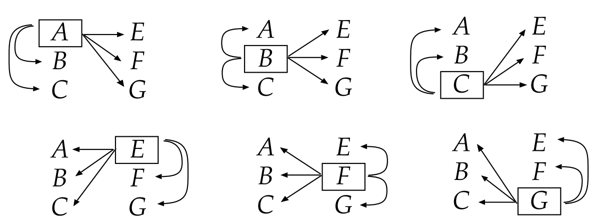


Figure 3.1: Comparison of different candidates for base status

- The most informative form is the one that permits the most *accurate* or *efficient* subgrammars (under a form definition of accuracy and efficiency, to be given below)
- Once the learner has compared the relative informativeness of various forms in the paradigm, the most informative one is selected as the base, and the grammar that is chosen is the one which operates on that base form to derive the remaining forms in the paradigm

Before we can find the optimal mappings for a language, we need two things: (1) a rule discovery system to learn the sub-grammars originating from different candidate base forms, and (2) a set of metrics that can score the resulting sub-grammars as more or less optimal. For the first task, I will use the *minimal generalization* rule induction system developed by Albright and Hayes (1999a, to appear). After giving a brief outline of this model in section 3.2, I will propose a set of metrics to evaluate the goodness of sub-grammars in section 3.3. Before describing the implemention, however, it is useful to review some of the considerations that led to the choice of this particular system.

3.1 Desiderata for an automated learner

Not all of the currently available systems for rule discovery produce grammars that are equally suitable for comparing their efficiency or certainty. Moreover, the choice of model crucially determines the types of comparisons which will be possible. The model of morphological learning that I will adopt is one developed by Albright and Hayes (1999a, to appear). This choice is not merely a matter of convenience, however, and in this section I will argue briefly why some other models are in fact unsuited to the task at hand.

Consider the problem of evaluating how well a model is able to perform on a particular mapping X→Y. One test that we might want to use is whether the model is generating the right output (Y) for each input (X). However, recall from the discussion in section 1.1 that there are potentially two different ways to generate outputs: either by retrieving the form from memory, or by synthesizing it using a set of generalizations (such as a grammar). If the model is evaluating its own performance on X→Y based on words that it already knows, then in some sense the easiest and most reliable strategy would be to memorize every form, so that the correct output can always be generated no matter how accurate the grammar is. This puts us in a bit of a bind; on the one hand, we want morphological models to be able to memorize output forms in order to handle irregulars, but we also want to be able to test the accuracy of the synthesized forms. In essence, we want to know how much of the input data the model considers irregular, and how much it could reproduce productively using just the grammar. Thus, our first desideratum is that a model must make a clear distinction between outputs that have been retrieved from memory and outputs that are synthesized productively.

A prominent class of models that do not meet this criterion are connectionist models (Rumelhart and McClelland 1987; MacWhinney and Leinbach 1991; Daugherty and Seidenberg 1994; Westermann 1997), since these intentionally blur the distinction between rote memorization and generalization. Even though it may be possible to use various techniques to diagnose to what extent a neural network is using generalization or word-specific memorization for each word, interpreting the end state of networks can be difficult and laborious (Clark and Karmiloff-Smith 1993; Hutchinson 1994). For this reason, connectionist models are not especially well suited to the current problem.[1]

Another framework that often does not make a distinction between lexicalized and grammatically generated patterns is Optimality Theory. In fact, one of the original claims of OT is that it is *impossible* to memorize a lexical entry that is not transformed into the grammatically preferred pattern on the surface (*Richness of the Base*, Prince and Smolensky 1993, Smolensky 1996). Versions of OT that assume Richness of the Base are not especially well suited to the task at hand, and I will not use OT for the analyses of Latin or Lakhota phonology in the next two chapters. However, it must be noted that one could probably recast at least portions of these analyses into some modified version of OT, like that proposed by (Zuraw 2000), in which speakers can reason about constraints and lexical entries and allows us to distinguish between lexically listed forms and grammatically derived forms.

A second way in which we might try to diagnose the performance of grammars is by seeing if they derive their outputs unambiguously and confidently. If a particular sub-grammar X→Y is attempting to project from a form with many neutralizations to a form with more contrasts, there will be many cases in which the input is ambiguous, and there are two or more possible outputs. In order to evaluate this, we need a model that can generate multiple outputs for each input (where appropriate). In addition, it is helpful if the model can offer some form of confidence or well-formedness scores, allowing us to diagnose which forms the model prefers, how close the competition is, etc. There are several models that are similar in spirit to the Albright and Hayes minimal generalization model, but which do not provide such scores; these include Ling and Marinov's inductive decision tree approach (Ling and Marinov 1993), and Neuvel's whole word morphology model (Neuvel and Singh, 2001; Neuvel, to appear; Neuvel and Fulop, to appear).

One last desideratum is that the model should operate incrementally, and that it should be able to provide an analysis at every stage of the learning process. The reason is that if we want base identification to simplify the task of morphological learning, then we need to be able to do it early on, so we do not have to keep exploring all sub-grammars in all directions. Therefore, we must be able to identify the best base using a small subset of the data, and for this reason, we want grammars based on a few forms to be reasonably similar to grammars based on the whole data. (I will discuss the issue of size of learning set again at the end of section 3.2.) Many approaches to morphological learning do not meet this criterion, taking instead a top-down approach that tries to find the most general rules first. Examples of such models include the First Order Inductive Decision List (FOIDL) framework (Mooney and Califf 1996) along with other Inductive Logic Programming (ILP) approaches, Kazakov's "Naïve Theory of Morphology" employing genetic algorithms and ILP (Kazakov and Manandhar 2001), and Minimum Description Length (MDL) approaches (Brent, Murthy, and Lundberg 1995; Goldsmith 2001). Unfortunately, in order to have a good idea of what is the most general pattern in the language, one must have access to all of the forms in language; therefore, these systems tend to require the

[1]Hare and Elman (1995) discuss a similar problem, of evaluating whether connectionist networks are easier to train on some patterns than on others, as an explanation for why certain patterns in Old English verb classes were eliminated while others were extended. They trained their models in such a way that the model was still making some errors at the end of the training period, and these were fed into the next "generation"; thus, their models were not always able to reproduce the input data perfectly, but the model does not "know" that it is making an error. In the task of base identification, we want the model to be able to diagnose how often its grammar will produce the wrong output. The easiest way to do this is to let the model memorize the correct output for every single word, and then compare its memorized and synthesized outputs.

entire learning set in advance.[2] This requirement is incompatible with the goal of discovering early on which mappings to focus on, and would only be useful for making post-hoc decisions about which mappings yielded the cleanest generalizations.

The Albright and Hayes learning model satisfies all of these requirements, in ways that will be explained in the next section.

3.2 The minimal generalization learner

The minimal generalization learner takes a bottom-up (smallest generalization first), incremental approach to morphological learning. As its input, it takes ordered pairs of morphologically related words, such as the following:

(27) Sample input for minimal generalization

absolutive		ergative
lag	∼	*lagi*
basag	∼	*basagi*
sub	∼	*subi*
lot	∼	*loti*

Given the first pair of forms (*lag,lagi*), it constructs a morphological rule which could apply to the first to yield the second. It does this by locating the structural change (in this case, $\varnothing \rightarrow$ *i*), and considering the entire rest of the word as the structural environment (*lag*___).[3] The result is shown in (28).

(28) $\varnothing \rightarrow$ i / lag___#

Generalization begins when a second pair of forms is encountered with the same structural change. For example, the second pair of forms in (27) also contains the structural change $\varnothing \rightarrow$ *i*, and yields a word-specific rule with the structural environment *basag*___#. The generalization algorithm then compares the structural descriptions of these two rules, factoring them into shared portions and non-shared portions. The shared portion has two parts: a continuous string of shared segments strictly adjacent to the structural change, and the phonological features shared by the first segment in which the two forms differ. A new rule is formed by retaining the shared portions as the structural description, and converting the non-shared portions into variables. I will call such rules *generalized rules*, in contrast to word-specific rules like the one in (28).

[2]There is no reason why we could not apply a top-down system to progressive subsets of the data, but I am not aware of any attempts to do this, and it is not clear to me how the results would differ from simply using a bottom-up algorithm.

[3]In its current implementation, the parsing component of the learner is able to consider only structural changes which involve prefixation, suffixation, and infixation, but not simultaneous combinations of these. It can also consider structural environments which involve only contiguous strings on both sides of the change, plus featural descriptions of the segments immediately before and after the change. These limitations should not be an insurmountable problem for any of the languages which I propose to examine.

(29) Minimal generalization across two word-specific rules:

	change	residue	shared features	shared segments	change location
comparing A:	∅ → i /		l	ag	___#
with B:	∅ → i /	ba	s	ag	___#
yields C:	∅ → i /	X	$\begin{bmatrix} +\text{cons} \\ +\text{cont} \\ +\text{cor} \\ \text{etc}\ldots \end{bmatrix}$	ag	___#

="suffix *-i* after *ag*, preceded by any coronal continuant"

When the input pairs are quite similar, as is the case in (29), the generalized rule that results from this procedure is quite specific. However, when generalization is iterated across the entire lexicon, comparing forms with a variety of phonological shapes, the resulting rules can also be quite general. Continuing with the forms in (27), the pair (*sub,subi*) tells us that *-i* suffixation can occur not only after *ag*, but after the class of voiced non-coronal stops. The next pair (*lot,loti*) further shows us that *-i* can be suffixed not just after voiced non-coronal stops, but in fact after any stop.[4]

The forms in (27) show a consistent pattern—all of the words take the suffix *-i* in the absolutive. Real languages are rarely so regular, however; there are often several different processes competing to express the same morphological category. Suppose, for example, that in addition to the forms in (27), the target language had the following two forms:

(30) A competing suffix

absolutive		ergative
rag	∼	*ragu*
tip	∼	*tipu*

When the first of these forms is heard (*rag*∼*ragu*), it spawns its own word-specific rule, just as the previous words did:

(31) ∅ → u / rag___#

However, this structural change (∅ → *u*) is not shared by any of the words in (27); therefore, the generalization algorithm can not compare it to any other form to create a more general rule of *-u* affixation at this point.

Once the second *-u* pair (*tip* ∼ *tipu*) is learned, a word-specific rule of ∅ → *u* / *tip*___# is spawned, and then this is compared with the ∅ → *u* / *rag*___# rule to yield a more general *-u* affixation rule:

[4]Since minimal generalization always tries to find the tightest fit to the data seen so far, the exact natural classes which are hypothesized will depend somewhat on the feature set; if the features distinguish some type of stops other than voiced or voiceless (such as, say, aspirated), or has more places of articulation (such as pharyngeal or palatalized), then minimal generalization over this input set will contain feature specifications to exclude them. In addition, the types of generalizations that are possible will also depend on whether various features are treated as binary or privative. If [nasal] is privative, for example, it would be impossible to construct a generalization that applies just in the context of non-nasals. It is an open empirical question as to what the correct set of features is, and whether the feature combinations that are necessary to describe phonological phenomena are the same as those that are needed to group words for the purposes of morphology.

(32) $\varnothing \rightarrow u \, / \; \begin{bmatrix} \text{C} \\ -\text{continuant} \\ -\text{coronal} \\ -\text{lateral} \\ -\text{strident} \\ \text{etc.} \end{bmatrix} \_\_\#$

Thus we see that when there are competing processes present in the language, each serves as an independent structural description, and the phonological environments where each applies are considered separately.

What does a grammar constructed in this way look like? This type of generalization is not terribly economical; given an input file with thousands of forms, there are possibly tens of thousands of rules to consider. One strategy for rule acquisition is to throw out intermediate hypotheses, keeping only the "best" rule so far. For example, among the forms listed in (27), the good phonology student would realize immediately that the rule we want to learn is *-i* suffixation, and the intermediate hypothesis that *ag* was relevant in the environment can be discarded once we recognize the larger generalization. Indeed, this is what Pinker and Prince (1988) seem to have had in mind when they sketched a similar learning strategy. However, it is not always the case that the most general rule is the best one. Linguistic analyses often make use of intermediate generalizations; for instance, the best description of the distribution of the [t] allomorph of the English past tense suffix is that it occurs after voiceless obstruents, even though the existence of forms like *learnt*, *burnt*, or *dwelt* for some speakers means that strictly speaking, in those dialects it may actually occur after any consonant. (See Albright and Hayes (1999b) for further discussion of such cases, and a procedure for detecting them.)

The answer is that we need a more sophisticated metric of the "goodness" of rules. In the minimal generalization learner, this is done by making rules prove their effectiveness in the lexicon. In particular, rules are annotated with a *reliability* score, which is defined as the ratio of the number of input forms the rule derives correctly (its *hits*) over the number of forms it could potentially apply to (its *scope*):

(33) Definition of a rule's reliability:

$$\text{Reliability} = \frac{\text{number of forms the rule correctly derives (= \textit{hits})}}{\text{number of forms included in the rule's structural description (= \textit{scope})}}$$

For example, the generalized rule in (32) attempts to add *-u* after any non-coronal stop, yielding the correct result for the words in (30) ($\surd$*ragu*, $\surd$*tipu*), but the wrong results for the first three words in (27) (**lagu*, **basagu*, **subu*). Thus, the reliability of this rule is 2/5 = .4.

Once these reliability ratios have been calculated, they are then adjusted using confidence limit statistics, as suggested by Mikheev (1997), yielding a *confidence* score for the rule.[5] The

[5] The lower confidence limit of a reliaiblity ratio is calculated as follows: first, the reliability (probability) ratio, which we may call $\hat{p}$, is adjusted to avoid zeros in the number or denominator, yielding an adjusted value $\hat{p}^*$: $\hat{p}^* = \frac{hits+0.5}{scope+1}$. This adjusted value is then used to calculate an estimate of the true variance of the sample:

$$\text{estimate of variance} = \sqrt{\frac{\hat{p}^* \times (1-\hat{p}^*)}{n}}$$

This value is then used to calculate the lower confidence limit (π_{Lower}), at a particular confidence value α:

rationale for this is that we are more certain about occurrence rates for things that have had more opportunities to occur, and less certain about things that happen more rarely. The result is that unambitious rules which cover only a few forms are penalized substantially, while more ambitious rules are barely penalized at all; for example, a rule with a reliability of 5 out of 5 is downgraded from 1 to .825, while a rule with a reliability of 1000 out of 1000 is downgraded only a minuscule amount, to .999. Various experimental results have shown that confidence values calculated in this way correlate well with human intuitions about the relative well-formedness of morphological processes in different phonological environments (Prasada and Pinker 1993; Albright 1998; Albright, Andrade, and Hayes 2001; Albright and Hayes 2002).

We see that there are therefore two factors which contribute to the goodness of rules in this system: ambitiousness in including substantial numbers of forms in the structural description, and accuracy in deriving the correct outcomes for those forms. In order to collect all of these statistics, the learner needs to keep around all of the rules it hypothesizes, at least until some critical number of forms have been learned. Thus, grammars constructed in this manner can be enormous.

Although the grammar must be learned from existing forms, the strongest argument for the existence of grammar is that we are able to derive forms of words that we have never heard before. Grammars learned by minimal generalization can be used to derive novel forms, as follows: given a novel form of category X and a (possibly quite large) sub-grammar of rules for the mapping from category X to the target category Y, the sub-grammar is searched for all rules for which the novel form meets the structural description. For example, suppose we wanted to use the sub-grammar induced by the forms in (27) to derive the absolutive of a new word *mag*. By inspecting the sub-grammar, we would find that several different rules could apply to this form, including adding *-i* after *ag*, adding *-i* after voiced stops, and adding *-i* after stops in general. Although these rules all produce the same output (*magi*), each carries its own reliability score. When many rules all point to the same output, the output is assigned the confidence score of the best available rule.

When a language has competing structural changes (such as $\varnothing \rightarrow i$ and $\varnothing \rightarrow u$), each one gets to apply in this fashion. The result is a set of competing outputs, each assigned a reliability value. It is assumed that the output with the highest reliability is the one which is selected in a forced choice situation, and is the one which is judged best in an acceptability rating task.

The implementation of a minimal generalization learner presented by Albright and Hayes (1999a) has one last feature which is relevant here: a limited capacity to discover phonological rules. In order to do this, it makes use of prior knowledge of what sequences are phonotactically illegal in the language.[6] The way it discovers a rule is this: upon learning two competing structural changes (A→B and A→C), it checks to see whether the choice of B vs. C is motivated phonotactically. For example, given the English past tense suffixation rule $\varnothing \rightarrow d$ / ___# (as in

---

$$\pi_{Lower} = \hat{p}^* - z_{(1-\alpha)/2} \times \sqrt{\frac{\hat{p}^* \times (1-\hat{p}^*)}{n}}$$

The confidence value α ranges from $.5 < \alpha < 1$, and is a parameter of the model; the higher α is, the greater the penalty for smaller generalizations. In the simulations reported here, I will always assume an α of .75. The value z for a particular confidence level α is found by consulting a standard statistics look-up table.

[6]Providing the learner with this information ahead of time may seem like a cheat. However, there is increasing evidence that infants as young as 10 months old already have an idea of what is phonotactically legal in their language (Jusczyk, Friderici, Wessels, Svenkerud, and Jusczyk 1993; Friederici and Wessels 1993; Jusczyk, Luce, and Charles-Luce 1994); for a review, see Hayes (to appear).

rub∼*rub[d]*), and the competing suffixation rule ∅ → ə*d* / ___# (as in *need*∼*needed*), the learner tries applying ∅ → *d* to *need*, yielding the output *nee*[dd]. The learner also knows that [nidd] is unpronounceable in English, since there is an inviolable ban word-final geminates. Therefore, it can posit a phonological rule that epenthesizes [ə] in this environment (∅ → ə / *d*___*d*#. (For a more detailed description of the mechanisms it uses to do this, the reader is referred to Albright and Hayes, to appear). With this phonological rule in place, the simple *d*-suffixation rule (∅ → *d* / ___#) yields the correct output *need*[əd].

Phonological rules can help to improve the reliability of morphological rules, since they allow more forms to be derived correctly. The reason this is relevant is that there are two types of alternations which may exist within paradigms: those which are phonologically motivated, and those which are not. When alternations have a purely phonological explanation, they arguably should not contribute to the measure of how difficult a morphological mapping is. Some toy languages which involve phonological neutralizations are discussed in sections 3.4.1 and 3.4.2.

It should be clear from the description thus far that the minimal generalization learner has various limitations that restrict the types of processes it can discover and the types of generalizations it can form. The rules it posits, for example, involve just one change (A → B), so it would have difficulty learning morphological processes that involve multiple simultaneous changes (such as circumfixing, and so on). The rule format also requires that the context for a rule be strictly adjacent to the change (C___D, not C X___D), so it would not be able to discover non-local conditioning of a morphological alternation. Most of these limitations are irrelevant to the cases discussed here, and can safely be ignored. There is one limitation that is potentially relevant, however, and that is the fact that its input representations contain only strings of segments, with no place to store additional information such as prosodic information, grammatical gender, or other morphosyntactic distinctions. There is abundant evidence that such factors do play a role in conditioning morphological alternations, and it would be impossible to describe some languages (such as Latin) without taking them into account. Although the present version of the learner has no ability to calculate prosodic features of words or infer that it would be useful to consider factors like grammatical gender, it is nonetheless possible to provide it with these features and allow it to form generalizations with them. Since the input to the learner is strictly a string of segments, this is done (as a hack) by including in each input pair a "segment" that encodes whatever prosodic or morphosyntactic information the analyst wishes to provide to the learner (e.g., '1' for masculine, '2' for feminine, '3' for neuter, etc.). The learner can then generalize over these segments in the same way that it generalizes over phonological segments in the context (e.g., change X occurs only in masculine nouns, change Y only in feminine nouns, change Z in both, and so on). This strategy was employed in the Latin simulations described in chapter 4.

From this presentation of the minimal generalization approach, we can see that the model meets all of the criteria laid out in the previous section: it is specifically designed to learn grammars to project between ordered pairs of forms (present→past, 1sg→infinitive, etc.), so it is perfectly capable of exploring each of these sub-grammars independently. The grammars it learns can include many, competing changes (A→B, A→C, etc.), so it can potentially produce multiple outputs for each input. Each output is associated with the rule that derives it, so it is easy to tell whether an output has been derived by a generalized rule from the grammar, and when it has been derived using a word-specific rule (which could be seen as analogous to resorting to memorization). The confidence values of rules allow the model to assign predicted

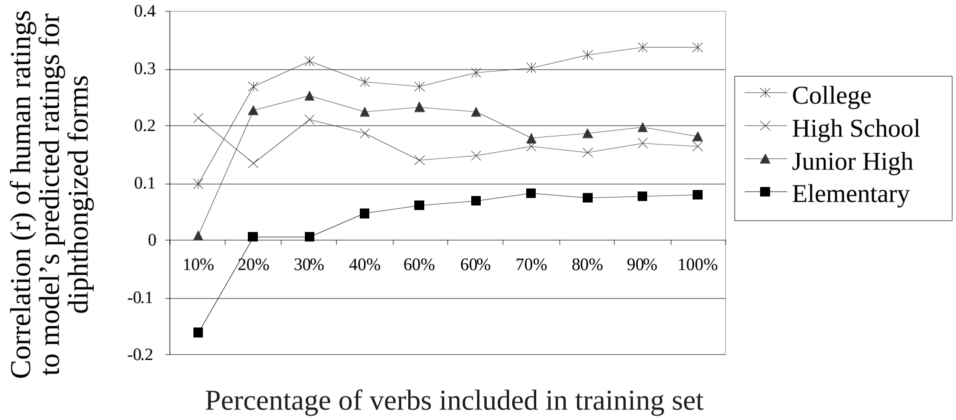


Figure 3.2: Time course of acquisition of the minimal generalization model (from Albright, Andrade, and Hayes 2001, p. 140)

well-formedness scores to its outputs. Finally, the model is incremental, and at every stage it has a working grammar which covers at least as many forms as it has heard. Furthermore, work on modeling various languages has shown that grammars learned by minimal generalization are quite "stable" throughout the training period—that is, they arrive at a basic analysis quickly, and the predictions for novel forms do not change much after that. For example, Albright, Andrade, and Hayes (2001) found in modeling the intuitions of adult Spanish speakers with varying degrees of education that adding more words does not make the predictions significantly better or worse, once 300 out of a training set of 1,698 verbs had been learned (in order of decreasing frequency). This result is reproduced in Figure 3.2.

This finding is also relevant for the task of base identification, because the goal here is to use an initial subset of the training data to make a decision that will reduce the amount of work needed to learn the remainder of the data. In principle, one might imagine that the initial batch of training data could be misleading, and the learner could be tricked into picking a suboptimal base because the initial batch of words was not representative. This is especially plausible because we know that irregular forms tend to be among the most frequent forms in the language (indeed, that is how their irregular status is preserved), and therefore learners would be exposed to disproportionately many irregulars. It turns out, however, that this fact makes the time course of learning *more* stable, not less. The reason is that once the learner has learned enough forms to identify the regular pattern as the predominant, productive one, the rest of the training set will continue to contain more and more regulars. In other words, as long as we do not restrict the learner to just the very top, most irregular batch of extremely frequent words, it is unlikely that it will be misled by a subset of the training data.

3.3 Metrics for sub-grammars

The focus of this thesis is not on *how* to learn morphological mappings, but rather on *which* morphological mappings to learn. The hypothesis is that we are searching for the sub-grammars which are *easier*, or more effective. In this section, I will consider a variety of metrics for measuring the complexity of a sub-grammar.

Given the statistics which the minimal generalization learner collects about its rules, there are a variety of possible ways to measure the complexity of a sub-grammar, enumerated below.

Accuracy

A very common way to test the performance of machine-learned grammars is by testing it on its ability to reproduce the training set. In this case, given that the model also has word-specific rules at its disposal, we are really interested in whether the generalized (i.e., not word-specific) rules are able to reproduce the training set correctly without resorting to the word-specific (memorized) forms. The accuracy of a sub-grammar can be calculated as follows: first, use the subgrammar to derive outputs for all of the real words using only the generalized rules, and take the highest-scoring output for each word. Then, compare these outputs with the actual outputs, to see if the correct output can be synthesized by the subgrammar. The percentage of correct outputs is the *accuracy* of the sub-grammar.[7]

Mean confidence of rules in the grammar

Grammatical rules in the minimal generalization model have confidence scores attached to them, as described above. When the learner is able to find generalizations that include more forms, and have fewer exceptions, then the rules in the grammar will have higher confidence values. Therefore, we can use the *mean confidence* of the generalized rules in a sub-grammar as a metric of its complexity.

---

[7] This is similar in spirit to the Minimum Description Length (MDL) approach of comparing grammars (Rissanen 1986; Rissanen 1989; Rissanen and Ristad 1994), but with some crucial differences: The principle behind the MDL approach is to reduce the amount of information that must be memorized by formulating rules that can derive as much of the surface data as accurately as possible. The MDL principle is generally used to compare various possible grammars of the same data; in this case, the grammars have been constructed in exactly the same way, and we are comparing how susceptible different data was to grammatical description, when the grammatical construction procedure is held constant. I am not aware of other attempts to use MDL to compare the complexity of data instead of grammars, but it seems like a useful tool for this task because it explicitly takes into account the sizes of both the grammar and the data. It should be noted that there are some obstacles to using the MDL approach for the current problem, however. First, it is necessary to find a way of calculating the cost of a grammar. Cost is usually thought of as the size of the grammar, but this in itself can be difficult to quantify. (Also, as discussed above, there is a certain amount of evidence that people do not actually use what we might think are the most economical or efficient grammars possible, at least in terms of detailed information about particular environments). In addition, the cost of an MDL grammar includes a penalty for storing whole forms, so the only forms that are stored are irregulars that could not be synthesized by any more economical means. Limiting storage to irregulars is certainly a traditional idea in theoretical linguistics, and has been the focus of a good deal of research in the dual mechanism approach—see, e.g., Pinker 1999, Clahsen 1999. However, various results in recent years have shown that regulars, too, may sometimes be stored (Stemberger and MacWhinney 1986; Sereno and Jongman 1997; Baayen, Dijkstra, and Schreuder 1997; Eddington and Lestrade 2002). Pinker and Prince themselves admit that regulars cannot be excluded categorically from the lexicon, even if the same forms could also have been synthesized productively using the grammar (Pinker and Prince 1994, p. 331). Therefore, I merely mention the parallel with MDL in passing, but I am not convinced that adopting one of the current MDL approaches to morphology would provide an accurate model of human learning.

Mean confidence of winning outputs

Another intuitive idea is that grammars are better when they produce their results unambiguously and confidently. One way to evaluate this is to use the generalized rules to wug-test the sub-grammar on all of the forms known so far, as above, and then collect the winning outputs and calculate their *mean confidence* scores. (Recall that the confidence score of an output is equal to the confidence score of the best rule in the grammar that could derive that output.) A higher mean confidence means that the outputs are supported by more data, and more consistent data.

Average winning margin

In addition to being confident about outputs, we would like to be sure that there are not other, almost as good outputs that we have to agonize over. Therefore, we may be interested in the *average winning margin* of outputs, as a measure of the degree of angst incurred when using a sub-grammar. As before, we can wug-test the grammar on the known words, collecting the outputs with the highest confidence. Then, for each winning output, we can compare its confidence with the confidence of the next best distinct output—this is the margin by which the best output won. We can then average these margins, to get an estimate of how often the sub-grammar is forcing us to agonize over close competitions between two outputs; a low average margin means that the grammar is often deciding between close competitors.

Summary

These are just a few of the many metrics that could be used to measure the complexity of a sub-grammar, but they should provide a starting point of testing the approach. With these metrics in hand, the next logical step is to see how they work on a few cases. I will begin with a few synthetic languages modeled on typical patterns found in introductory problem sets, and then consider two more realistic cases.

3.4 Results for synthetic languages

3.4.1 Synthetic language 1: Neutralization in suffixed forms

It is always a good idea to start testing models on very small and schematic cases. I will keep the first few cases simple by tackling only the problem of comparing two directions along the same mapping (i.e., *form1* → *form2* vs. *form2* → *form1*).

The first case to consider is one in which there is phonological neutralization in a suffixed form.[8] For example, consider a language with palatalization of *k* → *t͡ʃ* before *i*, as in (34). In such a language, underlying *k* and *t͡ʃ* are neutralized in this context— a pattern which is found in Fe'fe'-Bamileke (Hyman 1975, p. 70) and many other languages.

[8] I will try to avoid using the term *derived form*, since it presupposes that the suffixed form is built on the unsuffixed form, and the idea here is to diagnose when precisely the opposite is true.

(34) Neutralize I: Palatalization in suffixed forms

a. Stems ending in segments other than *k*

absolutive		ergative
dap	∼	*dapi*
lot	∼	*loti*
gub	∼	*gubi*
sat͡ʃ	∼	*sat͡ʃi*
rut͡ʃ	∼	*rut͡ʃi*
lag	∼	*lagi*
ban	∼	*bani*
yul	∼	*yuli*

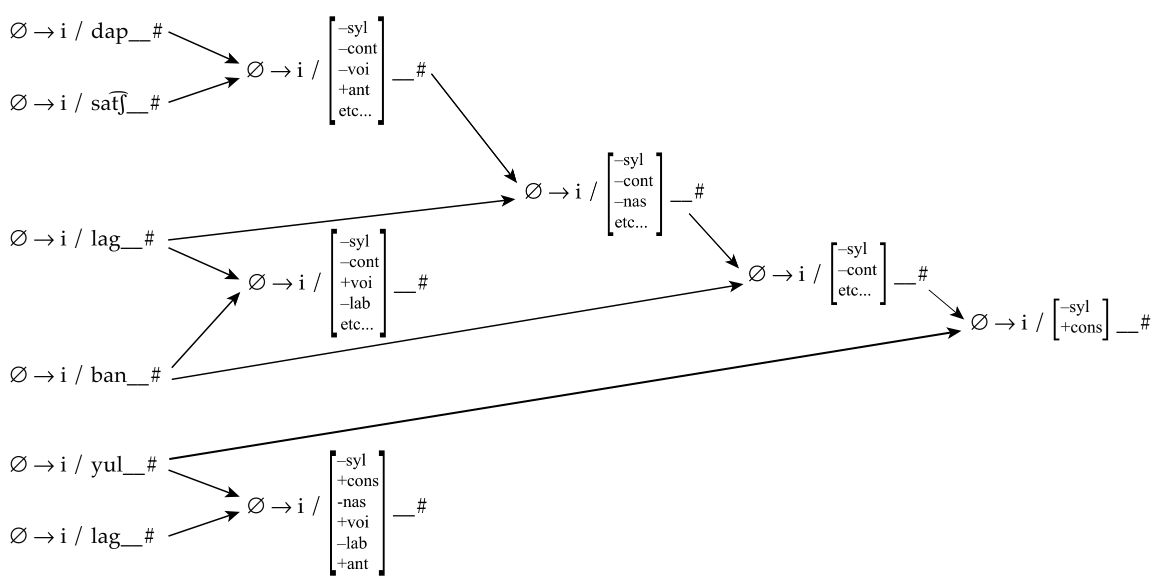
b. Stems ending in *k*

absolutive		ergative
ʔak	∼	*ʔat͡ʃi*
muk	∼	*mut͡ʃi*
lok	∼	*lot͡ʃi*

In this case, the correct answer to the problem set is that the base, or UR, is the unsuffixed (absolutive) form, since this is the form which contains the most distinctive information about the word (in particular, the /k/ vs. /t͡ʃ/ distinction). Therefore, we hope that the metrics suggested above are able to choose the absolutive→ergative mapping as the better one.

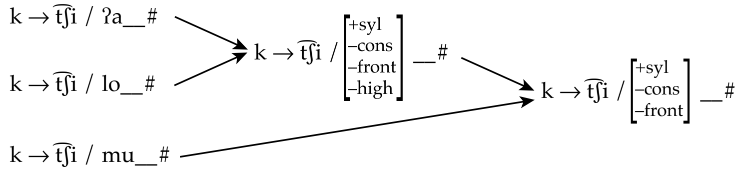
It is instructive to outline how the minimal generalization learner will handle this language, which involves a strictly phonological alternation. Going in the ergative to absolutive direction, it generalizes over the forms in (34a) to find the rule ∅ →*i* after consonants, as shown in (35). (The full set of generalizations, along with the details of the calculation of the metrics for this synthetic language, are given in appendix A).

(35) Minimal generalization over the forms in (34a):



The *k*-final roots in (34b) are somewhat more complicated: the first hypothesis which the learner explores is that there is a substitutive morphological process (*k*→*t͡ʃi*):

(36) Minimal generalization over the forms in (34b):



The learner does not stop with this parochial generalizations, however; it also employs its limited ability to discover phonological rules (p. 41). The generalizations in (35) produced a ∅ → *i* rule. Although this rule by itself cannot produce the correct output for words like *ʔak* or *muk*, the learner tries to discover phonological rules to convert the incorrect outputs (**ʔaki*, **muki*) into the correct ones (*ʔat͡ʃi*, *mut͡ʃi*). It first scans **ʔaki* and **muki*, and notes that they contains an illegal *ki* sequence (furnished ahead of time, as described above). It then compares **ʔaki* with *ʔat͡ʃi*, and posits a phonological rule that fixes the illegal sequence to make it match the observed outcome: *k* → *t͡ʃ* / ___*i*. This phonological rule is sufficient to allow simple -*i* suffixation to derive the forms in (34b) correctly as well. Therefore, the suffixation process applies with 100% reliability in the absolutive→ergative direction.

The grammar for the ergative→absolutive direction is not as clean. For the forms in (34a), minimal generalization results in a rule stripping off -*i* after a consonant. As before, the forms in (34b) are considered first as substitutive morphology (*t͡ʃi*→*k*). However in this case, there is no surface-true phonological rule which can allow *ʔat͡ʃi*→*ʔak* to be viewed as simply removing the final -*i*. Therefore, each of the forms in (34b) must be counted as an exception to the *i* → ∅ rule, and must be handled by a competing (substitutive) process. Thus, we can see that the reliability of rules in this direction will be lower, reflecting the unpredictability of "de-palatalization".

In fact, the absolutive → ergative mapping comes out better than the ergative → absolutive mapping for *all* of the metrics:

(37) Comparison of metrics for the Neutralize I language

	abs→erg	*erg→abs*
avg. confidence of generalized rules	.762	.585
avg. confidence of best output	.916	.743
accuracy (=ratio correct)	100%	73%
avg. winning margin	1.000	.552

The metric which turns out to be clearest in this example (and will be for all of the synthetic languages discussed here) is the average winning margin, which is designed to measure how much uncertainty is incurred by using the grammar. However, all of the metrics agree on the choice of base, and in general they turn out to be highly correlated in all of the cases that have been examined thus far.

3.4.2 Synthetic language 2: Neutralization in suffixless forms

The next step is to show that the same principles can be used to identify a suffixed form as the base (or UR), when there is a neutralization which affects just the unsuffixed form. This is often caused by final devoicing, or other processes that affect word-final segments.

A synthetic language with this property is given in (38):

(38) Final devoicing in suffixless forms

a. Underlyingly voiceless C's

absolutive		ergative
tak	∼	*taki*
dap	∼	*dapi*
yup	∼	*yupi*
lot	∼	*loti*
mut	∼	*muti*

b. Underlyingly voiced C's

absolutive		ergative
lak	∼	*lagi*
bak	∼	*bagi*
sak	∼	*sagi*
gup	∼	*gubi*
trep	∼	*trebi*
flap	∼	*flabi*
fut	∼	*fudi*
sat	∼	*sadi*

For this language, the correct problem set answer is that the UR is the absolutive form minus the *-i* suffix. In a model of morphology which employs only surface form-to-surface form mappings, the correct answer is that the absolutive is formed from the ergative by removing the *-i* suffix and applying final devoicing; it could not be done in the opposite direction, because we would have to guess for each word what the voicing of the consonant must be.

It should not be difficult to see why the same factors that made the ergative → absolutive mapping unreliable in the language with intervocalic voicing should make the absolutive → ergative mapping unreliable in this language. Unsurprisingly, the metrics come out exactly reversed, favoring the ergative as the base:

(39) Comparison of metrics for the Neutralize II language

	abs→erg	*erg→abs*
avg. confidence of generalized rules	.382	.659
avg. confidence of best output	.450	.928
accuracy (=ratio correct)	62%	100%
avg. winning margin	.136	.928

3.4.3 Synthetic language 3: Morphological neutralization

The first two synthetic languages have shown the role that phonological rules can play in making the grammar for a mapping "cleaner" in one direction than in the other. In these cases, the choice of a base for the morphological rule is essentially the same problem as the choice of the UR for phonological purposes.

In many cases, however, there is no phonological reason why one form should be considered the morphological base. One situation where this could happen is when there is "morphological neutralization"—two or more morphological classes share the same marking in a particular inflection. In the following example, the two classes of nouns (which are labelled *o*-stems and *a*-stems in honor of the vocalism in their absolutive forms) have the same lack of marking in their absolutive forms. (I will assume for the purposes of this discussion that an apocope analysis is ruled out by the existence of other words in the language with final unstressed vowels.)

(40) Neutralize III: morphological neutralization

a. *o*-stems

absolutive		ergative
lag	∼	*lagos*
ʔak	∼	*akos*
gub	∼	*gubos*
dap	∼	*dapos*
fud	∼	*fudos*
lot	∼	*lotos*
ran	∼	*ranos*

b. *a*-stems

absolutive		ergative
sag	∼	*sagas*
bak	∼	*bakas*
reb	∼	*rebas*
yup	∼	*yupas*
sad	∼	*sadas*
mut	∼	*mutas*

In this case, the correct "textbook' answer is not so clear. Since there is no phonological process involved, it is perhaps better to speak of the *stem* (or *root*) rather than the UR. One analysis would be to say that the stem is the same as the absolutive form, and there are two classes of nouns, distinguished by their endings (*-os* vs. *-as*) in the ergative. An alternate account would be to say that the stem is the ergative form minus the *-s* (*lago-*, *ʔako-*, *saga-*, etc.) and the case forms are the result of dropping the theme vowel or adding an *-s*, respectively. (Both approaches have been common in the analysis of noun and verb classes in Indo-European languages, among other places.)

From the point of view of having to map one form from the other (rather than positing an abstract stem), we see that the absolutive → ergative mapping forces us to provide an (unknown) vowel, whereas the ergative → absolutive direction merely involves taking off a well-defined amount of phonological material. Notice that since there is no phonological motivation for the alternation between *-as* and *-os*, the two processes cannot be unified by any overarching rule; the most general rules in either direction will be to add/remove *-as* and to add/remove *-os*.[9]

Even without phonology, we see that almost all of the metrics considered here agree that the ("backwards") mapping from ergative to absolutive is the better one:

[9]This points out a limitation of the syntax of morphological rules used by the minimal generalization learner; if we could use features in the structural change as well as in the environment, then the ergative → absolutive direction could be captured by a single rule peeling off a vowel plus *s*. This could turn out to be important in some especially close case, but is not necessary here.

(41) Comparison of metrics for the Neutralize III language

	abs→erg	*erg→abs*
avg. confidence of generalized rules	.353	.673
avg. confidence of best output	.439	.862
accuracy (=ratio correct)	54%	100%
avg. winning margin	.071	.862

3.4.4 Synthetic language 4: Lexical exceptions

One last factor which could lead one mapping to be better than another is the existence of lexical exceptions in a particular part of the paradigm. There are, of course, many different patterns of exceptions, but what I will consider here is the case where there is one general (“regular”) process, and a few isolated words which take a different (“irregular”) pattern.

A language with this property is given in (42):

(42) Exceptions I: a few lexical exceptions with a completely different marking

a. “Regulars” (suffix -*os*)

absolutive		ergative
lag	∼	*lagos*
ʔak	∼	*akos*
gub	∼	*gubos*
dap	∼	*dapos*
fud	∼	*fudos*
lot	∼	*lotos*
sag	∼	*sagos*
bak	∼	*bakos*
reb	∼	*rebos*
yup	∼	*yupos*

b. Lexical exceptions

absolutive		ergative
sad	∼	*sed*
mut	∼	*myt*

In this language, the generalized rule of adding -*os* after stops works for almost all of the forms (10/12), but the other two forms must be handled by separate ablaut rules (*a*→*e* and *u*→*y*). The minimal generalization learner was in fact designed to learn general patterns (like -*os* suffixation) in the face of lexical exceptions, so the focus until now has always been on this direction. However, it is interesting to see whether this is the direction which would in fact be picked as the optimal mapping according to the criteria employed here.

Looking through the forms in (42), we can see that it would actually be easier to guess the absolutive given the ergative than the other way around. Any absolutive form with an *a* in it should be eligible for ablaut (e.g., *bak*→**bek*). An ergative form without -*os*, on the other hand, could only be lacking -*os* because it is irregular, and an ergative form with -*os* is always regular.[10]

[10]This is *not* always the case in real languages—for example, English and German have “mixed” irregulars like *keep*∼*kept* or *brennen*∼*brannten*.

The comparison metrics reflect this, indicating that the absolutive would be the better base for this language:[11]

(43) Comparison of metrics for the Exceptions I language

	erg→abs	*abs→erg*
avg. confidence of generalized rules	.631	.738
avg. confidence of best output	.834	.908
accuracy (=ratio correct)	83%	83%
avg. winning margin	.834	.908

This leads to the rather counter-intuitive result that the best base in a language with lexical exceptions may often turn out to be the form in which the exceptionality is clearest. For example, in the case of English, it is much easier to tell if a verb has an irregular past tense by looking directly at the past than by looking at the present: if the verb ends in anything other than -*t*, -*d*, or -*əd*, it is guaranteed to be irregular (*sung*, *rang*). Furthermore, it is guaranteed to be irregular it ends in -*t* after a vowel or sonorant (*beat*, *wrote*, *caught*, *dwelt*), or -*d* after a lax vowel (*shed*). The only really ambiguous cases are those like *kept* or *rode* which have the voicing-appropriate dental suffix and would be phonologically legal without the final *t* or *d*. However, such cases constitute only a minority of the irregular past tense forms of English, so the problem would be reduced greatly by choosing the past tense as the base. I believe that the reason why English speakers do not seem to do this has to do with the enormous frequency difference between the present and the past tense. I will return to this question in more detail in section 6.2.1, addressing the question of why learners might sometimes be driven to use mappings with many exceptions, even if they could have been avoided using another mapping.

There are, of course, many other patterns of neutralizations and exceptions in languages, many of them more difficult than those exemplified by the synthetic languages in this section. In many cases, neutralizations and exceptions are sprinkled throughout the entire paradigm, so it can be difficult to identify a single slot in the paradigm as the clear best choice, as we could in the synthetic languages. Often, it is necessary to compare the various neutralizations and exceptions, considering also how many lexical items are affected by each, in order to decide which are the least serious. In the next chapter, I will consider a more complicated and realistic example, from Latin noun paradigms. I will show that although neutralizations affected every part of the paradigm, these neutralizations were more serious in the nominative than in any other form. The model, predicts, therefore, that a non-nominative form should be chosen in base, and subsequent historical changes show that this prediction seems to be true.

[11]The exact computation of the average winning margin depends on the treatment of word-specific rules, since this language requires that we sometimes derive forms using these rules. Intuitively, we might think that our confidence in a word-specific rule should be 1, since that rule does a single job and it does it unambiguously. However, the adjustment using confidence limits which is employed here does not allow this, since the confidence limit for an event which occurs just once is undefined. In calculating the average winning margin for this example, I have simply excluded the two cases where the word-specific rule was used, and averaged the margin for the remaining ten cases.

Chapter 4

The Latin *honor* analogy

In older stages of Latin, many nouns exhibited paradigmatic *s* ∼ *r* alternations created by rhotacism in suffixed forms, as in (44a). In the period immediately before Classical Latin, these *s* ∼ *r* alternations were eliminated by extending the *r* to the nominative form. At approximately the same time, an independent change shortened long vowels before word-final sonorants, resulting in the paradigm in (44b):[1]

(44) Elimination of *s* ∼ *r* alternations in Pre-classical Latin

a. *Pre-leveling*

	sg.	*pl.*
nom.	honoː**s**	honoː**re**ːs
gen.	honoː**ri**s	honoː**r**um
dat.	honoː**ri**ː	honoː**ri**bus
acc.	honoː**re**m	honoː**re**ːs
abl.	honoː**re**	honoː**ri**bus

b. *Post-leveling*

	sg.	*pl.*
nom.	hono**r**	honoː**re**ːs
gen.	honoː**ri**s	honoː**r**um
dat.	honoː**ri**ː	honoː**ri**bus
acc.	honoː**re**m	honoː**re**ːs
abl.	honoː**re**	honoː**ri**bus

As mentioned in chapter 1, the switch from [honoːs] to [honor] has attracted the attention of numerous linguists over the years (Hoenigswald 1960, pp. 107-111; Hooper 1976, 95-96;

[1]There is some evidence for an intermediate stage in the late Pre-Classical stage in which [honoːr] with long [oː] was at least one possible variant of the nominative form. Early poets such as Plautus and Ennius still allowed final long [o:r], using it to satisfy metrical requirements for heavy syllables, and also sometimes used [-o(ː)r] instead of [-oːs], e.g., Ennius 545 (Skutsch 1985) *Clā́m[ō:]r ād cā́elūm vṓlvēndŭs pĕr ā́ethĕră vā́gĭt*; also Ennius 409, 428. As far as I know, there is no evidence concerning how "clean" this intermediate stage was – the shortened variant [honor] may also have been used from the very beginning of the [s] > [r] change. The analysis that I propose here is compatible with the existence of an intermediate [honoːr] stage, but does not rely on it. It would, however, make a crucial difference for some other analyses, such as the Uniform Exponence analysis sketched below.

Wetzels 1984; Hock 1991, pp. 179-190; Barr 1994; Kiparsky 1997; Kenstowicz 2002; Hale, Kissock and Reiss 1998; and others), because the direction of leveling runs counter to so many other changes. Why was the nominative singular rebuilt on the basis of oblique forms in Latin,[2] while in so many other languages (including the Polish diminutives example in chapter 1), entire paradigms have been rebuilt on the basis of nominative singular forms?

There have been numerous attempts to explain why Latin speakers might have rebuilt the nominative singular instead of the oblique forms; indeed the problem has never been in coming up with hypotheses about what would have made [honor] an appealing form. Looking at the paradigms in (44), various possibilities come to mind: perhaps it was the predominance of forms in [r], or perhaps one of the oblique forms (such as the accusative) was more common than the nominative singular, or perhaps it was the existence of many other words with *-or* nominatives. The real challenge is to explain why one of these factors was able to cause a leveling of the nominative in Latin, but not in Polish or other cases. In this chapter, I will propose that what distinguishes Latin from many other cases is not some difference in the frequency or markedness of forms, but rather in the form that serves as the base. Employing the base identification algorithm from the previous chapter, I will show that neutralizations and exceptions make the nominative singular an uninformative form in Latin, and favor the choice of an oblique form as base (it does not matter which one). This explains how learners could have set up a different base in Latin than in other languages (a problem noted by, among others, Wetzels 1984, pp. 582-583), and helps to account for the unusual direction of the analogical change. Then, in section 4.3, I will show that when an oblique form is chosen as base, the grammar that the model learns makes the correct predictions for the *honor* analogy. Finally, in sections 4.4.1-4.4.3, I will show that this analysis provides a more accurate explanation for the details of the change—in particular, by differentiating words that changed from those that did not—than many of the factors that have previously been proposed.

4.1 Statement of the problem

As Hock points out, the *honor* analogy can be described in at least two different ways. The first is as a *four-part analogy*, in which [hono:s] was influenced by words that already had an [r] in the nominative, such as [soror] 'sister':

(45) Extension of *r* by four-part analogy:
[soro:ris] : [soror] :: [hono:ris] : X (X = [honor])

Although the four-part analogy notation expresses the change as the influence of one particular lexical item (in this case, [soror]), it is generally recognized that such changes are actually due to the collective influence of many words, such as [soror], [kruor] 'blood', and also the numerous agentive nouns ending in *-or* ([o:ra:tor] 'speaker', [gladia:tor] 'gladiator', etc.). However, even when we recognize that the four-part analogy notation is just a shorthand for the influence of a larger pattern, it is far from an explanation of the change. How many words does it take for speakers to construct such an analogy? How similar do they have to be? More importantly, it has often been noted that the four-part analogy notation cannot tell us why

[2] I will use the term *oblique* to denote all cases other than the accusative, which in the case of Latin, includes genitive, dative, accusative, ablative, and dative.

the influence was not from [honoːs] to [soror] ([honoːris]:[honoːs] :: [soroːris]:X, X=[soroːs]) (Hock 1991; Barr 1994; Kiparsky 1998), or why the change was not in the opposite direction, undoing rhotacism (something like [soror]:[soroːris] :: [honoːs]:X, X = [honoːsis]). This last question is especially puzzling, because the actual change in (45) violates the tendency for analogical change to extend from more "basic" or unaffixed forms to less basic, or affixed forms (Kuryłowicz 1947).

The four-part analogy notation can equally well capture the leveling of alternations ([honoːs] > [honor]) or the extension of alternations ([soror] > [soroːs]). An alternative description of the spread of [r] to the nominative in [honor] is as paradigm leveling, with the nominative form changing to match the remainder of the paradigm ([honoː**r**is], [honoː**r**iː], [honoː**r**em], etc.). The pressure to level paradigms has been formalized in Optimality Theory (OT, Prince and Smolensky 1993) as constraints on paradigm uniformity or uniform exponence (Kenstowicz 1997b; Kenstowicz 2002; Steriade 2000); for example, Kenstowicz (2002) schematizes the [honoːs] > [honor] change as the promotion of a constraint demanding uniformity in noun paradigms (UE). In the first stage, shown in the tableau in (46), the ban on intervocalic [s] (*VsV) outranks Faithfulness for /s/ (Faith(*s*)), yielding rhotacism in suffixed forms. Faith(*s*) in turn outranks UE, meaning that rhotacism does not overapply in the nominative form. The result is a paradigm with *s* ∼ *r* alternations, as in (46a). Note that candidate (c) contains a vowel length alternation ([or] ∼ [oːris] – Kiparsky 1997; Hale et al. 1998; Baldi 1999, p. 323), for which Kenstowicz does not record a UE violation. It is entirely possible that there are separate UE constraints for different alternations, with UE for *s* ∼ *r* alternations ranked higher than UE for vowel length alternations. If that is the case, then UE(*s* ∼ *r*) is the relevant constraint here, and this what I will assume.

(46) Stage 1: *VsV ≫ Faith(*s*) ≫ UE(*s* ∼ *r*)

/honoːs/ (nom.), /honoːs-is/ (gen.), /honoːs-em/ (acc.), …	*VsV	Faith(*s*)	UE(*s* ∼ *r*)
a. ☞ [honoːs], [honoːr-is], [honoːr-em], …	√	**	*
b. [honoːs], [honoːs-is], [honoːs-em], …	*(gen.)!*(acc.)	√	√
c. [honor], [honoːr-is], [honoːr-em], …	√	***!	√ (?)

Under this analysis, the change consists of promoting UE over Faith(*s*), so that rhotacism overapplies in the nominative, as in (47). This leaves two possible candidates for the nominative: [honoːr] (candidate (c)) and [honor] ((d) and (e)). The first of these violates a high-ranking phonotactic constraint against final [oːr] in Latin, favoring a paradigm with shortening in the nominative ((d) or (e)); of these, the paradigm that preserves long [oː] in the oblique forms is more faithful to the long [oː] of the input (Faith(Vː)), and thus candidate (d) is selected.

(47) Stage 2: *VsV, UE(s∼r) ≫ Faith(s), Faith(Vː)

/honoːs/ (nom.), /honoːs-is/ (gen.), /honoːs-em/ (acc.)		*oːr#	*VsV	UE(s∼r)	Faith(s)	Faith(Vː)
a.	[honoːs], [honoːr-is], [honoːr-em]	√	√	*(nom.)!	**	√
b.	[honoːs], [honoːs-is], [honoːs-em]	√	*(gen.)! *(acc.)	√	√	√
c.	[honoːr], [honoːr-is], [honoːr-em]	*!	√	√	***	√
d. ☞	[honor], [honoːr-is], [honoːr-em]	√	√	√	***	*
e.	[honor], [honor-is], [honor-em]	√	√	√	***	**!*

This analysis avoids many of the problems pointed out by Hale et al. (1998) by limiting UE to $s \sim r$ alternations. An alternative, however, is that the Uniform Exponence analysis actually captures the creation of the intermediate variant [honoːr] (see fn. 1), with perfect UE satisfaction, and that the underlying form of this word had already been reanalyzed as /honoːr/ (as Hale et al. claim) by the time final vowel shortening occurred, so Uniform Exponence was irrelevant by that stage.

The uniform exponence account formalizes the intuition that paradigm leveling is due to a pressure for nonalternating paradigms, and that the resulting paradigm is one which (in this case) satisfies both paradigmatic constraints and also general phonotactic constraints of the language (such as *VsV, and *oːr#). However, it leaves many details unaccounted for. Why, for example, was the old $s \sim r$ alternation suddenly intolerable, at the same time that a new $o \sim o$ː alternation was being created? We might have expected the UE constraint to move above both the *oːr# and *VsV constraints, since there was no crucial ranking between them before the change. This should have yielded a uniform paradigm with overapplication of both rhotacism and shortening: [honor], [honoris], [honorem], etc. Furthermore, we might expect an increased drive for paradigm uniformity to level other alternations in noun paradigms, but in fact these remained by and large intact:

(48) Alternations preserved in Latin noun paradigms

gloss	‘honor’	‘city’	‘art’
alternation	[o] ∼ [oː]	[p] ∼ [b]	∅ ∼ [t]
nom.	[honor]	[ur**p**s]	[ars]
gen.	[honoːris]	[ur**b**is] (*[ur**p**is])	[ar**t**is] (*[aris])
dat.	[honoːriː]	[ur**b**iː] (*[ur**p**iː])	[ar**t**iː] (*[ariː])

Another unexplained mystery is why UE was promoted over Faith(*s*) to extend rhotacism, rather than Faith(*s*) being promoted over *VsV to eliminate rhotacism. As mentioned above, the rhotacism constraint (*VsV) plays a crucial role in Kenstowicz's analysis, ensuring that the resulting paradigm will have uniform [r] and not uniform [s]. However, we could just as easily have used the same analytical device (historical reranking of two constraints) to predict the opposite change, promoting the IO Faith(*s*) constraint to yield uniform [s]. Thus, the mere existence of an active rhotacism constraint at one stage in the grammar is not sufficient to explain why it should continue to be true at the next stage in the grammar. A proposal by McCarthy (1998) that output-to-output constraints start at the top of the grammar in the initial state might help to explain the tendency for uniform exponence to move up in grammars over time. If this is correct, then we need only assume that learners sometimes fail to demote UE below the relevant markedness constraints – but why would one generation of Latin learners suddenly fail to apprehend the correct ranking of Faith(*s*) $\gg$ UE? If uniform exponence is to have any explanatory force in accounting for paradigm leveling, we would ideally like to be able to predict when such UE $\ll\!=\!\gg$ Faithfulness flips are likely to occur, and which faithfulness constraints will be demoted.[3]

A uniform exponence analysis also fails to capture various other details about the [hono:s] > [honor] change. Most notably, the spread of [r] to the nominative form was complicated by the fact that it was restricted primarily to polysyllabic, non-neuter nouns such as [hono:s] 'honor' (masc.) and [arbo:s] 'tree' (fem.), shown in (49a) (Hock 1991; Barr 1994; Kiparsky 1998). Monosyllabic nouns, such as [flo:s] 'flower' (masc.) were not affected (49b), nor were polysyllabic neuter nouns, such as [korpus] 'body' (neut.) (49c).

(49) Leveling restricted to masc. and fem. polysyllables

a.	[hono:s]	$\Rightarrow$	[honor]	'honor' (*masc.*)	
	[arbo:s]	$\Rightarrow$	[arbor]	'tree' (*fem.*)	
	[odo:s]	$\Rightarrow$	[odor]	'odor' (*masc.*)	
	(augus)	$\Rightarrow$	[augur]	'omen' (*masc.*)	
b.	[flo:s]	$\Rightarrow$	[flo:s]	'flower' (*masc.*)	($\not\Rightarrow$ [flor])
	[o:s]	$\Rightarrow$	[o:s]	'mouth' (*neut.*)	($\not\Rightarrow$ [or])
	[mo:s]	$\Rightarrow$	[mo:s]	'custom' (*masc.*)	($\not\Rightarrow$ [mor])
c.	[korpus]	$\Rightarrow$	[korpus]	'body' (*neut.*)	($\not\Rightarrow$ [korpor], [korpur])
	[tempus]	$\Rightarrow$	[tempus]	'time' (*neut.*)	($\not\Rightarrow$ [tempor], [tempur])
	[onus]	$\Rightarrow$	[onus]	'burden' (*neut.*)	($\not\Rightarrow$ [oner], [onur][4])

Furthermore, the new [r] forms appear to have replaced the older [s] forms relatively slowly; in the historical period we find both [hono:s] and [honor], [odo:s] and [odor], [arbo:s] and

[3] Dresher (2000, p. 60) makes this same point, and likewise concludes that UE-type constraints do not provide a adequate explanation for paradigm leveling.

[4] We might expect an [e] on the basis of the oblique forms of this word, which have [oner-]. A [u] might also be possible, on analogy with words like [femur] 'thigh'.

[arbor], [lepo:s] and [lepor] 'charm', [labo:s] and [labor] 'labor' (Leumann 1977, p.179).[5] Both of these problems could be handled by various means – UE could be restricted to non-neuter polysyllables in some way, for example, and the free variation could be accomplished by the gradual promotion of stochastically ranked constraints (Boersma and Hayes 2001). As above, the real problem is not in finding theoretical machinery that can describe the change; it is in understanding why the change should have occurred in this direction, to these particular words, and why Latin behaved differently from so many other languages.

In sum, there are a variety of questions that must be answered if our understanding of the Latin *honor* analogy is to move beyond description to actual explanation:

(50) a. Why was a basic, "unmarked" isolation form (the nominative) rebuilt on the basis of more marked suffixed forms, contrary to the usual direction of analogical change? (Lahiri and Dresher 1984; Bybee 1985, chap. 3)

b. What role (if any) did similar words, like [soror], play in the change? Is there a minimum number of such words that are necessary to effect such a change?

c. Why did [hono:s] change to [honor], and not [soror] to [soro:s]?[6]

d. Why were monosyllables and neuters generally not affected?

e. Why might both [o:s] and [or] variants have persist for so long?

In this chapter, I propose that the change of [hono:s] to [honor] was driven by more than just a phonological change involving paradigm uniformity constraints. I suggest that it was actually the result of the way that the morphological grammar of Latin noun paradigms was organized, with an oblique form serving as the base, and rules deriving the remaining forms (including the nominative singular). When the base identification of the previous chapter is applied to Latin, it emerges that the preferred base is an oblique form, not the nominative form. Furthermore, when an oblique form is used to project nominative forms, the system makes essentially the right predictions for the *honor* analogy: [-o:s] nominatives are strongly favored for monosyllabic and neuter nouns, [-or] nominatives are preferred for polysyllabic masculine and feminine nouns with [-o:s] remaining a strong second choice, and [-or] is strongly favored for agentive nouns.

This analysis is similar in spirit to the four-part analogy explanation of the change, but the gradient nature of the rules in this system gives us a quantitative expression of the influence of other lexical items (50b), and also helps predict which pattern should win out for each class of words (50cd). Finally, the close competition between [-o:s] and [-or] for many forms suggests an intriguing interpretation for the persistence of [-o:s] nominatives into the attested period: I conjecture that perhaps they were not merely a conservative retention of memorized archaic

[5]It is difficult to know in many cases whether the occurrence of forms like *hono:s* alongside *honor* reflects free variation, or simply literary archaism. Cicero, for example, systematically used the form [hono:s] instead of [honor], in both philosophy and oratory texts (elevated styles), as well as in letters (potentially less elevated/archaic); at the same time, he used [labor] instead of [labo:s] in all three contexts.

[6]One possible answer is that the paradigm of [hono:s] already had [r] forms in it, while the paradigm of [soror] never had [s] forms. Following Steriade (1994) we might call this the "lexical conservatism" analysis, in which speakers may only use or extend allomorphs that are already attested. It is possible that a lexical conservatism analysis could explain this part of the asymmetry ([honor] but not *[soro:s]), but it would tell us nothing about the effect of syllable count or gender, since monosyllables and neuters also had an available [r] allomorph which could have been extended to the nominative singular.

forms, but that they may also have supported to a certain extent by the synchronic grammar of Classical Latin. In other words, when Latin speakers heard an archaism like [honoːs], even if it was not the synchronically preferred form, it may have struck them as moderately grammatical, allowing [-oːs] forms to persist as an archaism much longer than some other archaic features. Clearly more philological work is needed to support this hypothesis, but it is a good example of how the current model that makes predictions not only about how forms can be innovated, but also about how they may be retained or lost.

4.2 Applying the model to Latin noun paradigms

In the small, hypothetical examples discussed in chapter 3, the neutralizations affected just one form in the paradigm, making the mappings in one direction clearly easier than those in the opposite direction. In real languages, however, the situation is rarely so clear. Neutralizations typically affect only a subset of the segments in the language, so the uncertainty that they cause may only affect a small number of words. Furthermore, neutralizations often affect different parts of the paradigm for different words. Thus, it is not always easy to intuit whether a mapping is easier in one direction than the other, or the magnitude of the asymmetry.

The question of interest for this chapter is whether Latin nouns were easier to project in the oblique → nominative direction than vice versa. Latin nouns are traditionally divided into five classes or declensions, each of which was inflected for five major cases: the nominative, genitive, dative, accusative, and ablative. (Two additional cases, the vocative and locative, were almost always identical to other cases.) A full description of all of the declensions and their subclasses is clearly beyond the scope of this chapter – see Leumann (1977), Kühner (1912), or Allen (1903) for in-depth discussions. What is important here is that the distinctions among many of these classes were neutralized or nearly neutralized in various parts of the paradigm.

As with Yiddish, it is possible to consider what neutralizations affected each slot in the paradigm. In the nominative singular, several morphological classes were neutralized by having the same case endings. For example, at least four classes of nouns all had the ending [-us] in the nominative singular: 2nd declension masculines like [populus] 'people'; 3rd declension neuters, some with [-oris] ([korpus], [korporis] 'body') and some with [-eris] ([genus], [generis] 'kind'); and 4th declension masculines, like [manus]. These are shown in the shaded row in (51a). Another ambiguous ending was the nominative singular [-er] (51b), which included both second declension (genitive [-iː]) and third declension ([-is]) nouns. Nouns with [-er] also exhibited a neutralization between [e]'s that alternated with ∅ in the oblique forms ([ag_riː]) and those that did not ([gen**e**riː]). Finally, the noun [iter] exhibited a relatively idiosyncratic alternation between ∅ and [in].

(51) Morphological neutralizations in the nominative

a. Nouns ending in *-us* in the nominative

	'people'	'body'	'kind'	'hand'
nom.sg.	[popul**us**]	[korp**us**]	[gen**us**]	[man**us**]
gen.sg.	[populiː]	[korporis]	[generis]	[manuːs]
dat.sg.	[populoː]	[korporiː]	[generiː]	[manuiː]
acc.sg.	[populum]	[korpus]	[genus]	[manum]
abl.sg.	[populoː]	[korpore]	[genere]	[manuː]

b. Nouns ending in *-er* in the nominative

	'field'	'father-in-law'	'brother'	'prison'	'journey'
nom.sg.	[ag**er**]	[sok**er**]	[fraːt**er**]	[kark**er**]	[it**er**]
gen.sg.	[agriː]	[sokeriː]	[fraːtris]	[karkeris]	[itineris]
dat.sg.	[agroː]	[sokeroː]	[fraːtriː]	[karkeriː]	[itineriː]
acc.sg.	[agrum]	[sokerum]	[fraːtrem]	[karkerem]	[iter]
abl.sg.	[agroː]	[sokeroː]	[fraːtre]	[karkere]	[itinere]

In addition to these (and other) morphological neutralizations, the nominative singular was also affected by several phonological neutralizations induced by the *-s* suffix. For example, the voiced stops [b] and [g] became devoiced in this environment, merging with [p] and [k], as in (52a,b). The coronal obstruents [t], [d], and [s] were deleted altogether in this environment, as in (52c).

(52) Devoicing before *-s*

a. Neutralization of *g* with *k*

	'murder'	'flock'
nom.sg.	[neks]	[greks]
gen.sg.	[nekis]	[gregis]
dat.sg.	[nekiː]	[gregiː]
acc.sg.	[nekem]	[gregem]
abl.sg.	[neke]	[grege]

b. Neutralization of *b* with *p*

	'feast'	'city'
nom.sg.	[daps]	[urps]
gen.sg.	[dapis]	[urbis]
dat.sg.	[dapiː]	[urbiː]
acc.sg.	[dapem]	[urbem]
abl.sg.	[dape]	[urbe]

c. Deletion of *t*, *d*, and *s*

	'foot'	'lawsuit'	'penny'
nom.sg.	[peːs]	[liːs]	[aːs]
gen.sg.	[pedis]	[liːtis]	[assis]
dat.sg.	[pediː]	[liːtiː]	[assiː]
acc.sg.	[pedem]	[liːtem]	[assem]
abl.sg.	[pede]	[liːte]	[asse]

Another phonological neutralization was caused by restrictions on word-final clusters, which resulted in the deletion of final coronals in clusters: *[kord] > [kor] 'heart', *[lakt] > [lak] 'milk':[7]

[7]There appear to be no neutralizations in noun paradigms caused by *[kt] > [k], since nouns ending in simple [k] end in *-s* in the nominative: [paks]/[paːkis] 'peace-nom./gen.'.

(53) Neutralizations caused by deletion of coronals

	'heart'	'sister'
nom.sg.	[kor]	[soror]
gen.sg.	[kordis]	[soroːris]
dat.sg.	[kordiː]	[soroːriː]
acc.sg.	[kor]	[soroːrem]
abl.sg.	[korde]	[soroːre]

The nominative was not the only form with neutralizations, however. Oblique forms also suffered from both morphological and phonological neutralizations.

The (morpho)phonological rhotacism process that yielded the [s] ∼ [r] alternation in [honoːs] ∼ [honoːris], for example, created a neutralization in oblique forms, between alternating [s] ∼ [r] and non-alternating [r]:

(54) Neutralization of /s/ and /r/ caused by rhotacism

	'honor'	'sister'
nom.sg.	[honoːs]	[soror]
gen.sg.	[honoːris]	[soroːris]
dat.sg.	[honoːriː]	[soroːriː]
acc.sg.	[honoːrem]	[soroːrem]
abl.sg.	[honoːre]	[soroːre]

In addition, a process of vowel reduction in non-final unstressed syllables caused several vowels to surface as [i] in oblique forms:

(55) Neutralization of /e/, /u/, /i/ to [i] in oblique forms

	'soldier'	'head'	'pyramid'
nom.sg.	[miːles]	[kaput]	[pyramis]
gen.sg.	[miːlitis]	[kapitis]	[pyramidis]
dat.sg.	[miːlitiː]	[kapitiː]	[pyramidiː]
acc.sg.	[miːlitem]	[kaput]	[pyramidem]
abl.sg.	[miːlite]	[kapite]	[pyramide]

These are just a few of the more serious neutralizations that would have occurred in the nominative and oblique forms in Latin. Unlike Yiddish verbs, which involved just a few neutralizations, it would be difficult in the case of Latin nouns to give a comprehensive list of all of the possible neutralizations, because there were many more inflectional classes (including nouns of three genders), and more irregular morphophonological processes involved. This sampling of the complexity of the system should at least serve to show that the choice of a base form is not nearly as straightforward as it was in the synthetic languages of chapter 3, or in Yiddish verbs. What we must take into account, therefore, is how serious these neutralizations were, by considering the number of lexical items involved.

Some of these neutralizations affect relatively large numbers of words. The neutralization between masculine ([-us]) and neuter ([-um]) second declension nouns in the genitive (both [-iː]) involves two very large (and productive) classes of nouns. The neutralizations caused by voicing agreement in final obstruent clusters in (52a,b), on the other hand, affected relatively fewer words. An additional complication is that other factors, such as grammatical gender, could help the speaker know which suffix to use in a potentially ambiguous situation – so, for example, the use of [-us] or [-um] in the nominative of a word with [-iː] in the genitive is almost completely predictable given the gender of the word.[8] Thus, if there is an asymmetry in predictability between nominative and other forms, it would be because of differences in the "severity" of the neutralizations involved, and the ability to predict the correct form using gender. Were the neutralizations in the nominative in fact more severe than in other cases?

In order to answer this question, I started with a database of fully inflected classical Latin nouns, prepared in 1997-1998 by a group working under the supervision of Bruce Hayes at UCLA. This database contained all of the nouns with five or more tokens in a lemmatized frequency count from classical texts (Delatte, Evrard, Govaerts, and Denooz 1981), based on a corpus of approximately 800,000 words (582,000 from prose, 212,000 from poetry). Nouns beginning with the letters R through Z were omitted from the simulations because the database was found to have incomplete information for many paradigms in this section of the alphabet. Nominative forms were listed in their forms prior to the [honoːs] > [honor] change; in cases of uncertainty, words were listed with a final [s]. The rationale for this was that we are interested in seeing if the model will favor [r] forms in spite of numerous [s] forms in the training data, and we do not want this to be the result of the influence of spurious [r] forms.

The model of base selection being tested here is that learners evaluate the usefulness of prospective bases early in the learning process. Therefore, the only input data which would be available to the learner for comparisons would be the most common words. As an idealization, words with 50 or more tokens in Delatte, Evrard, Govaerts, and Denooz (1981) were selected, for a total of 494 input nouns. Six forms were considered as possible bases: the nominative, genitive, dative, accusative, and ablative singular, and the nominative plural.[9] For each possible base form, training data files were then constructed to project each of the remaining forms (*nom.→gen.*, *nom.→dat.*, *nom.→acc.*, etc.), yielding 30 (=6×5) training sets in total.[10]

Nouns in the input files were listed in phonemic transcription. In order to take into account the effect of phonological processes, a list of illegal sequences was also prepared, including final clusters disagreeing in voicing (*bs#, *gs#, *ds#), final geminates (*ll#, *dd#, *ss#), the clusters *rts, *lts, and *nts, and a few other illegal sequences whose repair caused alternations in the nominative (*oːr#, *kt, *iiː). Because some of these illegal sequences refer crucially to word boundaries, word boundaries were also marked explicitly in the input files with brackets. Since rhotacism is not surface-true in this stage of Latin (cf: [kaːsus] 'fall', [rosa] 'rose'), intervocalic

[8]There were a few isolated second declension neuters ending in [us] in the nominative, including [vulgus] 'people', [pelagus] 'sea', and [virus] 'poison'.

[9]The remaining plural oblique cases were not considered because they never preserved distinctions that were neutralized elsewhere, and in fact some cases (especially the dative and ablative) involved massive neutralizations between different declension classes.

[10]All of the training input sets and results files for the simulations discussed here, as well as the original database of nouns, can be downloaded from http://www.linguistics.ucla.edu/people/grads/aalbrigh/papers/latin.html.

[s] was not included as an illegal sequence (Hoenigswald 1960, pp. 106-107).[11] Finally, each noun was provided with a numeric code indicating the grammatical gender and the number of syllables (monosyllabic vs. polysyllabic), since the current implementation of the minimal generalization learner does not have an independent capacity for considering general prosodic properties of words. These codes allowed words to be categorized as masculine, feminine, or neuter, and as monosyllabic or polysyllabic. Token frequencies were also included in the input files, but they were not employed in the simulations reported here.

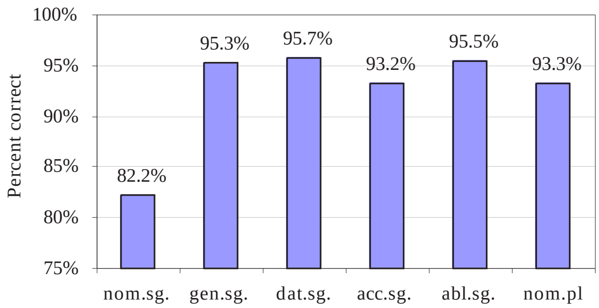
The input files were submitted to the minimal generalization learner, yielding subgrammars of rules with confidence values. The word-specific rules were then eliminated, and the resulting subgrammars were tested on the input forms. The metrics proposed in section 3.3 were calculated for each subgrammar, to obtain an estimate of the usefulness of each slot in the paradigm for predicting the remainder of the paradigm. The results, given in Appendix B, show that the predictability between *all* forms is quite high (over 80%); it is not the case that any part of the paradigm suffers from neutralizations that affect the majority of nouns in the input set. Nevertheless, the oblique forms tend to be substantially better than the nominative form on almost all of the metrics considered. In Figure 4.1, the candidates for base status are compared in terms of their mean effectiveness in projecting the five other forms in the paradigm.

As can be seen, the criteria proposed in section 3.3 generally agree on the relative effectiveness of the various forms as possible bases. The combined results from all five remaining criteria (excluding number of rules) are shown in Figure 4.2.

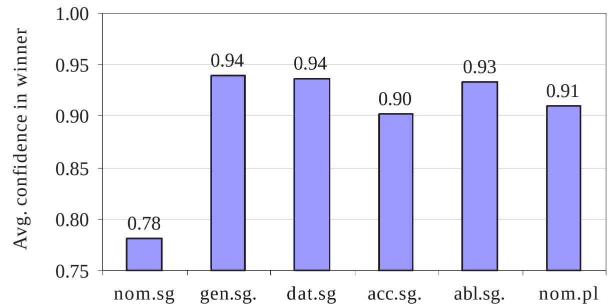
The nominative is the worst choice of base under all criteria, and thus receives the lowest rank for all metrics. This reflects the fact that the nominative suffers from more neutralizations, affecting both more words and more segments, than the oblique forms. Interestingly, the accusative also fares relatively poorly, because it is the same as the nominative for all neuter nouns, and thus shares many of the same neutralizations. Among the remaining forms, the dative comes out slightly ahead of the genitive and ablative.

An additional factor that has not been discussed here is the relative frequency of the different forms in the paradigm. As an idealization, I have assumed that learners have access to the six candidate forms in equal proportions for all nouns. Clearly this is not true in real life however;

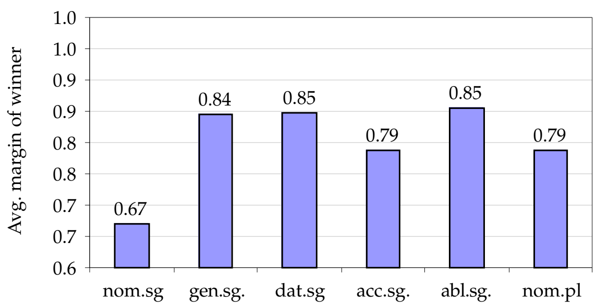
[11] The presence of intervocalic [s] does not necessarily preclude the possibility that rhotacism continued to be a synchronically active process in Latin, possibly restricted to a particular morphological environment, such as /V___+V, where '+' indicates a morpheme boundary. For the purposes of the current model, the synchronic status of rhotacism actually makes very little difference. Including a *Vs+V constraint would improve the reliability of nominative → oblique grammars slightly, because the model could learn to apply rhotacism in mappings like [hono:s] → [hono:ris] (instead of [hono:sis]); however, *-o:s* nominatives make up only a small fraction of the language as a whole, so improving the model's predictions for this subset of the vocabulary does not make a substantial difference in the calculations reported below. Note also that including a rhotacism constraint does not help the model at all in the oblique → nominative direction, since an oblique form with *-o:ris* could come from either underlying /s/ or underlying /r/. More generally, assuming that rhotacism was synchronically active in Latin can help to explain why the paradigm of 'honor' was not leveled to [hono:s], [hono:sis], etc., but it cannot explain other facts, like why speakers did not assume that forms like [soro:ris] were also the result of rhotacism (predicting the incorrect nominative [soro:s]), or why speakers did not continue to tolerate the rhotacism alternation. It could also be added that many authors have tried to make use of the exact opposite intuition, arguing that since rhotacism had been obscured by numerous exceptions caused by borrowings, dissimilations, and degemination of **ss* > *s*, it was no longer synchronically active in Latin. If we assume that rhotacism was no longer productive, then we can understand why [hono:s] and [hono:ris] could no longer be related to one another by an automatic phonological process, and why the alternation was then open to leveling (Hoenigswald 1960, pp. 108-109; Klausenburger 1979; Wetzels 1984; Barr 1994, pp. 519-524; and others).



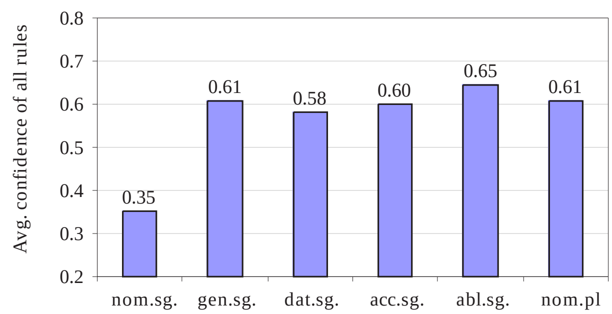
(a) Percent correct



(b) Avg. confidence in winner



(c) Avg. margin between winner and nearest competitor



(d) Avg. confidence of rules

Figure 4.1: Comparison of potential bases according to the criteria in §3.3

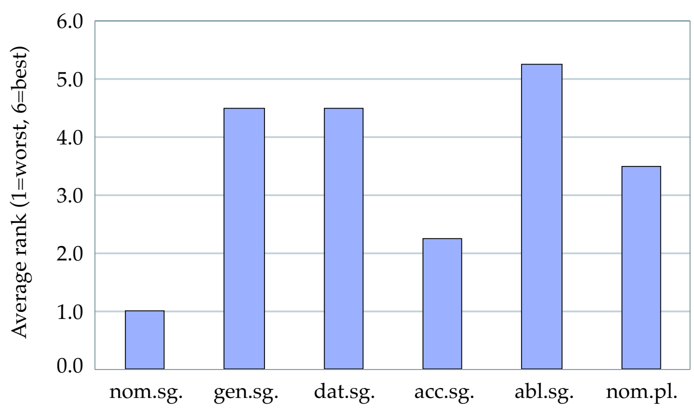


Figure 4.2: Average rank of forms as potential bases

some cases are more frequent than others, and the frequency of cases may differ from word to word. As mentioned on page 40, the Albright and Hayes implementation of minimal generalization uses confidence statistics to estimate the effectiveness of rules, so that rules covering a few forms are penalized more than rules covering many forms. Therefore, with more realistic input data, including different amounts of data about different cases, subgrammars involving less frequent cases would be penalized because their rules would be based on fewer forms. I will return to this issue section 6.2.1.

A simulation taking this into account would require more detailed frequency information about Latin noun paradigms than is currently available to me. Nevertheless, intuitively, it seems that there are substantial differences in the frequency of the oblique cases, and this could be the decisive factor in choosing a base from among the oblique forms that are more or less equivalent by all other criteria. (In section 4.4.2, I will show that at least for a small sample of words in a small text corpus, the ablative and genitive are both relatively frequent forms.) For the purposes of the [honoːs] > [honor] change, it is sufficient that the model proposed here select something other than the nominative as the base form; in the discussion that follows, I will use the genitive singular as the base for deriving the nominative, but the same result could be achieved using the ablative or dative singular.[12] This proposal is in line with Hooper's claim (Hooper 1976, p. 95) that an oblique form was the basic form in Latin, although Hooper claims (without much discussion) that it was in fact the accusative, not the genitive or dative. A possible solution that is compatible with both of these claims is that the global base of Latin noun paradigms was the dative or accusative, while the nominative was derived more locally from the accusative; I will discuss this possibility in more detail in section 6.3.

4.3 Projecting nominatives from the genitive

Choosing an oblique form as the base in Latin noun paradigms gives us only half of the explanation for the [honoːs] > [honor] change. In particular, it explains the "backwards" direction of the change (oblique forms affecting nominatives). This answers the question in (50a), of why it was the nominative that changed in Latin. What remains to be shown, then, is that once an oblique form has been chosen as the base, the model makes the right predictions for nominative forms: namely, that polysyllabic non-neuter *-oːs* nouns changed to *-or*.

Recall that an assumption of the current model is that bases are selected early in the learning process, but learners continue to fine-tune their grammars to derive the remainder of the paradigm. Therefore, in order to test the predictions of the model for nominatives using an oblique form as the base, the model was trained on the full set of 1,687 nouns in the gen.→nom. direction. The resulting grammar was then used to generate possible nominatives for all genitive forms ending with sequences that could potentially arise from rhotacism: [-oːris], [-oris], [-uris], [-eris] (157 in all). The grammar derived several possible nominatives for each noun, each with its own confidence value. For example, for [-oːris] and [-oris] genitives, the possible nominatives typically included an [-oːs] nominative, an [-or] nominative, and various other possibilities, such as [-oːris] (on the basis of words like [kanis] 'dog-nom./gen.sg.', which were

[12] A common intuition is that the [honoːs] > [honor] change may be due to the collective influence of *all* of the oblique forms combined, and not the effect of a single oblique form on the nominative singular. I will discuss this possibility further in section 4.4.1.

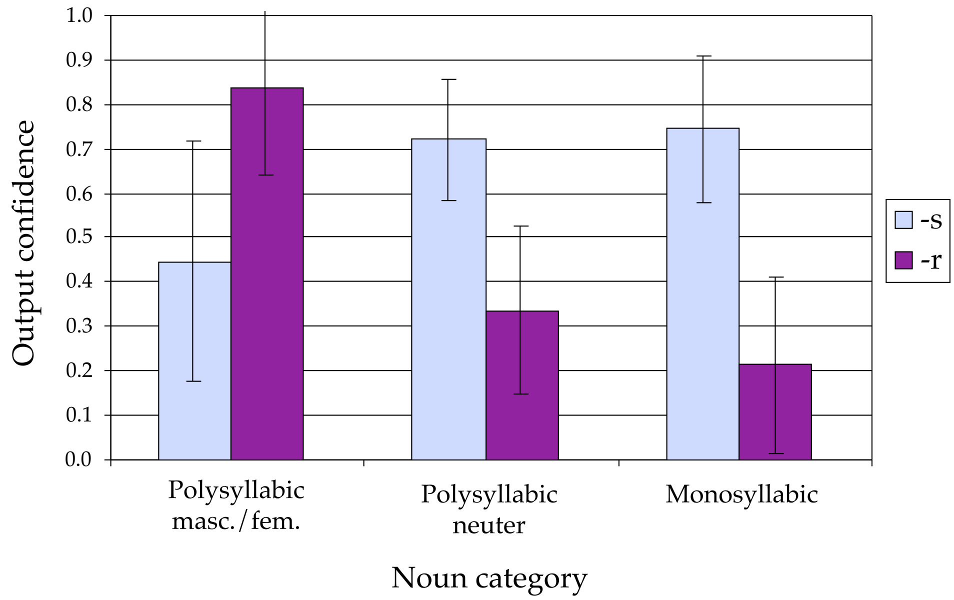


Figure 4.3: Preference for *-r* or *-s* in different categories of nouns

identical in the nominative and genitive), [-us] (like [korpus] ‘body-nom.sg.’), etc. For each noun, the best possible *-r* nominative was compared against the best possible *-s* nominative, in order to gauge the model’s preference for *-r* nominatives. As expected, the preference for *-r* or *-s* varied substantially from word to word. For masculine and feminine polysyllabic nouns, *-r* was generally favored, while for neuters and monosyllables, *s* was prefered (Figure 4.3), mirroring the observed outcome of the *honor* analogy. (Note that in the graph, bars indicate the size of a standard deviation, not the standard error.)

Why does this pattern emerge? The differing strength of *-r* and *-s* for different words is due to the fact that the system employs multiple rules, with different confidence values in different contexts. Among polysyllabic non-neuter nouns, genitives in [-o:ris] frequently have nominatives in *-or*. Thus, the rule of [o:ris] → [or] / $[X]_{polysyl,-neut}$___# has a relatively high confidence (.727), correctly deriving words like [soror] and [cruor], and all agentives, but failing for words like [hono:s]. Among these forms, then, there is a slight preference for *-or* in the nominative.

Furthermore, within the masculine and feminine polysyllabic nouns, two groups can be distinguished (Figure 4.4). The first are the agentive nouns, which strongly favor *r* in all cases. In fact, these words contained *-r* etymologically, and continued to have *-r* in the nominative with no variation or hypercorrections. The second are the non-agentive masculine and feminines, which show a slight tendency to favor *-r*, but with strong competition from *-s*. This is the *honor* class of words, which were etymologically *-s* but changed to *-r*, with some attested variation and occasional hypercorrections of etymological *-r* to *-s* (Neue-Wagener 1902, p. 265).

The model can capture this difference because it is able to posit a more specific rule that covers just the agentives; these are not only polysyllabic and non-neuter, but they also all have

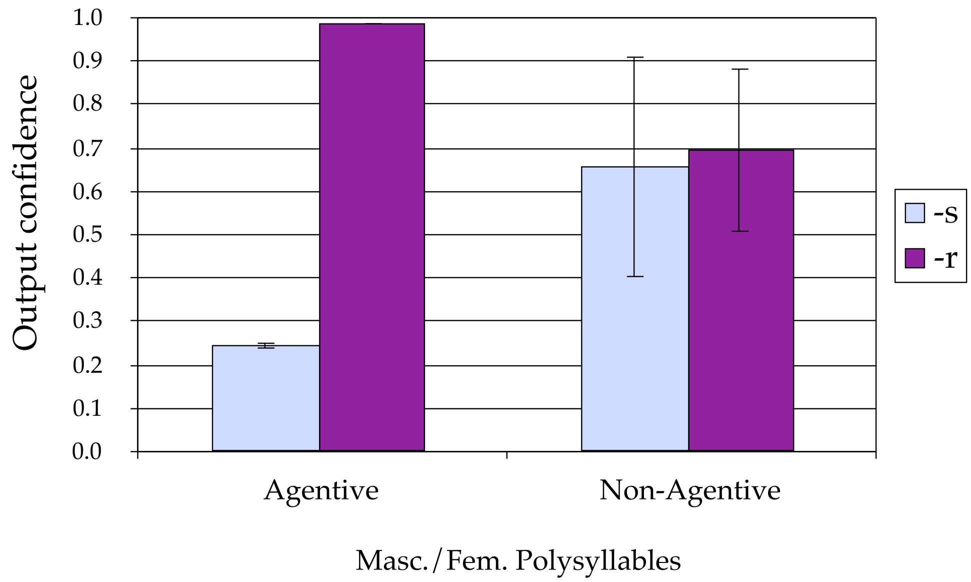


Figure 4.4: Comparison of agentive vs. non-agentive nouns

a stem-final [s] or [t]: *doctor, audītor, cēnsor,* etc.[13] Thus, the more specific rule [o:ris] → [or] / [X {s,t}]$_{polysyl,-neut}$___# is able to describe the agentives quite narrowly, and has an extremely high reliability (.980).

Outside the class of polysyllabic non-neuter nouns, the reliability of *-r* is much lower. Among polysyllabic neuters, very few nominatives end in *-r*, so the general rule [ris] → [r] / [X]$_{polysyl,+neut}$___# has a rather low confidence (.196). Among this group of words, the competing rule [ris] → [s] / [X]$_{polysyl,+neut}$___# has a high reliability, meaning that for these forms, non-uniformity (anti-correspondence) prevails, and there is no change. There are, however, a few local pockets of *-r* nominatives among the neuters, especially among those with *-aris* and *-eris* in the genitive ([kalkar]/[kalka:ris] 'spur-nom./gen.', [nektar]/[nektaris] 'nectar-nom./gen.', [aker]/[akeris] 'maple-nom./gen.') This is the reason why the model disfavors innovative *-r* nominatives among neuters, and why there is also a good deal of item-by-item variation among them.[14] Finally, among monosyllables, almost all *-ris* genitives had nominatives ending in *-s* – a notable exception being [fu:r] 'thief'. Thus, the model correctly learns that in this environment, the *s* ∼ *r* alternation is extremely robust, and final *-r* cannot compete with it.

It should be reiterated that the slight preference of the model for *-r* in words like [honor] emerges in spite of the fact that they were listed with *-s* in the training data. In other words, the grammar produces an output which is different from the existing form. Therefore, under this analysis, pre-change forms like [hono:s] would have been considered irregular, and would have

[13]This common phonological trait is not a coincidence: agentives were formed from the perfect passive participle (4th stem), which was generally formed by adding a [t], or in some phonological contexts, by changing a stem consonant to [s] (e.g., [kad-] → [ka:s-] 'fall').

[14]In fact, Kieckers (1960) points to one, possibly quite isolated example of a neuter noun with etymological *-r* being written with an *-s*: *femus* 'femur' (vol 2, §22).

had to have been listed as exceptions to the *-ris* ∼ *-r* pattern. Of course, if learners had perfect memories and access to all forms of all words, then they could perfectly well have memorized [hono:s] and continued to produce it, and the language would not have changed. However, in real life this is not the case, and speakers must sometimes synthesize new forms. The model is intended to predict what forms a speaker would produce in such situations, and in this case it correctly predicts errors, or overregularizations, like *honor*. It cannot, however, predict when existing forms would be unavailable—either because they have never been encountered, or because lexical access has failed—forcing speakers to use their grammars. What is missing from this model, then, is a production mechanism which uses both the lexicon and grammar to produce forms. Even so, the result in (4.3) is still significant, because we can assume that speakers do sometimes make overregularization errors (Marcus, Pinker, Ullman, Hollander, Rosen, and Xu 1992; Pinker 1999), and the errors that the model makes correctly mirror the attested historical change.

It is worth mentioning that the [hono:s] > [honor] change was just one of many changes that affected nominatives in the history of Latin. Numerous nouns with highly irregular nominative forms were regularized; for example, the nominative of [juppiter]/[jowis] 'Jupiter-nom./gen.' was eventually replaced by [jowis] (cf: [kanis]/[kanis] 'dog-nom./gen.'), and the nominative of [bo:s]/[bowis] 'cow-nom./gen.' is attested as [bowis] (Kieckers 1960, vol. 2, §I.21.3; Kühner §63.2). The converse change, of regularizing nouns by fixing the oblique forms, generally did not occur (*[juppitris] or *[juppitri:], *[bo:ris], etc.).[15] Furthermore, the form of nouns in modern Romance languages can generally be traced back to oblique forms in Latin. For example, Latin [pe:s]/[pedis] 'foot-nom./gen.' has yielded Italian [pje**d**e], instead of the expected *[pe] (cf: Latin [tre:s] > Italian [tre] 'three'); similarly [ars]/[artis] 'skill' > Ital. [ar**t**e], [flo:s]/[flo:ris] 'flower' > Ital. [fjo**r**e], [niks]/[nivis] 'snow' > Ital. [ne**v**e], and so on. The analysis of Latin presented here helps to explain why nominatives were subject to many changes in the later history of Latin, and not just leveling of *s* ∼ *r* alternations. The Uniform Exponence account of the Latin change presented above, on the other hand, does not predict leveling of alternations other than *s* ∼ *r*.

4.4 Discussion

This analysis captures two common intuitions about Latin nouns, and about the [hono:s] to [honor] change. The first is that oblique forms are "more revealing" about the declension of a noun than the nominative – seen, for example, in the common practice of listing both the nominative and genitive forms in dictionary entries, as the nominative alone is not considered

[15]There seems to be at least one case in which a property of the nominative was extended to the remainder of the paradigm: (1) the paradigm of words like [vo:ks] 'voice' originally had a long [o:] and [k] in the nominative, and a short [o] and [k^w] elsewhere ([vo:ks], [vo$k^w$is], etc.; see Meiser 1998, p. 141 regarding vowel length, Leumann 1977, p. 148 and Kieckers 1931, p. II.13 regarding [k^w]). The long [o:] and simple [k] of the nominative were subsequently extended to all forms: [vo:ks], [vo:kis], etc. It would be interesting to compare the relative chronology of these changes, since these nominative-driven changes may have occurred at an older stage of the language in which nominatives suffered from fewer phonological reductions, while changes driven by obliques occurred throughout the Classical and Late Latin period.

"informative enough" to predict the entire paradigm.[16] This intuition is reflected in the current analysis by the fact that an oblique form is chosen as the base, and the remainder of the paradigm is derived from an oblique form.

The second intuition is that the change from [hono:s] to [honor] involved replacing a small, irregular, morphologized alternation with the more general default pattern of non-alternation. This analysis shares with Barr (1994) the idea that this can be captured with competing rules, but differs with respect to why different classes of words were treated differently. Barr claims that monosyllabic nouns and neuter nouns retained *s* for fundamentally different reasons. The monosyllabic nouns, she argues, retained *s* because in a shorter word, the alternating $s \sim r$ segment constituted a larger proportion of the word, and was therefore more salient and more easily retained. The neuters, on the other hand, retained *s* because it occurred in two slots in the paradigm (both the nominative and the accusative) instead of just one, and was more salient for this reason. Both of these points seem to involve arbitrary thresholds. Even if we grant that the final segment of *flo:s* is "a larger portion of the word" than the final segment of *odo:s* (which both have four phonemes, but do differ in their syllable count), is the extra vowel in *odo:s* really enough to distract learners from reliably noticing the final $s \sim r$? And in the case of neuters, why was occurring in two slots in the paradigm sufficient to guarantee that the *s* would be preserved, when occurring in two slots was not enough to guarantee that [ɔ] would be preserved in Polish feminine diminutives (5b) (among many other cases)?

In the current system, the difference between different genders and word lengths is attributed to the existence of multiple versions of the rules in question, at varying levels of generality, and with differing reliability in different contexts. The use of multiple overlapping rules might be seen as unwanted redundancy in the model, but in fact cases like Latin are taken as evidence that speakers, too, have at least a certain amount of detailed knowledge about the reliability of different processes in different environments. Furthermore, an ability to assess the reliability of rules in different environments is required in any event in order for learners to locate the best morphological and phonological rules to describe the patterns of their language.

Although this analysis makes use of several intuitions about the factors that are thought to drive paradigm leveling, it ignores certain other factors that have been proposed in the literature. Some notable factors that do not play a role in this analysis are the frequency of an allomorph within the paradigm, the token frequency of various surface forms, or the semantic naturalness of different nouns in different cases. It is useful to consider, therefore, the extent to which these other factors could provide an alternative explanation of the [honor] analogy, and whether the current model would benefit from incorporating any of these factors.

4.4.1 Frequency of occurrence within the paradigm

It is often suggested that the [hono:s] to [honor] change was encouraged by the fact that every form in the paradigm except the nominative singular contained [r]. I will refer to this as the "majority rule" hypothesis. Under the model proposed here, each paradigm has a single unique base, and forms are derived by grammars relating individual pairs of forms. If the most

[16]The relative uninformativeness of the nominative in Latin is due, in part, to the fact that the nominative suffix for one large class of nouns lacked a vowel (*-s*), creating coda clusters that resulted in phonological simplifications (e.g., **arts* > *ars*). The oblique forms always provided a prevocalic context for the stem, resulting in far fewer neutralizations; rhotacism is a rare exception.

informative form had turned out to be the nominative singular, the prediction of this model is that all of the remaining forms could have been rebuilt on the basis of a single form. Thus, this model has no way to capture the majority rule intuition. It is not clear to me, however, that there is evidence that paradigm leveling is truly driven by majority rule. There are numerous cases in which a single form seems to have driven a paradigmatic change—for example, the leveling in Polish masculine diminutives in chapter 1 and the Yiddish change discussed in chapter 2 both involve the extension of a form that is vastly outnumbered in the paradigm. Conversely, it is difficult to prove that a leveling like [hono:s] > [honor] would not have happened if [s] had occurred in more slots in the paradigm. Barr (1994, p. 543) points out there was a difference between masculines and feminines (like [hono:s]) and neuters (like [tempus]): masculines and feminines had [s] in just one slot in the paradigm (the nominative), while neuters had [s] in two slots (the nominative and accusative). She suggests that this alone was enough to produce the difference between these classes of words. However, it seems unlikely that this threshold would work in general. Furthermore, this majority rule hypothesis does not actually explain the Latin facts. First, even if we accept that having two slots with [s] was enough to protect the neuter nouns from leveling, there is still the problem that monosyllabic masculine and feminine nouns had only one [s] in the paradigm (the nominative singular), but they did not change either. Furthermore, a few neuter nouns did change, or at least acquired [r] variants (e.g., [kinus]/[kiner], 'ash' (neuter)), in spite of the fact that they had more than one [s] variant. From the point of view of paradigm-internal pressures, there is no reason why these words should have behaved differently. In addition, there are many other noun paradigms in which the nominative had a different form from the rest of the paradigm, but was not leveled (e.g., [iter] ∼ [itineris] 'road-nom./gen.sg.'). Therefore, frequency of occurrence within the paradigm does not seem to add anything to the account of the change.

4.4.2 Token frequency of different paradigm members

A natural hypothesis, pursued by Mańczak (1958) and others, is that less frequent forms are often rebuilt on the basis of more frequent forms within the paradigm. Could it be the case that the nominative was significantly less frequent than the oblique forms in Latin? This is especially relevant in Latin because many or most of the words affected by the [honor] analogy were inanimate or abstract nouns, which are perhaps more frequent in oblique forms than in the nominative.

In order to get a rough (and very informal) estimate of the relative frequency of case forms for different nouns, I performed some counts on the complete works of Cicero, as found in the Perseus Digital Library (http://www.perseus.tufts.edu/). As (4.1) shows, it is true that among singular forms, nominatives rarely constitute the majority of tokens for any noun. This might possibly help to explain why nominative forms were open to rebuilding in Latin – perhaps they were not frequent enough to be reliably memorized and retrieved.[17]

[17]This explanation is not really convincing without a more explicit theory of how frequent a form must be before it can be reliably memorized and retrieved. In fact, 20% of the tokens for a relatively frequent noun seems like sufficient exposure to remember and maintain the form. I am simply granting for the sake of argument that perhaps the lower token frequency of nominatives in Latin could have made them susceptible to leveling.

[18]A problem arises in counting frequencies for neuters, since the nom. and acc. forms are identical. The hypothesis being tested here is that the frequency of [s] forms in the paradigm determines their susceptibility to leveling, so I have counted all *s* forms in the nom. column, to facilitate comparison with the masc. and fem. nouns.

Table 4.1: Distribution of singular tokens for some Latin nouns

Noun	Total Sg.	Nom.	Gen.	Acc.	Abl.
Polysyllabic, non-neuter					
hono:s/ *honor* 'honor'	285	19%	25%	18%	32%
labo:s/ *labor* 'work'	163	17%	21%	37%	35%
odo:s/ *odor* 'odor'	4	50%	0%	0%	50%
Monosyllabic					
flo:s 'flower' (masc.)	16	25%	0%	56%	19%
mo:s 'custom' (masc.)	146	18%	2%	17%	63%
o:s 'mouth' (neut.)	65	29%	9%	6%	55%
Neuter[18]					
corpus 'body'	174	21%	47%	–	28%
onus 'burden'	40	45%	40%	–	15%
tempus 'time'	935	32%	15%	–	51%
Masculine, agentive					
re:x 'king'	207	23%	21%	27%	18%
homo: 'man'	1049	19%	23%	35%	12%
sena:tor 'senator'	43	33%	23%	28%	14%

What these counts cannot explain, however, is why the change should have been restricted only to the non-neuter polysyllabic nouns. The nominative does not seem to be less frequent in this class of nouns that in any other class. Furthermore, there is apparently not even a difference between masculine agentive nouns like 'king', 'man' and 'senator', and inanimate, abstract nouns like 'honor' or 'custom'. Thus, a frequency-based account can explain only the direction, but not the details of the [hono:s] > [honor] change.

4.4.3 Semantics and local markedness

Another intuition, related to token frequency, but logically distinct from it, is that the semantics of particular lexical items make them more "natural" in some case forms than in others. Tiersma (1982), for example, shows that singular forms in Frisian have been rebuilt on the basis of plural forms, but just for those nouns which occur more naturally in the plural than in the singular (such as 'teeth' or 'geese'). He refers to this phenomenon as *local markedness*. On the whole, we would expect local markedness to be reflected in token frequency, which is much easier to measure, but does not provide an adequate explanation of the Latin change (see above). However, pursuing the Jakobsonian view of markedness, one might attempt to come up with a definition of case markedness as distinct from token frequency; I assume that among the least marked nominatives, in this case, would be things like agentive nouns, since they would be most likely to act as (nominative-marked) agents. I have no estimate of the naturalness of the nominative forms which changed from [-o:s] to [-or], but I see no reason why this would fare any better than token frequency as an explanation of the change. Nouns like [onus] 'burden', [korpus] 'body', and [flo:s] 'flower' seem to me to be just as "non-agentive" as [hono:s] 'duty', [odoːs] 'odor', or [arboːs] 'tree'. It appears that the class of nouns that changed is best defined by prosodic and morphological properties, and adding a sensitivity to frequency or semantics would not improve the model's predictions in this case.

4.4.4 Leveling vs. extending alternations

The analysis of paradigm leveling proposed here relies on a strong pre-existing pattern of non-alternation in the lexicon – in this case, the non-alternation of [r]. This proposal immediately raises two related questions: first, if paradigm uniformity is really just the extension of an existing pattern of non-alternation, then what happens when the dominant pattern is alternation? Why does there seem to be a universal tendency towards leveling?

As an example of a language with a dominant pattern of alternation, consider a previous stage of Korean (Martin 1992):

(56)

/__#	Example	/__V	(ACC -*ɨl̥*)	gloss
[t̚]	[pat̚]	[tʰ]	[patʰɨl̥]	'field'
[t̚]	[tʃʌt̚]	[dʒ]	[tʃʌdʒɨl̥]	'milk'
[t̚]	[k'ot̚]	[tʃʰ]	[k'otʃʰɨl̥]	'flower'
[t̚]	[ot̚]	[s]	[osɨl̥]	'clothing'

As (56) shows, all stem-final coronal obstruents alternate with [t̚] word-finally. As with Latin, this alternation could be expressed as the result of a markedness constraint against manner and laryngeal specifications in coda position (favoring [t̚]) outranking faithfulness constraints (which preserve underlying contrasts). If there was a universal pressure for uniform exponence constraints to move above IO-Faithfulness constraints, then we would expect that paradigmatic changes in Korean should bring Korean closer to non-alternating paradigms, perhaps as in (57). (The intervocalic voicing of /t/→[d] is a completely predictable process in Korean.) Note that although the phonotactics of Korean rule out a completely non-alternating paradigm ([nat̚] ∼ *[natɨl̥], or *[nad] ∼ [nadɨl̥]), we may assume that the relatively minor, predictable allophonic alternation between [t̚] and [d] better satisfies Uniform Exponence than a [t̚] ∼ [s] or [t̚] ∼ [tʃʰ] alternation, just as the shortening of final /oːr/→[or] in Latin is assumed to be a less serious violation of Uniform Exponence than a [s] ∼ [r] alternation is.

(57) Expected Korean paradigm leveling:

/__#		/__V		
[t̚]	[pat̚]	[d]	[padɨl̥]	'field'
[t̚]	[tʃʌt̚]	[d]	[tʃʌdɨl̥]	'milk'
[t̚]	[k'ot̚]	[d]	[k'odɨl̥]	'flower'
[t̚]	[ot̚]	[d]	[odɨl̥]	'clothing'

In fact, the attested change in Korean noun paradigms is quite different. As it turns out, the majority of coronal obstruent-final stems contained [s] or [tʃʰ] etymologically (i.e., most were like [k'ot̚]/[k'otʃʰɨl̥] or [ot̚]/[osɨl̥]), and many Korean noun paradigms are being rebuilt to contain [s] or [tʃʰ]:

(58) Actual change in Korean Paradigms: (Martin 1992; Hayes 1995, 1998)

/___#		/___V		
[t˺]	[pat˺]	[tʰ],[tʃʰ],[s]	[patʰɨɭ], [patʃʰɨɭ], [pasɨɭ]	'field'
[t˺]	[tʃʌt˺]	[s],[dʒ]	[tʃʌsɨɭ], [tʃʌdʒɨɭ]	'milk'
[t˺]	[k'ot˺]	[tʃʰ],[s]	[k'otʃʰɨɭ], [k'osɨɭ]	'flower'
[t˺]	[ot˺]	[s]	[osɨɭ]	'clothing'

Although there is a considerable amount of word-by-word and speaker-by-speaker variation, it is clear that the restructuring underway in Korean is introducing, not eliminating, alternations. For the most part, the dominant alternations of [t˺] ∼ [s] and [t˺] ∼ [tʃʰ] are coming to replace other, arguably less drastic alternations like [t] ∼ [d].

The explanatory challenge, therefore, is to explain why in some cases a pattern of alternation is extended (as in Korean), while in other cases, alternations are eliminated (as in Latin). The model of paradigm learning advocated in this chapter always extends the strongest pattern, regardless of whether it is alternating or uniform. The reranking of paradigm uniformity constraints, on the other hand, can explain only leveling; the spread of alternations would have to be handled by other means, such as anti-correspondence constraints (Hayes 1999), leaving us with no explanation for why sometimes paradigm uniformity wins out, and sometimes anti-correspondence wins out. I will discuss another example in which an alternation was extended in the next chapter. In chapter 6, I will also return to the question of why leveling may be somewhat more common than anti-correspondence typologically.

4.4.5 Local summary

In this chapter, I have argued that the Latin [honoːs] > [honor] change was caused by more than simply a sporadic pressure for paradigm uniformity or uniform exponence constraints to assert themselves over IO-Faithfulness constraints. I have shown that the spread of [r] to nominative forms did more than just create uniform paradigms; it also extended a pattern of non-alternation that was already dominant in the lexicon. Details of the change, such as its restriction to polysyllabic nouns and non-neuters reflect the fact that these were especially strong contexts for [r] stems. Furthermore, the "backwards" direction of the leveling, which is the most puzzling aspect of the change, can be explained by the model of base identification that was proposed in chapter 3. More generally, this result provides further evidence for a model of paradigm learning in which learners choose the base form that is "the most informative" – i.e., that preserves the most distinctions between classes of words, and allows the remainder of the paradigm to be predicted with the greatest accuracy and confidence. This echoes a proposal by Lahiri and Dresher (1984) that certain forms in the paradigm "matter more than others" to learners when they are determining what class a word belongs to. However, what we see from examples like the Latin *honor* analogy is that the most important part of the paradigm is not universal; learners can pay attention to which part of the paradigm would make the best base, and use that as a base to derive the remainder of the paradigm. The prediction, then, is that distinctions that are preserved in the base form will be easily learned and maintained, whereas distinctions that are neutralized in the base form may be lost by leveling or regularization.

The question of maintaining lexical distinctions has implications for how underlying forms are discovered. In traditional models of phonology (e.g., Chomsky and Halle 1968) as many surface contrasts as possible are given unique underlying representations, so that ideally all forms of a word can be derived unambiguously from a single UR. In the case of Latin, this would lead us to posit some underlying difference between words like *hono:s* and words like *soror* – perhaps *hono:s* ends in underlying /s/ and is marked with a diacritic to take a rhotacism rule in its suffixed forms, or perhaps *hono:s* has some underspecified archiphoneme (/Z/), or some other difference. We would need to compare various surface forms of each word (e.g., *hono:**s*** and *hono:**ris***) in order to learn that the word has an alternation, and set up the appropriate UR. If learners are paying more attention to some surface forms than others, however, then this constitutes a restriction on how underlying forms can be inferred. In the next chapter, I will pursue this hypothesis even further, arguing that underlying forms must be established on the basis of just one single surface form.

Chapter 5

Extension of *ablaut* in Lakhota

In the preceding chapters, I have outlined a model of base identification that compares various slots in the paradigm, assessing their effectiveness in projecting the remainder of the paradigm. I showed that in the case of both Yiddish (chapter 2) and Latin (chapter 4), one form in the paradigm preserved more contrasts than any other form, and moreover, in subsequent paradigm levelings, contrasts that were maintained in these forms were preserved, while contrasts that were neutralized in these forms were lost.

Most of the neutralizations discussed so far have been *asymmetrical*; that is, the contrast between two segments is better preserved in one form than in another, making mappings in one direction obviously easier than mappings in the other direction. This is illustrated in Figure 5.1. There are many cases that do not fit this pattern, however; often, a neutralization is *symmetrical*, in the sense that neither surface form shows the full range of possibilities. One common type of symmetrical neutralization occurs when a language has three surface patterns involving two phonemes: non-alternating [A], non-alternating [B], and alternating [A] ∼ [B]. In this case, neither direction is obviously better than the other, since each of the surface forms has one ambiguous phoneme (Figure 5.2).

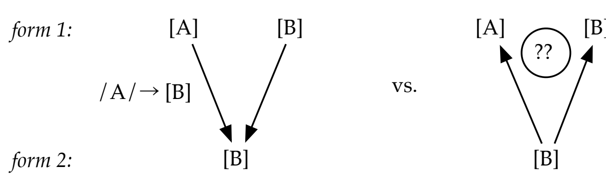


Figure 5.1: An asymmetrical neutralization

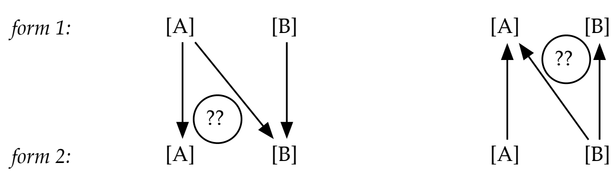


Figure 5.2: A symmetrical neutralization

Lakhota is an example of a language with this type of symmetrical neutralization. There are two types of verbs in Lakhota. The first has invariant final vowels, of any quality; examples of such verbs, with a variety of final vowels, are shown in (59), in the 3sg and 3pl.[1]

(59) Invariant final vowels

3sg	3pl	gloss
*gleshk**a***	*gleshk**a**-pi*	'be spotted'
*low**ã***	*low**ã**-pi*	'sing'
*washt**e***	*washt**e**-pi*	'be good'
*man**i***	*man**i**-pi*	'walk'
*naj**ĩ***	*naj**ĩ**-pi*	'stand'
*man**u***	*man**u**-pi*	'steal'
*nax'**ũ***	*nax'**ũ**-pi*	'hear'
*th**o***	*th**o**-pi*	'be blue'

A second type of verb has a variable final vowel, which surfaces as *-e* in unsuffixed forms (such as the 3sg), and *-a* in the suffixed forms (such as the 3pl) (60). This alternation is known in the Siouanist literature as *ablaut*.

(60) Variant final vowels (*ablaut*)

3sg	3pl	gloss
*chep**e***	*chep**a**-pi* (**chep**e**-pi*)	'be fat'
*kagh**e***	*kagh**a**-pi* (**kagh**e**-pi*)	'do, make'
*khat**e***	*khat**a**-pi*	'be hot'
*naphop**e***	*naphop**a**-pi*	'pop'
*yatk**e***	*yatk**ã**-pi*	'drink'

The puzzle, therefore, is how to distinguish the three-way contrast between invariant [a], invariant [e], and variant ablaut [e]∼[a]:

(61) Three surface patterns

*gleshk**a***	*gleshk**a**-pi*	(invariant [a])
*washt**e***	*washt**e**-pi*	(invariant [e])
*chep**e***	*chep**a**-pi*	(alternating [e]∼[a])

Lakhota ablaut is an "everywhere ambiguous" or symmetrical neutralization. On the face of it, such patterns pose a challenge for the hypothesis that the base must match a single surface form, and must come from the same part of the paradigm for all lexical items. The fact that the verb *chepe∼a* has [e] in some forms and [a] in others cannot be recovered from any single surface form; it is only by comparing two forms that the learner can come to the conclusion that a particular verb exhibits the *ablaut* alternation.

A traditional approach to this type of problem is to encode the difference in the URs of the words, by positing some sort of three-way underlying distinction. Under the standard approach to UR discovery (Chomsky and Halle 1968; Kenstowicz and Kisseberth 1977), it is assumed that all surface forms of a word are derived from a single UR, and wherever possible, surface contrasts should be derived in a lawful way from underlying distinctions. In other words, if it

[1] All Lakhota examples are given in a practical orthography, to be described in section 5.1.1, p. 78.

is possible to distinguish surface distinction ([A] vs. [B]) by using an underlying difference (/A/ vs. /B/), this is preferable to using some other technique, such as marking all B's with a diacritic (/A/ vs. $/A/_{[+A \rightarrow B]}$ or using one UR (/A/) and listing all Bs as lexical exceptions.

In addition to this bias for using distinct URs for contrasting surface patterns wherever possible, it is also generally assumed that for any given word, learners can compare various parts of the paradigm, observe whatever alternations occur, extract all of the unpredictable information, and set up URs that maintain all of the observed contrasts. To take a trivial and uncontroversial example, consider the pattern of alternations caused by final devoicing in German:

(62) German final devoicing

	[ra**t**]	[ra**t**-e]	'advice-nom./dat.'
vs.	[ra**t**]	[ra**d**-e]	'wheel-nom./dat.'

In this case, learners would be able to compare the word for 'advice' and the word for 'wheel', observe that the voicing of the root-final segment is unpredictable in the dative form, and encode this unpredictable voicing specification as part of the UR: /rat/ vs. /rad/. Given these URs, it is also straightforward to formulate a rule of final devoicing that neutralizes the underlying contrast in the nominate form.

A slightly more complicated example comes from Turkish. Like German, Turkish has a general process of final devoicing; however, Inkelas (1994) claims that in addition to words with non-alternating [t] and alternating [t]∼[d] (like German), there are also words with non-alternating [d], such as [etyd]:[2]

(63) Three-way contrast in Turkish final devoicing

	[sana**t**]	[sana**t**-ɨ]	'art-nom./acc.'
vs.	[kana**t**]	[kana**d**-ɨ]	'wing-nom./acc.'
vs.	[ety**d**]	[ety**d**-y]	'etude-nom./acc.'

Turkish final devoicing is therefore a symmetrical neutralization: there are three surface patterns ([t], [d], and [t]∼[d]), so a simple two-way underlying contrast (/t/ vs. /d/) is inadequate. One solution that has often been adopted in the literature is to use underspecification to create an underlying phonemic difference between alternating and non-alternating segments (Inkelas 1994; Inkelas, Orgun, and Zoll 1997; Krämer 2000). For example, following the Prague School practice of including in underlying forms only those specifications are common to all surface forms, we might say that the non-alternating [t] and [d] of Turkish are underlyingly /t/ and /d/, whereas alternating [t] ∼ [d] is an archiphoneme (/D/), with no underlying voicing specification (Trubetzkoy 1962; Anderson 1985, pp. 107-113). The [±voice] specification of underlying /D/ would then be filled on the surface by rules or by markedness constraints, such as no final voiced obstruents ($*[+voi,-son]/\_\_\_]_{\sigma}$) and no intervocalic voiceless obstruents (*[-voi,-son]/V__V).

The underspecification/archiphonemic analysis is consistent with the basic tenets of generative phonology, but it is important to remember that an alternative solution is also available in such cases. In particular, if we relax the requirement that all forms in the language must

[2] Not all Turkish speakers seem to agree on whether the nominative singular of 'etude' should be pronounced [etyd] or [etyt]; it is possible that the pattern described by Inkelas represents an especially formal or educated speech style, in which French words are pronounced as faithfully as possible, even if this means violating final devoicing.

be rule-governed, we may use just two underlying phonemes (/t/, /d/), and list some forms as exceptions. In this scenario, we could set up non-alternating [t] as underlying /t/, alternating [t] $\sim$ [d] as underlying /d/ (with a rule of final devoicing), and non-alternating [d] as underlying /d/, marked in some fashion as an exception to the final devoicing rule. It is this line of analysis that the single surface base restriction forces us to.

The outline of the rest of this chapter is as follows: first, I will show that Lakhota is like Turkish, in that it has a three-way contrast ([a], [e], [e]$\sim$[a]), but only two surface phonemes. Thus, it is a good candidate for a underspecification analysis – in fact, a better candidate than Turkish, because underspecification can account for not only final vowel alternations, but other processes in the language as well. I will then show that the underspecification analysis is nonetheless inadequate for Lakhota, and there is data that it cannot account for. In particular, historical changes show that many invariant [a]'s have switched to variant [e]$\sim$[a], but other logically possible changes have not occurred (invariant [e] $\not\Rightarrow$ [e]$\sim$[a], and [e]$\sim$[a] $\not\Rightarrow$ invariant [a] or [e]). The result is new forms that are inconsistent with any UR in the old system, for reasons that will be explained in section 5.2. This is unexpected under an approach in which learners can compare various parts of the word to posit a UR that can neatly derive all of the surface forms. However, I will show that it follows straightforwardly if we assume that the single surface hypothesis holds not only for bases in output-output effects (paradigm leveling, word-based morphology), but also for the underlying forms that are the inputs to phonology. The end result is that the single surface base restriction appears to be relevant not only for models that consider the relations between surface forms, but also for models that use potentially abstract URs of stems.

I begin with a brief overview of the Lakhota segment inventory, and the processes involved.

5.1 Background on Lakhota

Lakhota is a Siouan language, spoken by roughly 6,000 speakers today in the Dakotas and surrounding areas (Grimes 2000). I draw my Lakhota data from the following sources, differentiating them where necessary: Boas and Deloria's grammar (1941), Buechel's Lakhota dictionary (1970), a verb list compiled in field work by Munro (1989), and notes from my own field work from 1999-2001 with Mary Rose Iron Teeth, a native speaker from the Pine Ridge Reservation in South Dakota.

5.1.1 Phoneme inventory and phonotactics

The Lakhota phoneme inventory is given in Tables 5.1 and 5.2; the practical orthography that I will be using here is given in italicized letters, and the IPA (where different) is given in brackets.

A phonotactic fact about Lakhota that will be relevant for this discussion is that there is a relatively large set of permissable CC onsets (including sequences like [kt], [xt], [mn], and so on), but codas are generally not allowed, especially in word-final position.

5.1.2 Final vowel alternations ("ablaut")

As described above, some Lakhota verbs have final vowel alternations between [e] and [a], in a process known as umlaut; the basic problem is to differentiate the following three types of

Table 5.1: Lakhota consonant inventory

unaspirated	*p*	*t*	*c* [tʃ]	*k*
aspirated	*ph* [pʰ]	*th* [tʰ]	*ch* [tʃʰ]	*kh* [kʰ]
ejective	*p'*	*t', s'*	*c'* [tʃ'], *sh'* [ʃ']	*k', x'*
fricatives		*s, z*	*sh* [ʃ], *j* [ʒ]	*x, gh* [ɣ]
nasals	*m*	*n*		*ng* [ŋ]
liquid		*l*		
glides			*y* [j]	*w*

Table 5.2: Lakhota vowel inventory

Oral		Nasal	
i	*u*	*ĩ*	*ũ*
e	*o*		
	([ɔ]*)		
([æ]*)	*a*		*ã*

*[æ] and [ɔ] are derived from /aya/, /awa/

words:

(64) Three surface patterns for final [a], [e]

*gleshk**a***	*gleshk**a**-pi*	'be spotted'	(invariant [a])
*washt**e***	*washt**e**-pi*	'be good'	(invariant [e])
*chep**e***	*chep**a**-pi*	'be fat'	(alternating [e]∼[a] = "ablaut")

Pursuing an underspecification approach along the lines of Inkelas (1994), we would start by inferring that alternating verbs like *chep*{*e*∼*a*} must end in something other than [e] or [a]. Using the strategy of creating an archiphoneme with just the shared feature specifications, this would lead us to conclude that such verbs end in an abstract segment that I will write as /A/:[3]

[3]An alternative analysis, suggested by Kim (2002), is that alternating *e*∼*a* is not underspecified, but rather *overspecified*, including not only [a]-features, but also a floating dorsal feature which combines with a floating coronal (front) feature to yield [e] before certain suffixes. This suggestion, which is in line with Lieber's autosegmental approach to morphologically-conditioned mutations (Lieber 1987; Lieber 1992), is problematic in various respects. First, using Kim's feature system, we might expect the combination of [a] with coronal and dorsal features to produce [æ] rather than [e], particularly since the language already has a surface [æ] that results from coalescence of /aye/ and /aya/. Second, the representation with floating features is supposed to unify the ablaut alternation with another coalescence process, of /ai/ to [e]. However, /ai/ to [e] coalescence is not a productive process in the language—surface [ai] sequences can easily be created by combining, for example, the valence-adding prefix *a-* with the instrumental/locative prefix *i-*. In addition, there is another process, not discussed here, in which alternating *e*∼*a* raises to [i] before certain morphemes (such as the future marker *-(n)kte* and the conjunctive clitic *-na*), so we would need to find some other floating feature to attach to these morphemes, and also provide a mechanism to delete the place features of /a/ so that it can raise to [i] in this context. Finally, the floating feature representation cannot explain why ablaut verbs also behave differently in reduplication (section 5.1.3).

(65) Feature specifications of /a/, /e/, /A/

/a/	/e/	/A/
+syllabic	+syllabic	+syllabic
-high	-high	-high
+low	-low	
+back	-back	

Underspecification can be used to derive surface alternations quite naturally in OT; underspecified segments have less to be faithful to, so general principles of markedness (that are needed in the grammar anyway) can play a greater role in determining their surface realization without incurring faithfulness violations. In this case, the crucial markedness constraint is "no word-final [a]", which is admittedly rather language-particular. However, the pattern falls out easily with the following rankings: first, all of the IDENT constraints for vowel features are ranked at the top of the grammar, forcing the surface form to preserve whatever feature values have been specified underlyingly (66a). Second, the general markedness constraint banning [e] is ranked above the constraint banning [a], and the language-particular constraint banning word-final [a] is ranked above both of these, forcing underspecified vowels to be realized as [e] word-finally, and [a] elsewhere (66b).

(66) Ident(V) ≫ *[a]/___# ≫ *[e] ≫ *[a]

a. Ident(V): violated when an underlyingly specified vowel is changed (/a/→[e], /e/→[a])

/gleʃka/ 'spotted'	Ident(V)	*[a]/___#	*[e]	*[a]
☞ a. [gleʃka]	√	*	*	*
b. [gleʃke]	*!	√	**	√

/waʃte/ 'good'	Ident(V)	*[a]/___#	*[e]	*[a]
a. [waʃta]	*!	*	√	**
☞ b. [waʃte]	√	√	*	*

b. Ident(V) satisfied by both /A/→[a] and /A/→[e]; realization falls to markedness constraints

/tʃʰepA/ 'do, make'	Ident(V)	*[a]/___#	*[e]	*[a]
a. [tʃʰepa]	√	*!	√	**
☞ b. [tʃʰepe]	√	√	*	*

/tʃʰepA-pi/ 'do, make'	Ident(V)	*[a]/___#	*[e]	*[a]
☞ a. [tʃʰepa-pi]	√	√	√	**
b. [tʃʰepe-pi]	√	√	*!	*

This analysis captures the pattern of final vowel ablaut, but requires positing an abstract, underspecified archiphoneme. Under a traditional approach to UR discovery, the mere existence of a three-way contrast is sufficient evidence for learners to infer that they need an abstract segment (either an archiphoneme, or a fully specified segment that never surfaces as such). Ideally, however, we might like some external evidence confirming this analysis, such as an indication that final ablaut vowels behave differently from nonalternating [e] and [a] in other respects as well. In fact, there is such evidence, in the form of differences in reduplication patterns.

5.1.3 Reduplication

Verbs can reduplicate in Lakhota, with a variety of meanings. In many cases, reduplication marks plurality, especially with stative verbs (67a). In other cases, it marks intensivity/iterativity/durativity (67b), while in other cases, the meaning is not so clear (67c).

(67) Meaning of reduplication in Lakhota

a. Plurality (mainly statives)

sha-sha	'red-pl.'	
washte-shte	'good-pl.'	

b. Intensive/iterative/durative

yushna-shna	'sprinkle'	(cf: *yushna* 'drop')
naphã-phã	'trample'	(cf: *naphã* 'stomp')
lowã-wã hiyaye	'went along singing'	(cf: *lowã* 'sing')

c. Meaning not so clear

gleshka-shka	'checkered/plaid'	(cf: *gleshka* 'spotted')

The basic pattern of reduplication is to copy the final syllable, as seen in (68):

(68) Reduplication of the final syllable

3sg	redup.	gloss
gleshka	*gleshka-shka*	'be spotted'
washte	*washte-shte*	'be good'
lowã	*lowã-wã*	'sing'
naxcha	*naxcha-xcha*	'blossom'
shakpe	*shakpe-kpe*	'be six in number'
yamni	*yamni-mni*	'be three in number'
zaptã	*zaptã-ptã*	'be five in number'
shakowĩ	*shakowĩ-wĩ*	'be seven in number'
wikcemna	*wikcemna-mna*	'be ten in number'

However, verbs with final *e* ∼ *a* alternations generally copy the "maximal penult",[4] sometimes with accompanying segmental changes (devoice fricatives, change /t/ → [l], etc.), as seen in (69):

(69) "Non-final reduplication"

3sg	redup.	gloss
chepe	*chep-chepe* (**chepe-pe*)	'be fat'
kaghe	*kax-kaghe* (**kaghe-ghe*)	'do, make'
khate	*khal-khate* (**khate-te*)	'be hot'
naphope	*na-pho-phope* (**naphope-pe*)	'pop'

The traditional analysis of this difference (Boas and Deloria 1941; Shaw 1980) is that final alternating (ablaut) vowels are completely absent underlyingly: /tʃʰep/, /kaɣ/, etc. Under this

[4] The 3sg form *chepe* is syllabified *che.pe*, so reduplicating just the penult should yield *che-chepe*. In this and many other cases, reduplication ignores syllabification of the base form, and copies as much as it can fit into a syllable.

analysis, the URs of these words have codas, which are prohibited on the surface (section 5.1.1). The illegal codas are then fixed by a process of epenthesis, which inserts an [e] word-finally ([tʃ^{h}ep**e**]), and an [a] word-internally ([tʃ^{h}ep**a**-pi]).[5] This allows us to say that reduplication is always final, and precedes epenthesis:[6]

(70) Rule ordering: reduplication precedes epenthesis (after Shaw 1980)

a. Simple forms

UR	/waʃte/	/gleʃka/	/tʃ^{h}ep/
REDUPLICATION	—	—	—
EPENTHESIS	—	—	tʃ^{h}ep**e**
SR	[waʃte]	[gleʃka]	[tʃ^{h}epe]

b. Reduplicated forms

UR	/waʃte-RED/	/gleʃka-RED/	/tʃ^{h}ep-RED/
REDUPLICATION	waʃte-**ʃte**	gleʃka-**ʃka**	tʃ^{h}ep-**tʃ^{h}ep**
EPENTHESIS	—	—	tʃ^{h}ep-tʃ^{h}ep**e**
SR	[waʃte-ʃte]	[gleʃka-ʃka]	[tʃ^{h}ep-tʃ^{h}epe]

Treating ablaut vowels as epenthetic is a more radical version of the underspecification analysis sketched above. The claim is that not only the distribution of [a] and [e] but the very occurrence of the vowel is predictable based on surface markedness considerations. The analysis of ablaut alternations would be much the same as in (66b) above, with the addition of a high-ranking *CODA constraint, and constraints ruling out the insertion of vowels other than [e] and [a].[7]

The epenthesis analysis has some obvious advantages. First, it captures the co-occurrence of two properties of words like *chepe*: they have final vowel alternations, and they have non-final reduplication. Furthermore, all words are completely rule-governed. If a speaker knows that there is an epenthesis process (resulting in [e] word-finally and [a] before a morpheme boundary), a final reduplication process (rendered opaque by epenthesis), and two types of URs (those with final consonants and those with final vowels), then it is possible to use the grammar to derive all of the surface forms correctly (70).

Let us now consider the various possible sources of acquisition-related error under this analysis, as we did for Yiddish in section 2.5. Suppose that a learner is faced with a new word, whose forms are not completely known. For example, suppose she hears a new 3sg form *pughe*

[5] This analysis recapitulates the history of verbs with ablaut alternations. It appears that Siouan did originally have consonant-final and vowel-final verbs, but at some point two post-verbal clitics (*-a* and *-e*) were reanalyzed as part of the verb stem, or as epenthetic vowels inserted to fix word-final codas: *chep-e* ⇒ *chepe* (Rood 1983).

[6] It is not easy to recast this analysis of the reduplication facts into OT. Intuitively, we want to penalize copying an epenthetic vowel, but Base-Reduplicant (BR) correspondence constraints do not know which base segments have incurred IO faithfulness violations (such as a DEP violation). The only other possibility is to use Input-Reduplicant (IR) correspondence to penalize having an epenthetic vowel in the reduplicant; however, this would require ranking DEP-IR above DEP-IO, which leads to undesirable typological consequences (McCarthy and Prince 1995, pp. 114-117). I will not pursue this problem here, since I will ultimately be arguing that the "ablaut vowel as epenthesis" analysis is wrong in any case.

[7] This could be accomplished either by faithfulness, with DEP-IO(i,u,o,æ,ĩ,ũ,æ̃), or else by markedness, with *[i], *[u], *[o], etc. The former approach looks more promising, since it seems questionable to claim that [i] is a more marked vowel than [e], which would be required in the ranking *[i] ≫ *[e].

'he snorted' – what might she conclude? One possibility is that she may assume that the [e] is underlying, setting up a UR /puɣe/ and predicting a plural form *pughe-pi* and a reduplicated form *pughe-ghe*. Another possibility is that she may assume that the [e] is not underlying (/puɣ/), and predict a plural form *pugha-pi* and a reduplicated form *pux-puɣe*. Conversely, suppose that the learner has heard a new 3pl form *puza-pi* 'they are dry.' In this case, she may either assume that the [a] is underlying (predicting 3sg *puza*, reduplicated *puza-za*), or she may assume that the [a] is epenthetic (predicting 3sg *puze*, reduplicated *pus-puze*).

As in chapter 2, it is difficult to make exact predictions about which errors we expect under a traditional model without an explicit theory of how learners reason about URs with incomplete information. A reasonable default assumption would be that learners do not posit underlying underspecification unless they have heard evidence that the word actually alternates. This is the principle behind the Prague School's use of archiphonemic underspecification, and it is also the principle behind Lexicon Optimization in OT (Prince and Smolensky 1993). In the present case, that would mean that learners with incomplete information would always set up a fully specified vowel, but that vowel should sometimes be /a/ and sometimes /e/, depending which form had been learned. A more subtle assumption is that speakers know the predominant patterns of their lexicon, and if the dominant pattern is alternation, then they are able to set up underspecified URs without actually hearing the alternation. This has been proposed by Inkelas (1996) as Alternant Optimization, and by Harrison and Kaun (2000) as Pattern-Responsive Lexicon Optimization. In the present case, if ablaut is the dominant pattern, then Pattern-Responsive Lexicon Optimization might lead learners to assume that partially-known words are underspecified in such a way that produces ablaut alternations and non-final reduplication.

Crucially, all of these theories share a common prediction: no matter what principles the learner uses to set up a UR using incomplete information, the result should resemble a valid existing paradigm. In particular, if she assumes that the final vowel of a word is underlying, then it should be invariant, and the word should have final reduplication. If, on the other hand, she assumes that the final vowel is underspecified, it should exhibit the ablaut alternation, and have non-final reduplication. In the next section, I will show that this prediction is wrong. As it turns out, two new "inconsistent" paradigm types have been created in Lakhota, both of which are incompatible with the analysis laid out thus far. After presenting the data, I will show that although these new paradigm types are unexpected under any version of the traditional analysis, they are in fact predicted by a single surface base approach.

5.2 Innovative paradigms in Lakhota

The verb types that I have discussed thus far are those that have a straightforward historical origin. In addition to the two paradigm types discussed so far (invariant vowels with final reduplication, and ablaut alternations with non-final reduplication), there have also arisen two innovative paradigm types in Lakhota. The first are paradigms with variant final vowels (ablaut), but with final reduplication, as in (71):[8]

(71) Innovative paradigm type 1: ablaut plus final reduplication

3sg	*3pl*	*redup*		*gloss*
hãske	*hãska-pi*	*hãska-ska*	(**hã-hãske*[9])	'be tall'
hĩshme	*hĩshma-pi*	*hĩshma-shma*	(**hĩ-hĩshme*)	'be fuzzy'
ixat'e	*ixat'a-pi*	*ixat'at'a*	(**i-xa-xat'a*)	'laugh'
hoxpe	*hoxpa-pi*	*hoxpa-xpa*	(**hox-hoxpe*)	'cough'
naxme	*naxma-pi*	*naxma-xma*	(**nax-naxme*)	'hide'
kaxpe	*kaxpa-pi*	*kaxpa-xpa*	(**kax-kaxpe*)	'knock down'
katke	*katka-pi*	*katka-tka*	(**kal-katke*)	'choke'

The second innovation is a paradigm type with invariant final vowels, but non-final reduplication, as in (72):

(72) Paradigms with invariant final V, but non-final reduplication

3sg	*3pl*	*redup*		*gloss*
thokca	*thokca-pi*	*thok-thokca*	(**thokca-kca*)	'be different'
topa	*topa-pi*	*top-topa*	(**topa-pa*)	'be four in number'
ota	*ota-pi*	*ol-ota*	(**ota-ta*)	'be many'

It appears, then, that there have been been two changes, leading to the creation of two new paradigm types:

- The *-a*/*-e* alternation has been extended to some verbs that used to have invariant *-a* (**hinshma* ⇒ *hinshme* 'fuzzy-3sg')
- Nonfinal reduplication has been extended to some verbs that should have had final reduplication

These innovations are significant for two reasons. The first reason is that words belonging to the new paradigms are incompatible with any UR in the old system. The contradiction is

[8]There are several sources of evidence that these patterns are in fact innovative, and that the *-a*/*-e* alternation has been extended to forms which originally did not have it. First, there are verbs whose only vowel is an ablaut vowel (e.g., *t'e*∼*a* 'die'), and if ablaut vowels originated as reanalyzed clitics (fn. 5), then we would be forced to infer that these verbs were originally just a single consonant (*t'*). It seems more plausible to say that these verbs were originally CV (*t'a*), and that the ablaut alternation has been extended to them analogically – especially since there are sometimes words that appear to be etymologically related and have invariant *-a*, such as *t'at'a* 'listless, lazy.' In addition to this, some forms listed with *-a* in Boas and Deloria (1941) are now more common with *-e*/*-a* (e.g., *naxma* 'fled-3sg' ⇒ *naxme*). Finally, I have observed a fair amount of synchronic uncertainty or variation in whether a final *-a* should alternate or not, including even the use of both *-a* and *-e*/*-a* on the same verb in the same session. It should be noted, however, that some "impossible" forms also seem to be rather old – for example, *yatkan/e* is found in all sources and shared with other dialects, but appears to be a relatively local innovation in this branch of Siouan (Shaw 1980; Rood 1983). While it is interesting that this pattern is spreading to more and more verbs over time, what I am really concerned with here is what mechanism allowed the very first inconsistent paradigms to be created.

[9]It should be noted that in all cases, the non-occurring reduplications are phonotactically legal – so although on first glance, we might think of trying to explain the nonoccurrence of forms like *hã-hãska* as avoidance of sequences like [hãhã], perhaps due to its intervocalic [h], in fact such sequences are permitted in other words, like *hũke-shni* ∼ *hũhũka-pi-shni* 'weak'.

illustrated in Fig. 5.3; the fact that this word has a final vowel alternation would lead us to conclude that the final vowel is not specified underlyingly, while the fact that the final syllable reduplicates would lead us to conclude that the final vowel is present underlyingly.

The second reason is that the changes leading to new paradigm types have been *asymmetrical*; they have affected only words with original *-a* throughout the paradigm, and not *-e*. Thus, there are plenty of words like *hanske* which have switched from invariant *a* to alternating *e* ∼ *a* (73a), but no words that have switched from invariant *e* to alternating *e* ∼ *a* (73b).

(73) Changes have been asymmetrical

a. Attested:
naxma ⇒ *naxm**e***   *naxma-pi*   *naxma-xma*

b. Not attested:
washte   *washte-pi* ⇏ **washt**a**-pi*   *washte-shte*

Thus, the Lakhota change poses two mysteries: first, how were new, "internally inconsistent" paradigms created? (This was not predicted in any of the incomplete learning scenarios discussed in the previous section.) Second, why were only [a]-final verbs affected? In the next section, I will show that both of these mysteries can be explained under a model that limits learners to choosing URs that match a particular surface form (the single surface base hypothesis). Under this restriction, learners are not always able to set up a UR that preserves all surface contrasts; in fact, in the case of symmetrical neutralizations, neither form alone can predict the paradigm of a word. As with Yiddish and Latin, I will compare the various forms in the Lakhota verb paradigm and see whether there is a form that, while not preserving *all* contrasts, at least preserves more contrasts than any other form. It will emerge that once both phonological and morphological neutralizations are taken into consideration, there is such a form (a second person form). Moreover, when we consider the grammar that would be needed to derive the remainder of the paradigm from the second person, it predicts two types of overregularization: extending ablaut and non-final reduplication to [a]-final verbs.

5.3 Restricting UR discovery to a single surface form

Let us now go back to the beginning, this time operating under the single surface base restriction. Recall that the basic analytical problem in (61) (repeated below), is that there are three

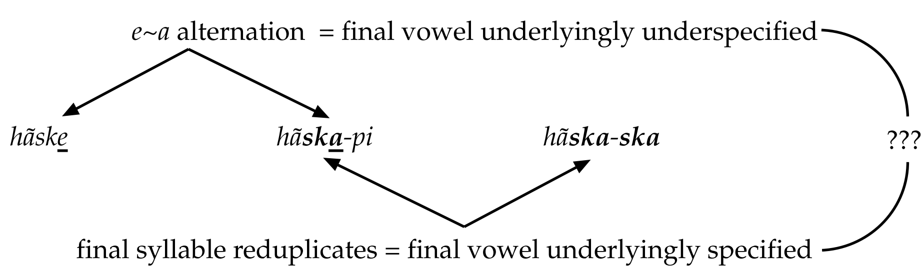


Figure 5.3: Innovative forms are incompatible with any UR

surface patterns, but only two phonemes involved. The challenge, therefore, is to come up with an underlying form for *chep*{*e*∼*a*}.

(61) Three surface patterns

*gleshk**a***	*gleshk**a**-pi*	(invariant [a])
*washt**e***	*washt**e**-pi*	(invariant [e])
*chep**e***	*chep**a**-pi*	(alternating [e]∼[a])

Under the single surface base restriction, we are now limited to choosing either /tʃʰepe/ or /tʃʰepa/. This leaves us with a number of possible analyses. We could, for example, choose /tʃʰepa/ with underlying /a/, and then posit a final raising rule (or its OT equivalent), as in (74):

(74) FINAL RAISING (ablaut): /a/ → [e] / ___#

This analysis would correctly derive [tʃʰepe] and [kaɣe] from their underlying forms /tʃʰepa/ and /kaɣa/, but it would fail for [gleʃka], incorrectly predicting the raised form *[gleʃke]. Thus, under this analysis, we would have to list [gleʃka] as an exception, which would block grammatically expected form [gleʃke]. Conversely, we could assume that there is no default final raising rule, and then list words with raising as exceptions ([tʃʰepe], [kaɣe]), or make Final Raising a lexically restricted rule, and mark /tʃʰepa/ and /kaɣa/ with [+Final Raising] diacritics. This is not an exhaustive list of all of the possible analyses, but it should be clear that no matter which UR we pick (/tʃʰepe/ or /tʃʰepa/), there will be some exceptions. The reason is that now we have only two URs available (/a/, /e/) to represent three surface patterns ([a], [e], [a]∼[e]). Some unpredictable information is going to have to be stored somewhere else, and that somewhere is the exception handling mechanism.

Once we recognize that exceptions are unavoidable, we can at least try to mitigate the problem by finding the set of URs and rules that requires the fewest listed exceptions. In order to do this, we will want to base the UR on the part of the paradigm that is "most informative" — that is, that has the fewest neutralizations, affecting the fewest lexical items. In order to assess this for Lakhota, we need to consider the neutralizations that might affect Lakhota verbs, and how many verbs are affected by each.

5.3.1 What is the most informative part of the Lakhota paradigm?

In order to evaluate the seriousness of various neutralizations in Lakhota, I selected a database of "simple" verbs. I began with the list of verbs compiled by Munro (1989), and then removed all entries that were morphologically complex according to one of the following criteria. First, I removed all "compound" entries, consisting of a combination of a verb plus verb, noun plus verb, preposition plus verb, and so on (e.g., *akan ishtima/e* 'sleep on', from *ishtima/e* 'sleep'). Next, I removed all entries derived by the valence-changing prefix *a-* (adds one argument), the causative suffixes *-ye* and *-khiye*, the reflexive marker *-c'i-*, and the possessive object marker *ki-* ('X one's own ___'). For example, the verb *akipsica/e* 'to jump over one's own' is derived by prefixing the valence-changing *a-* and possessive object *ki-* to *psica/e* 'jump'). I left in verbs containing derivational prefixes that are identifiable, but not predictable, such as *pa-* 'using hands', *na-* 'using feet', *ya-* 'using the mouth', etc. These prefixes are analogous to English *trans-* or *dis-*, in that they are easily segmented out as prefixes, but they are not productive,

Table 5.3: Number of words in each ablaut category (out of 545 total in database)

Pattern	*Count*
invariant -*a*	83
ablaut -*e*/-*a*	199
invariant -*e*	65
nasal ablaut -*e*/-*ã*	9
invariant -*ã*	41

they do not occur with all roots, and some verb roots that occur with them are bound roots. When these criteria were applied, a database of 545 simple verbs remained.

There are various sources of systematic unpredictability in Lakhota verbs. These include phonological unpredictability, such as whether or not a verb has ablaut alternations, and also morphological unpredictability, such as the location of person agreement. In addition to these wide-scale, systematic sources, there are also other sporadic irregularities that affect just a few verbs, and will not be discussed here.

Phonological unpredictability: ablaut

One major unpredictable property of a Lakhota verb is whether or not it has the ablaut alternation that has been the focus of discussion up until this point. In the examples thus far, I have limited the data to the three major patterns: invariant *a*, invariant *e*, and ablaut *e* ∼ *a*. There are, however, also a handful of words that display what I will call a *nasal ablaut* alternation between *e* and *ã*. Thus, the full range of possible surface patterns is as in (75).

(75) Ablaut alternations:

3sg (unsuffixed)	3pl (suffixed)	gloss	category
*gleshk**a***	*gleshk**a**-pi*	'spotted'	invariant *a*
*chep**e***	*chep**a**-pi*	'fat'	ablaut
*washt**e***	*washt**e**-pi*	'good'	invariant *a*
*yatk**e***	*yatk**ã**-pi*	'drink'	nasal ablaut
*yat**ã***	*yat**ã**-pi*	'light (a cigarette)'	invariant *ã*

Comparing the words in (75), we can see that both the unsuffixed 3sg and the suffixed 3pl suffer from neutralizations. The 3sg form neutralizes 3 types of words: invariant -*e*, ablaut -*e*/-*a*, and nasal ablaut -*e*/-*ã* all have -*e* in this form. Turning to the 3pl form, we see that 2 pairs of word types are neutralized: invariant -*a* and ablaut -*e*/-*a* are both -*a* in this form, and invariant -*ã* and ablaut -*e*/-*ã* are both -*ã*. Neither form is obviously better than the other in allowing us to predict which surface pattern a word should take; thus, as with Latin, we must compare the seriousness of the neutralizations by considering how many lexical items are affected by each.

The numbers of words instantiating each of the patterns in (75) are given in Table 5.3. As can be seen from the table, the (non-nasal) ablaut pattern is well represented, with almost 40% of verbs participating in it. There are also a fair number of invariant *a* and *e* verbs, with relatively fewer invariant *ã* verbs, and just a handful of nasal ablaut verbs.

Given these counts, let us now consider how informative the 3sg and 3pl forms are in practice in predicting the remaining of the paradigm. If we use the singular (unsuffixed) form, we will have the following URs for the words in (75): /gleʃka/, /tʃʰepe/, /waʃte/, /jatke/, and /jatã/. The problem here is the three forms with underlying /e/, which belong to three different surface classes. The majority of words with *e* in the 3sg are ablaut verbs with *a* in the plural (e.g., [tʃʰepa-pi]), so if our goal is construct a grammar that can cover a majority of forms, we need to posit some sort of derived environment non-final lowering rule: /e/ → [a] /___+C. With this rule in place, plurals with [a] like [tʃʰepa-pi] will be accounted for, and we just need to list non-lowerers like [waʃte-pi] (65 of them) and nasalizers like [jatkã-pi] (9 of them) as exceptions. Thus, choosing the 3sg as the UR would require 65 + 9 = 74 exceptions.

If, on the other hand, we were to use the plural (suffixed) form as the UR, we would have the following: /gleʃka/, /tʃʰepa/, /waʃte/, /jatkã/, and /jatã/. In this case, there would be two problems: the two verbs with underlying /a/, and the two with underlying /ã/. Among those with underlying /a/, the majority have [e] like [tʃʰepe] in singular, so we would need to posit a final raising rule (/a/ → [e] /___#). This would correctly derive /tʃʰepa/ → [tʃʰepe], but it would incorrectly predict *[gleʃke] for [gleʃka]. Therefore, we would need to list non-raisers like [gleʃka] as exceptions (83 exceptions). Among the underlying /ã/ words, the majority are invariant like [jatã], so we would not want to extend the final raising rule to cover nasalized vowels as well; rather, we would just list the nasal ablaut verbs like [jatke] as exceptions (9 exceptions).[10] Thus, choosing the 3pl as the UR would require 83 + 9 = 92 exceptions.

What we see from this comparison is that the unsuffixed (3sg) form is slightly better in predicting the final vowel of the suffixed (3pl) form than vice versa, requiring 18 fewer exceptions for this set of verbs (= 92 - 74). This advantage is rather small, however, and choosing the 3sg form as the UR relies on a rather questionable phonological rule (non-final lowering of /e/→[a] only before a suffix) in order to make ablaut verbs rule-governed. What I conclude from this section, therefore, is that the ablaut neutralization really is quite symmetrical, and any advantage that one form may have over the other will have to come from whatever other contrasts they may preserve.

Morphological unpredictability: person agreement

Another important unpredictable property of Lakhota verbs is the position of person agreement. Lakhota verbs fall into two classes, based largely (but not entirely) on whether they are active or stative. The subject markers for these two classes of verbs are given in Table 5.4. Note that *-pi* is a plural suffix for animate subjects, and therefore shows up in all of the plural cells; *-he*/*-ho* is a second person suffix, used in questions and second person declarative sentences (-*he* by female speakers, *-ho* by male speakers). Therefore, the the 2sg, 1pl, and 2pl forms usually occur with a suffix, as does the 3pl if it has an animate subject.

Membership of a verb in the active or stative class is more or less predictable given the meaning of the word; the position of the person agreement within the verb, on the other hand, is not. Subject markers in Lakhota may occur either as prefixes or as infixes, depending on the

[10]Shaw (1980) also treats nasal ablaut verbs as exceptions, marking them diacritically to take the /a/ → [e] ablaut rule even though they do not strictly provide the input for this rule, which is [a].

[11]When the 1sg marker *wa* occurs before a *y*, there is a morphophonological process that turns the *wa-y* sequence into *bl* – e.g., *wa-yatke* → *blatke* 'I drink'.

[12]When the 2sg marker *ya* occurs before a *y*, the *ya-y* sequence becomes *l* – e.g., *ya-yatke* → *latke* 'you drink'.

Table 5.4: Lakhota subject markers

a. Active (Munro's Type I)

	sg	*pl*
1st	wa[11]	un(k) ... pi
2nd	ya[12] ... he/ho	ya ... pi he/ho
3rd	∅	∅ ... pi

b. Stative (Munro's Type II)

	sg	*pl*
1st	ma	un(k) ... pi
2nd	ni ... he/ho	ni ... pi he/ho
3rd	∅	∅ ... pi

verb. Although I am not aware of any actual minimal pairs that differ only in the placement of person agreement, the verbs for 'to be lost' and 'to walk' in (76) are very similar phonologically, but get their subject markers in different positions.

(76) Variable position of subject markers

a. Sometimes prefixed

'be lost'	*sg*	*pl*
1st	**wa-nuni**	**un-nuni-pi**
2nd	**ya-nuni he**[13]	**ya-nuni-pi he**
3rd	nuni	nuni-**pi**

b. Sometimes infixed

'walk'	*sg*	*pl*
1st	ma-**wa-ni**	ma-**un-ni-pi**
2nd	ma-**ya-ni he**	ma-**ya-ni-pi he**
3rd	mani	mani-**pi**

The unpredictable location of person agreement is complicated even further by the fact that a small number of verbs take infixed person agreement in general, but prefixed agreement in the 1pl; for example, the verb *ahi* 'to bring someone somewhere':

(77) Mismatched location of person agreement:

'bring someone somewhere'	*sg*	*pl*
1st	a-**wa-hi**	**unk-ahi-pi (*a-un-hi-pi)**
2nd	a-**ya-hi he**	a-**ya-hi-pi he**
3rd	ahi	ahi-**pi**

In addition, there is occasionally free variation in the position of agreement for a single verb (e.g., *un-nawizi-pi* ∼ *na-un-wizi-pi* 'we are jealous', 3sg *nawizi*). Finally, there are a few words that take agreement in two locations simultaneously in the 1sg, 2sg, and 2pl (but not the 1pl). These complications affect relatively few forms, however, and including them would not influence the choice of base. Therefore, I will omit them from this discussion.

What does the variable position of person agreement mean for base or UR selection? The number of verbs with prefixing or infixing person agreement are summarized in Table 5.5; as it turns out, there are significant numbers of both prefixing and infixing verbs, so this is a

[13] For simplicity, I will list second person forms with just *-he*, as they would be said by a female speakers. The male version simply substitutes *-ho* instead.

Table 5.5: Number of words in each agreement pattern

position	*count*
prefixed	347
infixed	183
infixed, 1pl prefixed	12

serious neutralization. If we were to choose a third person form as the base, we would lose all information about where subject marking should go. We could then assume that agreement is prefixing by default, but this would force us to list 195 exceptions for the verbs in which it is infixed.

Fortunately, forms other than the third person reveal the position of person agreement more clearly, to varying degrees. The 1sg form unambiguously reveals the location of agreement and would allow us to project all other forms, except in two cases. The first is when the 1sg marker happens to be identical with the beginning of the verb root, as in *wawachi* 'I dance'. In these cases, it is impossible to tell whether the subject marking is the first *wa* (***wa**-wachi*) or the second *wa* (*wa-**wa**-chi*). This ambiguity, which I will call the *wawa* problem, is more pervasive than one might imagine; it affects 24 verbs in the database of 545 "basic" verbs. Furthermore, although I am unable to quantify it, the *wawa* problem probably affects many more verbs than this in practice, because *wa-* is a productive prefix used to mark indefinite objects. The other case for which the 1sg form may be misleading is for the 12 "mismatch" verbs (77). For these verbs, the 1sg form would lead one to believe the person agreement should be infixed in the 1pl, but in fact it is exceptionally prefixed in this form. The upshot is that the 1sg form is much more informative about the position of person agreement than a third person form, but it is not perfect.

In the 1pl, there is an ambiguity analogous to the *wawa* problem, which occurs when the 1pl marker *un(k)* is added to a verb that already begins with *un(k)*, such as *unk-unpa-pi* 'we smoke' (the "*unkun(k)*" problem). This problem affects only six verbs in the database, which is probably an accurate estimate, because unlike *wa-*, there are no prefixes homophonous to *unk-* in the language. However, there are two other problems with the 1pl as a potential base form. The first is the set of "mismatch" verbs discussed above; these are prefixed in the 1pl, but infixed in the remainder of the paradigm. The second problem is that the 1pl subject marker is identical for the active verbs (5.4a) and the stative verbs (5.4b). As previously discussed, this is not a serious problem in most cases, because it is usually possible to predict which class a verb belongs to based on its semantics. Nonetheless, there will still be a residue of verbs that require memorization, and listing the 1pl form would not help in these cases. This number is small, and I will leave it unquantified, since quantifying it would require a specific semantic analysis of the distinction between these two series of verbs, and a word-by-word count of which verbs fit the analysis and which do not.

Finally, let us consider the second person forms. In theory, one would expect these forms to suffer from a *yaya* problem, exactly analogous to the *wawa* and *unkun(k)* problems. However, there is a morphophonological process turning /*ya-y*/ into [l] (see fn. 12), so prefixing *ya* to a *ya*-initial root does not yield an ambiguous *yaya* sequence. Therefore, the only case in which second person forms are ambiguous with respect to the position of person agreement is for the 12 mismatch verbs, which have a different location for marking in the 1pl. This makes the

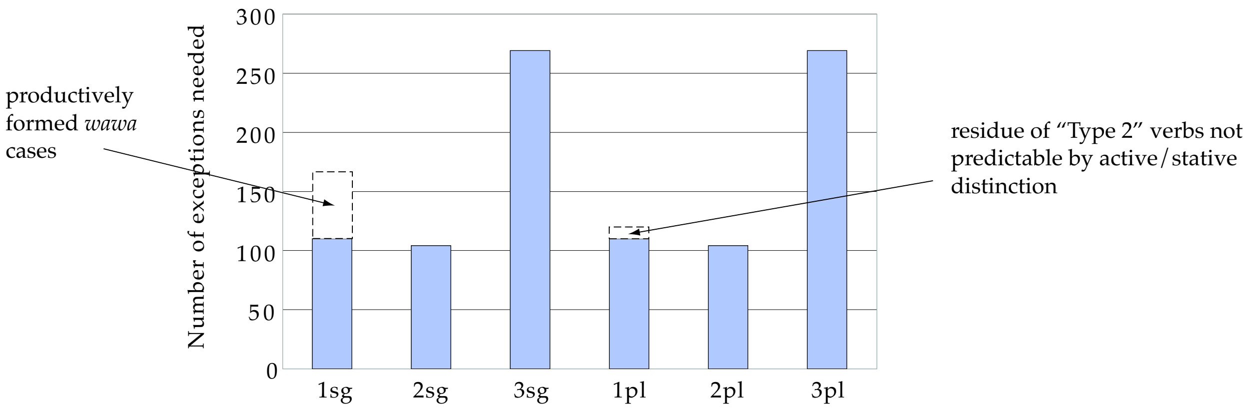


Figure 5.4: Comparing exceptions needed for each possible source of URs

second person forms the most informative, by a small margin, for purposes of predicting the location of subject marking.

Summary of unpredictability

When we compare the problems of predicting ablaut alternations and predicting the position of person agreement, we see that different forms have different advantages. The unsuffixed forms have a small advantage for maintaining ablaut contrasts (in particular, the contrast between invariant *a* and ablaut *e*∼*a*). However, this small advantage is far outweighed by the need to choose a base that reveals the position of person agreement. Unpredictable infixation favors choosing a first or second person form; moreover, the *wawa* and *unkun(k)* problems make the 1sg and 1pl forms problematic, while accidental facts about the language mean that there is no equivalent *yaya* problem affecting the second person. Therefore, this leads us to select a second person form as the all-around most informative part of the paradigm. This comparison is summarized in Figure 5.4.

Typologically, second person forms do not seem to serve as bases as often as third or first person forms; in fact, Bybee and Brewer (1980) hypothesize that second person forms might never serve as bases. However, I am aware of several other cases in which a second person form is claimed to be the base form. One is a change that occurred in the history of Eastern Scandinavian, discussed by Kuryłowicz (1947), in which a regular sound change made some verbs homophonous in the 2sg and 3sg, and this homophony was subsequently extended to all verbs in the language by replacing the 3sg forms with 2sg forms. Another type of evidence that second person forms can sometimes be bases is the fact that grammars sometimes describe the verbal inflection of a language by starting with a statement like "the root of the verb is the 2sg imperative"; this is the case in Tamil (Saravanan 2000), among others. Thus, positing that a second person form is the base in Lakhota does not seem completely anomalous.

The purpose of this section has been to show that a second person form is the most "informative" member of the Lakhota verb paradigm, and would thus be selected as the base, or UR, by a model that operates under the single surface base restriction. What remains to be shown, then, is that this makes the right prediction for the subsequent paradigmatic innovations discussed in section 5.2.

5.3.2 Consequences of using a 2nd person form as the UR

Suppose that you are a Lakhota learner, seeking the form in the paradigm with the most information about phonological and morphological properties of words. For verbs, this turns out to be the form found in the second person. We are now in a position to construct a grammar to derive the rest of the paradigm.

First, we must consider what the bases will be on this analysis. The second person forms are suffixed with the clitic *-he*, so alternating verbs have [a] (or [ã]) in this form. Therefore, the bases of alternating words will have /a/ or /ã/:

(78) Bases for Lakhota, under the single surface base restriction:

alternants	base	gloss
gleshka	ni-gleʃka-he	'be spotted'
chepe ~ *chepa-*	ni-$t\int^{h}$epa-he	'be fat'
washte	ni-waʃte-he	'be good'
yatke ~ *yatkã-*	latkã	'drink'
yatã	latã	'light (a cigarette)'

Note that this could also be translated into a model that seeks to discover underlying forms of verb roots by factoring out the person marking (removing *ni*, changing *l* to *y*, and removing the *he* suffix):

(79) URs under the single surface base restriction:

alternants	UR	gloss
gleshka	/gleʃka/	'be spotted'
chepe ~ *chepa-*	/$t\int^{h}$epa/	'be fat'
washte	/waʃte/	'be good'
yatke ~ *yatkã-*	/yatkã/	'drink'
yatã	/yatã/	'light (a cigarette)'

If the bases or URs of ablaut verbs like 'to be fat' have an underlying /a/, then we will also need a raising rule (or its OT equivalent) to derive the unsuffixed forms (/$t\int^{h}$epa/ → [$t\int^{h}$epe]), as in (74) above. This rule does not apply to /e/-final words like *washte*, and it correctly derives *chepe* from an underlying /a/. Words with invariant *a*, on the other hand, will need to be listed as exceptions to raising, to prevent incorrect unsuffixed forms like **gleshke*.

In addition to final raising, we will need two separate reduplication rules, since the difference between final and non-final reduplication (*chep-chepe* vs. *gleshka-shka*) can no longer be analyzed as a difference in their underlying forms, which both end in /a/ (/$t\int^{h}$ep**a**/, /gleʃk**a**/). It appears that there are simply two competing reduplication processes: one favoring final reduplication, and one favoring non-final reduplication.[14] Among /a/-final words, non-final reduplication is predominant, since there are more verbs like *chep-chepe* than like *gleshka-shka*; therefore, the grammar should be set up so that non-final reduplication applies by default in this environment, and forms like *gleshka-shka* must be listed as exceptions. In other environments, final reduplication prevails, and is the default rule.

What are the predicted errors if a word is not fully known? Suppose that a speaker has heard a 2sg form *ya-hoxpa he* 'you are coughing/are you coughing?', and has learned it as a base form, or has set up the UR /hoxpa/ for this verb. Since there is no evidence on the basis of this form alone that the verb is an exception to the Final Raising (ablaut) rule, the speaker will incorrectly apply raising to this verb, deriving the (etymologically) incorrect 3sg form **hoxpe*. Suppose, on the other hand, that the speaker has heard only a 3sg form of a verb, such as *kaze* 'he scoops'. In this case, the base form is not available, so the speaker simply memorizes this surface form and sets up no base or UR for the verb. (In the next section, I will discuss at greater length the idea of inferring nothing from non-basic forms.) Without a base, there is no way to derive an incorrect "undoing" of final raising, to predict incorrect suffixed **kaza-pi* (3pl) or **ya-kaza he* (2sg). There is no way to extend the [e]∼[a] alternation to invariant /e/ verbs (**washt**a**-pi*), since they have [e] in the base form, and [e] in the base always corresponds to [e] in the rest of the paradigm. Thus, there is an asymmetry: the only predicted error is on /a/-final verbs, by failing to learn that they are exceptions to final raising, and regularizing them to have final *e*∼*a* alternations. This is in fact the first innovation, shown in (71) on p. 84.

There is a similar asymmetry in the predicted reduplication errors. Suppose that a speaker has heard a ambiguous verb only in the 2sg, such as the (hypothetical) 2sg form *ya-t'apha he*. In this case, she would set up a base, or infer a UR /t'ap^{h}a/. The default reduplication pattern for /a/-final verbs is non-final reduplication, so in the absence of evidence that this verb takes final reduplication, she will apply the default (*t'ap-t'apha*). Suppose, on the other hand, that the speaker has heard an ambiguous verb only in the 3sg, such as the hypothetical 3sg form *sophe*. In this case, no base form has been learned, meaning there is no way to derive any reduplicated form (*sop-sophe* or *sophe-phe*). There is no way to apply incorrect final reduplication to ablaut verbs, since verbs with /a/ take penultimate reduplication, and ablaut verbs have /a/ in the base form. Furthermore, there is no way to derive incorrect non-final reduplication for invariant /e/ verbs, because they have /e/ in the base form, and final reduplication is the default for verbs that do not end in /a/. As a result, the only predicted reduplication error is for invariant /a/-final verbs, by failing to learn that they are exceptions to non-final reduplication, and incorrectly regularizing them to have non-final reduplication. This is the second innovation, shown in (72) on p. 84.

We see, then, that restricting bases or URs to a single surface form predicts only two types of errors, and both are attested in the new paradigm types in section 5.2. Furthermore, this ap-

[14]Nelson (to appear) points out that word-medial reduplication is a problem for OT because it does not satisfy either ANCHOR-L or ANCHOR-R. She goes on to argue that non-final reduplication patterns in cases like Lakhota are actually to be analyzed as stressed-syllable reduplication. This works for a majority of the Lakhota data, since most verb roots are di- or tri-syllabic, and stress is generally peninitial unless the second vowel is an ablaut vowel, meaning that the non-final syllable is usually (but not always) the stressed one for ablaut verbs. However, this analysis does not work completely; there are a number of verbs with nonfinal stress but final reduplication—e.g., [ˈhãske] ∼ [ˈhãska-ska] 'be tall', [wikˈdʒɛmna] ∼ [wikˈdʒɛmna-mna] 'be ten in number', etc. I do not have an alternative OT analysis of non-final reduplication at this time, but trust that it could be formulated somehow, perhaps using Nelson's insights about stress, or perhaps in some other fashion.

proach also explains the "de-coupling" of final vowel alternations from non-final reduplication, which were once predictably linked. In particular, final vowel raising and non-final reduplication are treated as the result of separate rules, rather than being derived from a common fact about underlying representations (underspecified final vowels). Since these are separate rules, they may each have their own lists of exceptions, and the fact that words like *gleshka* are an exception to both is purely an accident from the point of view of this analysis. If learners have evidence that a word is exceptional with respect to only one process, they may still regularize it with respect to the other. The result is "inconsistent paradigms", such as those that have arisen in Lakhota.

5.4 Inferring nothing from non-basic forms

A strong and perhaps uncomfortable assumption that was needed in the previous section was that if a speaker happens to have heard only non-basic forms of a word, she will memorize them as surface forms, but she will not infer a base form that can be used to derive other forms. As a consequence, there may be times when a speaker in some sense knows the word, but is unable to produce new forms of it. This assumption is potentially quite controversial—is there any way around it?

Consider a weaker version of the current theory, in which learners establish a base form by comparing the effectiveness of different forms in projecting the paradigm, but in which they retain the subgrammars needed to do mappings in *all* directions. Under this theory, learners prefer to derive forms using a base form as the input, since it is more reliable, but in the absence of such a form, they are able in a pinch to use a non-basic form as the input. This theory has some intuitive appeal, but it makes incorrect predictions about possible errors. In particular, it predicts that if Lakhota speakers happened to know only a 1sg or 3sg form, as must occur not infrequently, and that form ended in an *-e*, they would be able to reason backwards to infer that the suffixed form should end in *-a*, predicting errors like **washta-pi* instead of *washte-pi*. Similarly, in the case of Yiddish, if a speaker had heard an umlaut verb in only the 2sg or 3sg, she would have been able to project backwards to a 1sg with **e*, producing unattested errors like 1sg **fer* instead of *fɔr*.

Certainly, it would be difficult to argue that there is no such thing as backformation. However, asymmetries like these may show that it is not part of the ordinary, automatic workings of the synchronic morphological system.[15] The assumption that speakers infer nothing from

[15]Kiparsky (1982, pp. 21-22) makes the same claim for derivational morphology, following Marchand (1969). Given the fact that back-formations like *air-condition* do arise, Kiparsky and Marchand are forced to admit that back-formation exists, but only as a diachronic process. Kiparsky claims that synchronically, *air-condition* is the product of a N+V compounding process, which arose through reanalysis of N+N compounds ([*air* + [*condition*+*er*]]) as N+V+*er* compounds ([[*air* + *condition*] + -*er*]). This analysis is not totally satisfying, however, without a theory of possibly reanalyses; what allowed speakers (or learners) to reanalyze this form based on an unattested constituent? Crucially, whatever mechanism allows this reanalysis must *not* allow the reanalysis of [waʃte] as [waʃta] with final raising.

non-basic forms is needed here in order to explain the data.[16]

5.5 Other examples of inconsistent paradigms

Some readers may be wondering to what extent the changes discussed here are a result of the fact that Lakhota is an endangered language. The implied hope is that perhaps inconsistent paradigms arise only when the learning data is reduced or imperfect. Certainly, languages that are endangered experience far more radical and rapid changes than languages in which learners have access to a large sample of fluent monolingual speakers (see, e.g., Richards 2001 for a discussion of this in Lardil). Nevertheless, I believe that such factors merely facilitated the later stages of the Lakhota change, and that inconsistent paradigms can arise even in more stable environments.

For one thing, it appears that the changes discussed in this chapter probably began well before Lakhota was endangered. For example, the inconsistent paradigm of the verb *yatkã* 'drink' (*yatke*, *yatkã-pi*, *yatkã-tkã*) could have arisen only as an analogical extension of ablaut,[17] but it occurs in several related dialects that diverged before Lakhota was an endangered language (Rood 1983).

Furthermore, there seem to be examples of mixed behavior words in languages spoken more widely in monolingual environments. Tranel (1996) discusses one such case in French, in which a handful of indeclinable words behave like feminine forms in isolation, with their final consonants pronounced, but like masculine forms before a consonant-initial word, with the consonant deleted, as in (80). (See also L'Huiller 1999, p. 597.)

(80) Mixed behavior in French *huit* 'eight'

context	*petit* 'small' (masc.)	*huit* 'eight'	*petite* 'small' (fem.)
/ ___#V	*peti*[t]	*hui*[t]	*peti*[t]
/ ___#	*peti*[∅]	*hui*[t]	*peti*[t]
/ ___#C	*peti*[∅]	*hui*[∅]	*peti*[t]

Kenstowicz and Kisseberth (1977, p. 121) discuss a similar example from Chi-Mwi:ni, in which one exceptional verb behaves in some forms like it ends in a final /g/, and in others, like a final /k/. In chapter 6, I will discuss another possible example from Korean, which it appears that some nouns are more likely to appear with [t^h] in some forms, and with [s] in others.

I do not have an analysis of how such inconsistencies arose in French or Chi-Mwi:ni, nor do I have an estimate of how common such mixed-behavior or inconsistent words are in the world's languages. For present purposes, however, it suffices to note that the Lakhota case is not completely isolated, not does it appear to be a result of its current endangered status.

[16] A possible modification that would still explain the data would be to assume that whenever a speaker learns a new word in a non-basic form and does not know the base, she works her way backwards through the grammar to generate a set of possible base forms that could have yielded that derived form. If the set of possible bases has just one member, she infers it, otherwise she waits. Such a theory would allow speakers to set up underlying or base forms more rapidly, but strikes me as a rather perplexing strategy: why are speakers generally willing to guess about derived forms in the face of potential ambiguity, but not about base forms?

[17] The fact that the final syllable reduplicates and also the fact that it is nasalized indicate that it is etymologically an "underlying" vowel; if it had always been epenthetic, it would not be nasalized.

5.6 Local summary

In this chapter, I have shown that Lakhota presents an example of a three-way contrast ([e], [a], [e]∼[a]) that can be neatly described using archiphonemes or underspecification. This analysis is also supported by other facts in the language, since it can explain the co-occurrence of final vowel alternations and non-final reduplication. However, under this analysis, learners should always posit URs that produce "valid" paradigms, with the same set of properties as existing paradigms. This prediction is disproved by subsequent historical changes in Lakhota, which have resulted in the creation of two new paradigm types, inconsistent with any UR in the old system. These changes are puzzling not only because they have created novel paradigm types, but also because they have been asymmetrical: they have affected only verbs originally ending in invariant /a/. In section 5.3, I showed that by restricting learners to choosing a UR that matches a single surface form, and using the strategy of selecting the most informative part of the paradigm as the UR, we predict exactly these two errors and no others.

A consequence of this restriction is that learners are unable to capture certain generalizations about their language, such as the fact that ablaut and final reduplication are predictably linked with one another, since this cannot be deduced on the basis of any single form in the paradigm. It does allow them to capture other generalizations that the underspecification analysis does not allow, however, such as the fact that verbs that end in *-a* in suffixed forms tend to have *-e* in unsuffixed forms, and also tend to have penultimate reduplication. The historical evidence shows that these are in fact the generalizations that have been extended over time.

The Lakhota example complements the cases discussed in the previous chapters. In both Yiddish and Latin, alternations were leveled on the basis of other forms within the paradigm. In Lakhota, on the other hand, an alternation was extended, introducing new *e*∼*a* alternations into paradigms that did not originally have them. The difference between these cases, I have argued, is simply a difference in which pattern was predominant in the lexicon prior to the change. When the majority of words do not alternate, the best grammar to describe the language will not include a productive rule deriving the alternation, and exceptional alternating forms are open to replacement by overregularization to nonalternating forms. By contrast, when the majority of words alternate, it is more efficient to set up a rule producing alternations by default; in this case, we expect regularization to extend the alternation, rather than eliminating it.

This is a simple intuition, but as discussed at the end of the previous chapter, it is one that is not captured by the analysis of paradigm leveling as a universal preference for nonalternation. It is similar to the idea behind Harrison and Kaun's Pattern-Responsive Lexicon Optimization (discussed on p. 83), in that it allows speakers to assume that a new word has an alternation even if they have not actually heard it. Under Harrison and Kaun's proposal, however, there is still no reason to expect an asymmetry in the case of Lakhota. Alternating *-e*∼*-a* verbs outnumber both invariant *-e* and invariant *-a* verbs. Therefore, even if we assume that speakers may set up underspecified URs in response to the dominant patterns of the language, why would they do this only for *-a* verbs and not for *-e* verbs? What is missing is a theory of how "patterns" are defined; in the present case, the pattern is not merely a paradigm type with *-e* in some forms and *-a* in others, but rather a directional implication stating that if a word has *-a* in some forms, it will have *-e* in others. The single surface base restriction provides us with a theory of which patterns will be available to the speaker, namely, those involved in the mapping from the base form to the remainder of the paradigm.

If this analysis of Lakhota is correct, then the requirement that URs must obey the single surface base restriction has widespread implications for phonological analysis. In particular, it calls into question an assumption that dates back at least to Bloomfield and Trubetzkoy, that speakers may respond to patterns of alternation by setting up lexical representations with abstract phonemes that are unlike any surface realization. I will consider some of issues raised by this proposal in the next chapter.

Chapter 6

Discussion

In the preceding chapters, I have laid out a model of paradigm acquisition that starts by identifying a base form within the paradigm, and then develops a grammar of rules that take the base form as an input and derive the rest of the forms in the paradigm. I have pursued two specific hypotheses about how this is done: first, I have restricted the model to a single base form within the paradigm, and required further that the base form match a surface form from the same part of the paradigm for all lexical items (the "single surface base" hypothesis). Second, I have developed the idea that the learner selects the base form that is the most informative – that is, that permits the most efficient grammars for deriving the remainder of the paradigm. I have shown that this approach yields the right results in three cases. In chapter 2, I showed that for an older stage of Yiddish it selects the 1sg as the verbal base and correctly predicts the subsequent paradigm leveling. In chapter 4, I showed that for pre-classical Latin noun paradigms it selects an oblique form as the base, and correctly predicts the details of the *honor* analogy. Finally, in chapter 5, I showed that for Lakhota verb paradigms it selects a second person form as the base, and correctly predicts subsequent analogical changes.

This model is designed with two related tasks in mind. The first is to discover the inputs to morphology (the bases, or underlying forms) given a set of surface alternants. The second is to identify which form will serve as the base of analogical change in a particular language at a particular time. I have taken the strong position that the asymmetries we observe in analogical change are intimately related to the asymmetry between base forms and derived forms. In particular, I have assumed that analogy is a form of overregularization, in which exceptional forms are replaced with productively derived, grammatically expected forms.

It must be recognized that neither of these tasks is easy; both are fraught with problems, and both have been the subject of substantial discussion over the years. Furthermore, in both areas, the general consensus has been that it is impossible to provide an all-purpose discovery procedure. In the case of URs, a broad array of cases has been amassed showing that allowing URs to stray from a single surface alternant enables us to capture many generalizations that would have been missed otherwise (e.g., Kenstowicz and Kisseberth 1977; chap. 1; 1979, chap. 6). In the case of analogical bases, it has been shown that there are various generalizations that can be made about which forms are most likely to act as bases (isolation forms, morphosyntactically "unmarked" forms, forms with high token frequency, etc.), even if they are true only at the typological level (Hock 1991, pp. 234-237).

On the face of it, then, the current model appears to be inadequate on both fronts. It cannot construct abstract URs, and it does not have any built-in biases for selecting unmarked forms, frequent forms, or any other particular form. Thus, there are two types of data that it is in danger of not being able to handle satisfactorily. The first is the set of cases in which it has been argued that underlying forms must go beyond single surface forms, or basic alternants. The second is the typological data regarding bases of analogical change, discussed by Kuryłowicz (1947), Mańczak (1958), Bybee (1985) and others.

In the sections that follow, I will consider how the proposed model might nonetheless be able to handle each of these types of data.

6.1 URs

It has been accepted since at least the 1970s that there may be no general, automatic discovery procedure to find the UR that unifies all of the surface alternants and allows them to be projected using a reasonably natural and efficient set of phonological rules or constraints (Hyman 1975, pp. 90-98; Kenstowicz and Kisseberth 1977, chap. 1; Kenstowicz and Kisseberth 1979, chap. 6). The reason that formalizing the UR discovery process is so hard is that several of the necessary steps require human insight or intuition.[1] We must be able to determine which forms should be derived from others, comparing each of the possibilities and taking into account how elegant, natural, or simple the rules that would be needed are. We must sometimes have the insight that abstraction is necessary, either by combining surface elements from different forms, or even sometimes by positing abstract structure that is never visible on the surface. In such cases, we must have some intuition about the right types of structure to posit, and whether the pay-off for positing the abstract structure is sufficient to motivate it. Finally, we must sometimes recognize that an alternation defies grammatical description, and certain alternants must simply be memorized as listed allomorphs or as lexical exceptions.

The current model does only a fraction of this: it can determine which alternants should be derived from one another by considering the efficiency of the rules in both directions, but it does not (at present) have any sense of naturalness, nor does it abstract away from the surface alternants to create a UR that does not appear as such somewhere in the paradigm. Thus, it could be said that the model is really discovering "basic alternants", and not URs as they are known and used in generative phonology. For this reason, there are a great many analyses in the literature that it would never be able to discover. This restriction does have historical precedent, however. Kenstowicz and Kisseberth point out that the "basic alternant" restriction on URs, either with or without the additional constraint that we use the same part of the paradigm for all lexical items, is appealing because it is a strong and concrete hypothesis (Kenstowicz and Kisseberth 1977, pp. 28-29), and in fact Sapir seems to have operated under something like this

[1] A similar sentiment is to be found in Hockett (1955): "We know of no set of procedures by which a Martian, or a machine, could analyze a phonologic system. . . " (p. 147)

restriction (McCawley 1967).[2]

The reason why this constraint is so strong is because it severely limits the possible analyses in cases of symmetrical neutralization, or in cases where there are multiple neutralizations working simultaneously. As discussed at length in the previous chapter, in the case of a symmetrical neutralization between *a*, *b*, and $a \sim b$, it means that the basic alternant cannot keep $a \sim b$ distinct from both *a* and *b*, but it can at least keep it systematically distinct from one of the two. This is what allows the model to predict that invariant [a] should be merged with alternating [e]∼[a] in Lakhota, but kept distinct from invariant [e]. In this case, the device that the model is unable to use is the archiphoneme. However, it is not the case that the model cannot learn the difference between invariant [e] and alternating [e]∼[a]—it simply learns it in a different way, relying on listed exceptions rather than on a distinct underlying representation to preserve the minority pattern.

The real challenge comes in the case of multiple neutralizations, such as the following: suppose *a* and *b* are always neutralized to *a* in form 1, while *c* and *d* are always neutralized to *d* in form 2. Under the usual assumptions of generative phonology, learners can abstract away from these forms and combine their information to yield an UR in which both distinctions are preserved: /ac/, /ad/, /bc/, /bd/. The single surface base restriction, on the other hand, means that the decision about which form to choose as the base for the *a*/*b* neutralization also affects the analysis of the *c*/*d* neutralization. If we choose form 1, *a* and *b* are distinct but *c* and *d* are merged, and the reverse for form 2.

Latin rhotacism is a case of this sort. Nominative forms in Latin had a number of neutralizations, including neutralizations in morphological classes (e.g., *-us* for both second and fourth declension nouns), phonological neutralizations caused by devoicing or deleting obstruents before the *-s* suffix (e.g., *ur*[p]*-s* ‘city’, *ar-s* ‘art’), and so on. One of the few neutralizations that went in the opposite direction was rhotacism, in which the [s]∼[r] distinction was maintained in the nominative but merged in the remaining forms. If learners had been able to combine information from multiple forms—or even if they had just been able to use forms from different parts of the paradigm for different words–then they should have had no trouble setting up underlying /s/ to keep words like [honoːs] distinct from words like [soror].Thus, the single surface base restriction plays a role in constraining the predictions of the model even when archiphonemes and underspecification are not involved.

Unfortunately, Latin rhotacism is very mild as neutralizations go, affecting only two segments in a restricted class of words. In some cases, each of the multiple neutralizations involves a large class of segments, potentially affecting a huge number of lexical items. A famous case of this, discussed by Kenstowicz and Kisseberth (1977, pp. 18-19 and 26-27), occurs in Russian. Russian, like German, Dutch, Turkish, and many other languages, has a process of final devoicing, by which voiced obstruents become devoiced in final position. Final devoicing can create alternations within noun paradigms, since the case endings are null for some noun classes

[2]Zellig Harris, although usually associated with the approach of distributional learning of recurring partials, also proposes something like this towards the end of his *Methods in Structural Linguistics.* He suggests that in languages with multiple inflectional classes, such as Latin, it is necessary to list words with some sort of inflectional ending that will tell what class the word belongs to. In considering which form to choose as the default, or base, Harris proposes a model that is strikingly similar to the one I have implemented here: “The criteria for selecting a basic alternant are not meaning or tradition, but descriptive order, i.e. resultant simplicity of description in deriving the other forms from the base.” (Harris 1951, p. 308, fn. 14).

in some forms (the nominative singular and inanimate accusative singular of consonant-final nouns, and the genitive plural of a certain set of nouns). Examples of words with and without voicing alternations are given in (81).

(81) Russian final devoicing: *ryčag* vs. *rybak*

a. Underlying /g/ → [k] in nom./acc. sg. of *ryčag* 'lever'[3]

	sg.	*pl.*
nom.	ryčák	ryčagí
gen.	ryčagá	ryčagóf
dat.	ryčagú	ryčagám
acc.	ryčák	ryčagí
instr.	ryčagóm	ryčagámi
loc.	ryčagé	ryčagáx

b. Underlying /k/ always [k] in *rybak* 'fishmonger'

	sg.	*pl.*
nom.	rybák	rybakí
gen.	rybaká	rybakóf
dat.	rybakú	rybakám
acc.	rybaká[4]	rabakí
instr.	rybakóm	rybakámi
loc.	rybaké	rybakáx

From the point of view of final devoicing, then, we would need to take a suffixed form as the underlying form of a Russian noun. There is, however, a competing process that affects suffixed forms: vowel reduction of stressless /e/ and /o/ to [i] and [a], respectively. Since some nouns in Russian have stress alternations (sometimes on the root, sometimes on the suffix), this can create alternations in vowel quality. For example, there is a class of nouns in which stress falls on the first syllable of the suffix, if there is one; in forms with no suffix, the stress has no choice but to fall on the root. This means that the final syllable of the root surfaces intact in unsuffixed forms, but is reduced to [i] or [a] in suffixed forms (82a), causing a neutralization with underlying /i/ and /a/—cf. *rybaká* (81) vs. *sjedaká* (82a). The fixed-stress form in (82b) shows that vowel reduction really is conditioned by stress, since it does not occur when the stress remains on the root.

[3] I will use the transcription <y> for the Russian vowel <ы>.

[4] The difference between [račák] and [rabak-á] is due to a difference in animacy—animate nouns have overt accusative marking, and inanimate ones do not.

(82) Russian vowel reduction in words with stress shift

a. Reduction in *sjedok* 'rider' (stress shift)

	sg.	*pl.*
nom.	sjedók	sjedakí
gen.	sjedaká	sjedakóf
dat.	sjedakú	sjedakám
acc.	sjedaká	sjedakí
instr.	sjedakóm	sjedakámi
loc.	sjedaké	sjedakáx

b. No reduction in *pr'itok* 'influx' (fixed stress)

	sg.	*pl.*
nom.	pr'itók	pr'itóki
gen.	pr'itóka	pr'itókof
dat.	pr'itóku	pr'itókam
acc.	pr'itók	pr'itóki
instr.	pr'itókom	pr'itókami
loc.	pr'itóke	pr'itókax

We see, then, that among nouns with a stress shift in suffixed forms, the processes of final devoicing and vowel reduction affect mutually exclusive sets of forms: final devoicing affects the suffixless forms, while vowel reduction affects the suffixed forms. This means that there can exist words in which no part of the paradigm reveals the full underlying form:

(83) Devoicing in unsuffixed forms, vowel reduction in suffixed forms

a. *pirog* 'pie'

	sg.	*pl.*
nom.	pirók	piragí
gen.	piragá	piragóf
dat.	piragú	piragám
acc.	pirók	piragí
instr.	piragóm	piragámi
loc.	piragé	piragáx

b. *sapog* 'boot'

	sg.	*pl.*
nom.	sapók	sapagí
gen.	sapagá	sapagóf
dat.	sapagú	sapagám
acc.	sapók	sapagí
instr.	sapagóm	sapagámi
loc.	sapagé	sapagáx

This and similar cases compelled Kenstowicz and Kisseberth and others to conclude that URs must absolutely be able to combine information from multiple surface forms, since there is no single place in the paradigm where the vowel quality of the final vowel and the voicing status

Table 6.1: Distribution of terminations and accent patterns

Termination	Total Count	Suffix-accenting
/-ak/	196	100
/-ag/	46	4
/-ok/	948	9[6]
/-og/	178	5

of final obstruents are both unambiguously revealed. As previously discussed, the assumption that drove this conclusion was that we would like to be able to derive all surface forms unambiguously from a single UR using a simple set of case endings and phonological rules, and in order to do this, we need to create underlying distinctions between all of the observed surface patterns.

What if we relaxed these assumptions, however? Suppose that we allowed the grammar to have fancier rules than simple affixation, including multiple, competing generalizations about specific phonological environments—such as changing X*aga* → X*ok* to form the nominative, and so on. And suppose further that instead of unambiguously deriving all forms of all words, we merely required that the grammar get as many forms right as it could given the single surface base restriction, and then we allowed the remaining forms to be listed as exceptions. How much headway could the learner make on Russian using such a grammar, and how much would need to be memorized?

In principle, we might expect this to be a hopeless task, since there are so many underlying possibilities, and so many neutralizations. There could, for example, be nouns of all of the following four types (labeled, for convenience, with the UR they would receive under the traditional analysis):

(84) Four possible types of Russian nouns

	/…ak/	/…ag/	/…ok/	/…og/
nom. sg.	[…ák]	[…ák]	[…ók]	[…ók]
gen. sg.	[…aká]	[…agá]	[…aká]	[…agá]
dat. sg.	[…akú]	[…agú]	[…akú]	[…agú]
acc. sg.	[…ák]	[…ák]	[…ók]	[…ók]
…				

In practice, however, not all of these possibilities are equally attested. Using Zaliznjak's reverse dictionary of Russian (Zaliznjak 1977), I did some rough counts of words ending in /-ak/, /-ag/, /-ok/, and /-og/, shown in Table 6.1. First, one may note that there are relatively few /ag/ stems, and there is a preponderance of /ok/ stems. What is more significant, though, is the relative distribution of stem vs. suffix-accenting nouns. Among stems ending in /-ag/, /-ok/, and /-og/, virtually all have fixed stress on the stem (like *pr'itók* in (82b) above), and vowel reduction is not an issue. Words like Kenstowicz and Kisseberth's *pirog* are vanishingly rare; they include only *pirog* 'pie', *sapog* 'boot', *batog* 'thick stick', *podog* (no gloss), and *tvoróg* 'curd'.[5] Almost all of the stress-shifting stems in this portion of the lexicon belong to the /-ak/ group.

[5]Zaliznjak also lists a variant of this word with fixed penultimate stress: [tvórok], [tvóroga], etc.

The result of all of this is that at least for this set of words, it is not at all difficult to learn to project the nominative from an oblique form (that is, a form other than the nominative or accusative). If the word has stress on the stem, one need only peel off the case suffix and devoice the final consonant. If the word ends in suffix-stressed [...aká], it is virtually certain to end in [ák] in the nominative, and the nine exceptions could be easily listed. If the word ends in suffix-stressed [...agá], it is a toss-up whether it will end in [...ák] or [...ók] in the nominative, but there are only four of the former and five of the latter, so listing these does not seem overly costly.

One might wonder, given the preponderance of /-ok/ stems in general, why there are not more stress-shifting /-ok/ stems that could be neutralized with the stress-shifting /-ak/ stems. In fact, a sizeable majority of /-ok/ stems (787 out of 948) participate in a different alternation, caused by deletion of stressless mid vowels in the final syllable of the root—the so-called *yers*. Without going into the details of yer alternations (see Lightner 1972, Yearley 1995, and Hermans 2001 for discussion), we can note simply that words with /-ok/ and /-ak/ remain distinct, even in stress-shifting forms, because the /-ok/ words generally delete the /o/ rather than reducing it:

(85) /-ak/ vs. /-ok/ remain distinct

a. *sudak* 'pike perch'

	sg.	*pl.*
nom.	sudák	sudakí
gen.	sudaká	sudakóf
dat.	sudakú	sudakám
acc.	sudák	sudakí
instr.	sudakóm	sudakámi
loc.	sudaké	sudakáx

b. *sudok* 'cruet-stand, dinner-pail'

	sg.	*pl.*
nom.	sudók	sudkí
gen.	sudká	sudkóf
dat.	sudkú	sudkám
acc.	sudók	sudkí
instr.	sudkóm	sudkámi
loc.	sudké	sudkáx

The forms in (85) show that if we take an oblique form as the base in Russian, as in Latin, then it is not difficult to predict what vowel a reduced [a] should correspond to in the nominative—it is almost always [a]. The forms with *yer* deletion in (85b) show that there may be other problems, such as predicting whether a word should have a vowel inserted between the consonants in the nominative, and which vowel should be inserted ([o] or [e]). However, even this may be somewhat predictable. Various scholars working on Polish have suggested that *yer* deletion

[6] I am excluding here the large number of nouns that have suffix accentuation, but also have final *yer* alternations (see (85b) below). The reason is that in these words, there is no issue of trying to recover the underlying quality of a neutralized vowel, since the relevant vowel is deleted altogether.

might actually be analyzable as an epenthesis process to break up word-final clusters (*[sudk] > [sudok]) (Gorecka 1988; Czaykowska-Higgins 1988, and others). Yearley has argued that this is harder in Russian than in Polish (p. 538), but it does not seem impossible if we allow exceptions. For example, the *-ok* nominatives listed in Zaliznjak far outnumber the *ek* nominatives and *-Ck* nominatives, so a genitive form with *-Cka* most likely corresponds to a nominative in *-ok*.

It appears, then, that for a particular subset of the consonant-final nouns in Russian, the final vowel of the root is not nearly as hard to predict as Kenstowicz and Kisseberth's discussion would imply.

It is important to consider whether this might just be an "easy" region of the vocabulary, or whether it is representative of the language as a whole. As it happens, there are reasons to believe this is not an especially easy sample, and that the *-ak*/ *-ok*/ *-ag*/ *-og* nouns are actually some of the hardest nouns to predict. Recall that vowel reduction affects only mid vowels, so the only relevant types of stems are [...oC] and [...eC] stems—which happen to also be the environments for yer deletion. Furthermore, a quick glance through Zaliznjak's dictionary reveals that many final consonants appear to have few, if any stress-shifting nouns. (For example, there are virtually no [...ad] or [...od] nouns with shifting stress.) For nouns that do not shift stress, the vowels are constant throughout the paradigm, and there is no predictability problem. Thus, the region of the vocabulary that I have considered here may actually have *more* neutralizing vowel reduction than average. The single surface base analysis of Russian, in which it would probably be an oblique form that serves as the base, appears to be worth pursuing.

My conclusion from this section is that it is not obvious that a model operating under the single surface base restriction would be completely unable to learn languages like Russian, or by extension, perhaps also Tonkawa, Pengo, or any of the other cases in Kenstowicz and Kisseberth (1977) in which it was argued that it is necessary to combine information from different forms. It seems that languages tend to have a surprising amount of surface predictability between forms; even if the possible underlying forms could theoretically have been distributed evenly and randomly throughout the vocabulary, this appears not to happen.[7]

If an analysis using a single part of the paradigm as the base could be made to work for Russian—that is, if a suitably detailed grammar could manage to productively derive the correct outputs for a reasonably large proportion of the vocabulary—it would have the advantage of making strong predictions about possible errors and analogical changes, as in Lakhota. In particular, we would expect that stem vowels of stress-shifting nouns might sometimes be overregularized—e.g., *[pirák] instead of [pirók]—but we should not get any changes to oblique forms, such as changing the voicing of obstruent, or leveling the placement of stress (*[piróga] instead of [pirogá]). It would be interesting to see if changes of the former type have already been going on in Russian, helping to create and solidify the regularities noted here.

6.2 Bases of analogical change

Most of the evidence that I have used here about bases has come from historical analogical changes, either in the form of paradigm leveling (Yiddish, Latin) or in the form of extending alternations (Lakhota). I have shown that the proposed model makes the right predictions in each case. These three changes are all "hard cases" — they involve typologically unusual

[7]This phenomenon has also been noted for Yidiɲ by Hayes 1999, pp. 11-14.

changes, in that forms that are typically "more basic" (nominative singulars, third person singulars) have been changed on the basis of more marked forms. The model accomplishes this in part by ignoring the factors that are often said to play a typological role in determining the base of an analogical change: the degree of suffixation (unsuffixed, single consonant suffix, full syllable suffix, etc.), the relative markedness of forms, type frequency, and so on. However, even if these factors are not the right explanation for why analogical effects take place, the typological approach has uncovered a number of tendencies that must be accounted for. If learners are really able to search the entire paradigm to find whatever form is the optimal base form, why should the same forms be chosen in language after language? Furthermore, if a lexically predominant pattern of alternation can be extended just as easily as non-alternation leveling, then why is there such a strong tendency for paradigm leveling?

I will consider each of these issues in turn.

6.2.1 Why a nominative/3sg/etc. preference?

It is often noted that certain forms tend to act as bases over and over again. In noun paradigms, it is the nominative singular that usually drives paradigmatic changes (e.g., Lahiri and Dresher, 1984), while in verb paradigms, it is most often the 3sg present form, although the 1sg is also common (Bybee and Brewer 1980; Bybee 1985, chap. 3). Intuitively, there are a number of factors that could conspire to favor these forms: they are often the forms with the highest token frequency, they often have null affixes making them "isolation forms", and morphosyntactically, they may in some sense be the "least marked" members of the paradigm. As Bybee points out, these factors are highly correlated; the morphosyntactically least marked forms also tend to be the most frequent, and also tend to have the least overt morphological marking. Thus, in many cases, it is impossible to say which of these is responsible in making one form the most basic, and in fact, we might rather say that it is the collective influence of all of them together.

To the extent that we can sometimes differentiate these factors in languages where they are *not* correlated, this does not seem to help in identifying a universal principle of what privileges forms as bases. For example, sometimes the 3sg present form has an overt affix, but it acts as a base anyway; one example is the operation of "Watkins' law" in Provençal discussed by Bybee (1985, pp. 55-56), in which the 3sg form replaced all other forms, including the relatively less suffixed 1sg form. In other cases, such as the Yiddish leveling discussed in chapter 2, the suffixless 1sg form acted as the base, taking precedence over the 3sg form. Thus, it is generally concluded that each of these factors is merely a typological tendency, and it is impossible to predict which will win out in a given case. They are, however, tendencies with a great deal of empirical support, and an adequate model of analogical change has to be able to explain why these factors seem to increase the chances that a particular form will serve as base.

I believe that the answer does not lie in how learners seek to find contrasts and organize their grammar, but is rather an epiphenomenal effect of how they receive the input data for morphological learning. In the discussion up until this point, I have been operating under the idealized (and unrealistic) assumption that learners have all parts of the paradigm available in equal proportions when considering their effectiveness as bases. This is clearly not true in real life; some parts of the paradigm are much more frequent than others. This means that learners do not have equal amounts of data about the possible subgrammars. Suppose, for example, that you are learning verb paradigms and comparing the relative effectiveness of three forms: the

1sg, the 2sg, and the 3sg. Suppose further that 3sg forms are by far the most frequent, followed by 1sg forms, and then 2sg forms. This means that on average, you will have heard 3sg forms for most words, and 1sg and 2sg forms for relatively fewer words. Moreover, the number of words that you have heard in both the 3sg and 1sg will be greater than the number heard in both the 3sg and 2sg, and greater yet than the number heard in both the 1sg and 2sg. Figure 6.1 shows a hypothetical input of this sort, listing the attested forms that a learner has encountered for 100 verbs. The most frequent form (the 3sg) is on the left, while the least frequent form (the 2sg) is on the right. Arrows indicate that a verb has been heard in at least two forms, and can thus contribute an input pair for morphological learning on the mapping between those two forms.

How could these frequency differences be used to derive a typological prediction? In section 3.2 (p. 40), I mentioned that the minimal generalization model uses lower confidence limit statistics to favor rules that are based on larger numbers of forms. The original intent of including this adjustment was to encourage the model to use large-scale generalizations, and to be wary of generalizations based on just a few forms that may turn out to be exceptional. This adjustment may have another practical consequence, however, which is to favor subgrammars that are based on larger numbers of input pairs. In particular, if the learner has encountered 60 pairs leading from the 3sg (43 to the 1sg, 17 to the 2sg), but only 29 pairs leading from the 2sg (17 to the 3sg, 12 to the 1sg), this will make the rules in the 3sg-based subgrammars on average more reliable than those in the 2sg-based subgrammars. This captures the intuition that you want to choose a base that you are sure will be reliable in the language as a whole, and you also want to choose a base that will generally be available in order to produce the remaining forms.

The effect would be that sometimes a less common form may preserve contrasts slightly better than a more common form, but the lower confidence limit adjustment would make its subgrammars less reliable, forcing us to select a somewhat less informative (but more frequent) form as base. This bias to chose more readily available forms should be seen not as part of the grammatical learning mechanism, but rather as an incidental external factor that interferes with the ability of the learner to select the truly optimal base for the language. This leads us to wonder, naturally, how suboptimal a base could we be forced to choose, given a large difference in the availability of forms?

A possibly telling example is the change currently underway in Korean, discussed at the end of chapter 4. Korean nouns occur in a variety of forms. Frequently, they occur with no overt case marking (i.e., as just a bare stem), and are affected by word-final neutralization processes caused by severe phonotactic restrictions on what sequences and segments are allowed to occur in word-final codas. For example, Korean has eight coronal obstruents ([t, t^h, t'[8], t͡ʃ, $\widehat{t\int}^h$, t͡ʃ', s, s']), but only one coronal obstruent is allowed in word-final position: [t̚] (Kim 1987). Non-coronal obstruents involve similar, but less dramatic neutralizations—in particular, all three laryngeal types are reduced to unreleased unaspirated stops, but there are no fricatives or affricates at the other places of articulation. Thus, unsuffixed forms involve a potentially massive neutralization, as illustrated in Figure 6.2.

When nouns are overtly marked, they may appear in one of several cases: topic (*-ɨn* after consonants/ *-nɨn* after vowels), nominative (*-i*/ *-ka*), accusative (*-ɨɭ*/ *-lɨɭ*), dative (*-ey*/ *-ekey*), genitive (*-ɨy*), and so on. These suffixes protect the noun from the word-final obstruent neutralizations, but they may cause phonological modifications of their own. The nominative suffix *-i*,

[8]I will follow the practice of using [C'] to indicate a tensed obstruent in Korean.

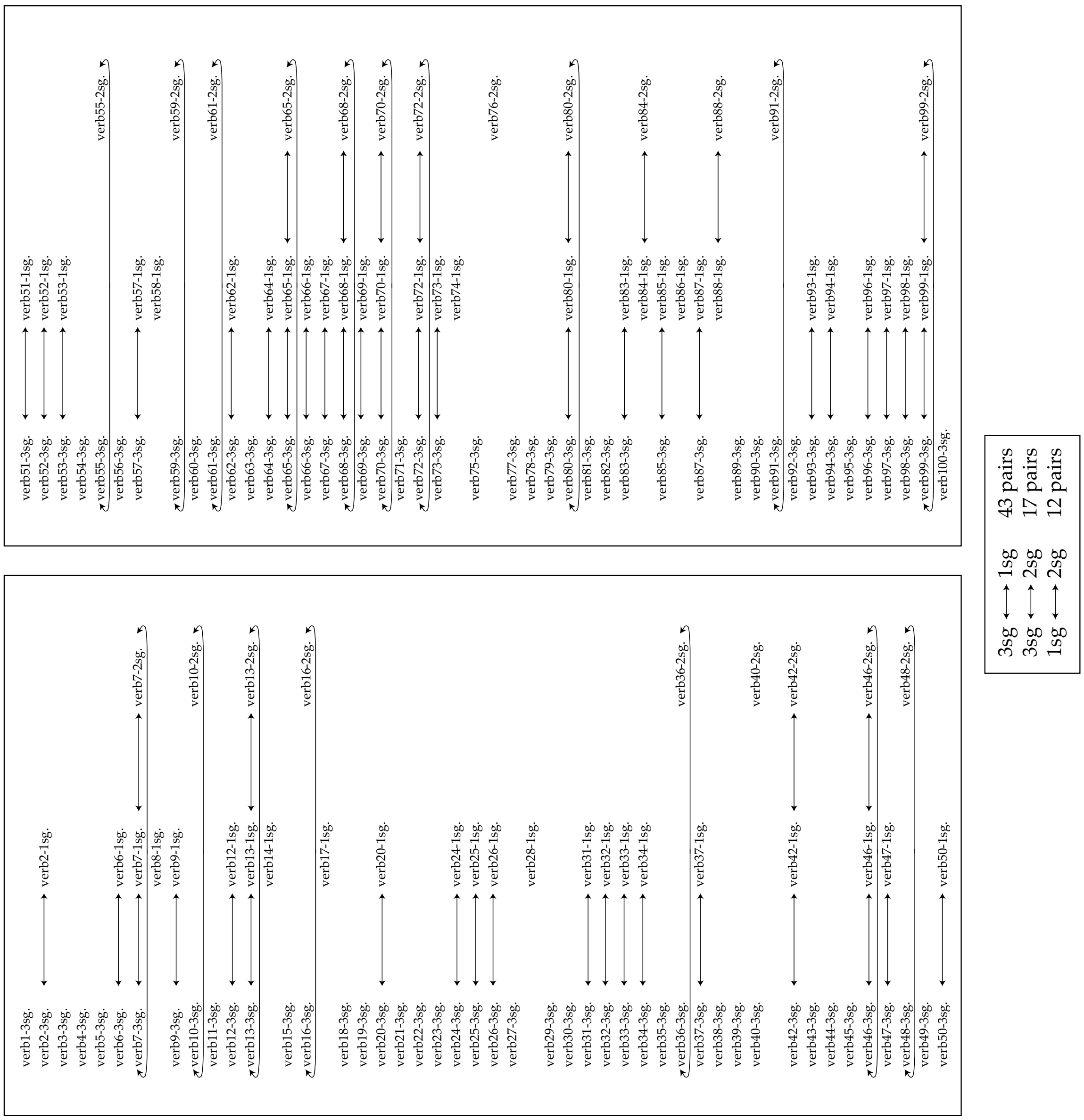


Figure 6.1: Sample input assuming unequal frequencies

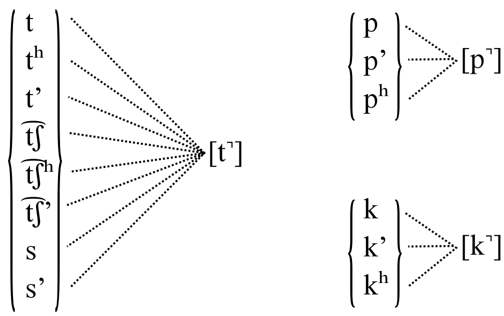


Figure 6.2: Theoretical neutralization of obstruents in word-final position in Korean

for example, causes palatalization of preceding coronal segments (/t/ → [t͡ʃ], /s/ → [ʃ], /n/ → [ɲ]), neutralizing the contrast between /t/ and /t͡ʃ/ in this form. The topic and accusative case markers begin with a non-front vowel ([ɨ]), so in this context, all underlying contrasts would be preserved (though there is one non-neutralizing modification whereby lax stops become voiced intervocalically). As with the other languages discussed in the previous chapters, we see that different forms in the paradigm involve different degrees of neutralization.

The reason why Korean is a problem is that in this case, the base of the analogical change appears to be the unmarked isolation form, even though that form is the most neutralizing form of the entire paradigm. Recall from section 4.4.4 (p. 72) that a change currently underway in Korean is replacing the final coronal obstruents of noun stems with a *s* in inflected forms, apparently because speakers are attempting to project inflected forms from the massively neutralizing isolation form. This would not be a problem for a more typological approach, since we could simply say that in Korean, the preference for leveling to isolation forms has won out over whatever other preferences might have favored other forms. How could this change be reconciled with the informativeness approach?

I speculate that there are two factors that might allow the current model to select the isolation (unmarked) form as the base in Korean, in spite of its neutralizations. The first is that the picture in Figure 6.2 is not quite accurate. In fact, not all obstruents occur in stem-final position in Korean; in particular, there are no stems ending in a tensed obstruent ([p', t', t͡ʃ', s']), and there are no nouns ending in unaspirated [t]. Furthermore, among the remaining four coronal obstruents, [c] seems to be relatively rare, [t^{h}] and [t͡ʃ^{h}] are moderately well attested, and [s] outnumbers all of the others (Figure 6.3). We see, then, that the actual neutralization is not quite as bad as it could have been; given an unsuffixed noun ending in [t̚], the set of possible suffixed forms is far more constrained than the full set of [t, t^{h}, t', t͡ʃ, t͡ʃ^{h}, t͡ʃ', s, s'] at equal probabilities of 1/8 each.

These asymmetries lessen the severity of the word-final neutralizations, but they do not eliminate them altogether. Even if the unsuffixed form is not as terrible as we might have imagined, it is still the least informative member of the paradigm—what could possibly have

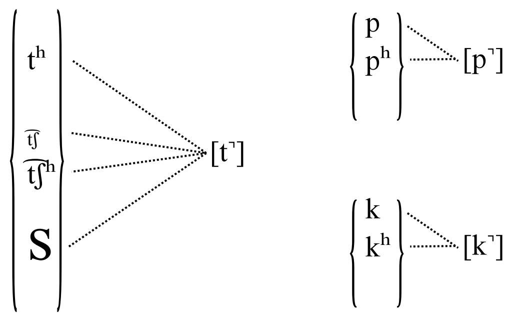


Figure 6.3: Actual neutralization of obstruents in word-final position in Korean

led learners to choose it as the base?

What we have not taken into account is the relative availability of different forms in Korean. Unlike many Indo-European languages, case marking in Korean is optional in many contexts; in fact, marking appears to be omitted in a majority of tokens. In corpus counts of various case forms in child speech and in the child-directed speech of mothers, I. Lee (1999) found that case marking was omitted by mothers 70% of the time when talking to their children. Furthermore, among forms with overt case marking, nominatives far outnumbered accusatives: nominative *-i*/*-ka* occurred on 25% of child-directed tokens, while accusative *-ɨɭ*/*-lɨɭ* occurred on only 5% of tokens. Thus, the hypothetic proportions shown in Figure 6.1 probably *overestimate* the availability of accusative forms in the learning data. The accusative may be a perfect base in Korean from a predictability point of view, but it is rare enough that learners would have a hard time finding enough input pairs to learn that this generalization can be trusted.

In sum, the two factors that could conspire to select a less informative form as the base in Korean are: (1) the predictability asymmetry between the suffixed and unsuffixed forms is smaller than we might have expected, owing to gaps and asymmetries in the Korean lexicon, and (2) small differences in predictability may be overwhelmed by substantial differences in the availability of forms to construct input pairs for the morphological learner. Forced by these considerations to choose the unsuffixed form as the base, speakers must then learn to predict which obstruent an unreleased [t̚] corresponds to on a noun by noun basis, leaving the oblique forms open to overregularization. Since the predominant pattern in the lexicon is for [t̚] to correspond with [s] (or [ʃ] in palatalizing environments), this is the pattern that is winning out.

An interesting (but unverified) prediction of the current approach is that new, previously impossible noun paradigms could arise, parallel to the change we saw in Lakhota and to the cases discussed by Hayes (1998). Suppose that a particular noun is relatively frequent in a particular case—for example, a word for a duration may occur often in accusative adverbial constructions, or a word for a place may occur often in locative constructions. This would allow it to be memorized as an exception to the [t̚] → [s] pattern for the accusative or locative. When adding the nominative marker *-i*, however, the speaker is unable to use the accusative form to project the correct consonant, because the accusative is not the base form. Thus, we predict that "inconsistent" paradigms of the following sort might arise:

(86) A potential inconsistent paradigm in Korean

unsuffixed	[...t̚]
nom.	[...ʃ-i]
acc.	[...tʰ-ɨl]

Stated differently, the prediction is that there may be nouns for which a speaker prefers the more conservative, etymologically correct form for some cases, but the newer, analogically created form for other cases.

An example of this sort may occur with the word for 'field'. This word is [pat̚] in its isolation form, and its etymologically expected final segment is an aspirated *t* (/patʰ/). In an experimental study on variability in the treatment of final consonants, H.-J. Lee (Lee, in progress) asked ten native Korean speakers to produce nouns before a variety of different case markers: some beginning with [ɨ], such as *-ɨn* (topic), some with [i], such as *-i* (nominative), and some with [e], such as *-e* (dative). In the topic form, none of Lee's participants used the etymologically expected [tʰ]; all of the responses were analogical, half with [t͡ʃʰ] and half with [s]. In the dative, however, which is also sometimes used with a locative meaning, eight out of ten of the participants used the more conservative [tʰ] form. This suggests the following, inconsistent paradigm:

(87) An inconsistent paradigm in Korean

unsuffixed	[pat̚]
top.	[pat͡ʃʰ-ɨn] / [pas-ɨn]
dat.	[patʰ-e]

Inspection of Lee's data shows that such discrepancies seem to exist for many words, to varying degrees. These inconsistent paradigms are problematic for a theory that attempts to derive all surface forms from the same UR whenever possible, in the same way that the Lakhota examples in the previous chapter are. What UR can we set up for the new [t̚] ∼ [s] ∼ [tʰ] paradigm? And more importantly, what UR were the first speakers who uttered [paʃ-i] using that allowed them to change the nominative and not the locative?[9] Such a change is to be expected, though, under a model in which speakers choose a single form in the paradigm as the base, and then derive the remaining forms using separate morphological rules, each with a potentially different set of stored exceptions.

It is worth mentioning that another famous analogical change that may be similar to the Korean change is one that has occurred in Maori passive (Hohepa 1967; Hale 1973; Hyman 1975, pp. 184-185; Hock 1991, pp. 200-202; Barr 1994, pp. 468-477; Kibre 1998). In this case, a historical word-final deletion of consonants caused a massive neutralization in unsuffixed forms (all consonants → ∅). As in Korean, a suffixed form like the passive would be much more informative, since it preserves the identity of the stem-final consonant.

[9] I cannot preclude the possibility that the first speakers to introduce analogical forms used analogical forms in all cases (*pat͡ʃʰ-ɨn*/*pas-ɨn* 'field-TOP', and also *pat͡ʃʰ-e*/*pas-e* 'field-DAT'), but that the analogical forms failed to catch on in the dative. This scenario does not seem likely, however, since it is unclear why a subsequent generation of speakers would adopt the innovative nominative form but reject the innovative dative, choosing an inconsistent paradigm when a more consistent one was available in their input data.

(88) Older stage of Maori

base	passive	gloss
afi	*afi**t**ia*	'embrace'
hopu	*hopu**k**ia*	'catch'
tau	*tau**r**ia*	'come to rest'
waha	*waha**ŋ**ia*	'carry on back'

What has happened, however, is that the passives of many words have changed, so that they now use -*tia*, or have at least gained -*tia* variants (e.g., *wahaŋia* > *wahatia*). It appears, therefore, that rather than using the more informative passive form as the base, Maori speakers instead used the unsuffixed form as the base, and were forced to memorize or guess what consonant should appear in the passive (-*tia*, -*kia*, -*ria*, etc.). As shown by Sanders (1990), -*tia* is the statistically predominant pattern, and would be the most reliable rule for creating passives. Like the Korean change, the Maori change could be handled in the informativeness approach if there is a substantial difference in the frequency or availability of verbs in these two forms. Whether or not this is actually the case in Maori remains an open question; Kibre (1998) notes that Maori passives are used in a considerably wider array of contexts than English passives (see also Chung 1978), but I do not know whether this means that they actually occur as frequently as actives, or whether it is just a smaller asymmetry than in English.

Crucially, the skewing based on token frequency proposed in this section would most likely not change the results in the previous chapters for Yiddish, Latin, or Lakhota. In the case of Yiddish, the base was the 1sg form, which, although not as frequent as the 3sg form, is probably the next most frequent form. There are no available corpus counts for Yiddish,[10] but Bybee (1985, p. 71) lists some equivalent counts from Spanish, which we may take as a rough estimate. She finds that from a written corpus of adult Spanish (Juilland and Chang-Rodríguez 1964), the 3sg is the most frequent form (44% of the tokens), while the 1sg is second (23%); approximately the same proportion also holds for a corpus of Spanish spoken by children (Rodríguez-Bou 1952) (3sg: 41%, 1sg: 24%). Thus, although the 3sg is significantly more frequent, the 1sg also occurs frequently enough to allow generalizations based on more than just a few input pairs, and could still be chosen as a base form.

In the case of Latin, the sample counts in section 4.4.2 (p. 70) showed that on the whole, oblique forms are at least as frequent in Latin as the nominative form is. If frequency counts on larger corpora and larger numbers of words continued to show the same pattern as in Table 4.1, then the confidence limit adjustment would not affect the results for Latin, either.

Finally, in the case of Lakhota, I argued that it was a second person form that acted as the base in verbal paradigms. This result seems most in danger of being affected by the proposed skewing for confidence. Bybee's counts for Spanish show that the 2sg form is quite a bit lower in token frequency, with 16% of the tokens in the written adult corpus, and 11% in the spoken child corpus. However, I believe that morphological differences between Lakhota and Spanish would make the 2sg form more frequent in Lakhota than in Spanish. In particular, the 2sg form is used not only for statements and questions, but also for imperatives, making 2sg sentences

[10]The equivalent counts from the German portion of CELEX are also uninformative, because they fail to differentiate homophonous forms. For example, an ambiguous form like *geben* is counted in the token frequency of the infinitive, the 1pl present indicative, the 3pl present indicative, and so on, meaning that all homophonous forms are listed with the same token frequency.

three-ways ambiguous, as in 89:

(89)	*Shunka*	*he*	*walaka-he*
	dog	DET.	see-2sg. (female speaker)
	You see the dog. / Did you see the dog? / Look at the dog!		

Bybee's Spanish counts do not include imperative forms, but they are probably quite frequent, especially in child-directed speech.[11] Berman (1985, p. 268) notes that in Hebrew, children acquire imperative forms very early, with boys using the masculine imperative and girls using the feminine imperative, in spite of the fact that they often involve difficult consonant clusters. Thus, as with Yiddish, even if 3sg forms are more frequent, it appears that second person forms are not so uncommon that they would be penalized too severely by the confidence limit adjustment.

6.2.2 Local Markedness

An argument that is often put forth in favor of the role of markedness in analogical change is the fact that when the meaning of individual lexical items gives them unusual markedness relations, analogy can go in the opposite direction. This has been argued most explicitly by Tiersma (1982), who called this phenomenon *local markedness*. Tiersma showed that in particular dialects of Frisian, Dutch, and German, nouns denoting objects that occur most frequently in the plural (such as eyes, arms, geese, etc.) have sometimes undergone paradigm leveling to the plural form. This is taken as evidence that any explanation of analogical change must take into account the markedness or frequency of forms not only in general, but also on a word-by-word basis. The single surface base hypothesis, on the other hand, precludes such word-by-word effects, because the organization of the grammar requires that the same base be used for all lexical items. How can these two viewpoints be reconciled?

We might start by noting that there is very little evidence of local markedness effects in any of the cases discussed here. In Yiddish, paradigm leveling affected virtually every verb in the language, with only one possible case of local markedness: the verb *gefeln* 'to be pleasing' is apparently derived from a 3sg instead of a 1sg form. This could have been caused by local markedness in Tiersma's sense, but it may also be caused by a far more drastic version of local markedness in which lexical items do not occur at all in certain forms (defectiveness). If a mere imbalance is enough to reverse the direction of leveling, then why is this the only such form? Similarly, the Latin *honor* analogy was remarkably complete; all of the words of the right phonological shape underwent the $s > r$ change in the nominative, with no evidence that more "agentive" nouns retained *s* merely because the nominative was (even) less marked for these forms. Furthermore, words that did not meet the criteria consistently failed to undergo leveling; the only possible exceptions were neuter nouns with masculine doublets, but never words with unusual markedness relations. Finally, in the case of Lakhota, there is no evidence of words in which an unmarked 3sg caused an analogy based on the 3sg, such as extending the *-e* (*chepe*, **chepe-pi*) or extending ablaut to an invariant *-e* (*washte*, **washta-pi*). In these, and

[11] One may speculate further that children probably pay more attention to utterances that are addressed directly to them, both because of the attention, and because they need to understand what is required of them. Anecdotally, it is interesting to note that the vast majority of the Lakhota sentences that Ms. Iron Teeth could recall verbatim from the more archaic speech of older people (such as her grandparents) were commands for her to do various things.

many other cases, paradigm leveling is across the board, affecting all words of the appropriate morphological and phonological classes in the same way. If paradigm leveling could choose different bases for different lexical items, then such clean changes should never happen.

Even in the case of Frisian, it is not clear how good local markedness is as an explanation for the behavior of the particular words that Tiersma discusses. It is true that arms and geese do occur most naturally in the plural, but so do hundreds of other objects in the world. In fact, the nouns discussed by Tiersma seem to be frequent enough that we would expect learners to have had ample exposure to them in both the singular *and* the plural; surely the singular of 'arm' is attested often enough that in Bybee's terms, it should be strong enough to protect itself from leveling. The words that one should really expect to see on Tiersma's list of levelings are words with both reversed markedness and lower frequency, like 'kidney', 'termite' or 'barnacle'.[12]

The model I have proposed here makes the strong prediction that local markedness should not play a role in analogical change. This leaves us without an explanation for Tiersma's Frisian, Dutch, and German cases. However, it appears that such cases may be quite rare (and inadequately explained by local markedness in any event), so one might hope to find alternative explanations for them.

6.2.3 Why a leveling preference?

Another typological preference that is sometimes mentioned is a preference for paradigm leveling (eliminating alternations) over analogical extension (introducing alternations). Mańczak (1958), for example, claims that it is a fundamental tendency of analogical change that root alternations are generally leveled, not extended.[13] The model that I have proposed here, however, does not have any built-in preference to level; it derives leveling simply as the extension of a dominant pattern of non-alternation. Why might this tend to lead to leveling more often than analogical extension?

A comparison of Latin and Korean may help to shed light this problem. In Latin, rhotacism affects just one segment (/s/), so the only words that exhibit the alternation are those with stem-final [s]. In Korean, on the other hand, coda neutralizations affect all obstruents, so many lexical items exhibit alternations. I conjecture that morphophonemic alternations typically affect only a small subset of the phonemic inventory; rhotacism affects just /s/, umlaut affects only back vowels, palatalization tends to affect just coronals (and often just coronal stridents), and so on.

When the minimal generalization learner is confronted with a language that has such an alternation, it will therefore end up with two competing processes: the rules that produce the morphophonemic alternation (such as [t̚] → [tʃ-i]), and the general rules that cover the rest of the segments (such as ∅ → [–i]). When the alternation can be learned as automatic phonology, this is no problem—using the general rule to suffix [-i] would produce incorrect *[. . . t–i], but this would be fixed by the phonology to yield [tʃ-i]. If, on the other hand, the process cannot be learned as automatic phonology (for example, if sound change or borrowings have reintroduced

[12]The frequency and number-markedness of this set of words is a guess on my part. Also, it is important to bear in mind that the leveling discussed by Tiersma could only occur in words that meet the right environment for the phonological "breaking" process that caused singular∼plural vowel alternations in Frisian.

[13]Hock (1991, pp. 235) appears to be rather non-committal about what we should conclude from this fact, pointing out that there are nevertheless a large number of cases in which alternations have been extended. He concludes his discussion by saying that both occur frequently, and must be explained.

surface [ti] into the language), then the [t˺] → [tʃ-i] rule and the ∅ → [-i] rules will yield different outputs.

The preference for leveling, then, comes from the fact that when an alternation can be produced only using the smaller, more local rule ([t˺] → [tʃ-i]), the "regular" output produced by the more general process (∅ → [-i]) will be a strong competitor, and will provide the impetus for leveling. Whenever alternations affect only a subset of segments and a minority of the words in the language, they will be threatened the more general non-alternating pattern. The end result is that this model captures common traditional insight about paradigm leveling: once phonological alternations are no longer surface true and must be learned as morphophonology, they will be open to leveling because the general, dominant pattern of simple affixation cannot produce the alternation. Furthermore, since phonological alternations, even when surface true, often involve just a subset of the segment inventory, non-alternation is often the dominant pattern in languages. I conjecture that this suffices to predict whatever preference there may be for leveling over analogical extension.

6.2.4 Frequency, forgetting, and overregularization

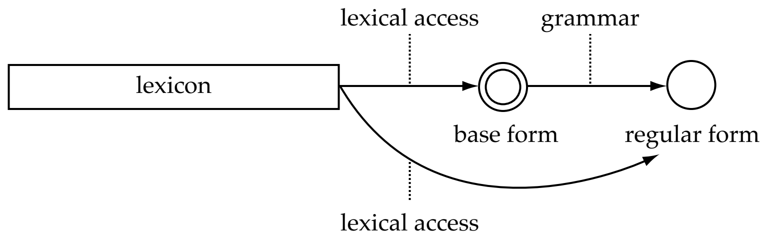
Throughout this thesis, I have been assuming that speakers have two means of producing non-basic forms: they may either use a lexically listed form (encoded here as a word-specific rule), or they may use their grammar to derive a form. For regular forms, these two routes yield the same answer, but for irregular (exceptional) forms that the grammar cannot derive, they yield different results. This was illustrated in Figure 1.1 in chapter 1, and is repeated in Figure 6.4.

In the usual case, it is assumed that a lexically listed form takes precedence over, or *blocks* its grammatically derived competitor (Aronoff 1976). The mechanism for change under this model is *overregularization*—exceptional forms that could be produced only by retrieving them from memory are replaced with their regular, grammatically derived rivals (Marcus, Pinker, Ullman, Hollander, Rosen, and Xu 1992).

There are numerous reasons why blocking might fail, leading to overregularization. For example, the speaker might not know the exceptional form, or she might know it but fail to retrieve it in time to block the grammatically derived form. It is widely agreed that the more frequent an exceptional form is, the easier it is to retrieve, and the better it is able to block the synthesized output (Bybee 1985; Baayen, Dijkstra and Schreuder 1997; Pinker 1999, pp. 129-131). In principle, it would not be difficult to incorporate these "accessibility" effects directly into the current model to produce a more explicit production model of overregularization and language change. One way to do this would be to redefine the probability of using a word-specific rule (currently assumed to be 1) to be related to the token frequency of the output of the rule. This would mean that very frequent outputs would be produced with near 100% reliability, while infrequent outputs would be more likely to resort to the output derived by the grammar of generalized rules.

Such a model would be useful in capturing the overregularizations errors seen in children (and even adults), but it would not be sufficient to capture the historical changes that I have discussed here. In both Yiddish and Latin, paradigm leveling affected all of the relevant lexical items, not just the less frequent ones. What mechanism allowed speakers to replace even very high frequency forms like 'he takes' (**nimt* > *nemt*) or 'they know' (**visən* > *veysən*)? The usual assumption, dating back at least to Paul (1920), is that such changes should be attributed not

a. Two (identical) ways to produce non-basic, regular forms:



b. Two (non-identical) ways to produce non-basic, exceptional forms

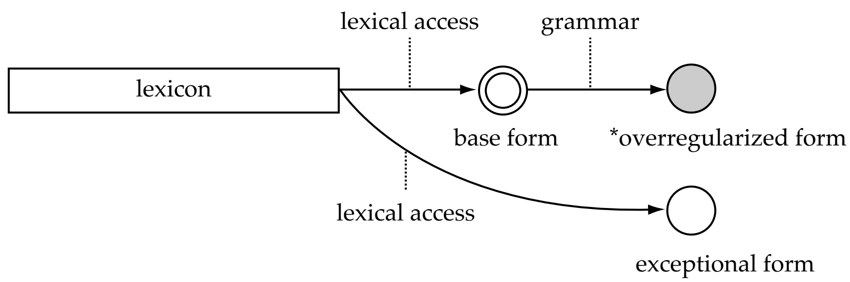


Figure 6.4: Routes for deriving different types of words

to the adult production system, but rather to the errors that children make. For children, no words have extremely high token frequency yet, and it is far more likely for a child to make an error on 'they know' than for an adult to do so. However, as Marcus et al. (1992) show, the number of child errors on high frequency verbs is never all that high, even at the height of the overregularization stage (though cf. Maratsos 2000 for a challenge to these counts). Furthermore, one may question to what extent children's errors are really able to propagate themselves; adults seem impervious to them, and even if younger children are impressed by them, they still constitute only a fraction of their linguistic input.[14] If imperfect acquisition by children was really the cause of such whoesale paradigm levelings, we would expect to sometimes encounter older children (or even adults) who had never emerged from the overregularizations. However, this seems not to happen, at least in English; every child examined by Marcus et al. eventually stopped producing more than the occasional overregularization. I believe that there must be some factor other than low accessibility and imperfect memorization contributing to the widespread levelings seen in Yiddish or Latin.

In the case of Yiddish, it is quite likely that sociolinguistic factors encouraged speakers to favor overregularized forms. If my conclusions about the dating of paradigm leveling in section 2.1.2 are correct, then leveling must have occurred at a time when Yiddish was beginning to develop as a written language with its own literature, distinct from German. This may have

[14]One possible response to this would be to posit that wholesale paradigm levelings can only occur at moments of social upheaval or violent discontinuities, in which access to adult forms is limited and child errors assume abnormal importance. This seems not to have been the case in either Yiddish or Latin, however.

provided exactly the environment necessary for speakers to be willing to produce and accept overregularized forms at a far greater rate than in, say, present day English. As distinguishing features go, such overregularizations would be easy to produce and remember, since, under this analysis, they were the grammatically preferred forms.

The suggestion, then, is that at least for adults, morphological blocking and the failure to block involve more than just lexical access; they involve sociolinguistic factors like normative pressure, the desire to maintain conservative forms, the desire to mark sociological affiliation through linguistic innovations, and so on. A more complete model that takes all of these factors into account is far beyond the scope of this thesis.

6.3 Scaling up: discovering local bases

As I have reiterated throughout this thesis, the part of the proposed model that is responsible for predicting many of the observed asymmetries is the single surface base hypothesis. According to this hypothesis, learners are restricted to selecting just one form as the base within the paradigm, and all other forms must be derived from the same base. As discussed in chapter 1, this assumption is appealing from a learnability point of view, because it greatly reduces the number of possible grammars that must be explored. I did admit at that time, however, that the paradigms under discussion here have all been relatively small and local—a single tense or a small set of case and number forms. In this section, I will present data that appears to require a more sophisticated paradigm structure, composed of subparadigms each with its own local base. I will then discuss several possible ways in which the procedure proposed in chapter 3 could be extended to accommodate such cases.

6.3.1 The need for local bases

When we look at larger paradigms, involving more tenses, moods, and so on, it often appears that we need local bases for each sub-paradigm (something like the traditional idea of *principal parts*, or multiple stems). Furthermore, Bybee (1985) and others have noted that analogical changes tend to operate within a particular subpart of a paradigm (a particular tense, aspect, or mood, or in the Polish example from chapter 1, just within diminutives), providing further evidence that there must be local bases. The set of forms that will require a local base cannot be assumed to fall within any particular tense, mood, or aspect, however. Aronoff (1994) points out that in Latin verbs, a certain set of forms seem to be derived from a particular local base (the so-called "third stem"), but these forms do not fall within any natural class.

A nice example of the need for local bases comes from the Spanish present subjunctive (Harris 1969; Butt 1997). In an earlier stage of Spanish, a phonological rule that deleted velar stops before front vowels created alternations within verb paradigms.

(90) *cono[s]er* 'to know'

imper.	cono[s]e	
pres. ind.	**cono[sk]o** cono[s]es cono[s]e	cono[s]emos cono[s]éis cono[s]en
pret.	cono[s]í cono[s]iste cono[s]ió	cono[s]imos cono[s]isteis cono[s]ieron
pres. subj.	**cono[sk]a** **cono[sk]as** **cono[sk]a**	**cono[sk]amos** **cono[sk]áis** **cono[sk]an**

Within the present indicative paradigm, this led to a neutralization of [k] with ∅ in most of the paradigm, and created what looks like unpredictable insertion of a [k] in the 1sg form. In spite of this neutralization, the optimal base within the present indicative paradigm turns out to be the 1pl form,[15] meaning that the mapping to derive the 1sg form often involves needing to decide whether or not to insert a velar. Some support for this analysis comes from the fact that historically incorrect velars have been added to new verbs, such as in the verb *venir* 'to come':

(91) *venir* 'to come'

imper.	ven	
pres. ind.	**ven[g]o** vienes viene	venimos venís vienen
pret.	vine viniste vino	vinimos vinisteis vinieron
pres. subj.	**ven[g]a** **ven[g]as** **ven[g]a**	**ven[g]amos** **ven[g]áis** **ven[g]an**

One might think, given this data, that the analysis that Spanish speakers had come up with was to use a form like the infinitive as the base, and then the mappings to the forms with back vowels each included a set of rules adding a velar (1sg indicative: *imos* → *go* / ___#; 1sg subjunctive: *imos* → *ga* / ___#; etc.). There is another way to describe the same facts using a local base, however. In particular, we could say that the 1sg indicative is derived from the infinitive, and then that the subjunctive is derived from the 1sg indicative (1sg indicative: *imos* → *go*; 1sg subjunctive: *o* → *a*). There is historical evidence that this is in fact the analysis that speakers

[15] I have used the model proposed here to verify this computationally, using a database of Spanish verbs taken from the 5.5 million word LEXESP corpus (Sebastián, Cuetos, and Carreiras, in press). It turns out that the problem of predicting velar insertion and other changes that take place in the singular, such as diphthongization and raising, are all much easier than the problem of predicting which conjugation class a verb belongs to (*-ar*, *-er*, or *-ir*). Therefore, the model chooses a base form that unambiguously reveals the conjugation class (namely, the infinitive, 1pl or 2pl), and has to learn how to predict velar insertion. Bybee (1985, p. 60) claims that the 3sg is the base of Spanish paradigms and that the remainder of the forms can be derived from it, but this is only true for the first conjugation example (*cantar* 'to sing') that she uses; in the other two conjugations, the theme vowel is neutralized to *e* in the 3sg. Much of Bybee's evidence in favor of the 3sg in Spanish can be reinterpreted as showing that the 3sg is a local base for the other singular forms, not the global base of Spanish verb paradigms.

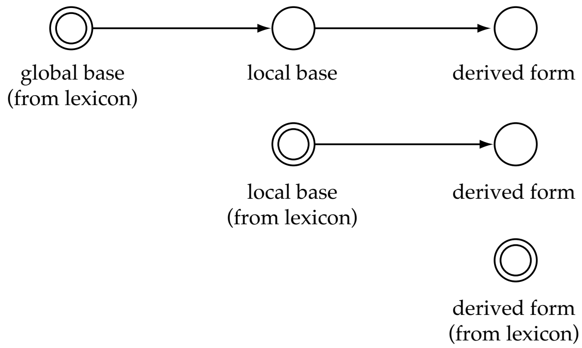


Figure 6.5: Three possible ways to produce a non-basic form using a local base

were using. There is an extremely irregular Spanish verb *caber* 'to fit', which has the 1sg form *quepo* [kepo], with a completely idiosyncratic vowel change and stop devoicing. This irregularity has been extended from the 1sg to the entire present subjunctive:

(92) *caber* 'to fit'

imper.	cabe	
pres. ind.	**quepo**	cabemos
	cabes	cabéis
	cabe	caben
pret.	cupe	cupimos
	cupiste	cupisteis
	cupo	cupieron
pres. subj.	**quepa**	**quepamos**
	quepas	**quepáis**
	quepa	**quepan**

This example, like Aronoff's Latin third stem example, shows that speakers know that some local paradigms may be productively derived from forms other than the global base. In other words, there is no question that we need local bases.

How would a system with local bases be used to produce forms? In the simpler case of a single paradigm, we saw that there are two ways to produce a form: it may either be memorized (provided that it is frequent enough), or it may be synthesized, sometimes yielding different results (Figure 6.4). The same is also true with local bases, but when there are local bases, there may be more steps in the derivation, as shown in Figure 6.5. The prediction, then, is that two types of regularization are possible: leveling of a subparadigm to its local base, and leveling of a subparadigm to the global base. Leveling to the global base requires that a lexically listed local base be inaccessible, and have been regenerated starting from the global base. What is excluded under this model, therefore, is a leveling in which some forms in the subparadigm are overregularized on the basis of the global base, while others continue to be derived from an irregular, listed local base. The reason is that if a local base is inaccessible for some of its derivatives, it should be (on average) inaccessible for all of them.

6.3.2 Extending the model to find local bases

In order to discover local bases, we need two things. First, we need a mechanism that finds not only the single optimal base form for the entire paradigm, but also other local pockets of strong predictability that might diagnose local bases. In addition, we need the model to be able to discover which forms should be derived from which bases, given that this may be somewhat arbitrary (as in the case of the Latin third stem). This is shown schematically in Figure 6.6, which indicates that a remote part of the paradigm that is strongly predictable from the 1pl remains attached to it, while another remote part of the paradigm with common unpredictable features "break off' for be derived by a more local base.

To a certain extent, the procedure for discovering local bases is a natural extension of the procedure described in chapter 3 for finding global bases; we compare numerous possible mappings and find the forms that provide the most reliable mappings to the remaining forms. There are several possible approaches.

The procedure outlined in chapter 3 is designed to find a single global base. It does this by determining which form has the best average score in deriving the remainder of the paradigm. Recall how this worked in the case of Latin; the scores for each form according to the "percent correct" metric are given in Table 6.2. The rows in the table represent all of the subgrammars based on a certain input form (all of the subgrammars based on the nominative, on the accusative, etc.), while the columns represent all the ways to derive a form as an output. As can be seen in the rightmost column, the dative performs, on average, best in deriving the other forms, and would be selected as the base (indicated by shading). This is not to say, however, that the dative is the best base for every single form in the paradigm. For example, in the nominative column, we see that the nominative can be predicted somewhat more accurately from the accusative form (94.5% accuracy) than from the dative form (91.7% accuracy), because the nominative and accusative are identical for all neuter nouns; this is shown in bold type.

What we want to do in establishing local bases, then, is to allow forms to delink from the global base and reassociate themselves to forms that could derive them more efficiently. In this case, the nominative and accusative forms could both find better bases elsewhere in the paradigm. So, one possible algorithm would be to allow a second "clean-up" stage of base identification: after the global base has been established (using the procedure in chapter 3), we then check each form in the paradigm to see if we could do better by deriving it more locally from some other form. If so, we "demote" the form so that it is derived by a local base. In the

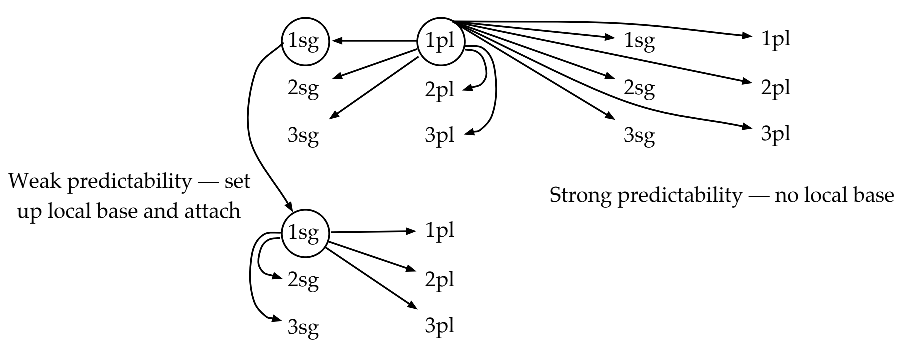


Figure 6.6: Setting up local bases by finding pockets of high reliability

Table 6.2: Percent correct in deriving Latin noun paradigms

↓In/Out→	nom.sg.	gen.sg.	dat.sg.	acc.sg.	abl.sg.	nom.pl.	average
nom.sg.		0.818	0.771	0.870	0.854	0.797	0.822
gen.sg.	0.929		**0.972**	0.964	0.957	0.945	0.953
dat.sg.	0.917	**0.982**		0.963	**0.966**	**0.957**	**0.957**
acc.sg.	**0.945**	0.937	0.921		0.931	0.924	0.932
abl.sg.	0.917	0.974	0.966	**0.968**		0.949	0.955
nom.pl	0.893	0.941	0.943	0.945	0.943		0.933

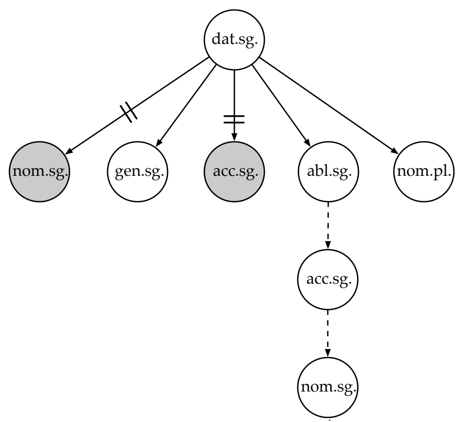


Figure 6.7: Result of "local demotion" for Latin noun paradigms

case of Latin, the nominative is best derived from the accusative, so it would get reassociated to the accusative as a local base, while the accusative in turn is best derived from the ablative. The resulting, somewhat more articulated paradigm is shown in Figure 6.7. (In this and subsequent figures, I borrow some notations from suprasegmental phonology: a double broken line indicates a form that has been "demoted", a dashed line indicates a derivation that has been added in the "demotion" stage. Suboptimally derived forms in search of a better base are shaded gray.)

When we apply this procedure to Spanish verbal paradigms, however, we run into problems. The equivalent accuracy scores for the Spanish present tense indicative paradigm are given in Table 6.3. Here we see that the infinitive, 1pl, and 2pl all do equally well in predicting the remainder of the paradigm, though the confidence limit adjustment discussed in section 6.2.1 would probably favor the infinitive over these other two forms. These forms are the best on average, but there are still a number of particular forms that we could do better for; the 1sg, 2sg, 3sg, and 3pl are all highly interpredictable because they share properties like mid-vowel diphthongization and raising. These "unhappy" forms are the ones shaded gray in Figure 6.8.

Now consider trying to demote or reassociate individual forms, as we did for Latin in Figure 6.7. Here we run into a problem of ties: the 1sg form could be equally well derived from the 2sg, the 3sg, or the 3pl; the 2sg form could be equally well derived from the 3sg and the 3pl, and so on. the 3sg and 3pl are 100% mutually predictable, leaving us unable to decide which to demote, while the 2sg is equally well predictable from both the 3sg and the 3pl, leaving us uncertain which to derive it from. What should we do in this case? If we were to allow multiple

Table 6.3: Percent correct in deriving Spanish present tense paradigms

↓In/Out→	1sg	2sg	3sg	1pl	2pl	3pl	infin.	average
1sg		0.895	0.896	0.867	0.867	0.896	0.867	0.881
2sg	**0.986**		0.995	0.740	0.740	0.995	0.740	0.866
3sg	**0.986**	**1.000**		0.750	0.750	**1.000**	0.750	0.873
1pl	0.902	0.908	0.909		**1.000**	0.909	**1.000**	**0.938**
2pl	0.902	0.908	0.909	**1.000**		0.909	**1.000**	**0.938**
3pl	**0.986**	**1.000**	**1.000**	0.750	0.750		0.750	0.873
infin.	0.902	0.908	0.909	**1.000**	**1.000**	0.909		**0.938**

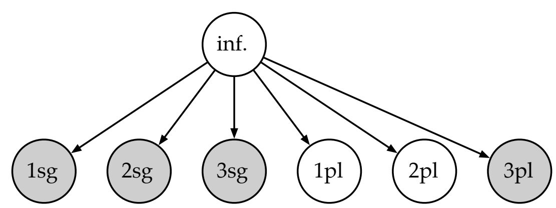
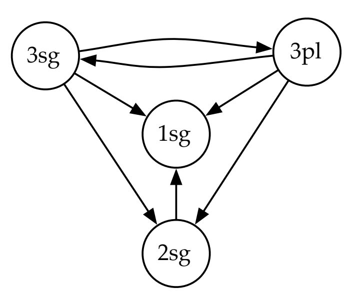


Figure 6.8: Results of global base selection for Spanish present tenses

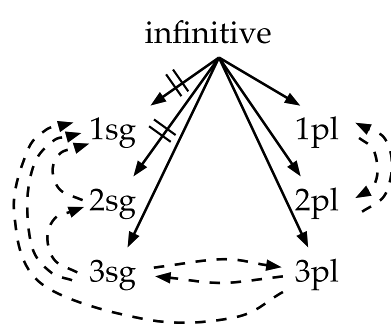
reassociations, we would end up with the rather complicated paradigm in Figure 6.9.

Paradigm structures of this type seem not only needlessly complicated, but also insufficiently predictive. The most interesting results in the preceding chapters came from the restriction that grammars be asymmetrical and have a single base; allowing paradigms like those Figure 6.9 removes many of these asymmetries. If we want to preserve the restrictiveness of the single base hypothesis as much as possible, we need some additional conditions on how local bases are assigned.

A sensible restriction would be to require that each form be derived from at most one other form, ruling out the multiple and circular derivations in Figure 6.9. Under this restriction, forms are reassociated to just one of their optimal local bases. (Note that this restriction also means

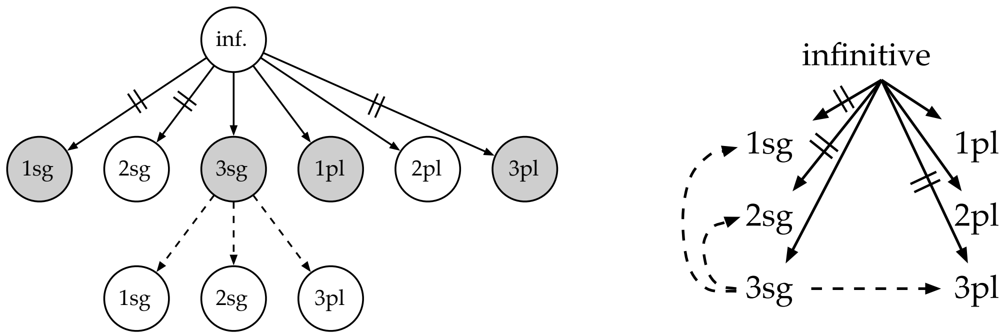


(a) Schematic representation



(b) Resulting paradigm

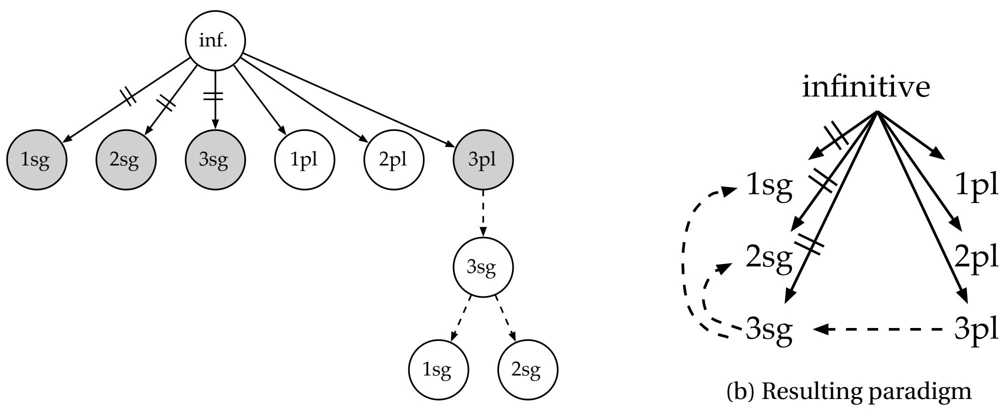
Figure 6.9: Circularity and ties in the Spanish present tense paradigm



(a) Demote 3pl below 3sg

(b) Resulting paradigm

Figure 6.10: 3sg as local base of Spanish paradigm



(a) Demote 3sg below 3pl

(b) Resulting paradigm

Figure 6.11: 3pl, 3sg as local bases of Spanish paradigm

that one of the unhappy forms must remain attached to the global base, in order to "break into the circle" of Figure 6.9.) This alone is not enough, however. The paradigms in Figures 6.10 and 6.11 both satisfy this restriction, but the one in Figure 6.10 seems simpler and more restrictive than the one in Figure 6.11. For ease of interpretation, each paradigm is shown in two notations: a tree-like notation showing demotions, and a more traditional paradigm-like notation. The problem is that the 3sg and 3pl forms are mutually predictable, so there is no way to decide which one to derive from the other.[16]

Intuitively, we want to find the grammar that allows each form to be derived as reliably as possible, but also makes use of the smallest number of local bases. The most promising way to do this is to let the base selection algorithm proceed recursively, rather than letting each suboptimally derived form float off individually in search of a better local base.

[16]Crucially, the confidence limit adjustment favoring better-instantiated forms could not help to decide here, because the number of (3sg,3pl) pairs is the same no matter which is the input and which is the output.

Table 6.4: Competition for basehood of 1sg,2sg,3sg,3pl subparadigm

↓In/Out→	1sg	2sg	3sg	3pl	average
1sg		0.895	0.896	0.896	0.896
2sg	**0.986**		0.995	0.995	0.992
3sg	**0.986**	**1.000**		**1.000**	**0.995**
3pl	**0.986**	**1.000**	**1.000**		**0.995**

The way this would work is as follows: first, the algorithm compares the entire paradigm, and selects a global base, as before. Then, each form is examined to see whether it can be derived more reliably by a form other than the global base. If it can, it is submitted to the second round, in which a secondary base is chosen. In the case of Spanish, this would mean that the 1sg, 2sg, 3sg, and 3pl must compete to become the local base of a subparadigm consisting of just these forms. The numbers for this competition are shown in Table 6.4. We see from the table that the 3sg and 3pl are equally good at deriving the other forms; in this case the higher token frequency of the 3sg would probably be the deciding factor favoring it over the 3pl, since there are more (3sg,2sg) and (3sg,1sg) pairs than there are (3pl,2sg) and (3pl,1pl) pairs available in the input set.

Once a local base is identified for the subparadigm, it stays where it is and the remaining forms are reassociated to it. At this point, all of the forms are being derived by one of their optimal bases, so there is no need to iterate further. This procedure yields the paradigm in Figure 6.10, and in general it will favor "flatter" paradigms because at each step in the process, it tries to find a single base form that can derive all of the remaining forms.

The procedure outlined here is a first pass at identifying local bases. Finding the optimal local bases given a set of predictability values is a hard problem, and there is clearly much more work to be done in understanding how local bases are identified. I anticipate that two additional restrictions may be helpful or even necessary in tackling the problem of larger paradigms, with multiple tenses, aspects, etc. The first is some sort of restriction on the batches of forms that are grouped together for the purpose of finding local bases. It is commonly observed (e.g., Bybee 1985, chap. 3) that forms within a particular tense/aspect/mood are more cohesive than forms from more distant parts of the paradigm. I have avoided making use of morphosyntactic features in choosing base form within relatively small paradigms, because it seems that no particular combination of morphosyntactic features guarantees that a form will serve as the base in a particular language. It would not be inconsistent with this approach, however, to limit the sets of forms that can be grouped together as subparadigms, by restricting the types of morphological rules that the system is allowed to consider. For example, it would be quite reasonable to prohibit the system from considering (2sg present indicative active, 3pl perfect subjunctive passive) pairs, since this would require a rule that simultaenously changes person, number, tense and mood features. The question of which mappings are possible is orthogonal to the question of how bases are selected from among a set of mappings. Nevertheless, in extended paradigms, these two factors would interact crucially in determining the final structure of the paradigm.

The second restriction on local bases that may be useful is some sort of predictability threshold for positing a local base. In the Latin noun paradigms above, the predictability advantage for deriving the accusative from the ablative (a local base) instead of from the dative (the global base) is tiny— 0.968 vs. 0.963. It seems unlikely that speakers complicate their paradigm structure enormously for such minuscule gains in predictability. It is an open question how large a threshold would achieve the desired balance between simplicity and effectiveness. An interesting sidenote, however, is that allowing more deeply embedded paradigm structure could help to alleviate the problem of "inferring nothing from non-base forms" mentioned at the end of the previous chapter. If Spanish 2sg forms are derived from a local base (the 3sg) which in turn is derived from a global base (the infinitive), then when the global base is not known, a speaker needing to produce a 2sg form would be less impaired, since it is likely that at least the local base (the 3sg) would be known. Thus, there may be a trade-off between choosing a paradigm with less structure, and a paradigm that is more likely to provide the input necessary to derive an unknown form.

The local bases algorithm proposed here is not likely to change the results in any crucial way for the cases discussed in the previous chapters. In the case of Latin, it would result in the nominative being derived locally from the accusative (mainly because they are identical for neuter nouns), which is still compatible with the result in chapter 4 that the nominative was rebuilt on the basis of oblique forms. In the case of Yiddish, the local base algorithm would likely create sub-paradigms for the plural forms (which would have always had identical vowels in pre-leveling Yiddish), and for the 2sg and 3sg (which were also always identical. In this case, too, the crucial aspect of the analysis (all forms ultimately derived from the 1sg) is preserved. Finally, in Lakhota, the unsuffixed forms (1sg, 3sg) would form their own local subparadigm, since they always share the same final vowel. However, the suffixed forms would still be derived from a global 2nd person base form, so in this case, too, the crucial asymmetries observed in chapter 5 would be preserved.

6.4 Relation to acquisition evidence

Another important source of evidence about bases that I have not made use of here is acquisition evidence from children's productions and errors. Bybee (1985) provides numerous pieces of evidence that 3sg forms, which she argues are the universally best base forms, are acquired first by children, and are also systematically substituted for other forms in the paradigm. A careful review of the acquisition evidence would be a logical next step in testing whether language learners really do use a strategy like the one proposed here for learning the organization of paradigms. It should be noted, however, that there are some types of child acquisition data that are probably *not* relevant to this question, including descriptions about which forms are uttered first by children. For example, Bybee (1985, p. 50) notes that Spanish-speaking children typically master 3sg forms first, while Berman (1985, p. 268) claims that for Hebrew-learning children, it is the imperative. The fact that a child can uttter a form means only that she has heard it enough to have memorized it, not that she has established it as a base form.

The errors made by children are a much more useful source of data about bases. Bybee cites data from Simões and Stoel-Gammon (1979) that shows that Portuguese-learning children incorrectly extend the vowel of the 3sg indicative to the 1sg. Similarly, Clahsen, Aveledo and Roca (to appear) show that Spanish-learning children make systematic errors in diphthongization, substituting the non-diphthongized form of the verb (found in the infinitive, 1pl, and 2pl) for the diphthongized version that should appear in the 1sg, 2sg, 3sg, and 3pl. It is this type of asymmetry that can help to shed light on the organization that children are imposing on paradigms in the course of acquisition.

6.5 Summary

In this chapter, I have tried to show how the proposed model, or a suitable extension of it, could handle not only the Yiddish, Latin and Lakhota cases discussed in the previous chapters, but also a wider variety of data, including problematic cases and typological tendencies. In some cases, such as Russian, I argued that the model's inability to arrive at the traditional analysis may not as much of a problem one might have thought. In other cases, such as Korean, I argued that more realistic simulations, taking into account not only the informativeness of forms, but also the amount of data available to learners, might help to explain why it is sometimes not the most informative form that acts as the base. Finally, I discussed the need to relax the single surface base hypothesis to accomodate local bases in extended paradigms, suggesting some possible (but tentative) approaches to the problem. In all of these cases, I have discussed ways in which the current approach makes strong, testable, and unique predictions about asymmetries in possible errors and historical changes.

Chapter 7

Conclusion

In this thesis, I have presented a computationally implemented model of paradigm learning in which learners identify one form as the base form, and then learn a grammar of morphological and phonological rules that operate on the base form to derive the remaining forms in the paradigm. I have pursued two related, but logically distinct hypotheses about how this is done. The first hypothesis is that learners are limited to selecting a single form as the base, and that the base form must be a surface form from somewhere within the paradigm. Furthermore, the choice of base is global, meaning that the same part of the paradigm must serve as the base for all lexical items. Such an approach is similar to the "basic alternant" approach of some pre-generative phonologists. The second hypothesis is that learners select the base form that is maximally informative, in the sense that it preserves the most contrasts, and permits accurate and productive generation of as many forms of as many words as possible. I have argued along the way that these hypotheses could be interpreted in different ways, depending on the model of morphology that is employed. If one adopts a word-based model of morphology, in which morphological operations take whole words as their inputs, then these are hypotheses about how learners figure out which forms in the paradigm to derive from which other forms. If, on the other hand, one adopts a stem-based model of morphology, in which morphological operations combine morphemes listed in their underlying forms, then these hypotheses constitute a restriction on the set of possible underlying forms that learners are allowed to consider.

The single surface base hypothesis is far more restrictive than the standard assumptions about how speakers arrive at the underlying representations of words. It is generally assumed in the phonological literature that speakers can combine unpredictable information from anywhere in the paradigm to construct underlying forms that ideally can be used to derive all surface properties of the word (Kenstowicz and Kisseberth 1977). When inputs are restricted to a single surface form, on the other hand, it is often impossible to find a single form that unambiguously reveals the surface properties of all of the members of the paradigm.

This situation can arise in two different ways. The first is when a language has multiple neutralizing phonological processes, with some affecting some slots in the paradigm, and others affecting other slots. This was the case in Latin, discussed in chapter 4: the nominative form preserved the $s \sim r$ contrast but suffered from various phonological neutralizations, while the oblique forms preserved most contrasts but merged stem-final *s* and *r*. This is also the case for stress-shifting nouns in Russian (section 6.1): unsuffixed forms preserve the contrast between *a* and *o* but suffer from final devoicing, while suffixed forms preserve the voicing of final obstruents but suffer from vowel reduction. In both Latin and Russian, every slot in the paradigm suffers from at least one phonological neutralization, so no single slot reveals all properties of all words.

The other situation in which there is no surface form that preserves all contrasts arises when a neutralization is *symmetrical* (chapter 5): the language contains non-alternating segments *a* and *b*, and an alternating segment $a \sim b$ that is neutralized to *a* in some forms and to *b* in others. This was the case in an earlier stage of Yiddish (chapter 2), since verbs with $a \sim e$ umlaut alternations were neutralized with invariant *a* verbs in some forms, and with invariant *e* verbs in others. The same was also true for final $a \sim e$ ablaut alternations in Lakhota (chapter 5). In all of these cases, the single surface base restriction means that the underlying form of a word cannot unambiguously encode all of the surface contrasts of the language.

If speakers cannot encode all of the unpredictable surface properties of a word in its underlying form, then some of this information must be stored elsewhere in the system. For example, if we cannot encode the contrast between root-final *s* and *r* in Latin as part of the underlying form, we need some other way to keep track of whether a particular word should end in *s* or *r* in the nominative. I proposed that in the absence of abstract underlying forms, speakers may nevertheless be able to overcome such neutralizations using a sophisticated system of morphological rules and exceptions. If we allow morphological rules to refer to details of the phonological and morphosyntactic environment—such as "after coronal obstruents", "in neuter nouns", and so on—then we can use these to formulate a grammar that captures intermediate and small-scale generalizations about groups of words that tend to behave alike. In the case of Latin, such a grammar includes rules specifying that masculine and feminine polysyllabic nouns ending in *-ris* in the genitive should end in *-r* in the nominative, while neuter nouns ending in *-ris* in the genitive should end in *-s* in the nominative, as should monosyllabic nouns of any gender. For words that do not conform to these generalizations, the nominative form must simply be stored as a lexical exception.

The single surface base hypothesis runs contrary to the standard use of underlying forms in phonology, but it has some advantages. First, it is computationally simpler as a learning strategy, since it greatly reduces the number of underlying representations that must be considered. It also means that there need not be a mechanism for positing abstract underlying representations that combine information from multiple forms, or that contain underspecified elements such as archiphonemes, since such elements never occur in any surface form. Moreover, I have shown throughout the course of this thesis that the single surface base hypothesis also has empirical advantages. In particular, it predicts asymmetries in which forms and which contrasts should be open to analogical change: analogy should affect only non-basic forms, and it should target only those contrasts that are not preserved in the base form. In other words, analogical changes happen only when speakers are forced to guess their way out of a neutralization.

There are several aspects of this proposal that could benefit from further inquiry. First, I have needed to suppose at various points that there is a bias towards selecting bases with higher token frequency, because they are more readily available as inputs to derive other forms. I have tried to maintain the view that this bias is not actually a formal grammar acquisition strategy, but is rather a side effect of the way that learners assess the reliability of generalizations. In particular, generalizations based on more data are more trustable, and higher frequency forms give the learner more input data from which to create generalizations. I suggested that the confidence limit adjustment already employed by the minimal generalization learner might provide an appropriate way to model this. Further modeling, with more realistic input sets in which various forms occur in their actual proportions, is needed to better understand the trade-off between informativeness and availability, and whether confidence limit statistics are the best way to model it.

Further research is also required concerning the implications of this proposal for the use of underlying forms in phonology. In all three cases discussed here (Yiddish, Latin, Lakhota), I argued that a model that can combined unpredictable information from various parts of the paradigm cannot predict the asymmetries observed in subsequent historical changes. In all three of these cases, however, the amount of recombination that would be necessary to arrive at the "traditional" underlying forms was quite small, and the expense of having to store the extra information as lexical exceptions was not all that great. In chapter 6, I considered whether this type of analysis could be extended to a case in which significantly more recombination is necessary (Russian), suggesting that perhaps even here, an analysis restricted to using a single surface form as the UR might not require an unreasonably large number of lexical exceptions. It would be worthwhile expanding this analysis beyond the sample of words that I have considered in that discussion, to see what would be required to analyze all of Russian under this system. More generally, phonological analyses that are restricted to using a single surface form as the UR may often look quite different from traditional analyses, and have the potential to make unique predictions about what patterns speakers may notice and extend.

Finally, I have mentioned at several points that data from language change is only one of the many areas in which base effects can be seen. Two other areas that seem particularly relevant are acquisition evidence from children, and experimental evidence from psycholinguistic studies. As discussed in section 6.4, the type of acquisition evidence that would most directly support or refute this approach concerns asymmetries between forms: we can test whether overgeneralization affects some parts of the paradigm but leaves others intact, extends some patterns but not others, and so on. In psycholinguistic experiments, too, we might look for asymmetries, such as differences in how inflectionally forms affect one another in a morphological priming task, or some type of dissociation between forms that are predicted to be productively derived by the grammar and those that must be stored as lexical exceptions. Converging evidence from historical changes, child data, and psycholinguistic tasks could play a crucial role in assessing the plausibility of the current approach as a model of how speakers learn to produce and comprehend paradigms of inflectionally related forms.

Appendix A

Results for Neutralize 1 (chapter 3)

A.1 Input forms

The Neutralize 1 language has palatalization of *k* before *i*, resulting in neutralization of *k* and *t͡ʃ* in the ergative.

a. Forms ending in segments other than *k*

absolutive		ergative
dap	∼	dapi
lot	∼	loti
gub	∼	gubi
sat͡ʃ	∼	sat͡ʃi
rut͡ʃ	∼	rut͡ʃi
lag	∼	lagi
ban	∼	bani
yul	∼	yuli

b. Forms ending in *k*

absolutive		ergative
ʔak	∼	ʔat͡ʃi
muk	∼	mut͡ʃi
lok	∼	lot͡ʃi

A.2 Absolutive to ergative direction

A.2.1 Grammar for the *absolutive→ergative* direction

The following table lists all of the rules created in the process of learning the mapping from the absolutive form to the ergative form, in order of discovery. The "parents" column lists the two rules which were compared to yield the generalization (or, in the case of word-specific rules, a '*', indicating that the learner has just received a new input pair). The right three columns list the reliability statistics of the rule; its *hits* (the number of forms it can derive correctly), its *scope* (the number of forms it could potentially apply to in the lexicon) and its *confidence* (the ratio of the hits to scope, adjusted with lower confidence limit statistics). The confidence value for a word-specific rule is undefined.

Note that the feature specifications discovered by the minimal generalization learner do not use contrastive underspecification (e.g., Steriade 1995), but rather define natural classes using *all* of the features that the segments have in common. Thus, the set of features that define the natural class $\{o,u\}$ includes not only [+syllabic], [−consonantal], [+back],[−low] and [+tense], but also features like [+voice], [−nasal], and so on.

The best rule in the grammar is rule number 59, which attaches *-i* after any consonant. This rule correctly accounts for all of the forms in the input data, achieving a confidence of .916.

a. Morphological rules in the absolutive to ergative direction

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
1	*	∅ → i /			dap	___#	1	1	*undef*
2	*	∅ → i /			lot	___#	1	1	*undef*
3	1,2	∅ → i /	X	[−syl +cons −son −cont −nas −voi +ant −dors] = {p,t}		___#	2	2	0.570
4	*	∅ → i /		=	gub	___#	1	1	*undef*
5	1,4	∅ → i /	X	[−syl +cons −son −cont −nas +lab −cor +ant +distr −dors] = {b,p}		___#	2	2	0.570
6	3,4	∅ → i /	X	[−syl +cons −son −cont −nas +ant −dors] = {b,d,p,t}		___#	3	3	0.718
7	*	∅ → i /		=	sat͡ʃ	___#	1	1	*undef*

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
8	1,7	∅ → i /	X	[−syl +cons −son −cont −nas −voi +distr −dors] = {t͡ʃ,p}		___#	3	3	0.718
9	2,7	∅ → i /	X	[−syl +cons −son −cont −nas −voi −lab +cor −dors] = {t͡ʃ,t}		___#	3	3	0.718
10	3,7	∅ → i /	X	[−syl +cons −son −cont −nas −voi −dors] = {t͡ʃ,p,t}		___#	4	4	0.786
11	4,7	∅ → i /	X	[−syl +cons −son −cont −nas +distr −dors] = {t͡ʃ,d͡ʒ,b,p}		___#	4	4	0.786
12	6,7	∅ → i /	X	[−syl +cons −son −cont −nas −dors] = {t͡ʃ,d͡ʒ,b,d,p,t}		___#	5	5	0.825
13	*	∅ → i /		=	rut͡ʃ	___#	1	1	*undef*

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
14	7,13	∅ → i /	X	[+syl −cons +son +cont −nas +voi −rnd −cor +ant +distr +dors −front +tense] = {a,o,u}	t͡ʃ	___#	2	2	0.570
15	*	∅ → i /		=	lag	___#	1	1	*undef*
16	1,15	∅ → i /	X	[−syl +cons −son −cont −nas −cor +ant +distr] = {b,g,k,p}		___#	6	6	0.852
17	2,15	∅ → i /	X	[−syl +cons −son −cont −nas −lab +ant] = {d,g,k,t}		___#	5	5	0.825
18	3,15	∅ → i /	X	[−syl +cons −son −cont −nas +ant] = {b,d,g,k,p,t}		___#	7	7	0.872
19	4,15	∅ → i /	X	[−syl +cons −son −cont −nas +voi −cor +ant +distr] = {b,g}		___#	2	2	0.570

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
20	7,15	∅ → i /	X	[−syl, +cons, −son, −cont, −nas, −lab, +distr] = {t͡ʃ,d͡ʒ,g,k}		___#	6	6	0.852
21	8,15	∅ → i /	X	[−syl, +cons, −son, −cont, −nas, +distr] = {t͡ʃ,d͡ʒ,b,g,k,p}		___#	8	8	0.887
22	9,15	∅ → i /	X	[−syl, +cons, −son, −cont, −nas, −lab] = {t͡ʃ,d͡ʒ,d,g,k,t}		___#	7	7	0.872
23	10,15	∅ → i /	X	[−syl, +cons, −son, −cont, −nas] = {t͡ʃ,d͡ʒ,b,d,g,k,p,t}		___#	9	9	0.898
24	*	∅ → i /		=	ban	___#	1	1	*undef*
25	1,24	∅ → i /	X	[−syl, +cons, −cont, +ant, −dors] = {b,d,m,n,p,t}		___#	4	4	0.786
26	2,24	∅ → i /	X	[−syl, +cons, −cont, −lab, +cor, +ant, −distr, −strid, −dors] = {d,n,t}		___#	2	2	0.570
27	3,24	∅ → i /	X	[−syl, +cons, −cont, +voi, +ant, −dors] = {b,d,m,n}		___#	2	2	0.570

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
28	7,24	∅ → i /	X	[−syl +cons −cont −lab +cor −dors] = {t͡ʃ,d͡ʒ,d,n,t}		___#	4	4	0.786
29	8,24	∅ → i /	X	[−syl +cons −cont −dors] = {t͡ʃ,d͡ʒ,b,d,m,n,p,t}		___#	6	6	0.852
30	15,24	∅ → i /	X	[−syl +cons −cont +voi −lab +ant] = {ŋ,d,g,n}		___#	2	2	0.570
31	16,24	∅ → i /	X	[−syl +cons −cont +ant] = {ŋ,b,d,g,k,m,n,p,t}		___#	8	8	0.887
32	17,24	∅ → i /	X	[−syl +cons −cont −lab +ant] = {ŋ,d,g,k,n,t}		___#	6	6	0.852
33	19,24	∅ → i /	X	[−syl +cons −cont +voi +ant] = {ŋ,b,d,g,m,n}		___#	3	3	0.718
34	20,24	∅ → i /	X	[−syl +cons −cont −lab] = {t͡ʃ,d͡ʒ,ŋ,d,g,k,n,t}		___#	8	8	0.887
35	21,24	∅ → i /	X	[−syl +cons −cont] = {t͡ʃ,d͡ʒ,ŋ,b,d,g,k,m,n,p,t}		___#	10	10	0.908
36	*	∅ → i /		=	yul	___#	1	1	*undef*
37	1,36	∅ → i /	X	[−syl +cons −nas +ant −dors] = {ð,θ,b,d,f,l,p,r,s,t,v,z}		___#	4	4	0.786

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
38	2,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ -\text{lab} \\ +\text{cor} \\ +\text{ant} \\ -\text{distr} \\ -\text{dors} \end{bmatrix}$ = {d,l,r,t}		___#	2	2	0.570
39	4,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ +\text{voi} \\ +\text{ant} \\ -\text{dors} \end{bmatrix}$ = {ð,b,d,l,r,v,z}		___#	2	2	0.570
40	7,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ -\text{lab} \\ +\text{cor} \\ -\text{dors} \end{bmatrix}$ = {t͡ʃ,ð,d͡ʒ,ʃ,θ,ʒ,d,l,r,s,t,z}		___#	4	4	0.786
41	8,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ -\text{dors} \end{bmatrix}$ = {t͡ʃ,ð,d͡ʒ,ʃ,θ,ʒ,b,d,f,l,p,r,s,t,v,z}		___#	6	6	0.852
42	15,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ +\text{voi} \\ -\text{lab} \\ +\text{ant} \end{bmatrix}$ = {ð,d,g,l,r,z,}		___#	2	2	0.570
43	16,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ +\text{ant} \end{bmatrix}$ = {ð,θ,b,d,f,g,k,l,p,r,s,t,v,x,z,}		___#	8	8	0.887
44	17,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ -\text{lab} \\ +\text{ant} \end{bmatrix}$ = {ð,θ,d,g,k,l,r,s,t,x,z,}		___#	6	6	0.852
45	19,36	$\varnothing \rightarrow$ i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ +\text{voi} \\ +\text{ant} \end{bmatrix}$ = {ð,b,d,g,l,r,v,z,}		___#	3	3	0.718

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
46	20,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \\ -\text{lab} \end{bmatrix}$ = {t͡ʃ,ð,d͡ʒ,ʃ,θ,ʒ,d,g,k,l,r,s,t,x,z,}		___#	8	8	0.887
47	21,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{nas} \end{bmatrix}$ = {t͡ʃ,ð,d͡ʒ,ʃ,θ,ʒ,b,d,f,g,k,l,p,r,s,t,v,x,z,}		___#	10	10	0.908
48	24,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ +\text{son} \\ +\text{voi} \\ -\text{lab} \\ +\text{cor} \\ +\text{ant} \\ -\text{distr} \\ -\text{dors} \end{bmatrix}$ = {l,n,r}		___#	2	2	0.570
49	25,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ +\text{ant} \\ -\text{dors} \end{bmatrix}$ = {ð,θ,b,d,f,l,m,n,p,r,s,t,v,z}		___#	5	5	0.825
50	26,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{lab} \\ +\text{cor} \\ +\text{ant} \\ -\text{distr} \\ -\text{dors} \end{bmatrix}$ = {d,l,n,r,t}		___#	3	3	0.718
51	27,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ +\text{voi} \\ +\text{ant} \\ -\text{dors} \end{bmatrix}$ = {ð,b,d,l,m,n,r,v,z}		___#	3	3	0.718
52	28,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{lab} \\ +\text{cor} \\ -\text{dors} \end{bmatrix}$ = {t͡ʃ,ð,d͡ʒ,ʃ,θ,ʒ,d,l,n,r,s,t,z}		___#	5	5	0.825
53	29,36	∅ → i /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{dors} \end{bmatrix}$ = {t͡ʃ,ð,d͡ʒ,ʃ,θ,ʒ,b,d,f,l,m,n,p,r,s,t,v,z}		___#	7	7	0.872

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
54	30,36	∅ → i /	X	[−syl +cons +voi −lab +ant] = {ð,ŋ,d,g,l,n,r,z,}		___#	3	3	0.718
55	31,36	∅ → i /	X	[−syl +cons +ant] = {ð,ŋ,θ,b,d,f,g,k,l,m,n,p,r,s,t,v,x,z,}		___#	9	9	0.898
56	32,36	∅ → i /	X	[−syl +cons −lab +ant] = {ð,ŋ,θ,d,g,k,l,n,r,s,t,x,z,}		___#	7	7	0.872
57	33,36	∅ → i /	X	[−syl +cons +voi +ant] = {ð,ŋ,b,d,g,l,m,n,r,v,z,}		___#	4	4	0.786
58	34,36	∅ → i /	X	[−syl +cons −lab] = {t͡ʃ,ð,d͡ʒ,ŋ,ʃ,θ,ʒ,d,g,k,l,n,r,s,t,x,z,}		___#	9	9	0.898
59	35,36	∅ → i /	X	[−syl +cons] = {t͡ʃ,ð,d͡ʒ,ŋ,ʃ,θ,ʒ,b,d,f,g,k,l,m,n,p,r,s,t,v,x,z,}		___#	11	11	0.916
60	*	k → t͡ʃi /		=	ʔa	___#	1	1	*undef*
61	*	k → t͡ʃi /		=	mu	___#	1	1	*undef*
62	60,61	k → t͡ʃi /	X	[+syl −cons +son +cont −nas +voi −cor +ant +distr +dors −front +tense] = {a,o,u}		___#	3	3	0.718
63	*	k → t͡ʃi /		=	lo	___#	1	1	*undef*

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
64	60,63	k → t͡ʃi /	X	$\begin{bmatrix}+\text{syl}\\ -\text{cons}\\ +\text{son}\\ +\text{cont}\\ -\text{nas}\\ +\text{voi}\\ -\text{cor}\\ +\text{ant}\\ +\text{distr}\\ +\text{dors}\\ -\text{hi}\\ -\text{front}\\ +\text{tense}\end{bmatrix}$ = {a,o}		___#	2	2	0.570
65	61,63	k → t͡ʃi /	X	$\begin{bmatrix}+\text{syl}\\ -\text{cons}\\ +\text{son}\\ +\text{cont}\\ -\text{nas}\\ +\text{voi}\\ +\text{lab}\\ +\text{rnd}\\ -\text{cor}\\ +\text{ant}\\ +\text{distr}\\ +\text{dors}\\ -\text{low}\\ -\text{front}\\ +\text{back}\\ +\text{tense}\end{bmatrix}$ = {o,u}		___#	2	2	0.570

b. Phonological rule: /k/ → [t͡ʃ] / ___i

A.2.2 Using the grammar to derive ergatives for each word

The highest confidence rule in the grammar (number 59) applies to all 11 words in this hypothetical language; therefore, if we were to use the grammar to derive ergatives for each known absolutive, this rule could always be used to attach *-i* with a confidence of .916. For words that end in *k* in the absolutive, the phonological rule *k* → *t͡ʃ* / ___i would automatically change the *k* into a *t͡ʃ*.

A.2.3 Calculation of metrics in the *absolutive→ergative* direction

Accuracy

Rule 59, together with the phonological rule *k* → *t͡ʃ* / ___*i*, productively generates the correct output for all 11 words of the language. Therefore, the accuracy of this grammar is 100%.

Mean confidence of rules in the grammar

The mean confidence of the rules in the grammar is found by averaging the *confidence* values of the grammar; this value is .762.

Mean confidence of winning outputs

When the grammar is used to derive outputs productively for each known word, the same rule (number 59) is used in all cases; its confidence is .916, so the mean confidence of winning outputs in this direction is .916.

Average winning margin

This grammar yields just one output per form (attaching -*i*, along with palatalization in the case of *k*). Thus, there are no other competing outputs, so I will assume that the confidence in the second best output using this grammar is 0 (no second choice). This means that the average difference between the best form and the second form is, in this case, identical with the average confidence of the winning form: .916.

A.3 Ergative to absolutive direction

A.3.1 Grammar for the *ergative→absolutive* direction

a. Morphological rules in the ergative to absolutive direction

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
1	*	i → ∅ /			dap	___#	1	1	*undef*
2	*	i → ∅ /			lot	___#	1	1	*undef*
3	1,2	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −voi, −rd, +ant, −dors] = {p,t}		___#	2	2	0.570
4	*	i → ∅ /			gub	___#	1	1	*undef*

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
5	1,4	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, +lab, −rd, −cor, +ant, +distr, −dors] = {b,p}		___#	2	2	0.570
6	2,4	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −rd, +ant, −dors] = {b,d,p,t}		___#	3	3	0.718
7	*	i → ∅ /			sat͡ʃ	___#	1	1	*undef*
8	1,7	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −voi, −rd, +distr, −dors] = {t͡ʃ,p}		___#	3	6	0.351
9	2,7	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −voi, −lab, −rd, +cor, −dors] = {t͡ʃ,t}		___#	3	6	0.351
10	3,7	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −voi, −rd, −dors] = {t͡ʃ,p,t}		___#	4	7	0.428

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
11	4,7	i → ∅ /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{son} \\ -\text{cont} \\ -\text{nas} \\ -\text{rd} \\ +\text{distr} \\ -\text{dors} \end{bmatrix}$ = {t͡ʃ,d͡ʒ,b,p}		___#	4	7	0.428
12	6,7	i → ∅ /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{son} \\ -\text{cont} \\ -\text{nas} \\ -\text{rd} \\ -\text{dors} \end{bmatrix}$ = {t͡ʃ,d͡ʒ,b,d,p,t}		___#	5	8	0.489
13	*	i → ∅ /			rut͡ʃ	___#	1	1	*undef*
14	7,13	i → ∅ /	X	$\begin{bmatrix} +\text{syl} \\ -\text{cons} \\ +\text{son} \\ +\text{cont} \\ -\text{nas} \\ +\text{voi} \\ -\text{cor} \\ +\text{ant} \\ +\text{distr} \\ +\text{dors} \\ -\text{front} \\ +\text{tense} \end{bmatrix}$ = {a,o,u}	t͡ʃ	___#	2	5	0.254
15	*	i → ∅ /			lag	___#	1	1	*undef*
16	1,15	i → ∅ /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{son} \\ -\text{cont} \\ -\text{nas} \\ -\text{rd} \\ -\text{cor} \\ +\text{ant} \\ +\text{distr} \end{bmatrix}$ = {b,g,k,p}		___#	3	3	0.718
17	2,15	i → ∅ /	X	$\begin{bmatrix} -\text{syl} \\ +\text{cons} \\ -\text{son} \\ -\text{cont} \\ -\text{nas} \\ -\text{lab} \\ -\text{rd} \\ +\text{ant} \end{bmatrix}$ = {d,g,k,t}		___#	2	2	0.570

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
18	3,15	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −rd, +ant] = {b,d,g,k,p,t}		___#	4	4	0.786
19	4,15	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, +voi, −rd, −cor, +ant, +distr] = {b,g}		___#	2	2	0.570
20	7,15	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −lab, −rd, +distr] = {t͡ʃ,d͡ʒ,g,k}		___#	3	6	0.351
21	8,15	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −rd, +distr] = {t͡ʃ,d͡ʒ,b,g,k,p}		___#	5	8	0.489
22	9,15	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −lab, −rd] = {t͡ʃ,d͡ʒ,d,g,k,t}		___#	4	7	0.428
23	10,15	i → ∅ /	X	[−syl, +cons, −son, −cont, −nas, −rd] = {t͡ʃ,d͡ʒ,b,d,g,k,p,t}		___#	6	9	0.537
24	*	i → ∅ /			ban	___#	1	1	*undef*

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
25	1,24	i → ∅ /	X	$\begin{bmatrix}-\text{syl}\\+\text{cons}\\-\text{cont}\\-\text{rd}\\+\text{ant}\\-\text{dors}\end{bmatrix}$ = {b,d,m,n,p,t}		___#	4	4	0.786
26	2,24	i → ∅ /	X	$\begin{bmatrix}-\text{syl}\\+\text{cons}\\-\text{cont}\\-\text{lab}\\-\text{rd}\\+\text{cor}\\+\text{ant}\\-\text{distr}\\-\text{dors}\end{bmatrix}$ = {d,n,t}		___#	2	2	0.570
27	4,24	i → ∅ /	X	$\begin{bmatrix}-\text{syl}\\+\text{cons}\\-\text{cont}\\+\text{voi}\\-\text{rd}\\+\text{ant}\\-\text{dors}\end{bmatrix}$ = {b,d,m,n}		___#	2	2	0.570
28	7,24	i → ∅ /	X	$\begin{bmatrix}-\text{syl}\\+\text{cons}\\-\text{cont}\\-\text{lab}\\-\text{rd}\\+\text{cor}\\-\text{dors}\end{bmatrix}$ = {t͡ʃ,d͡ʒ,d,n,t}		___#	4	7	0.428
29	8,24	i → ∅ /	X	$\begin{bmatrix}-\text{syl}\\+\text{cons}\\-\text{cont}\\-\text{rd}\\-\text{dors}\end{bmatrix}$ = {t͡ʃ,d͡ʒ,b,d,m,n,p,t}		___#	6	9	0.537
30	15,24	i → ∅ /	X	$\begin{bmatrix}-\text{syl}\\+\text{cons}\\-\text{cont}\\+\text{voi}\\-\text{lab}\\-\text{rd}\\+\text{ant}\end{bmatrix}$ = {ŋ,d,g,n}		___#	2	2	0.570
31	16,24	i → ∅ /	X	$\begin{bmatrix}-\text{syl}\\+\text{cons}\\-\text{cont}\\-\text{rd}\\+\text{ant}\end{bmatrix}$ = {ŋ,b,d,g,k,m,n,p,t}		___#	5	5	0.825

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
32	17,24	i → ∅ /	X	[−syl, +cons, −cont, −lab, −rd, +ant] = {ŋ,d,g,k,n,t}		___#	3	3	0.718
33	19,24	i → ∅ /	X	[−syl, +cons, −cont, +voi, −rd, +ant] = {ŋ,b,d,g,m,n}		___#	3	3	0.718
34	20,24	i → ∅ /	X	[−syl, +cons, −cont, −lab, −rd] = {t͡ʃ,d͡ʒ,ŋ,d,g,k,n,t}		___#	5	8	0.489
35	21,24	i → ∅ /	X	[−syl, +cons, −cont, −rd] = {t͡ʃ,d͡ʒ,ŋ,b,d,g,k,m,n,p,t}		___#	7	10	0.579
36	*	i → ∅ /			yul	___#	1	1	*undef*
37	1,36	i → ∅ /	X	[−syl, +cons, −nas, −rd, +ant, −dors] = {ð,θ,b,d,f,l,p,r,s,t,v,z}		___#	4	4	0.786
38	2,36	i → ∅ /	X	[−syl, +cons, −nas, −lab, −rd, +cor, +ant, −distr, −dors] = {d,l,r,t}		___#	2	2	0.570
39	4,36	i → ∅ /	X	[−syl, +cons, −nas, +voi, −rd, +ant, −dors] = {ð,b,d,l,r,v,z}		___#	2	2	0.570

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
40	7,36	i → ∅ /	X	[−syl, +cons, −nas, −lab, −rd, +cor, −dors] = {t͡ʃ,ð,d͡ʒ,ʃ,θ,Z,d,l,r,s,t,z}		___#	4	7	0.428
41	8,36	i → ∅ /	X	[−syl, +cons, −nas, −rd, −dors] = {t͡ʃ,ð,d͡ʒ,ʃ,θ,Z,b,d,f,l,p,r,s,t,v,z}		___#	6	9	0.537
42	15,36	i → ∅ /	X	[−syl, +cons, −nas, +voi, −lab, −rd, +ant] = {ð,d,g,l,r,z,}		___#	2	2	0.570
43	16,36	i → ∅ /	X	[−syl, +cons, −nas, −rd, +ant] = {ð,θ,b,d,f,g,k,l,p,r,s,t,v,x,z,}		___#	5	5	0.825
44	17,36	i → ∅ /	X	[−syl, +cons, −nas, −lab, −rd, +ant] = {ð,θ,d,g,k,l,r,s,t,x,z,}		___#	3	3	0.718
45	19,36	i → ∅ /	X	[−syl, +cons, −nas, +voi, −rd, +ant] = {ð,b,d,g,l,r,v,z,}		___#	3	3	0.718
46	20,36	i → ∅ /	X	[−syl, +cons, −nas, −lab, −rd] = {t͡ʃ,ð,d͡ʒ,ʃ,θ,Z,d,g,k,l,r,s,t,x,z,}		___#	5	8	0.489

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
47	21,36	i → ∅ /	X	[−syl, +cons, −nas, −rd] = {t͡ʃ,ð,d͡ʒ,ʃ,θ,Z,b,d,f,g,k,l,p,r,s,t,v,x,z,}		___#	7	10	0.579
48	24,36	i → ∅ /	X	[−syl, +cons, +son, +voi, −lab, −rd, +cor, +ant, −distr, −dors] = {l,n,r}		___#	2	2	0.570
49	25,36	i → ∅ /	X	[−syl, +cons, −rd, +ant, −dors] = {ð,θ,b,d,f,l,m,n,p,r,s,t,v,z}		___#	5	5	0.825
50	26,36	i → ∅ /	X	[−syl, +cons, −lab, −rd, +cor, +ant, −distr, −dors] = {d,l,n,r,t}		___#	3	3	0.718
51	27,36	i → ∅ /	X	[−syl, +cons, +voi, −rd, +ant, −dors] = {ð,b,d,l,m,n,r,v,z}		___#	3	3	0.718
52	28,36	i → ∅ /	X	[−syl, +cons, −lab, −rd, +cor, −dors] = {t͡ʃ,ð,d͡ʒ,ʃ,θ,Z,d,l,n,r,s,t,z}		___#	5	8	0.489
53	29,36	i → ∅ /	X	[−syl, +cons, −rd, −dors] = {t͡ʃ,ð,d͡ʒ,ʃ,θ,Z,b,d,f,l,m,n,p,r,s,t,v,z}		___#	7	10	0.579

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
54	30,36	i → ∅ /	X	[−syl +cons +voi −lab −rd +ant] = {ð,ŋ,d,g,l,n,r,z,}		___#	3	3	0.718
55	31,36	i → ∅ /	X	[−syl +cons −rd +ant] = {ð,ŋ,θ,b,d,f,g,k,l,m,n,p,r,s,t,v,x,z,}		___#	6	6	0.852
56	32,36	i → ∅ /	X	[−syl +cons −lab −rd +ant] = {ð,ŋ,θ,d,g,k,l,n,r,s,t,x,z,}		___#	4	4	0.786
57	33,36	i → ∅ /	X	[−syl +cons +voi −rd +ant] = {ð,ŋ,b,d,g,l,m,n,r,v,z,}		___#	4	4	0.786
58	34,36	i → ∅ /	X	[−syl +cons −lab −rd] = {t͡ʃ,ð,d͡ʒ,ŋ,ʃ,θ,Z,d,g,k,l,n,r,s,t,x,z,}		___#	6	9	0.537
59	35,36	i → ∅ /	X	[−syl +cons −rd] = {t͡ʃ,ð,d͡ʒ,ŋ,ʃ,θ,Z,b,d,f,g,k,l,m,n,p,r,s,t,v,x,z,}		___#	8	11	0.612
60	*	t͡ʃi → k /			ʔa	___#	1	1	*undef*
61	*	t͡ʃi → k /			mu	___#	1	1	*undef*
62	60,61	t͡ʃi → k /	X	[+syl −cons +son +cont −nas +voi −cor +ant +distr +dors −front +tense] = {a,o,u}		___#	3	5	0.420
63	*	t͡ʃi → k /			lo	___#	1	1	*undef*

no.	parents	change	residue	shared features	shared segments	change location	hits	scope	confidence
64	60,63	t͡ʃi → k /	X	[+syl, −cons, +son, +cont, −nas, +voi, −cor, +ant, +distr, +dors, −hi, −front, +tense] = {a,o}		___#	2	3	0.396
65	61,63	t͡ʃi → k /	X	[+syl, −cons, +son, +cont, −nas, +voi, +lab, +rd, −cor, +ant, +distr, +dors, −low, −front, +back, +tense] = {o,u}		___#	2	3	0.396

b. Phonological rules: *none*

A.3.2 Using the grammar to derive absolutives for each word

In this direction, the grammar can potentially produce multiple outputs, because there are some rules that simply remove the final *i* of the ergative, and other rules simultaneously remove the *i* and change the *k* to a *t͡ʃ*. When the grammar is used to derive absolutives for each word in the ergative form, two patterns emerge:

- Words ending in sequences other than -*t͡ʃi* have only one possible absolutive, which is the ergative minus the final -*i*—e.g., *dapi* → *dap*, *loti* → *lot*, and so on. (There are 6 such words in this hypothetical language.) The best way to derive these words is with a rule removing -*i* after anything other than an affricate (rule 55), with a confidence of .852.
- Words ending in -*t͡ʃi* have two possible outputs: one with -*t͡ʃi* (e.g., *lot͡ʃi* → *lot͡ʃ*), derived by a rule removing -*i* after any consonant (rule 59) with a confidence of .612. The second, less preferred output is one with *k* (e.g., *loki*), using the -*t͡ʃi* → -*k* change (rule 62), which works for 3 out of 5 of the relevant words in the vocabulary, and has a confidence of .420.

A.3.3 Calculation of metrics in the *ergative*→*absolutive* direction

Accuracy

The productively preferred output is always the one in which *-i* has simply been removed; this is correct for 8 out of 11 of the words in this language, but incorrect for the three which end in *k* in the absolutive. Therefore, the accuracy of the grammar in this direction is 8/11 = 73%.

Mean confidence of rules in the grammar

Averaging the confidence scores in the table above yields a mean confidence of .585 for the rules in the ergative to absolutive direction.

Mean confidence of winning outputs

As described above, six of the winning outputs are derived with a confidence of .852, while the remaining five are derived with a confidence of .612. This yields a mean confidence of .743 in the winning outputs.

Average winning margin

Six of the outputs have no competitor at all, so win by their full confidence of .852. For the five words ending in *-t͡ʃi*, the outputs with *t͡ʃ* beat the outputs with *k* by only .612 − .420 = .192. Thus, the mean winning margin in this direction is .552.

Appendix B

Metrics for base selection in Latin

Table B.1 lists all of the effectiveness measures of each of the six candidates for base status, based on the 494 most frequent Latin nouns. Rows indicate the input forms, and columns indicate the output forms; for example, the average winner confidence for the *nom.→gen.* mapping is 0.76, in the upper left.

Table B.1: Base selection metrics for Latin noun paradigms

	↓In/Out→	nom.sg.	gen.sg.	dat.sg.	acc.sg.	abl.sg.	nom.pl.	**sum**
Avg winner:	nom.sg.		0.76	0.70	0.86	0.82	0.76	**0.78**
Avg margin:			0.66	0.59	0.77	0.68	0.65	**0.67**
Percent correct:			0.82	0.77	0.87	0.85	0.80	**0.82**
Avg grammar:			0.35	0.31	0.45	0.33	0.33	**0.35**
Avg winner:	gen.sg.	0.88		0.97	0.94	0.96	0.95	**0.94**
Avg margin:		0.69		0.92	0.91	0.91	0.79	**0.84**
Percent correct:		0.93		0.97	0.96	0.96	0.94	**0.95**
Avg grammar:		0.45		0.70	0.67	0.69	0.53	**0.61**
Avg winner:	dat.sg.	0.86	0.97		0.94	0.96	0.95	**0.94**
Avg margin:		0.69	0.93		0.89	0.91	0.81	**0.85**
Percent correct:		0.92	0.98		0.96	0.97	0.96	**0.96**
Avg grammar:		0.43	0.81		0.59	0.65	0.43	**0.58**
Avg winner:	acc.sg.	0.89	0.91	0.90		0.91	0.90	**0.90**
Avg margin:		0.76	0.82	0.80		0.79	0.77	**0.79**
Percent correct:		0.95	0.94	0.92		0.93	0.92	**0.93**
Avg grammar:		0.55	0.73	0.60		0.60	0.53	**0.60**
Avg winner:	abl.sg.	0.86	0.96	0.96	0.95		0.94	**0.93**
Avg margin:		0.68	0.93	0.90	0.91		0.85	**0.85**
Percent correct:		0.92	0.97	0.97	0.97		0.95	**0.95**
Avg grammar:		0.45	0.78	0.72	0.72		0.56	**0.65**
Avg winner:	nom.pl	0.84	0.94	0.92	0.92	0.92		**0.91**
Avg margin:		0.69	0.76	0.81	0.86	0.81		**0.79**
Percent correct:		0.89	0.94	0.94	0.94	0.94		**0.93**
Avg grammar:		0.50	0.69	0.56	0.66	0.62		**0.61**

References

Albright, A. (1998). Phonological subregularities in productive and unproductive inflectional classes: Evidence from Italian. Master's thesis, University of California, Los Angeles. Available for download at http://www.linguistics.ucla.edu/people/grads/aalbrigh/papers.html.

Albright, A., A. E. Andrade, and B. Hayes (2001). Segmental environments of spanish diphthongization. In A. Albright and T. Cho (Eds.), *UCLA Working Papers in Linguistics, Number 7: Papers in Phonology 5*, pp. 117–151. Available for download at http://www.linguistics.ucla.edu/people/hayes/SegEnvSpanDiph.pdf.

Albright, A. and B. Hayes (1999a). An automated learner for phonology and morphology. Unpublished UCLA ms. Available for download at http://www.linguistics.ucla.edu/people/hayes/learning/learner.pdf.

Albright, A. and B. Hayes (1999b). Distributional encroachment and its consequences for morphological learning. UCLA ms.

Albright, A. and B. Hayes (2002). Rules vs. analogy in English past tenses: A computational/experimental study.

Albright, A. and B. Hayes (to appear). Modeling English past tense intuitions with minimal generalization. *To appear in the Association for Computational Linguistics (ACL-2002): Workshop on Morphological and Phonological Learning*.

Allen, J. H. (1903). *Allen and Greenough's New Latin Grammar for Schools and Colleges.* Boston: Ginn and Company.

Anderson, S. R. (1985). *Phonology in the Twentieth Century: Theories of Rules and Theories of Representations.* The University of Chicago Press.

Anderson, S. R. (1992). *A-Morphous Morphology.* Cambridge University Press.

Anttila, A. (2002). Morphologically conditioned phonological alternations. *Natural Language and Linguistic Theory 20*(1), 1–42. Available for download at http://roa.rutgers.edu/view.

php3?roa=425.

Aronoff, M. (1976). *Word formation in generative grammar.* Cambridge, MA: MIT Press.

Aronoff, M. (1994). *Morphology by Itself: Stems and Inflectional Classes.* MIT Press.

Baayen, R. H., T. Dijkstra, and R. Schreuder (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language 37*, 94–117.

Baldi, P. (1999). *The Foundation of Latin.* Berlin: Mouton de Gruyter.

Baroni, M. (1998). The phonetic nature of the Northern Italian allophones [s] and [z] in words with variable realization: electroglottographic and acoustic evidence. *UCLA Working Papers in Phonetics 96*, 166–174.

Barr, R. (1994). *A lexical model of morphological change.* Ph. D. thesis, Harvard.

Benua, L. (1997). *Transderivational Identity: Phonological Relations Between Words.* Ph. D. thesis, University of Massachusetts, Amherst.

Berman, R. A. (1985). The Acquisition of Hebrew. In D. I. Slobin (Ed.), *The Crosslinguistic Study of Language Acquisition, Volume 1: The Data*, pp. 255–372. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Boas, F. and E. Deloria (1941). *Dakota Grammar*, Volume 23 of *Memoirs of the National Academy of Sciences.* Washington: United States Government Printing Office.

Bochner, H. (1993). *Simplicity in Generative Morphology.* Berlin: Mouton de Gruyter.

Boersma, P. and B. Hayes (2001). Empirical Tests of the Gradual Learning Algorithm. *Linguistic Inquiry 32*(1), 45–86.

Brent, M. R., S. K. Murthy, and A. Lundberg (1995). Discovering morphemic suffixes: A case study in minimum description length induction. In *Proceedings of the Fifth International Workshop on Artificial Intelligence and Statistics, Ft. Laudersdale, FL.*

Buckley, E. (1999). Uniformity in extended paradigms. In B. Hermans and M. van Oostendorp (Eds.), *The derivational residue in phonology*, pp. 81–104. Amsterdam: Benjamins.

Buechel, E. (1970). *A Dictionary of Teton Sioux.* Pine Ridge, S.D.: Red Cloud Indian School, Inc.

Burnage, G. (1991). *CELEX—A Guide for Users.* Nijmegen: Centre for Lexical Information, University of Nijmegen.

Burzio, L. (1996). Surface constraints versus Underlying Representation. In J. Durand and B. Laks (Eds.), *Current trends in phonology: Models and methods*, pp. 97–122. CNRS, Paris, and University of Salford: University of Salford Publications.

Butt, J. (1997). *Spanish Verbs.* Oxford University Press.

Butterworth, B. (1983). Lexical representation. In B. Butterworth (Ed.), *Development, writing and other language processes.* London: Academic Press.

Bybee, J. (1985). *Morphology: A study of the relation between meaning and form.* Amsterdam: John Benjamins Publishing Company.

Bybee, J. L. and M. A. Brewer (1980). Explanation in morphophonemics: Changes in Proven cal and Spanish preterite forms. *Lingua 52*, 201–242.

Chen, M. (1970). Vowel length variation as a function of the voicing of the consonant environment. *Phonetica 22*, 129–159.

Chomsky, N. and M. Halle (1968). *The sound pattern of English.* New York: Harper and Row.

Chung, S. (1978). *Case marking and grammatical relations in Polynesian.* Austin: University of Texas Press.

Clahsen, H. (1999). Lexical entries and rules of language: A multidisciplinary study of German inflection. *Behavioral and Brain Sciences 22*, 991–1060.

Clark, A. and A. Karmiloff-Smith (1993). The cognizer's innards: A psychological and philosophical perspective on the development of thought. *Mind and Language 8*, 487–519.

Cole, J. and J. Hualde (1998). The object of lexical acquisition: A UR-free model. In *Proceedings of CLS 34.*

Czaykowska-Higgins, E. (1988). *Investigations into Polish morphology and phonology.* Ph. D. thesis, MIT.

Dammers, U., W. Hoffmann, and H.-J. Solms (1988). *Grammatik des Frühneuhochdeutschen, Vol. 4: Flexion der starken und schwachen Verben.* Heidelberg: Carl Winter Universitätsverlag.

Daugherty, K. and M. Seidenberg (1994). Beyond Rules and Exceptions: A Connectionist Approach to Inflectional Morphology. In S. D. Lima, R. L. Corrigan, and G. K. Iverson (Eds.), *The Reality of Linguistic Rules.* Amsterdam: J. Benjamins.

Delatte, L., E. Evrard, S. Govaerts, and J. Denooz (1981). *Dictionnaire fréquentiel et Index inverse de la langue latine.* Université de Liège: Laboratoire d'analyse statistique des langues anciennes.

Dinnsen, D. A. and J. Charles-Luce (1984). Phonological neutralization, phonetic implementation and individual differences. *Journal of Phonetics 12*(1), 49–60.

Dresher, B. E. (2000). Analogical levelling of vowel length in West Germanic. In A. Lahiri (Ed.), *Analogy, Levelling, Markedness: Principles of Change in Phonology and Morphology.* Mouton de Gruyter.

Eddington, D. and P. Lestrade (2002). Are plurals derived or stored? Unpublished University of New Mexico and Mississippi State University manuscript.

Ellison, M. (1994). Constraints, exceptions and representations. In *Proceedings of the First Meeting of the ACL Special Interest Group in Computational Phonology*, pp. 25–32. Available for download at citeseer.nj.nec.com/47799.html.

Evans, R. and G. Gazdar (1996). DATR: A language for lexical knowledge representation. *Computational Linguistics 22*(2), 167–216.

Ford, A. and R. Singh (1991). Propedeutique morphologique. *Folia Linguistica 25*(3-4), 549575.

Friederici, A. D. and J. E. Wessels (1993). Phonotactic knowledge of word boundaries and its use in infant speech perception. *Perception and Psychophysics 54*, 287–295.

Geyer, A. (1912). *Die starke Konjugation bei Johann Fischart.* Hofbuchdruckerei C. A. Kaemmerer & Co.

Goldsmith, J. (2001). Unsupervised learning of the morphology of a natural language. *Computational Linguistics 27*, 153–198.

Gordon, P. and M. Alegre (1999). Is there a dual system for regular inflections? *Brain and Language 68*, 212–217.

Gorecka, A. (1988). Epenthesis and the coda constraint in Polish. *MIT ms.*.

Green, A. D. (2001). The tense-lax distinction in English vowels and the role of parochial and analogical constraints. *Linguistics in Potsdam 15.* Available for download from the Rutgers Optimality Archive, http://roa.rutgers.edu/view.php3?roa=464.

Greenberg, J. (1966). *Language universals.* The Hague: Mouton.

Grimes, B. F. (Ed.) (2000). *Ethnologue, 14th Edition.* SIL International.

Gussmann, E. (1980). *Studies in Abstract Phonology.* Cambridge, MA: MIT Press.

Gvozdanović, J. (1985). Theories of sound change fail if they try to predict too much. In J. Fisiak (Ed.), *Papers from the 6th International Conference on Historical Linguistics.* John Benjamins.

Hale, K. (1973). Deep-surface canonical disparities in relation to analysis and change: An Australian example. *Current Trends in Linguistics 11*, 401–458.

Hale, M., M. Kissock, and C. Reiss (1998). Output-output correspondence in Optimality Theory. In E. Curtis, J. Lyle, and G. Webster (Eds.), *WCCFL 16*, Stanford. CSLI.

Halle and Marantz (1994). Some Key Features of Distributed Morphology. In A. Carnie and H. Harley (Eds.), *MIT Working Papers in Linguistics 21: Papers and phonology and morphol-*

ogy, Volume 21, pp. 275–288.

Hare, M. and J. L. Elman (1995). Learning and morphological change. *Cognition 56*, 61–98.

Harris, J. (1969). *Spanish Phonology.* Cambridge, MA: MIT Press.

Harris, Z. S. (1951). *Methods in Structural Linguistics.* The University of Chicago Press.

Harrison, K. D. and A. Kaun (2000). Pattern-responsive lexicon optimization. In M. Hirotani, A. Coetzee, N. Hall, and J. Kim (Eds.), *Proceedings of the Northeast Linguistics Society (NELS) 30.* Amherst, MA: University of Massachusetts, Amherst, Graduate Linguistic Student Association.

Hayes, B. (1995). On what to teach the undergraduates: Some changing orthodoxies in phonological theory. *Linguistics in the Morning Calm 3*, 59–77.

Hayes, B. (1998). On the richness of paradigms, and the insufficiency of underlying representations in accounting for them. Talk Handout, Stanford, April 2 [http://www.linguistics.ucla.edu/people/hayes/talks/RichnessOfParadigms%.pdf].

Hayes, B. (1999). Phonological Restructuring inYidiɲand its Theoretical Consequences. In B. Hermans and M. van Oostendorp (Eds.), *The derivational residue in phonology*, pp. 175–205. Amsterdam: Benjamins.

Hayes, B. (to appear). Phonological Acquisition in Optimality Theory: The early stages. In R. Kager, J. Pater, and W. Zonneveld (Eds.), *Fixing Priorities: Constraints in Phonological Acquisition.* Cambridge University Press.

Hermans, B. (2001). Yer vocalization in suffixed forms; the case of russian. In M. van Oostendorp and E. Anagnostopoulou (Eds.), *Progress in Grammar: Articles at the 20th Anniversary of the Comparison of Grammatical Models Group in Tilburg.* Published online at: http://www.roquade.nl/meertens/progressingrammar/toc.html.

Hock, H. H. (1991). *Principles of Historical Linguistics* (2nd ed.). Mouton de Gruyter.

Hockett, C. (1958). Two models of grammatical description. In M. Joos (Ed.), *Readings in linguistics*, pp. 386–399. Chicago: University of Chicago Press.

Hockett, C. F. (1955). *A Manual of Phonology.* Baltimore: Waverly Press.

Hoenigswald, H. (1960). *Language Change and Linguistic Reconstruction.* University of Chicago Press.

Hohepa, P. W. (1967). *A Profile Generative Grammar of Maori.* Indiana University publications in anthropology and linguistics, Memoir 20. Baltimore: Waverly Press.

Hooper, J. (1976). *An Introduction to Natural Generative Phonology.* New York: Academic Press.

Hutchinson, A. (1994). *Algorithmic Learning.* Oxford: Clarendon Press.

Hyman, L. (1975). *Phonology: Theory and Analysis.* Holt, Rinehart and Winston.

Inkelas, S. (1994). The consequences of optimization for underspecification. *NELS 25*, 287–302. Available for download at: http://roa.rutgers.edu/view.php4?roa=40.

Inkelas, S. (1996). Archiphonemic Underspecification: an optimization approach to the phonological description of morphemes. *Berkeley ms.*.

Inkelas, S., O. Orgun, and C. Zoll (1997). The implications of lexical exceptions for the nature of grammar. In I. Roca (Ed.), *Derivations and constraints in phonology*, pp. 393–418. Oxford: Clarendon.

Itô, J. and A. Mester (1995). The core-periphery structure in the lexicon and constraints on reranking. In D. Beckman and Urbanczyk (Eds.), *Papers in Optimality Theory*, Volume 18 of *University of Massachusetts Occasional Papers.* Amherst, Mass: GLSA.

Itô, J. and A. Mester (2002). The phonological lexicon. In N. Tsujimura (Ed.), *A Handbook of Japanese Linguistics.* Oxford: Blackwell.

Jakobson, R. (1939). Signe zéro. In *Mélanges de linguistique offerts à Charles Bally*, pp. 143–152. [Reprinted in Jakobson's Selected writings II, 211-219. The Hague: Mouton.].

Joesten, M. (1931). *Untersuchungen zu ahd. (as.)* ë, i, *vor* u *der Folgesilbe und zur 1. Pers. Sg. Präs. Ind. der starken* e-*Verben (Kl. IIIb, IV, V).* Number 28 in Gießener Beiträge zur deutschen Philologie. Gießen: Wilhem Schmitz Verlag. 1968 reprint, Swets & Zeitlinger N.V., Amsterdam.

Juilland, A. and E. Chang-Rodríguez (1964). *Frequency dictionary of Spanish words.* The Hague: Mouton.

Jusczyk, P. W., A. Friderici, J. M. Wessels, V. Y. Svenkerud, and A. M. Jusczyk (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of Memory and Language 32*, 402–420.

Jusczyk, P. W., P. A. Luce, and J. Charles-Luce (1994). Infants sensitivity to phonotactic patterns in the native language. *Journal of Memory and Language 33*, 630–645.

Kager, R. (1999). *Optimality Theory.* Cambridge University Press.

Kager, R. (2000). Stem stress and peak correspondence in Dutch. In J. Dekkers, F. van der Leeuw, and J. van de Weijer (Eds.), *Optimality Theory: Phonology, Syntax, and Acquisition.* Oxford:

Oxford University Press.

Katz, D. (1987). *Grammar of the Yiddish Language*. London: Duckworth.

Kazakov, D. and S. Manandhar (2001). Unsupervised learning of word segmentation rules with genetic algorithms and inductive logic programming. *Machine Learning 43*, 121–162. Available for download at http://citeseer.nj.nec.com/kazakov01unsupervised.html.

Kenstowicz, M. (1997a). Base-identity and uniform exponence: Alternatives to cyclicity. In J. Durand and B. Laks (Eds.), *Current Trends in Phonology: Models and Methods*, pp. 363–394. CNRS and University of Salford Publications. Available for download from the Rutgers Optimality Archive, http://roa.rutgers.edu.

Kenstowicz, M. (1997b). Uniform exponence: Extension and exemplification. In V. Miglio and B. Morén (Eds.), *University of Maryland Working papers in Linguistics 5: Selected papers from the Hopkins Optimality Workshop 1997*, pp. 139–54. Available for download on the Rutgers Optimality Archive: http://roa.rutgers.edu.

Kenstowicz, M. (1998). Uniform exponence: Exemplification and extension. MIT ms., available on the Rutgers Optimality Archive: http://ruccs.rutgers.edu/roa.html.

Kenstowicz, M. (2002). Paradigmatic uniformity and contrast. In A. Csirmaz, Z. Li, A. Nevins, O. Vaysman, and M. Wagner (Eds.), *MIT Working Papers in Linguistics #42: Phonological Answers (and their Corresponding Questions)*.

Kenstowicz, M. and C. Kisseberth (1977). *Topics in Phonological Theory*. New York: Academic Press.

Kenstowicz, M. and C. Kisseberth (1979). *Generative Phonology*. New York: Academic Press.

Kern, P. O. (1903). *Das starke Verb bei Grimmelshausen: Ein Beitrag zur Grammatik des Frühneuhochdeutsche* Chicago: University of Chicago.

Kibre, N. (1998). Formal property inheritance and consonant/zero alternations in Maori verbs. Unpublished ms. Available for download on the Rutgers Optimality Archive at: http://roa.rutgers.edu/view.php3?roa=285.

Kibrik, A. E. (1998). Archi (Caucasian – Daghestanian). In A. Spencer and A. Zwicky (Eds.), *The Handbook of Morphology*. Blackwell.

Kim, J. (2002). A feature-analysis of ablaut. In A. Csirmaz, Z. Li, A. Nevins, O. Vaysman, and M. Wagner (Eds.), *MIT Working Papers in Linguistics #42*.

Kim, N.-K. (1987). Korean. In B. Comrie (Ed.), *The World's Major Languages*, pp. 881–898. Oxford: Oxford University Press.

King, R. D. (1980). The history of final devoicing in Yiddish. In M. I. Herzog, B. Kirshenblatt-Gimblett, D. Miron, and R. Wisse (Eds.), *The Field of Yiddish: Studies in Language, Folklore, and Literature, Fourth Collection*, pp. 371–430. Philadelphia: Institute for the Study of Human Issues.

Kiparsky, P. (1965). *Phonological Change*. Ph. D. thesis, MIT.

Kiparsky, P. (1968). Linguistic universals and linguistic change. In E. Bach and R. T. Harms (Eds.), *Universals in Linguistic Theory*. New York: Holt, Rinehart, and Winston.

Kiparsky, P. (1982). Lexical morphology and phonology. In T. L. S. of Korea (Ed.), *Linguistics in the Morning Calm: Selected Papers from SICOL-1981*. Seoul, Korea: Hanshin Publishing Company.

Kiparsky, P. (1998). Covert generalization. In A. Ralli (Ed.), *Proceedings of the First Mediterranean Morphology Conference*.

Kiparsky, P. (to appear). *Paradigm Effects and Opacity*. Stanford: CSLI Publications.

Klausenburger, J. (1979). *Morphologization: Studies in Latin and Romance morphophonology*. Tübingen: Max Niemeyer.

Kopkalli, H. (1993). *A phonetic and phonological analysis of final devoicing in Turkish*. Ph. D. thesis, University of Michigan.

Krämer, M. (2000). Voicing alternations and underlying representations: The case of Breton. *Lingua 110*(9), 639–663.

Kraska-Szlenk, I. (1995). *The Phonology of Stress in Polish*. Ph. D. thesis, University of Illinois at Urbana-Champaign.

Kühner, R. (1912). *Ausführliche Grammatik der Lateinischen Sprache* (2nd ed.). Hannover: Verlag Hahnsche Buchhandlung.

Kuryłowicz, J. (1947). The nature of the so-called analogical processes. *Diachronica 12*(1), 113–145. (trans. Margaret Winters, 1995, Diachronica 12:1.113-145).

Lahiri, A. and B. E. Dresher (1984). Diachronic and synchronic implications of declension shifts. *The Linguistic Review 3*, 141–163.

Lass, R. (1980). Paradigm coherence and the conditioning of sound change. In J. Fisiak (Ed.), *Historical morphology*, pp. 251–272. The Hague: Mouton.

Lee, I. (1999). *A principles-and-parameters approach to the acquisition of (the morphosyntax of) IP in Korean*. Ph. D. thesis, University of Essex.

Lee, J. (in progress). An analysis of Korean stem final consonants. Unpublished UCLA manuscript.

Leumann, M. (1977). *Lateinische Laut- und Formenlehre*. Munich: C. H. Beck'sche Verlagsbuchhandlung (Oscar Beck).

L'Huiller, M. (1999). *Advanced French Grammar.* Cambridge University Press.

Lieber, R. (1987). *An Integrated Theory of Autosegmental Processes.* Albany, NY: State University of New York Press.

Lieber, R. (1992). *Deconstructing Morphology: Word Formation in Syntactic Theory.* University of Chicago Press.

Lightner, T. (1972). *Problems in the Theory of Phonology, Vol. 1: Russian Phonology and Turkish Phonology.* Edmonton: Linguistic Research.

Ling, C. and M. Marinov (1993). Answering the connectionist challenge: a symbolic model of learning the past tenses of English verbs. *Cognition 49*, 235–290.

MacWhinney, B. and J. Leinbach (1991). Implementations are not conceptualizations: Revising the verb learning model. *Cognition 40*, 121–157.

Mańczak, W. (1958). Tendences générales des changements analogiques. *Lingua 7*, 298–325 and 387–420.

Mańczak, W. (1980). Laws of analogy. In J. Fisiak (Ed.), *Historical morphology*, pp. 283–288. The Hague: Mouton.

Maratsos, M. (2000). More overregularizations after all: new data and discussion on Marcus, Pinker, Ullman, Hollander, Rosen & Xu. *Journal of Child Language 27*, 183–212.

Marchand, H. (1969). *Categories and types of present-day English word-formation.* Munich: Beck.

Marcus, G., S. Pinker, M. Ullman, M. Hollander, T. Rosen, and F. Xu (1992). *Overregularization in language acquisition.* Monographs of the Society for Research in Child Development.

Martin, S. E. (1992). *A Reference Grammar of Korean.* Tokyo, Japan: Charles E. Tuttle Company.

McCarthy, J. (1998). Morpheme structure constraints and paradigm occultation. In M. C. Gruber, D. Higgins, K. Olson, and T. Wysocki (Eds.), *CLS 32, vol. II: The Panels*, Chicago. Chicago Linguistic Society.

McCarthy, J. and A. Prince (1995). Faithfulness and reduplicative identity. In D. Beckman and Urbanczyk (Eds.), *Papers in Optimality Theory*, Volume 18 of *University of Massachusetts Occasional Papers*. Amherst, Mass: GLSA.

McCawley, J. (1967). Sapir's phonologic representation. *International Journal of American Linguistics 33*(2), 106–111.

Meiser, G. (1998). *Historische Laut- und Formenlehre der lateinischen Sprache.* Darmstadt:

Wissenschaftliche Buchgesellschaft.

Mieses, M. (1924). *Die jiddische Sprache: Eine historische Grammatik des Idioms der integralen Juden Ost- und Mitteleuropas.* Berlin/Vienna: Benjamin Harz.

Mikheev, A. (1997). Automatic rule induction for unknown-word guessing. *Computational Linguistics 23*(3), 405–423.

Mooney, R. J. and M. E. Califf (1996). Learning the Past Tense of English Verbs Using Inductive Logic Programming. In *Symbolic, Connectionist, and Statistical Approaches to Learning for Natural Language Processing.* Springer Verlag.

Munro, P. (1989). Lakhota verb list. Unpublished UCLA ms.

Nelson, N. (to appear). Deriving the "prefixing" preference in reduplication. In *WCCFL 21: Proceedings of the 21st West Coast Conference on Formal Linguistics.* Somerville, MA: Cascadilla Press.

Neue, C. F. and C. Wagener (1902). *Formenlehre der lateinischen Sprache,* (3rd ed.), Volume 1 (Das Substantivum). Leipzig: O.R. Reisland.

Neuvel, S. (to appear). Whole word morphologizer: Expanding the word-based lexicon: A nonstochastic computational approach. *Brain and Language*.

Neuvel, S. and S. Fulop (to appear). Unsupervised learning of morphology without morphemes. *To appear in the Association for Computational Linguistics (ACL-2002): Workshop on Morphological and Phonological Learning*.

Neuvel, S. and R. Singh (2001). Vive la différence! What morphology is about. *Folia Linguistica 35*(3-4).

Noyer, R. (1998). The basis of bases. Talk handout, April 24, University of Maryland, College Park.

Odden, D. A. (1981). *Problems in Tone Assignment in Shona.* Ph. D. thesis, University of Illinois, Urbana-Champagne.

Paul, H. (1920). *Prinzipien der Sprachgeschichte* (5th ed.). Halle: Niemeyer.

Paul, H., P. Wiehl, and S. Grosse (1989). *Mittelhochdeutsche Grammatik* (23rd ed.). Max Niemeyer Verlag Tübingen.

Philipp, G. (1980). *Einführung ins Frühneuhochdeutsche: Sprachgeschichte, Grammatik, Texte.* Heidelberg: Quelle und Meyer.

Pinker, S. (1999). *Words and Rules: The Ingredients of Language.* New York: Basic Books.

Pinker, S. and A. Prince (1988). On language and connectionism: Analysis of a Parallel Distributed Processing model of language acquisition. *Cognition 28*, 73–193.

Pinker, S. and A. Prince (1994). Regular and irregular morphology and the psychological status of rules of grammar. In S. D. Lima, R. L. Corrigan, and G. K. Iverson (Eds.), *The Reality of Linguistic Rules*, pp. 321–351. Amsterdam: J. Benjamins.

Port, R. and M. O'Dell (1986). Neutralization of syllable-final voicing in German. *Journal of Phonetics 13*, 455–471.

Port, R. F. and P. Crawford (1989). Incomplete neutralization and pragmatics in german. *Journal of Phonetics 17*, 257–282.

Prasada, S. and S. Pinker (1993). Generalization of regular and irregular morphological patterns. *Language and Cognitive Processes 8*, 1–56.

Prince, A. and P. Smolensky (1993). Optimality Theory: Constraint Interaction in Generative Grammar (Technical Report CU-CS-533-91). Technical report, Department of Computer Science, University of Colorado, Boulder.

Raffelsiefen, R. (2000). Constraints on schwa apocope in Middle High German. In A. Lahiri (Ed.), *Analogy, Levelling, Markedness: Principles of Change in Phonology and Morphology*. Mouton de Gruyter.

Richards, N. (2001). Leerdil Yuujmen bana Yanangarr (Old and New Lardil). In R. Pensalfini and N. Richards (Eds.), *MIT Working Papers on Endangered and Less Familiar Languages #2: Papers on Australian Languages*. Available for download at http://web.mit.edu/norvin/www/papers/Leerdil.pdf.

Rissanen, J. (1986). Stochastic complexity and modeling. *The Annals of Stochastics 14*(3), 1080–1100.

Rissanen, J. (1989). *Stochastic Complexity in Statistical Inquiry*. New Jersey: World Scientific.

Rissanen, J. and E. S. Ristad (1994). Language acquisition in the MDL framework. In E. S. Ristad (Ed.), *DIMACS Series in Discrete Mathetmatics and Theoretical Computer Science, Vol. 17: Language Computation*. American Mathematical Society.

Rockowitz, A. (1979). *201 Yiddish Verbs, fully conjugated in all the tenses*. Woodbury, NY: Barron's Education Series, Inc.

Rodríguez-Bou, I. (1952). *Recuento de vocabulario español*. Río Peidras: Universidad de Puerto Rico.

Rood, D. S. (1983). A preliminary consideration of Proto-Siouan ablaut. *Na'pao: A Saskatchewan Anthropology Journal (Special Issue: Proceedings of the Second Siouan Languages Conference, 1982) 13*.

Rumelhart, D. and J. McClelland (1987). Learning the past tenses of English verbs. In B. MacWhinney (Ed.), *Mechanisms of language acquisition.* Hillsdale, NJ: Erlbaum.

Sadock, J. M. (1973). Word-final devoicing in the development of Yiddish. In B. B. Kachru, R. B. Lees, Y. Malkiel, A. Pietrangeli, and S. Saporta (Eds.), *Issues in Linguistics: Papers in Honor of Henry and Renée Kahane*, pp. 790–797. Urbana, IL: University of Illinois Press.

Sanders, G. (1990). On the analysis and implications of Maori verb alternations. *Lingua 80*, 150–198.

Sapir, E. (1915). Notes on Judeo-German Phonology. *The Jewish Quarterly Review 6*, 231–266.

Saravanan, B. (2000). Predicting irregularity in Tamil verbs. In *WCCFL 19: Proceedings of the 19th West Coast Conference on Formal Linguistics*, pp. 453–465. Somerville, MA: Cascadilla Press.

Schirmunski, V. (1962). *Deutsche Mundartkunde.* Berlin: Akademie-Verlag.

Sebastián, N., F. Cuetos, and M. Carreiras (to appear). Lexesp: Creación de una base de datos informatizada del español. Technical report, DGICYT: APC93-0122.

Sereno, J. A. and A. Jongman (1997). Processing of English inflectional morphology. *Memory and Cognition 25*, 425–437.

Shaw, P. A. (1980). *Theoretical Issues in Dakota Phonology and Morphology.* New York: Garland Publishing, Inc.

Skutsch, O. (1985). *The Annals of Q. Ennius.* Oxford: Clarendon Press.

Smolensky, P. (1996). The initial state and 'richness of the base' in Optimality Theory. Technical report, Johns Hopkins University.

Sonderegger, S. (1987). *Althochdeutsche Sprache und Literatur: eine Einführung in das älteste Deutsch; Darstellung und Grammatik (2nd ed.).* Berlin: Walter de Gruyter.

Stemberger, J. P. and B. MacWhinney (1986). Frequency and the lexical storage of regularly inflected forms. *Memory & Cognition 14*(1), 17–26.

Steriade, D. (1994). Lexical conservatism. In *Linguistics in the Morning Calm, Selected Papers from SICOL 1997, Linguistic Society of Korea.* Hanshin Publishing House.

Steriade, D. (1995). Underspecification and markedness. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory*, pp. 114–174. Blackwell.

Steriade, D. (2000). Paradigm Uniformity and the Phonetics/Phonology boundary. In M. Broe and J. Pierrehumbert (Eds.), *Papers in Laboratory Phonology V: Acquisition and the Lexicon.* Cambridge University Press.

Tiersma, P. (1982). Local and general markedness. *Language 58*(4), 832–849.

Tihany, L. (1996). Number of Hungarian word forms. http://www.lpl.univ-aix.fr/projects/multext-east/

MTE2.number.html.

Tranel, B. (1996). Exceptionality in Optimality Theory and final consonants in French. In K. Zagona (Ed.), *Grammatical Theory and Romance Languages*. Benjamins: Benjamins. Available for download from the Rutgers Optimality Archive, http://roa.rutgers.edu/view.php3?roa=61.

Weinreich, U. (1990). *Modern English-Yiddish, Yiddish-English Dictionary*. New York: Yivo Institute for Jewish Research.

Westermann, G. (1997). A constructivist neural network learns the past tense of English verbs. In *Proceedings of the GALA '97 Conference on Language Acquisition*, Edinburgh. HCRC.

Wetzels, L. (1984). Paradigm leveing in Latin and Old French: A critical view of rule inversion. In P. Baldi (Ed.), *Papers from the XXth Linguistic Symposium on Romance Languages*, Volume 26(4) of *Amsterdam Studies in the Theory and History of Linguistic Science*, Amsterdam. John Benjamins Publishing Company.

Whitney, W. D. (1924). *Sanskrit Grammar* (5th ed.). Delhi: Motill Banarsidass Publishers.

Yearley, J. (1995). Jer vowels in Russian. In Beckman, Dickey, and Urbanczyk (Eds.), *Papers in Optimality Theory*, Volume 18 of *University of Massachusetts Occasional Papers*. Amherst, Mass: GLSA.

Zaliznjak, A. A. (1977). *Grammaticheskij Slovarj Russkogo Jazyka*. Izdatel'stvo Russkij Jazyk.

Zehetner, L. (1989). *Bairisches Deutsch*. Hugendubel.

Zonneveld, W. (1978). *A Formal Theory of Exceptions in Generative Phonology*. Lisse: The Peter de Ridder Press.

Zuraw, K. (2000). *Patterned Exceptions in Phonology*. Ph. D. thesis, UCLA.