UNIVERSITY OF CALIFORNIA

Los Angeles

An Effort-Based Approach to Consonant Lenition

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

by

Robert Martin Kirchner

1998

The dissertation of Robert Martin Kirchner is approved.

Bruce Hayes

Patricia Keating

Donka Minkova

Donca Steriade, Committee Chair

University of California, Los Angeles

1998

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Though few who work with me may have discerned it, beneath my mild-mannered, taciturn exterior lie innards of quivering gelatinous mush; and here's where I put them on public display. Brace yourselves.

First, to Donca Steriade: In making it possible for me to come to UCLA, in giving me all manner of vital professional and personal advice and support, in addition to your role as my thesis advisor, my debt to you is profound, as is my admiration for you. Though we agree on many theoretical issues, I've had more fun arguing with you about the areas of disagreement than I could derive from successfully persuading nearly anyone else. I hope our arguments continue.

To Bruce Hayes: For the same range of advice and support (not to mention help with the programming of the mass-spring model), and the same profundity of debt and admiration, ditto. In particular, I want to express my appreciation for your uncanny ability to take in my often half-baked ideas, figure out what I mean to say, and offer insightful comments on the rephrased ideas. This is a rare skill, which I hope I can develop in my own teaching.

To Pat Keating and Donka Minkova, thanks for helpful feedback on this dissertation. Thanks in particular to Pat for recognizing that I'm not an experimentalist (but who knows what the future may bring?). Thanks also to Ian Maddieson, who helped me as much as my official committee members, offering many suggestions and corrections that (I hope) have saved me from appearing a complete ignoramus to any phonetician who reads this work.

My fellow students at UCLA all helped create an atmosphere of creativity and discovery, for which I am grateful. But in particular, I wish to thank Edward Flemming, one of the most brilliant people I've ever met, without whose pioneering work this dissertation would be inconceivable; Richard Wright, whose brain I have shamelessly and repeatedly picked (I hope I can somehow eventually repay the debt); Rod Casali, for helping me pretend to know something about physics; and Peggy (sorry, *Margaret*) MacEachern, for general camaraderie, thought-provoking discussions on a variety of linguistic and philosophical issues, and drawings of horses for Miriam.

Thanks to Amy Weinberg, my first phonology teacher(!), who helped rescue me from a legal career (*shudder*), and to Marie Huffman, who got me excited about phonology and very wisely told me to go to UCLA.

To Donna Albino: thanks for helping me stay centered as a human being.

To my wife Suzanne: I started graduate school as an anti-sexist egalitarian idealist, and as this dissertation took over my life, I gradually degenerated into a typical nerd husband who spent all day hunched over the computer while you cooked, cleaned, and took care of the girls, as well as taught. Thank you. And now it's your turn.

To my daughters, Miriam and Naomi: OK, I'm ready to play with you now.

Finally, acknowledgment is given to Cambridge University Press for granting permission to reproduce, in revised form, Kirchner 1997, which appears as Chapter 3 of this dissertation.

VITA

February 4, 1961	Born, Bayshore, New York
1983	B.A., Political Science, Johns Hopkins University, Baltimore, Maryland
1986	J.D., University of Maryland School of Law, Baltimore, Maryland
1992	M.A., Linguistics, University of Maryland, College Park, Maryland
1989-1992	Teaching Assistant Department of Linguistics University of Maryland
1993-1996	Teaching Assistant Department of Linguistics University of California, Los Angeles
1996-1997	Visiting Lecturer Department of Linguistics University of Illinois, Urbana-Champaign
1998	Visiting Lecturer Department of Linguistics University of California, Irvine

PUBLICATIONS AND PRESENTATIONS

"Contrastiveness and Faithfulness," *Phonology* 14:1, 83-111.

"Synchronic Chain Shifts in Optimality Theory," *Linguistic Inquiry* 27:2, 341-350 (1996).

"Some New Generalizations Concerning Geminate Inalterability," paper presented at LSA, New York, January 9, 1998.

"The Importance of Being LAZY: an OT Approach to Lenition," poster presented at Hopkins Optimality Theory Workshop, Baltimore, Maryland, May 11, 1997.

"An Effort-Based Approach to Lenition Generalizations," presented at Mid-Continental Workshop on Phonology, Urbana, Illinois, November 1996.

"A Non-Representational Theory of Contrastiveness," presented at Arizona Phonology Workshop, Tucson, April 1995.

"Contrastiveness Is an Epiphenomenon of Constraint Ranking," *Proceedings of the Berkeley Linguistics Society* 21 (1995).

"Turkish Vowel Harmony and Disharmony: an Optimality Theoretic Account," presented at Rutgers Optimality Workshop (ROW-I), New Brunswick, N.J., October 23, 1993.

"Lardil Truncation and Augmentation: a Morphological Account," presented at Linguistic Society of America conference, Philadelphia, PA, January 11, 1992.

"Yidin^y Apocope and Vowel Lengthening: Phonological Processes without Phonological Rules," *Proceedings of NELS 21*, U.Mass. GLSA (1990).

ABSTRACT OF THE DISSERTATION

An Effort-Based Approach to Consonant Lenition

by

Robert Martin Kirchner
Doctor of Philosophy in Linguistics
University of California, Los Angeles, 1998
Professor Donca Steriade, Chair

Despite the pervasiveness of lenition in the sound systems of natural language, this class of patterns has eluded adequate characterization in previous theories of phonology. Specifically, previous theories have failed to capture formally the phonetic unity of the various lenition processes (e.g. degemination, voicing, spirantization, debuccalization, deletion), or to account for the environments in which lenition typically occurs.

I present a unified approach to consonant lenition, wherein particular lenition patterns arise from Optimality Theoretic conflict between a principle of effort minimization (which I style LAZY), and faithfulness to auditory features, in combination with (perceptually-based) fortition constraints, building upon the proposals of Jun (1995) and Flemming (1995). I further demonstrate that this effort-based approach straightforwardly accounts for a number of generalizations, drawn from a survey of 272 grammars:

- Geminate stops never lenite unless they concomitantly degeminate.
- Unaffricated stops never synchronically spirantize to strident fricatives.
- All else being equal, lenition occurs more readily the greater the openness of the flanking segments (the widely attested pattern of intervocalic lenition being a special case).
- Lenition occurs more readily the faster or more casual the speech.

The approach is illustrated with case studies of lenition in Tümpisa Shoshone and Florentine Italian.

Chapter 1:

Introduction

The term "lenition" (< L. *lenis*, 'weak') refers to synchronic alternations, as well as diachronic sound changes, whereby a sound becomes "weaker," or where a "weaker" sound bears an allophonic relation to a "stronger" sound. An explicit, unified characterization of this "weakening" has been a vexed question of phonological theory (see Bauer 1988); but the core idea, as applied to consonants, is some *reduction in constriction degree or duration*. The term thus uncontroversially includes:

- *degemination*, or reduction of a long consonant to a short one (e.g. tː - t);
- *flapping*, or reduction of a stop to a flap (e.g. t - ɾ);
- *spirantization*, or reduction from a stop (or affricate) to a fricative or approximant continuant (e.g. t - {θ, θ̞});
- reduction of other consonants to approximants (e.g. r - ɹ, s- s̞);
- *debuccalization*, or reduction to a laryngeal consonant (e.g. t - ʔ, s - h);

and, at its most extreme,

- complete *elision* (e.g. t - Ø).

Voicing (e.g. t - d), although ostensibly involving an adjustment in laryngeal specification rather than reduction of constriction, is also standardly included in this typology, for at least two reasons: (a) the contexts and conditions under which voicing occurs substantially overlap with those of the other lenition patterns, and may even occur in chain shifts with them (e.g. tt > t, t > d, d > ð, as in Gallo-Romance (Bourciez & Bourciez 1967)); and (b) as discussed in Chapter 2 section 4, voicing does in fact generally conform to the gestural reduction characterization above, upon a closer examination of the articulatory implementation of voiced vs. voiceless obstruents.

Lenition processes, including casual speech reductions, are ubiquitous in the sound systems of natural language. Nevertheless, this class of patterns has eluded adequate characterization in previous theories of phonology. Indeed, in comparison with better-understood phonological phenomena such as assimilation processes, phonotactic conditions, and stress patterns, lenition has been largely ignored in the theoretical literature. Specifically, no comprehensive explanation has been offered for the following elementary questions:

- How is this class of patterns to be formally characterized in a unified manner?
- Why does consonant lenition commonly occur in intervocalic position?
- Why do geminate stops fail to lenite unless they concomitantly degeminate?
- Why does spirantization overwhelmingly yield weak fricatives or approximants, rather than strident fricatives (e.g. [s]), which otherwise appear to be the unmarked fricatives?
- Why does lenition occur more readily in faster rates or lower registers of speech?
- What is the relation between such rate- and register-sensitive reduction and more stable, categorical lenition patterns?
- What is the relation between lenition and other classes of phonological patterns, such as assimilation?

In seeking answers to these and related questions, I have been led to certain conclusions concerning the nature of phonological representations and the constraint system. Specifically, I will argue that phonologically conditioned[1] lenition patterns are

[1]As opposed to morphologically conditioned consonant mutations or gradations, which are sometimes referred to as "lenitions" (particularly among Celticists) because the structural changes are often similar to lenition, and presumably have their historical origin in phonologically conditioned lenition processes. Being morphologically conditioned, such mutation processes do not involve considerations of effort

driven by a phonetic imperative to *minimize articulatory effort:* lenition is characterized as substitution of a less effortful set of gestures. I further argue that this effort minimization imperative can and should be incorporated directly into the phonological formalism, as a scalar, violable constraint, which I call LAZY. Phonological representations must therefore include considerably more detail concerning articulatory gestures, and their resulting effort costs, than has standardly been conceived. This demonstration of lenition's basis in effort minimization thus points in a similar direction to recent research, e.g. Steriade 1993, 1995, 1996; Jun 1995; Flemming 1995; Boersma 1997a,b,c,d, which seeks to capture language-specific sound patterns in terms of conflicts among phonetic constraints, expressing considerations of ease of articulation and ease of perception, the conflicts being resolved through Optimality Theoretic (Prince and Smolensky 1993) constraint ranking.

In the remainder of this chapter, I discuss the grammatical survey upon which my typological generalizations are based, consider previous approaches to lenition, outline my proposed approach, and provide an overview of the body of the dissertation.

1. OVERVIEW OF THE SURVEY

1.1. SOURCES

The generalizations discussed in this dissertation are supported by a composite of two surveys: (a) first, a partial search of the UCLA University Research Library for

---

minimization which drive the phonologically conditioned spirantization patterns. Rather, such mutations must be attributed to constraints of the form “*[-cont] / $\_\_\__{\text{mutation environment}}$,” where the environment is morphosyntactically defined. For a non-serial OT treatment of chain shifts, which characterize many mutation alternations (e.g. the Irish "lenition" series: (roughly) stop - fricative - Ø, see section 2), see Kirchner 1996.

grammars which contained some lenition phenomenon, supplemented by a number of grammars passed on to me by advisors and colleagues, yielding 107 languages; and further supplemented by (b) the more extensive lenition survey of Lavoie 1996, covering 165 additional languages, from an exhaustive search of the library of the University of California at Irvine, yielding a total of 272 languages for the composite survey. These are listed, with references, and a brief description of the pattern, in the Appendix.

1.2. Principal Generalizations

From the composite survey, I have extracted four principal generalizations, which form the empirical centerpiece of this dissertation.

I. Unaffricated stops never synchronically lenite to strident fricatives, such as [s] or [f]. See Chapter 4.

II. Geminates stops never lenite unless they concomitantly degeminate.[2] Partial geminates (i.e. homorganic nasal + stop or lateral + stop clusters) also conform to this generalization; however, partial geminates readily undergo voicing lenition. See Chapter 5.

III. The common restriction of lenition to intervocalic position is actually a special case of a broader generalization: ceteris paribus, if a consonant lenites when flanked by relatively highly constricted segments, it also lenites when flanked by more open segments. See Chapter 6.

IV. If a consonant lenites in some context at a given rate or register of speech, it also lenites in that context at all faster rates or more casual registers of speech. See Chapter 6.

---

[2]The relation of this generalization to the phenomenon of "geminate inalterability" is discussed in Chapter 4.

I defer documentation of these generalizations until the appropriate chapter. In the remainder of this section, I discuss a few empirical observations which are not taken up in depth elsewhere in the dissertation.

1.3. IS LENITION A UNIFIED PHENOMENON?

The traditional consensus that lenition *is* a unified phenomenon is aptly expressed in Theo Vennemann's dictum (cited in Hyman 1975: 165), "A segment X is said to be weaker than a segment Y if Y goes through an X stage on its way to zero": i.e. lenition is a progression along a unified dimension of "weakness." Indeed, to my knowledge, no linguist has ever explicitly maintained the contrary view, that "lenition" is merely an arbitrary collection of unrelated processes. Nevertheless, as a threshold matter, we must ask why, aside from tradition, the unified view of lenition should be maintained. First, as noted above, lenition processes have a unified phonetic characterization, in that they involve reduction of the magnitude or duration of articulatory gestures (primarily consonantal constriction gestures).[3] Moreover, a unified view is supported by the lenition pattern of Florentine Italian (see Chapter 8): in (roughly) intervocalic position, voiceless stops display increasing lenition, first to fricatives, then to approximants, then to voiced approximants or [h], and finally to Ø, the faster the speech rate or the lower the register. In light of the scalar nature of this lenition, and its sensitivity to the same conditions and contexts (fast/casual speech, flanking vowels) at every step along the scale, these alternations manifestly constitute a unified phenomenon. Patterns of increasing reduction in fast/casual speech are likewise reported for German by Kohler (1991). Further motivation for a unified approach is found in patterns such as Malayalam, in which

[3]In Chapter 4, section 6.2, however, I argue that lenition processes cannot be characterized *purely* as operations of gestural reduction; for such reduction is commonly accompanied by other gestural modifications, or even gestural insertion, typically for purposes of perceptual enhancement of the lenited output.

retroflex stops undergo flapping, while stops at the other places of articulation spirantize to voiced approximants, all in the context /[+son,-nas]__V (Mohanan 1986). Similarly, in Yindjibarndi (Wordick, 1982), in intervocalic position, /k/ deletes, while stops at other places of articulation spirantize to glides. Without a unified approach, we must treat each of these structural changes (spirantization, flapping, deletion) as unrelated phenomena, missing the generalization that these are all manifestations of gestural reduction, occurring in the same context, across all places of articulation, in each of these languages. Nor could we account for the striking fact that these different patterns of reduction happen to be conditioned by substantially the same context in Florentine, Malayalam, and Yindjibarndi (i.e. roughly intervocalic position; or more precisely, a two-sided context requiring some threshold of openness of the flanking segments).

1.4. No Place-of-Articulation Asymmetries

The survey appears to falsify the hypothesis that there exist broad implicational universals concerning lenition at particular places of articulation, e.g. Foley's (1977) claim that lenition of labials in a given language implies lenition of coronals, and that lenition of coronals implies lenition of velars. Rather, we find cases of lenition at any and all places of articulation, as exemplified in the following table (note that, in all these cases, corresponding consonants at other places of articulation occur in the relevant contexts):

Table 1-1. Places of articulation targeted by lenition processes:

Language	Reference	Description of process	What lenites?		
			labials	*coronals*	*dorsals*
Lomongo	Hulstaert 1961	b -w or Ø /V+___V	yes	no	no
Tamazight Berber	Abdel-Massih 1971	b ~ β in free variation	yes	no	no
Badimaya	Dunn 1988	d,d^j - ð,ʒ/V__V	no	yes	no
Purki	Rangan 1979	d̪,d - ð,ɾ/V__V	no	yes	no
Saek	Gedney 1993	g ~ ɣ in free variation	no	no	yes
Tigrinya	Kenstowicz 1982	k,q - x,χ/__V	no	no	yes
Kupia	Christmas & Christmas 1975	p - ɸ/V__Vʈ,ɖ - ɾ/V__V	yes	yes	no
Dahalo	Tosco 1991	b,d - β,ð/V__V	yes	yes	no
Georgian	Aronson 1989	q’ ~ χ in free variation, v - w/__#	yes	no	yes
Uzbek	Sjoberg 1963	p,b - ɸ,w medially, q - x non-initially	yes	no	yes
West Tarangan	Nivens 1992	g,dʒ - w,ʝ in medial unstressed position	no	yes	yes
Spanish	Harris 1969	b,d,g - β,ð,ɣ non-initially, except after a homorganic nasal or lateral	yes	yes	yes
Gujarati	Cardona 1965	b^h,d^h,g^h - β,ð,ɣ/V__V	yes	yes	yes

In short, all the logical possibilities are attested, and there are no asymmetries to account for. (Nor do any asymmetric *tendencies* regarding place of articulation strike the eye as one examines the Appendix. It is of course possible that some interesting statistical tendencies in this regard might emerge from a more controlled sampling of the world's languages, or from more careful instrumental studies of the articulatory movements underlying these patterns.)

1.5. Summary of Other Findings

1.5.1. Other conditioning environments. Aside from intervocalic position (and variants thereon, cf. chapter 6), lenition appears to be frequently conditioned by coda and final position.

Table 1-2. Cases of lenition in word-final position:

Language	Reference	Description of lenition processes
Afar	Bliese 1981	word-final degemination
Basque	Hualde 1993	k - ɣ word-finally
Carrier	Story 1984	ɢ,g,gʷ,ɣ - ɣ,j,w,Ø word-finally
Finnish	Sulkala & Karjalainen 1992	k, h delete word-finally
Guayabero	Keels 1985	d - θ word-finally
Haitian Creole	Tinelli 1981	ij - i, ʒ - j in final position
Lama	Ourso & Ulrich 1990	p - w word-finally
Maidu	Shipley 1963	ejectives deglottalize to voiceless stops word-finally
Mataco-Noctenes	Claesson 1994	ʔ - h word-finally
Pawnee	Parks 1976	r deletes word-finally
Pipil	Campbell	w - h word-finally
Tiberian Hebrew	Malone 1993	word-final degemination
Tojolabal	Furbee-Losee 1976	deletion of glides, h word-finally
Totonac	MacKay 1984	q - χ word-finally

I have encountered the following cases of coda lenition:[4]

[4]Note that lenition in coda position is also frequently bled by place assimilation to the following consonant, when such consonant is available, see Jun 1995. This may explain why explicit cases of medial coda lenition appear from the survey to be somewhat less frequent than word-final coda lenition.

Table 1-3. Cases of lenition in coda position:

Language	Reference	Description
Arbore	Harris 1990	debuccalization of coda ejectives
Hausa	Klingenheben 1928	b,d,g - w,r,w in coda
Quechua	Whitley 1979	k,q - x,χ in coda
Toba Batak	Hayes 1986	p,t,k - ʔ in coda
Uyghur	Hahn 1991	k,g - x, q - χ in coda

1.5.2. BLOCKING/FORTITION ENVIRONMENTS. The survey contains a large number of cases of blocking of lenition (which may, in some cases, be alternatively analyzed as fortition)[5] in word-initial position, and in the onset of a stressed syllable.

[5]The question of how to decide whether a given pattern relating "stronger" and "weaker" segments should be analyzed as lenition or fortition is taken up in Chapter 3 section 8.

Table 1-4. Fortition / blocking of lenition in word-initial position:

Language	Reference	Description
Burushaski	Lorimer 1935	x,ɣ,h - q,g,k word-initially
Carrier	Story 1984	w > b word-initially
Creole French	Goodman 1964	v - b word-initially
Efik	Dunstan 1969	blocking of spirantization, flapping in initial position
Hausa	Kraft & Kraft 1973, Dunstan 1969	ɸ - p word-initially
Ladakhi	Koshal 1976	voiced stops in free variation with fricatives except word-initially
Lamani	Trail 1970	flapping of retroflex d blocked word-initially
Mbabaram	Dixon 1991	blocking of stop voicing word-initially
Navaho	Kari 1976	x - h except word-initially
Nepali	Acharya 1991	ɖʰ - aspirated retroflex flap except word-initially
Pawnee	Parks 1976	w - p word-initially
Pennsylvania German	Kelz 1971	b devoiced word-initially
Southern Italian	Oftedal 1984	partial voicing of voiceless stops blocked word-initially
Spanish	Harris 1969	optional blocking of spirantization word-initially
Tauya	MacDonald 1990	k,kʷ - ʔ,ʔʷ except word-initially
Turkish	Bayraktaroglu 1992	v - w except word-initially
Uzbek	Sjoberg 1963	q - x except word-initially

A caveat: many of these grammars do not discuss phrasal phonology, the examples being limited to words in isolation; thus for many of the cases above it is unclear whether "word-initial" position is the proper characterization, or whether "initial" position in fact refers to some larger domain, e.g. phrase-initial, utterance-initial, or post-pausal. Explicit cases of phrase- or utterance-initial blocking are given below.

Table 1-5. Phrase- or utterance-initial blocking of lenition:

Language	Reference	Description
Nepali	Acharya 1991	spirantization of k^h except phrase-initially
Samoan	Mosel & Hovdhaugen 1992	s - ts utterance-initially
Spanish	Harris 1969	spirantization obligatorily blocked utterance-initially.
Tümpisa Shoshone	Dayley 1989	spirantization of stops and nasals blocked utterance-initially

Cases of stress-conditioned fortition or blocking of lenition are given below:

Table 1-6. Fortition / blocking of lenition in onset of stressed syllable:

Language	Reference	Description
American English	Kahn 1976	deaspiration of stops in unstressed medial or final syllable, flapping of t,d /+stressed V__V
Djabugay	Patz 1991	flapping of r, d /+stressed V__V
Kupia	Christmas & Christmas 1975	spirantization, flapping /V__V-stress
Southern Tati (Chali dialect)	Yar-Shater 1969	elision of /h/ blocked in onset of stressed syllable
Pattani	Sarma 1982	optional deaspiration of voiced and voiceless stops, medially and finally, particularly in unstressed syllables
Cardiff English	Collins and Mees 1990	voiceless stops strongly aspirated in onset of stressed syllable

2. Previous Approaches

2.1. Scope of this Review

In this section, I focus on previous approaches to the problem of formally characterizing, in a reasonably restrictive and insightful way, the class of lenition patterns. This section is not intended as an exhaustive presentation of previous work of lenition, particularly from the realm of historical linguistics. Indeed, much of this literature, to the extent that it raises theoretical issues at all, is concerned with the question of precisely which structural changes should be classified as "lenition" (see, e.g., Bauer 1988). In contrast, the effort-based approach presented herein largely sidesteps this debate, defining lenition not in terms of a definitive enumeration of structural changes, but as any substitution of a less effortful set of gestures. This characterization shifts the focus of inquiry away from cataloguing of structural changes, to the more phonetically concrete question, what is a less effortful set of gestures in the relevant context?

Relevant phonetic literature, including Articulatory Phonology approaches (e.g. Browman & Goldstein 1990, 1992), is discussed not in this section, but throughout the remainder of the dissertation.

2.2. Classic Generative Phonology

The notion that lenition is driven by considerations of articulatory effort is hardly novel. Hock (1991: 80), for example, expresses the naive, but apt, intuitiveness of such an idea:

> Among non-linguists, the perhaps most commonly cited cause for sound change is 'laziness'. While this is a dubious explanation for the great variety of changes that are found in the world's languages, it seems to be singularly appropriate for the class of changes which has been termed **weakening** or **lenition.**

However, this intuitive explanation has not standardly been incorporated into the formal characterization of these patterns. Rather, sound patterns have been standardly expressed in terms of language-specific rewrite rules which convert some class of underlying segments into a different class in a particular context; thus, intervocalic spirantization of oral stops may be expressed as the following rule:

(1-1) [-nas] - [+cont] / V__V

In the Generative tradition, however, it has been assumed that the formal simplicity of a rule should reflect its naturalness, thereby offering some insight as to why phonological phenomena such as lenition are widespread, whereas other conceivable rules are rare, or unattested (Chomsky and Halle 1968, chs. 8,9). By its own standards, then, the classic Generative Phonological formalism exemplified in (1-1) is inadequate: for an unattested rule, such as intervocalic stop formation, can be expressed with equal formal simplicity:

(1-2) [-nas] - [-cont] / V__V

Nor should this problem be dismissed as an artifact of Generative Phonology's preoccupation (in the view of its critics) with formalism, as opposed to functional explanation: for if explanatory principles such as "ease of articulation" are left in the realm of unformalized metatheory, their explanatory adequacy cannot be rigorously evaluated, and so they remain mere illusions of explanation.

2.3. NATURAL PHONOLOGY

The Natural Phonology program of Stampe (1972) and Donegan & Stampe (1979) attacked the phonetic arbitrariness of classic Generative Phonology. Anticipating much of the orientation of the effort-based approach, Donegan & Stampe invoke the twin functional principles of ease of articulation and ease of perception; and in fact use the term "lenition" to cover all patterns motivated by the former, including articulatorily-driven assimilations, such as /nb/ - [mb]. Moreover, they make a number of proposals concerning the *ordering* of lenition rules, relative to fortition rules. Unfortunately, the Natural Phonology program did not develop a restrictive, unified formal characterization of lenition processes. In the absence of a formalism capable of expressing violable conflicting principles, the functional insights of Natural Phonology remained unformalized metatheory.

2.4. AUTOSEGMENTAL PHONOLOGY: LENITION AS FEATURE-SPREADING.

An alternative reaction to the arbitrariness of classic Generative Phonology's feature-changing rules, Autosegmental Phonology (Goldsmith 1976) permitted a large class of natural phonological rules, most notably assimilation processes, to be expressed as operations on association lines, such as feature spreading. Within the framework of Autosegmental Phonology, then, an obvious move is to attempt to reduce lenition to autosegmental feature-spreading assimilation. Thus, Harris (1984) accounts for Spanish spirantization in terms of a rule that spreads [continuant] from an adjacent segment (see also Mascaró 1983 and Jacobs and Wetzels 1988; see Selkirk 1980, Mascaró 1987, Cho 1990, Lombardi 1991 for similar treatments of voicing).

This approach, however, faces several immediate problems. First, certain types of lenition, namely degemination, debuccalization, and elision, can only be expressed in autosegmental theory as deletion or delinking of phonological material, not as spreading. For example, there is no feature which can be spread onto a consonant to turn it into Ø, nor a feature which could turn a geminate into a singleton. Therefore the feature-spreading approach does not permit a unified characterization of lenition. Further note that processes such as deletion and debuccalization typically occur in the same sorts of contexts as voicing and spirantization, e.g. intervocalic position, as documented in Chapter 6 section 1.1. In the absence of an alternative explanation for the behavior of lenition as a unified phenomenon, this piecemeal approach must be considered inadequate. Second, this approach predicts that lenition may occur *whenever* a consonant is preceded (or followed) by any segment bearing a lenitional feature value ([+cont], [+voi], [+son], etc.), with equal likelihood. In particular, the feature-spreading approach fails to give a natural account of two-sided lenition contexts such as intervocalic position: it suffices to spread the relevant feature from *either* adjacent vowel, and so the role of the other vowel in conditioning the lenition is unexplained.[6] Inouye (1995) attempts to address this problem with respect to intervocalic flapping, by proposing that flaps are represented as tripartite contour segments (1-3a):

(1-3) a. [ɾ] as a contour segment: b. Intervocalic flapping as double spreading of [+cont]:

With this representation, intervocalic flapping can be captured as spreading of [+cont], from both adjacent vowels (1-3b). But this contour-segment treatment does not extend to intervocalic spirantization or voicing, as it is untenable to posit contour representations for fricatives and voiced obstruents with respect to continuancy and voicing. Nor can this

---

[6] As Flemming (1995: 116) has observed, this problem constitutes a general flaw of the autosegmental formalism: in this framework, coarticulatory processes are treated as feature-spreading, but coarticulations are frequently more extreme when they are two-sided. Flemming gives the example of Cantonese vowel fronting, which occurs between two coronal consonants.

approach offer a unified account of intervocalic flapping and flapping in other contexts, e.g. word-final flapping of /ɖ/ in Gujarati (Cardona 1965).

2.5. LENITION AS DEFAULT FILL-IN RULES.

Jacobs and Wetzels (1988) describe an approach to lenition in terms of underspecification and default fill-in rules. Analyzing post-vocalic voicing and spirantization in Gallo-Romance, they (initially) assume that [voi] and [cont] were underlyingly unspecified, and posit "feature-filling" processes of voicing (1-4a) and spirantization (1-4b), as well as a set of more general default fill-in rules for these features.

(1-4) a.

$$[0\ \text{voi}] \rightarrow [+\text{voi}]\ /\ \underset{\text{Nucleus}}{\underset{|}{X}}\ \ \overset{\begin{bmatrix} \_\_\_ \\ -\text{son} \end{bmatrix}}{\overset{|}{X}}$$

b.

$$[0\ \text{cont}] \rightarrow [+\text{cont}]\ /\ \underset{\text{Nucleus}}{\underset{|}{X}}\ \ \overset{\begin{bmatrix} \_\_\_ \\ -\text{son} \\ +\text{voi} \end{bmatrix}}{\overset{|}{X}}$$

Jacobs and Wetzels' stated goal is to capture the ostensible tendency of lenition processes to be non-neutralizing. If lenition rules are typically feature-filling, then no neutralization of underlying contrasts can occur. This formalism, however, fails to capture the naturalness of the spirantization and voicing rules (as contrasted, say, with occlusivization or devoicing in the same environment), in precisely the same way that the classical Generative formalism fails to do so. The assumption of underlying underspecification here merely serves to recast arbitrary feature-changing rules as equally unconstrained feature-filling rules. Furthermore, like the spreading approach, this approach cannot be extended to elision, degemination, or debuccalization, which involve operations of deletion of phonological material, not feature insertion operations.

2.6. Lenition as Scalar Promotion

The fact that lenition patterns often involve chain shifts, and the fact that lenition often involves a diachronic gradual "erosion" of stops into more reduced consonants, ultimately culminating in elision, make the scalar nature of lenition readily apparent. It is therefore an obvious move to attempt to reduce the weakening scale seen in lenition typology to the other phonological scale countenanced in the standard theory, viz. the sonority scale, from the domain of phonotactics and syllabification (e.g. Clements 1990).

However, the fit between the two scales is not particularly compelling:

(1-5) a. Sonority scale (Dell & Elmedlaoui 1985)
stops < voiceless fricatives < voiced fricatives < nasals < liquids < high vowels/glides < low vowels

b. Strength scale (composite, from Hock 1991 and Lavoie 1996)
geminate stops > voiceless stops > voiced stops > voiceless fricatives > voiced fricatives > liquids > laryngeals > glides > Ø

At the strong end of the lenition scale lie the geminate stops; but there is no evidence that these behave as less sonorous than singletons for phonotactic purposes; indeed, they typically do not pattern as single segments at all, but as clusters. Conversely, the weak end of the lenition scale is occupied by Ø; but it is impossible to speak of a deleted segment as having any sonority at all, let alone being maximally sonorous. Furthermore, going by the sonority scale, we would expect that fricatives can lenite to nasals; yet such alternations are unattested. Moreover, if we broaden the scope of our investigation to include vowel reduction (which like consonant lenition, occurs more readily the faster the speech rate and the lower the register, and is commonly blocked in stressed syllables, and hence would appear to warrant a unified treatment), we arrive at a paradox. Vowel reduction typically involves *raising* (and centralization), e.g. a - ə; but the higher the vowel, the *less* sonorous it is; thus vowel reduction appears to involve sonority *demotion* rather than the expected promotion. Finally, the sonority scale suffers from the same lack of an explicit, unified phonetic characterization (see Kawasaki 1982) that has plagued the "strength" scale of lenition theory (see Bauer 1988). Thus, even if it were successful, the strategy of unifying the two scales at best amounts to debt consolidation, not payment in full.

It is, of course, conceivable that lenition involves some abstract scale of "strength," which bears no straightforward relation to any phonetic dimension, which is distinct from sonority, and which may even vary from language to language. Lenition then is characterizable in terms of an operation of promotion on this scale. Such a position is explicitly adopted by Foley (1977). However, this view of the "strength" scale does not appear to offer anything more than a bare restatement of the facts. Moreover, in the context of rule-based frameworks, which attempt to characterize possible sound systems in terms of a maximally restrictive set of operations (such as Autosegmental Phonology's operations on association lines), the introduction of promotion operations constitutes a serious weakening of the theory.

Furthermore, we have been indulging in an oversimplification by characterizing lenition in terms of a simple *scale*: the unrestrictiveness of the promotion operation becomes strikingly apparent once we acknowledge that lenition in fact involves a lattice:

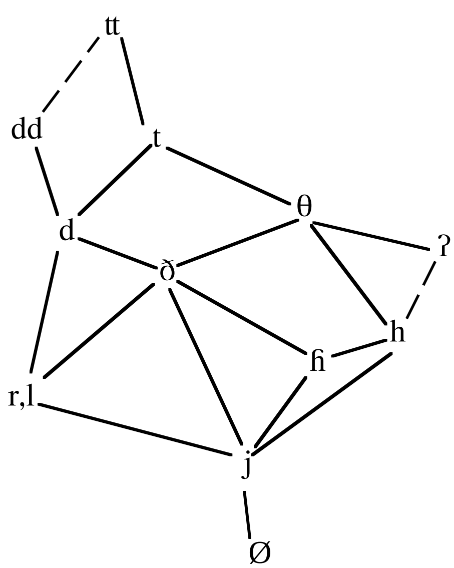


Figure 1-1. Possible lenition changes (from Hock 1991:83).

(Dotted lines in the figure above represent possible (in Hock's view) but unattested lenition patterns.) If promotion operations may refer to any and all of the downward paths within some such lattice, this approach becomes quite unconstrained, even with respect to the limited goal of characterizing the structural changes of lenition rules. Finally, the promotion operation says nothing about the contexts and conditions under which lenition occurs; hence it can afford no characterization insightfully linking the structural changes of lenition with the environments in which they naturally occur.

In fact, although the sonority-promotion approach to lenition has frequently been put forward in the theoretical literature (e.g. Foley 1977, Churma 1988, Clements 1990, Hock 1991, Ní Chiosáin 1991, Elmedlaoui 1993, Lavoie 1996), this proposal has rarely been fleshed out in explicit analyses of actual sound patterns. Ní Chiosáin's (1990) analysis of morphologically conditioned lenition in Irish (“eclipsis") illustrates the problem. Ní Chiosáin proposes a rule that promotes the target segment one step on a language-specific scale (nominally identified with "sonority," but without independent motivation):

(1-6) p,t,k - b,d,g - m,n,ŋ
f - v

However, there is a set of lenition mutations in Irish distinct from eclipsis:

(1-7)	p - f	t,s - h	f - Ø
	b,m - v	d - ɣ	
	k - x		
	g - ɣ		

Rather than posit promotion operations along two distinct scales, she accounts for the second type of mutation using default fill-in rules, similar to Jacobs and Wetzels' approach, as well as deletion and debuccalization rules, thereby giving up on a unified treatment of lenition.

2.7. Lenition as Loss of Phonological Material

A more comprehensive, principled approach to lenition typology is Harris' (1990) attempt to extend deletion operations, required to characterize elision and degemination, to lenition generally, within the framework of Government Phonology (Kaye, Lowenstamm, and Vergnaud 1985). In Government Phonology, all features are privative. Thus voicing lenition can be expressed as the loss of a h˚ element (corresponding to [-voice] in conventional feature inventories), and spirantization is the loss of a ʔ˚ (= [-continuant]) element. More generally, Harris characterizes lenition as *any deletion of phonological material*. (Cf. Grijzenhout's (1995) characterization of Celtic consonant mutations as deletion of aperture nodes (Steriade 1993a).)

To account for coda lenition, Harris posits a "Complexity Condition," which requires a "governed" position to be of lesser complexity (i.e. number of elements, which is directly related to "strength") than its governor, in this case the following onset. Consequently, a coda consonant which is underlyingly governed by a less complex (weaker) onset may jettison one or more features to avoid a Complexity Condition violation. Similarly, the nucleus of a governed (unstressed) syllable must have a less complex onset than that of its governor (the nucleus of an adjacent stressed syllable); this may then induce loss of features in the onset of the unstressed syllable.

However, this approach presents a number of problems. Firstly, the privative representation of voicelessness is inconsistent with the widely attested phenomenon of word-final devoicing. In Harris' feature system, such a rule must be expressed as the insertion of a h˚ element: not only is such feature insertion formally arbitrary; it occurs in coda position, where Harris predicts feature *deletion*. But without this assumption, Harris' unified approach to lenition collapses.[7] Moreover, the Complexity Condition incorrectly predicts that lenition of the onset of one syllable can be sensitive to the complexity of the onset of an adjacent syllable (the nucleus of an unstressed syllable must

[7]A solution to this apparent conundrum under the effort-based approach is discussed in Chapter 2, section 6.2.

have a simpler onset than that of a stressed syllable). We would therefore expect to find lenition patterns such as C - +cont /[__V]$_{\sigma}$[CCV]$_{\sigma}$: to my knowledge, such patterns are completely unattested. Nor is it clear how Harris' notion of onset lenition as relative strength of nucleus government translates into an intervocalic environment for lenition (although Harris asserts that it does): apparently under Harris' definition of government, a closed as well as an open syllable can govern a following syllable, in which case the onset of the following syllable is a lenition environment by Harris' definition, although it is not intervocalic. More seriously, Harris' lenition-as-deletion approach does not capture the intuition that lenition in intervocalic position is, in some sense, assimilatory: that it is the openness of the vowels which is making the intervening consonant less constricted. Finally, Harris' approach says nothing about the further typological generalizations noted above, such as sensitivity of lenition to speech rate and register, geminate inalterability under lenition, and the non-stridency of spirantization outputs.

2.8. ASSESSMENT

None of the foregoing approaches afford a reasonably restrictive, unified characterization of lenition processes. Genuine insights into lenition, I submit, are to be found in an examination of the phonetic motivations for lenition. Moreover, to incorporate such phonetic explanation into the formal theory, the Optimality Theoretic notion of constraint conflict is required.[8]

---

[8] A previous Optimality Theoretic treatment of lenition is Jacobs' (1994) reanalysis of Jacobs & Wetzels' (1988) treatment of Gallo-Romance spirantization and voicing. To account for spirantization, Jacobs posits a constraint on voiced non-continuants in a "lenition environment." The account of voicing is along similar lines. Unfortunately, Jacobs does not attempt any formal characterization of "lenition environments"; nor is there any explanation of why the specifications [-voice] and [-cont] might be disfavored in these environments. Consequently, Jacobs (1994), like Jacobs and Wetzels (1988), does not address the fundamental arbitrariness of the formal devices used to express lenition phenomena.

3. An Effort-Based Approach to Lenition

3.1. Optimality Theory

My approach is couched in the framework of Optimality Theory (OT) (Prince and Smolensky 1993), which I will now briefly sketch. (Readers already familiar with OT may skip this section.)

In OT, the basic formal device is *constraints* rather than *rules* (i.e. operations). The constraints are *violable*. Cross-linguistic variation lies not in the constraint set, but in their *ranking*. Input-output mappings are determined by two functions, GEN (candidate generation) and H-EVAL (harmonic evaluation): GEN takes the input and maps it to an (infinite) set of candidate outputs. H-EVAL takes the candidate set and maps it to the output (the "winner"), i.e. the candidate with the best score, vis-à-vis the constraint hierarchy. A candidate which violates a higher-ranked constraint loses to any candidate which violates only some lower-ranked constraints, regardless of the number of violations of the lower-ranked constraints by either candidate. A candidate which multiply violates some constraint loses to any candidate which violates that same constraint to a lesser degree. If several candidates *tie* w.r.t. (violation or satisfaction) of some constraint, the constraint on which they tie is then irrelevant to the choice between them, and so it passes to lower-ranked constraints to discriminate between them. Proof that a particular input-output mapping obtains under some ranking takes the form of a *tableau*:

(1-8)

Input: /an-ba/	Preserve (nasal)	*Heterorganic Clusters	Preserve (coronal)
anba		*!	
☞ amba			*
aba	*!		*
etc.			

Tableau violations are indicated with an asterisk. '!' marks a 'fatal' violation (one which eliminates the candidate). '☞' marks the winner. Cells are shaded to indicate that any violations which they may contain are irrelevant to the evaluation. Lack of a vertical border separating constraint columns indicates that the constraints are not (crucially) ranked with respect to each other.

In sum, OT is a framework in which phonological systems are expressed in terms of tradeoffs between conflicting universal constraints. Notions of well-formedness "supply the very substance from which grammars are built: a set of highly general constraints which, through ranking, interact to produce the elaborate particularity of individual languages" (Prince and Smolensky 1993:198). For a more detailed presentation of, and motivation for, this formalism, see Prince and Smolensky 1993.

3.2. PHONETICALLY BASED OT

Within this framework, then, an obvious move is to attempt to identify the OT constraint set, the "notions of well-formedness," with independently motivated, functionally based principles, such as those notions of phonetic optimality (informally) appealed to by Natural Phonology (see section 2.3) and other functionally oriented traditions of linguistic analysis. For example, the *HETERORGANIC CLUSTERS constraint in (1-8) plausibly reduces to the more general, physiologically grounded imperative to minimize effort. That is, by substituting a single extended labial closure gesture ([mb]) in place of the coronal + labial sequence of gestures in (nb), a more economical articulation is presumably achieved.

Since appeals to functional explanation are traditionally received with skepticism by generative phonologists, as typically lacking in formal rigor, it is perhaps helpful to address some of these concerns up front. It has sometimes been claimed, for example, that phonetic factors are mere tendencies, too "fuzzy" to be useful in formal phonological analyses of particular languages (Lass 1980, Anderson 1981). Moreover, phonetic principles refer to concrete, continuous representations, with gestures and formant frequencies and the like; whereas it is standardly assumed that phonological representations are more abstract, since they ought only to reflect phonologically significant distinctions. Finally, even highly "natural" processes may have phonetically arbitrary aspects to their behavior; therefore it would appear that any attempt to reduce them to phonetics is misguided.

However, the fact that languages differ in the extent to which they adhere to one phonetic principle vs. another is not, as e.g. Lass 1980 assumes, an indication of the formal inutility of such principles, but rather a basic prediction of OT's core assumptions: constraint violability and language-specific ranking (cf. McCarthy and Prince's (1994) discussion of the "fallacy of perfection"). As for representations, within the OT framework, the contrastiveness and categorical behavior of particular features within a sound system can be captured in terms of constraint ranking, particular with respect to faithfulness constraints; therefore phonological representations may be as phonetically concrete as the theory of phonetic implementation may require (see Chapter 3). Finally, the fact that some single phonetic principle does not explain the totality of some phonological process is not a bar to its insightful deployment, as e.g. Anderson 1981 assumes; for OT presupposes that sound patterns arise from *interactions* of principles. The modus operandi of phonetically based OT is to tease apart sound patterns into an interaction of conflicting principles, each of which directly expresses some notion of phonetic or cognitive functionality.

3.3. A SKETCH OF THE PROPOSAL

The core proposal of this dissertation is that lenition patterns are expressed in terms of conflicts between the effort minimization constraint, LAZY, on the one hand, and on the other hand a class of lenition-blocking constraints. The lenition-blocking constraints in turn are further divisible into "faithfulness" constraints (penalizing divergence from identity between underlying representation and corresponding surface form), and "fortition" constraints (which serve to enhance the salience and robustness of perceptual distinctions):

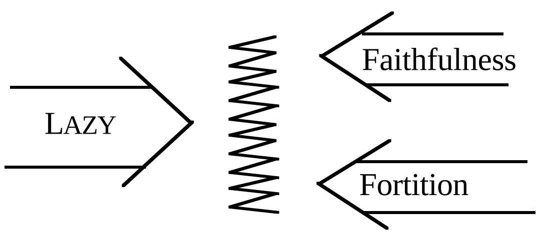


Figure 1-2. Opposition between LAZY and lenition-blocking constraints.

For example, assume a constraint PRESERVE(continuant), which penalizes any change in the underlying value of the feature [continuant]. Spirantization corresponds to rankings where LAZY » PRES(cont) (1-9a); under the opposite ranking (b), spirantization is blocked:

(1-9) a.

	/d/	LAZY	PRES(cont)
	d	**!	
☞	ð	*	*

b.

	/d/	PRES(cont)	LAZY
☞	d		**
	ð	*!	*

To capture the range of variation in lenition patterns, this naively intuitive idea can be elaborated as follows.

3.3.1. A UNIFIED CHARACTERIZATION OF LENITION

Trivially, the treatment of spirantization above, in terms of conflict between LAZY and faithfulness, can be extended to all manner of lenition phenomena. The type of structural change occurring in a given language depends upon which of the lenition-blocking constraints, if any, are ranked below LAZY: if PRES(length), then degemination; if PRES(voice), then voicing; if PRES(sonorant), then reduction of an obstruent to an approximant; if PRES(place features), then debuccalization; if PRES(segment), then elision; if no PRESERVE constraint, then no lenition at all. Lenition, then, receives a unified characterization, in terms of the ranking schema: LAZY » lenition-blocking constraint.

3.3.2. CONTEXT-SENSITIVE LENITION PATTERNS

Under the simple case of LAZY outranking some faithfulness constraint, such as PRES(cont), the result is context-free lenition, e.g. Berber (Saib 1977), in which all singleton obstruents are realized as fricatives, in all contexts.[9] However, with a few enrichments, the theory can capture context-sensitive lenition patterns as well, in terms of the same basic ranking schema.

3.3.2.1. INTERVOCALIC POSITION. Intervocalic lenition, and related context-sensitive lenition patterns, receive a straightforward effort-based treatment. Ceteris paribus, the more open the flanking segments, the greater the displacement (hence effort)

[9]The failure of the geminates to undergo spirantization receives an effort-based account in Chapter 5.

required to achieve a given degree of consonantal constriction. The primacy of intervocalic position as a context for lenition thus falls out from the natural assumption that the impetus to lenite more effortful gestures is stronger than the impetus to lenite easier gestures. A formal treatment of this observation, involving decomposition of LAZY into a series of effort thresholds, and interleaving of lenition-blocking constraints within this series, is presented in Chapter 6. It is further shown that in Chapter 6 that the typical sensitivity of lenition to speech rate and register falls out from this same treatment.

3.3.2.2. CODA, WORD-FINAL POSITION. Restriction of lenition to coda and word-final position can be understood in terms of the impoverished perceptual cues to a consonant's identity in phonotactic positions where it lacks an audible release, see Steriade 1993, 1995, 1996; Jun 1995. The greater perceptibility of consonants in positions where their release is audible can be formally expressed by breaking up faithfulness constraints according to context: the more salient position corresponds to a universally higher-ranked faithfulness constraint, thus PRES(F/released position) » PRES(F), as motivated by Jun 1995, cf. Beckman 1997. Coda and word-final lenition can then be obtained by ranking LAZY between these:[10]

---

[10]It is not necessarily the case that a consonant in word-final position lacks release, even if it is unreleased in word-medial coda position, as the above account appears to presume; for example, a consonant will be released if followed by a vowel-initial word within the phrase. However, stabilization of the word-final lenited outcome can be obtained through interaction with output/output faithfulness constraints. That is, the constraint ranking above causes /ak/ to be realized as [ax] in pre-pausal and pre-consonantal environments; and an undominated faithfulness constraint, PRES(cont,O/O), enforces identity between this output and the output of /ak/ in pre-vocalic phrase-medial position, resulting in uniform word-final spirantization. See Steriade 1996 for an application of this idea to word-initial devoicing; and see the further deployment of output/output faithfulness in connection with heteromorphemic geminates in Chapter 5, section 4.

(1-10)

	PRES(cont/ released)	LAZY	PRES(cont)
ak˺ - ak˺		**!	
☞ ak˺ - ax		*	*
ak˺ta - ak˺ta		**!	
☞ ak˺ta - axta		*	*
☞ aka - aka		**	
aka - axa	*!	*	*

Intuitively speaking, this treatment captures the insight that there is greater impetus to lenite in contexts where there is relatively little perceptual "bang" for the articulatory "buck."

Restriction of lenition to particular *places* of articulation may similarly be obtained in terms of context-sensitive faithfulness constraints: specifically, the operative context refers to specific place features. For example, the following ranking results in spirantization of coda dorsal consonants, but not coronals or labials, as in Quechua.[11]

[11]Alternatively, such patterns may be obtained by allowing place-specific LAZY constraints: the ranking {LAZY_{cor}, LAZY_{lab}} » IDENT(cont) » LAZY_{dors} yields the same pattern. It is not clear to me that any empirical difference follows from this place-specific LAZY proposal. I am inclined to favor the context-sensitive faithfulness proposal, however, based on the intuition that effort is an indivisible notion, and therefore it does not make sense to suppose that languages might arbitrarily disfavor effort involving some particular articulator; whereas it does make sense to suppose that speakers of languages differentially attend to particular auditory cues (cf. Chapter 3).

(1-11)

	PRES(cont/released)	PRES(cont/lab)	PRES(cont/cor)	LAZY	PRES(cont/dors)
Coda position: ak˺ - ak˺				**!	
☞ ak˺ - ax				*	*
ap˺ - ap˺				**	
☞ ap˺ - aϕ		*!		*	*
at˺ - at˺				**	
☞ at˺ - aθ			*!	*	*
Onset position: ☞ pa, ta, ka - pa, ta, ka				**	
pa, ta, ka - ϕa, θa, xa	*!	*	*	*	*

3.3.3. FORTITION CONSTRAINTS.

Note, however, that for cases of *complementary distribution*, e.g. no word-initial fricatives, and no non-initial stops, the use of faithfulness constraints as lenition-blockers is insufficient.

(1-12)

		PRES(cont/#__)	LAZY	PRES(cont)
☞	a. #ka - #ka		**	
	#ka - #xa	*!	*	*
	b. #aka - #aka		**!	
☞	#aka - #axa		*	*
	c. #xa - #ka	***!**	******	*****
☞	**#xa - #xa**		*****	
	d. #axa - #aka		**!	*
☞	#axa - #axa		*	

If, as in (1-12c), some word-initial obstruent is underlyingly [+cont] (and the OT tenet of Richness of the Base (Prince & Smolensky 1993, ch. 9) prevents us from excluding such an input), both faithfulness and LAZY favor the fricative candidate; thus it is impossible to rule out word-initial fricatives. An additional class of lenition-blocking constraints is therefore required: these must not only block lenition, but actively *induce fortition*, e.g.

requiring word-initial obstruents to be realized as stops (*[+cont,-son]/#__). It seems plausible that these fortition constraints are, like the context-sensitive faithfulness constraints, grounded in perceptual considerations. For example, the release burst of a stop contains salient place of articulation cues (e.g. Wright 1996); thus, by militating in favor of consonants with a release burst, this constraint can be viewed as enhancing the perceptibility of the consonant; and the allocation of more robust cues to word-initial position may be viewed as reflecting the greater importance of word-initial consonants in lexical access. More generally, I will assume that the fortition constraints which we appeal to for purposes of account for lenition typology are of the form $*\alpha F/[_D\_\_$, where D is some prosodic or morphological domain (including a stressed syllable), and αF refers to some feature specification which is less perceptually salient in the context $/[_D\_\_$ than is $-\alpha F$. The precise formulation of the fortition constraints, however, and their relation to broader perceptual considerations, is not central to the thrust of this dissertation, namely the role of effort in lenition; therefore I will not discuss these ideas at great length. For a more thorough treatment of perceptual enhancement in phonology, see Flemming 1995.

3.3.4. GENERALITY OF THE APPROACH.

This general constraint system is motivated not merely by lenition typology. Essentially the same set of constraints is deployed by Jun (1995) to account for place assimilation in consonant clusters. Jun demonstrates that casual speech gradient assimilation (e.g. /fon bʊk/ - fɸmbʊk]), attributed by Browman and Goldstein (1990) to gestural overlap, in fact involves gestural reduction of C_1, to the point where the percept of C_1's place of articulation is lost; moreover, categorical "phonological" assimilations can be analyzed in the same terms, where the reduction of the C_1 gesture is total. Local assimilations,[12] then, emerge as a special case of lenition, where gestural reduction is accompanied by temporal extension of the gesture of C_2, in order to preserve other underlying properties of the target segment, such as non-continuancy.

(1-13)

/atka/	PRES(seg)	PRES(cont)	LAZY	PRES(cor)
atka			****!	
☞ akka			***	*
ahka		*!	**	*
aka	*!	*	*	*

It can readily be seen from this tableau that, with higher ranking of LAZY, the manner as well as the place of the underlying /t/ would be lost, resulting in debuccalization or elision, i.e. lenition *tout court.*

More generally, the effort-based approach which I adopt may be viewed as part of an emerging research program, which weds the substance of functional phonetic explanation with the formalism of OT constraint interaction, in order to achieve more deeply explanatory accounts of phonological phenomena: this goal appears, to varying degrees, in such recent works as Steriade 1993, 1995, 1996; Kaun 1994; Flemming 1995, 1997; Jun 1995; Silverman 1995; Myers 1996; Beckman 1997; Boersma 1997a,b,c,d; Hayes 1997; Kirchner 1997; MacEachern 1997; and Gordon (in progress). Furthermore, the approach continues a line of research on phonetic explanation in phonology, associated with phoneticians such as Ohala (1981, 1983); Lindblom (1983, 1990); Browman & Goldstein (1990, 1992); and Kohler (1991).

---

[12]Harmonic (long-distance) assimilations, in contrast, appear to be perceptually driven, see e.g. Kaun 1994, requiring a distinct perceptually-based treatment.

Finally, note that my focus upon reduction of *consonants* in this dissertation does not reflect an assumption that vowel reduction is an unrelated phenomenon. On the contrary, vowel reduction, like consonant reduction, involves gestural reduction, tends to occur in unstressed syllables, and occurs to a greater extent in fast and casual speech. However, I only address vowel reduction tangentially, rather than undertaking a broad typological survey, because I anticipate that the crucial interactions between LAZY-driven reduction and prosodic constraints are more complex than in consonant lenition typology; and because reliance on transcriptions in published grammars, rather than instrumental phonetic measurements, seems riskier in the case of vowel reduction, because the distinction between, say [ɪ] and [ɨ], is often less clear-cut than manner distinctions in consonants.

4. STRUCTURE OF THE DISSERTATION

The structure of the remainder of the dissertation is as follows. In Chapter 2, I discuss the notion of "effort" to which the constraint LAZY refers. In Chapter 3, I take up the question of the phonetic richness of phonological representations under the effort-based approach. Chapter 4 discusses the non-stridency of spirantization outputs; and Chapter 5 examines the behavior of geminate consonants under lenition. In Chapter 6, an enrichment of the basic approach is motivated, decomposing the scalar LAZY constraint into a series of binary constraints, each of which refers to some threshold of effort: this enrichment allows us to capture the primacy of intervocalic position as a lenition environment, as well as the typical rate and register sensitivity of lenition processes. Chapter 7 contains a case study of voicing, spirantization, and flapping in Tümpisa Shoshone. Chapter 8 presents a case study of the partially rate- and register-sensitive pattern of lenition in Florentine Italian. Finally, Chapter 9 concludes the dissertation by addressing the problem of stable (i.e. rate- and register-insensitive) patterns of lenition under this approach.

Chapter 2:

Articulatory Effort

The notion of effort minimization has often been invoked as a source of explanation for particular sound patterns and sound changes, e.g. Grammont 1933; Zipf 1949; Stampe 1972; Lindblom 1983, 1990. Nevertheless, such appeals to effort minimization (and appeals to functionalist principles generally) have typically been received with skepticism by Generative phonologists, e.g. Anderson 1981, for three principal reasons.

First, prior to the development of Optimality Theory, linguists did not well understand how violable principles such as effort minimization could be incorporated into formal analyses. If a principle of effort minimization is active in phonology, the argument went, why does it not apply to the same extent and yield the same results in every language (e.g. Lass 1980, ch. 2)? Optimality Theory has shown us, however, that formal analyses of particular sound systems can be constructed from the interaction of a set of universal *violable* constraints; language-specific variation arises from variable ranking of the constraints. As noted in Chapter 1 section 3, Lass's argument can now be recognized as a case of the "fallacy of perfection" (McCarthy and Prince 1994), i.e. the assumption that notions of well-formedness can only be stated as inviolable principles. Phonetic principles such as effort minimization may be expressed in the formal theory as universal constraints, yet they may be violated in particular languages due to conflicting constraints.

Second, the actual effort involved in a given utterance may vary with speech rate, loudness, the size of the speaker's jaw, the amount of air in the speaker's lungs, the presence of chewing gum in the speaker's mouth, etc. Certain aspects of speech production are undoubtedly sensitive to such idiosyncratic, token-specific conditions; but other processes, including many documented lenition processes, are stable across tokens (see Steriade 1995).[13] Thus, direct conditioning of phonological processes by considerations of effort cost appears problematic. This problem is addressed in Chapter 9.

Finally, previous appeals to effort in phonological explanation have in some cases been insufficiently explicit about what “effort" means. Such appeals therefore are not obviously distinguishable from the naive, subjective impression that sound patterns which are unattested in one's own language are “hard to pronounce." Such difficulty may merely reflect unfamiliarity. Obviously, phonological systems cannot be explained by appealing to difficulty in this sense, since any unattested sound pattern is necessarily unfamiliar, rendering the "explanation" tautologous: the pattern is unattested because it is difficult, and it is difficult because it is unattested, hence unfamiliar. To avoid this pitfall, it is therefore crucial to provide an explicit, physically-based notion of articulatory effort. It is to this task which we turn for the remainder of this chapter.

[13]Though in some cases this may be an artifact of idealized grammatical descriptions.

1. The Neuromuscular Basis of Effort

1.1 Action potentials.

In principle, effort is a neuromuscular notion. A nerve impulse activates a group of voluntary muscle fibers (the “motor group") through an electrochemical reaction (the “action potential"), which causes a brief twitch of the fibers (Borden and Harris 1994: 57-58), and which consumes a quantity of the basic "fuel" of muscle tissue, adenosine triphosphate (ATP). Repeated firing of action potentials is needed for full contraction of the motor group (Clark and Yallop 1990: 19-20). If the *agonist* activity of some set of motor groups in the vocal tract is greater than the *antagonist* (opposing) activity of others, muscle contraction occurs, resulting in movement of some articulatory structure. If, however agonist and antagonist activity is balanced (or if agonist activity is opposed by some external force), the muscle isometrically tenses, without movement. Effort, then, is the extent of this activation: specifically, the total sum (integral) of the action potentials of a motor group, summed over each active motor group, for every muscle in the vocal tract; or, in neurochemical terms, the amount of ATP consumed by the muscles of the vocal tract.

1.2. Grammatical implications.

Information on the extent of this activation is available to the nervous system via sensors known as “muscle spindles," which together with joint receptors form the proprioceptive feedback system (Borden and Harris 1994: 147-148). It is therefore plausible that, as speakers acquire experience in making articulatory gestures, they develop knowledge of the effort required for those gestures, and can therefore anticipate the effort required to produce any given set of gestures. I assume that it is part of the speaker's phonological competence to appeal to this knowledge in evaluating the well-formedness of phonological representations with respect to the following constraint:

(2-1) LAZY: Minimize articulatory effort.

Formally, an *estimate* of effort cost is computed for each candidate representation as part of the candidate generating function, GEN. The candidate with the higher effort estimate fares worse with respect to LAZY:

(2-2)

		LAZY
	Candidate A	5 mg ATP!
☞	Candidate B	4 mg ATP

This grammatically relevant notion of an effort estimate, then, is, like all linguistic constructs, a mental notion (cf. Donegan and Stampe 1979), albeit corresponding to a physical reality. It is of course conceivable that the mental estimate of effort bears no straightforward relation to actual physical effort; but the null hypothesis is a transparent relation, therefore I will not take pains to distinguish the mental, grammatical notion of the effort estimate from the actual neuromuscular effort cost.

1.3. PROBLEMS OF OBJECTIVELY QUANTIFYING NEUROMUSCULAR EFFORT

Unfortunately, the hypothesized effort estimate is not directly accessible to external investigation; nor is the actual neuromuscular feedback. Proprioceptive feedback of muscle activity is typically unconscious (Borden & Harris 1980). And as noted above, simple introspection is probably an unreliable indicator of articulatory effort: for a subject's report of difficulty in articulating a speech sound may merely reflect relative unfamiliarity with the gesture, or set of gestures, associated with the sound.

Firing rates of individual muscle spindles can be measured by means of electromyography (EMG) techniques, i.e. attachment of electrodes to muscle spindles (Borden and Harris 1994: 244). However, to compare degrees of effort involved in particular utterances, activation information is required for *all* the muscle groups which are involved in the gesture; otherwise, one is in the position of guessing the shape of the proverbial elephant from measurements of its tail. Unfortunately, there are at present no techniques of measuring such global activity directly: obviously, it is not possible to insert electrodes into every muscle of the vocal tract simultaneously. Standard methods of measuring physical exertion in larger-scale tasks, such as measurement of heartbeat and breathing rates, are obviously useless in regard to the fine degrees of effort involved in speech.

2. A BIOMECHANICAL APPROACH TO EFFORT

2.1. EFFORT AS WORK OR ENERGY.

For these reasons, phoneticians have typically approached articulatory effort from a biomechanical rather than a direct neuromuscular perspective. Lindblom (1990) for example, defines effort as *work (= force × displacement) per unit time*. However, the equation of effort with work seems problematic. Without displacement, i.e. actual movement, zero work is performed; and if work is equated with effort, we must say that zero effort is expended. This seems incorrect: intuition suggests that considerable effort can be expended, for example, in pushing against a brick wall, or holding a heavy package, without actually moving anything. The same is true, on a smaller scale, of the effort expended in compressing an active articulator against a passive articulator, e.g. in a stop, to maintain closure. The notion of *energy* (the capacity for doing work, or that which is expended in the performance of work) derives from the notion of work, and faces the same objection.

2.2. EFFORT AS FORCE.

Therefore, *force*, rather than work (or derivative notions), seems the better mechanical analog of effort (e.g. Nelson 1980). The force involved in some articulatory displacement x at velocity v is:

(2-3) $$F(t) = m \times \frac{dv}{dt}$$

where m is the mass, and dv/dt (i.e. change in velocity over change in time) is the instantaneous acceleration at time t. Unlike the notion of work, however, no actual displacement is required in order for force to be exerted; for the positive force (which would ordinarily result in displacement) may be opposed by some negative (antagonistic) force:

(2-4) $$\begin{aligned} F_{pos}(t) &= -F_{neg}(t) \\ F_{net}(t) &= F_{pos}(t) + F_{neg}(t) = 0 \\ F(t) &= |F_{pos}(t)| + |F_{neg}(t)| = 2F_{pos}(t) \end{aligned}$$

If the negative force equals the positive force,[14] as in (2-4), no displacement is achieved (i.e. *net force* equals zero); but the *total force* exerted (*F(t)*) equals the sum of the (absolute values of the) positive and negative forces combined.[15] This accords with our neuromuscular observation that the activation of muscle groups may oppose one another, resulting in isometric tension rather than displacement.

---

[14]That is, if the *absolute value* of the two forces are equal. The two forces cannot, strictly speaking, be equal (except in the case where both = 0), since negative force is expressed as a negative number.
[15]Assuming that both are active forces, i.e. exerted due to contraction of the speaker's muscle fibers -- as opposed to passive forces, exerted on the speaker by gravity, elastic recoil, and other external forces.

2.3. FORCE THROUGHOUT A GESTURE.

However, we are not interested in the force applied at a single point in time, *F(t)*, during some gesture,[16] as in (2-3) and (2-4), but rather the force exerted throughout the gesture, i.e. the total sum (integral) of the force through time:

(2-5) $$F_j = \int_{ti}^{tj} |F(t)| d(t)$$

where F_j is the force exerted during a gesture beginning at time t_i and ending at t_j. Integration is required because the level of force, *F(t)*, typically varies during the course of the gesture. That is, the force of the gesture is equated with the *area bounded by the function* of force against time:

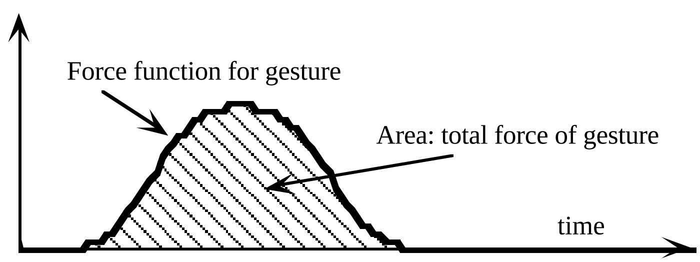


Figure 2-1. Schema of force of a gesture.

Finally, the effort involved in an entire utterance is simply a summation of the forces associated with each component gesture: i.e.,

---

[16]Note that I am using "gesture" here not in the task dynamic sense of the Articulatory Phonology literature (Browman and Goldstein 1990), but in a purely physical sense, as any voluntary displacement or tension of any organ in the vocal tract.

(2-6) $$F_{utterance} = \sum_{i=1}^{n} (F_i)$$

where F_i is the total sum of the force involved in some gesture i, and n is the number of gestures in the utterance.

2.4. INTRA- AND INTER-ARTICULATOR COMPARISONS.

From these assumptions, we obtain the intuitive result that, the bigger or faster the gesture, the greater the force thereof (all else being equal), as shown in the following schemata:

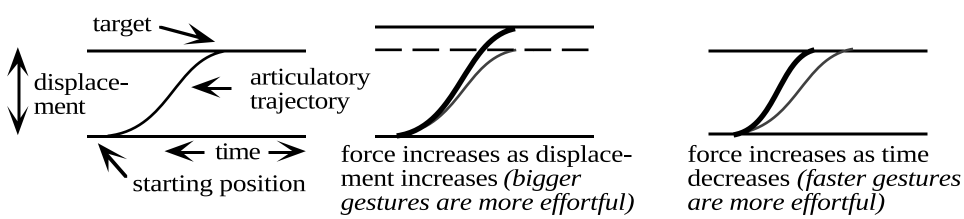


Figure 2-2. Inferences from equation of effort with force.

Velocity equals displacement over time. Therefore, to increase the displacement, while holding time constant, is to increase the velocity; the same is true if the time is decreased while holding displacement constant. Assuming an initial state (t_i) with the articulator at rest (i.e. $v(t_i) = 0$), the greater the velocity of the gesture at time t_j, the greater the acceleration (i.e. $v(t_j)-v(t_i)$) The greater the acceleration, the greater the force (holding mass constant), hence the greater the effort. Therefore, this biomechanical approach permits us to compare the effort required for gestures which involve the same articulators, where the principal variation in the gestures concerns the acceleration (i.e. intra-articulator comparisons).

In principle, we could likewise compare gestures involving different articulators, by varying the masses involved. However, too little is currently known concerning the masses of the articulators involved in particular gestures to permit cross-articulator comparisons with any reasonable confidence. For example, in comparing [k] vs. [t], the effect of greater jaw displacement in the [t] (see Keating et al. 1994) in the effort equation may be offset by the (presumably) greater mass of the active articulator (tongue body vs. tip) in the [k]. In this study of lenition, we are, for the most part, relying upon intra-articulator comparisons involving *manner* of articulation, where the gestures vary in displacement, or where the comparison is between the presence of some gesture and its absence -- this because lenition typically does not result in substitution of gestures involving different articulators, e.g. /s/ - [w], due (under the present approach) to considerations of perceptual faithfulness. Therefore the uncertainty attendant upon of cross-articulator comparison does not typically arise.

2.5. A COMPUTATIONAL MASS-SPRING MODEL

To make the relation between articulator movements (specifically, consonantal constrictions) and their respective effort costs more explicit, I present the following computational model of a mass-spring system, schematized in Figure 2-3a:[17]

[17]This computational model was developed with the generous assistance of Bruce Hayes.

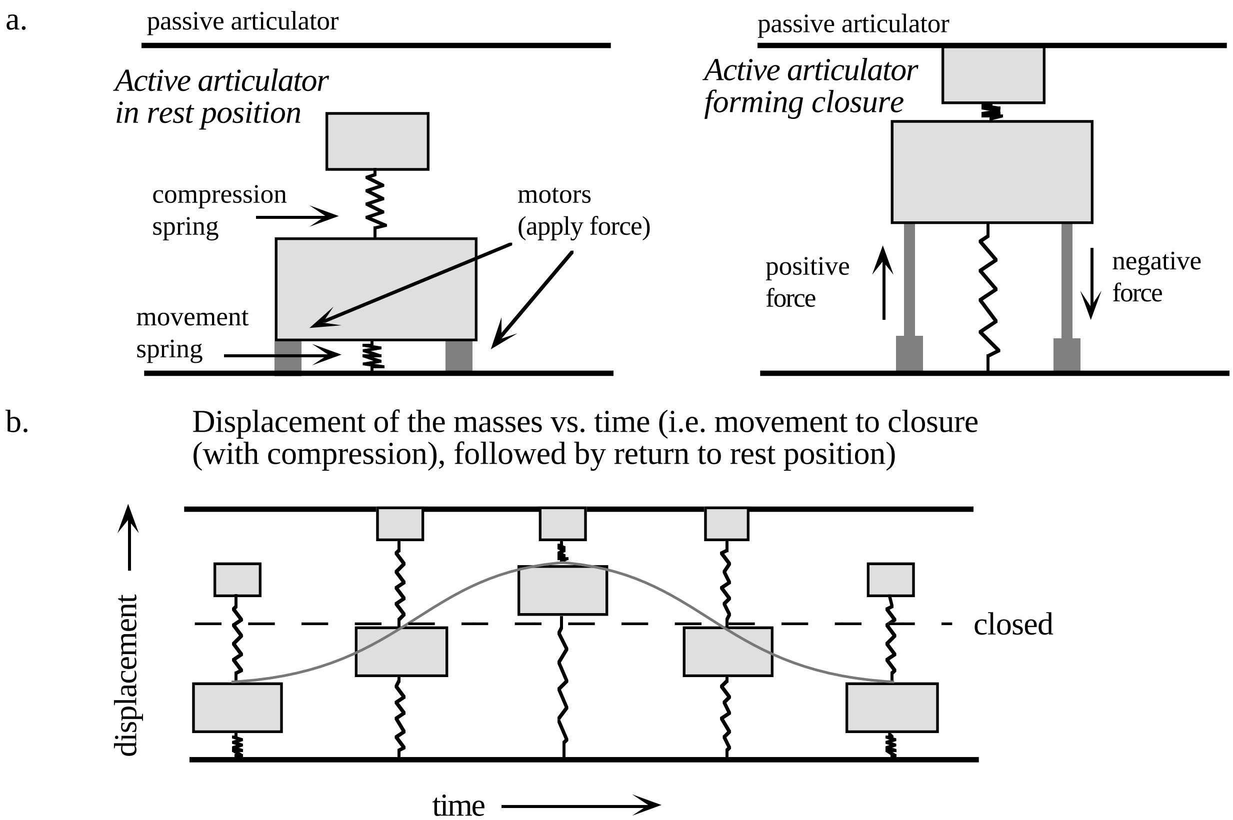


Figure 2-3. Schema of computational mass-spring model of consonant constriction.

That is, an (abstract) articulator is modeled as a spring-loaded mass, which moves upward or downward through time in response to force exerted upon it (Figure 2-3b). When no force is exerted upon it, either by the virtual motors or the retraction of the springs (i.e. "rest position"), the articulator is in a relatively open position. The assumption is that this corresponds to some greater-than-consonantal degree of aperture in the human vocal tract, and that this is the sort of oral constriction degree which occurs when the musculature of the vocal tract is fully relaxed.[18]

[18]Note that this use of mass-spring modeling is different from that of the task dynamics literature (e.g. Saltzman & Kelso 1987, Saltzman & Munhall 1989, Browman & Goldstein 1990). In the task dynamic mass-spring model, the system is at equilibrium when the articulator reaches its target; and there is no attempt to model the forces required to achieve the target; whereas this system is at equilibrium in a relatively open position, and force is required to move the articulator toward its consonantal target.

2.5.1. EQUATIONS IMPLEMENTED IN THIS MODEL. To determine the current position of an object moving along a given path, one needs to know where the object started from, the speed at which it is traveling, and how long it has been traveling. In the mass-spring model, time is divided into discrete, uniform timesteps. Thus, the position s at any timestep t (other than the initial timestep, at which we will assume the position = 0), can be expressed as a function of its position and velocity at the previous timestep:

(2-7) $$s(t) = s(t-1) + v(t-1)$$

Similarly, the velocity of the object at any timestep t (other than the initial timestep, at which, perforce, the velocity = 0), is a function of its velocity and acceleration (i.e. change in velocity) at the previous timestep:

(2-8) $$v(t) = v(t-1) + a(t-1)$$

It remains to determine the acceleration. Force = mass × acceleration: thus in the simplest case, the acceleration at any timestep t equals the force applied to the object at t divided by the mass, m (in the current version of the model, m is set to 1). In a mass-spring system, however, to determine the net acceleration resulting from the application of some force, one must take into account certain reactive forces as well. First, the spring exerts a force tending to pull the object back to its rest position (i.e. 0). The strength of this force is a function of the object's distance from rest position, times a constant, k, for the "stiffness" of the spring (the stiffer the spring, the faster the object snaps back). (The decision to model the articulator as a spring-loaded mass, as opposed to, say, a mass on a simple hinge, reflects the assumption that, in the human vocal tract, the tissues of the articulators, when pushed out of rest position, or when compressed against a passive articulator, exert such a spring-like force, passively drawing the articulator back to rest position.) Moreover, the force applied to an object may be partially dissipated in friction, rather than resulting in pure acceleration: the extent of this "loss" of force can be expressed as a function of the object's velocity times a constant, *b*, for the "damping" of the spring (a damped mass-spring system can be thought of as moving through a viscous medium: the higher the value of *b*, the more viscous the medium). Furthermore, upward movement is, of course, opposed by gravity (as is typically the case in oral constriction formation, assuming an upright orientation of the head).

Putting these observations together, we arrive at the following equation for acceleration at any timestep *t:*

(2-10) $$a(t) = \frac{U(t)}{m} - (g + b \cdot v(t) + k \cdot s(t))$$

where *U(t)* is the force applied at *t*, and *g* is gravity (set to .001).

The actual acceleration equation implemented in the mass-spring model contains a few further elaborations. First, force is broken down into positive force, $U_{pos}(t)$ (i.e. force applied in the direction of greater constriction), and negative force, $U_{neg}(t)$ (applied in the opposite direction). ($U_{pos}(t) \geq 0$, and $U_{neg}(t) \leq 0$.) Secondly, the model also contains a separate spring, with its own damping and stiffness values, for compression of the articulator during closure (the compression spring values are denoted as b_c and k_c, the values for the articulator *movement* spring as b_m and k_m). Thus, when the articulator is in a closed position (i.e. *s(t)* ≥ *h*, the height of the passive articulator), the acceleration equation is as follows:

$$(2\text{-}11)\quad a(t) = \frac{U_{pos}(t) + U_{neg}(t)}{m} - (g + b_m v(t) + k_m s(t) + b_c v(t) + k_c (s(t) - h))$$

(For the compression stiffness term, the relevant displacement is the displacement *above closed position*, i.e. *s(t)-h.*) And for a partial constriction (i.e. *s(t)* < *h*), the equation is simply

$$(2\text{-}12)\quad a(t) = \frac{U_{pos}(t) + U_{neg}(t)}{m} - (g + b_m v(t) + k_m s(t)$$

Or, unifying the two cases,

(2-13)

$$a(t) = \frac{U_{pos}(t) + U_{neg}(t)}{m} - (g + b_m v(t) + k_m s(t) + (b_c v(t) + k_c (s(t) - h) \cdot j))$$

If *s(t)* < *h*, then *j* = 0, else *j* = 1

In the current version of the model, b_m is set to .1, and k_m to .008: these values are chosen so that the articulator returns to rest position within 120 msec of completion of the constriction without any active "opening" (i.e. negative force). b_c and k_c are arbitrarily set to double the values for the movement spring.

Finally, the effort, *E*, involved in a gesture beginning at timestep 1 and ending at timstep *n* is computed by summing over the absolute values of the positive and negative forces for each timestep in the gesture:

(2-14) $$E = \sum_{t=1}^{t=n} \left(|U_{pos}(t)| + |U_{neg}(t)| \right)$$

2.5.2. FURTHER DESCRIPTION OF THE MODEL. The user specifies temporal and spatial targets for articulator movement. The model then finds, by a gradient-ascent learning algorithm (further described below), the function of force against time which achieves these targets with the least possible effort.[19] The output of the model consists of graphs of displacement vs. time and force vs. time, and a value for total effort.

In addition, the force function itself is subject to a constraint: the targets are achieved using at most one positive and one negative (bell-shaped) force “impulse," whose timing, magnitude, and breadth are determined by the learning algorithm. (The terms "positive" and "negative" in this context refer to the direction in which the force is applied: "positive" means in the direction of greater constriction, and "negative" means the opposite.) This restriction on the force function reflects the conjecture that the neuromuscular system is not capable of independently manipulating the precise force applied to an articulator at each instant in time (Figure 2-4a). Instead, we assume a more

---

[19]This learning algorithm simply serves to ensure that the output gesture, and its associated effort cost, represents an optimal gesture: that is, we can be sure that we are comparing, for example, the best possible singleton fricative against the best possible singleton stop. A comparison of the costs of optimal gestures is clearly more useful than a comparison of gestures plotted by hand, the efficiency of which is unestablished. The system is expressly *not* intended to model how speakers might actually compute gestures in speech production.

global control regime, which imposes a smooth contour on the force function (Figure 2-4b):

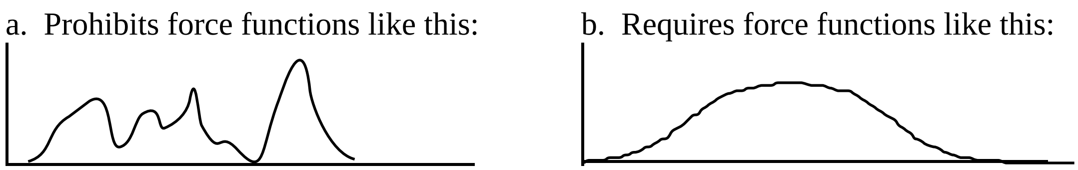


Figure 2-4. “Impulse" restriction on mass-spring model.

The learning algorithm perturbs the magnitude, time, and breadth of the positive and negative force impulses, and then checks to see if the perturbation reduced the error. If so, the new values are adopted; if not, the old ones are retained. Error is defined first as failure to meet the temporal and spatial criteria as defined above; once error in this sense falls below some threshold (set to 1.0 in the current version), error reduction takes the form of minimization of total effort. For all outputs of the mass-spring model presented elsewhere in the dissertation, the gradient-ascent learning algorithm was allowed to run through five million iterations, terminating early when it achieved no further error reduction after two hundred thousand consecutive iterations. Moreover, each simulation was run at least twice, to ensure that the learning algorithm converged on substantially the same result each time.

Admittedly, the human vocal tract is a vastly more complex physical system than what is modeled here, involving aerodynamic forces and fluid mechanical interactions which are not addressed in this simple up-and-down system. This model is not presented as a conclusive answer to issues of articulatory effort quantification, but as a modest first step, intended to shed light on the basic tradeoffs between displacement, intragestural timing, and precision discussed above.

2.5.3. RESULTS OF THE MODEL. Bearing these caveats in mind, the general assumption that reduction of oral constriction, e.g. by spirantization, constitutes a reduction of effort, implicit in the effort-based treatment of spirantization above, is substantiated by the mass-spring model:

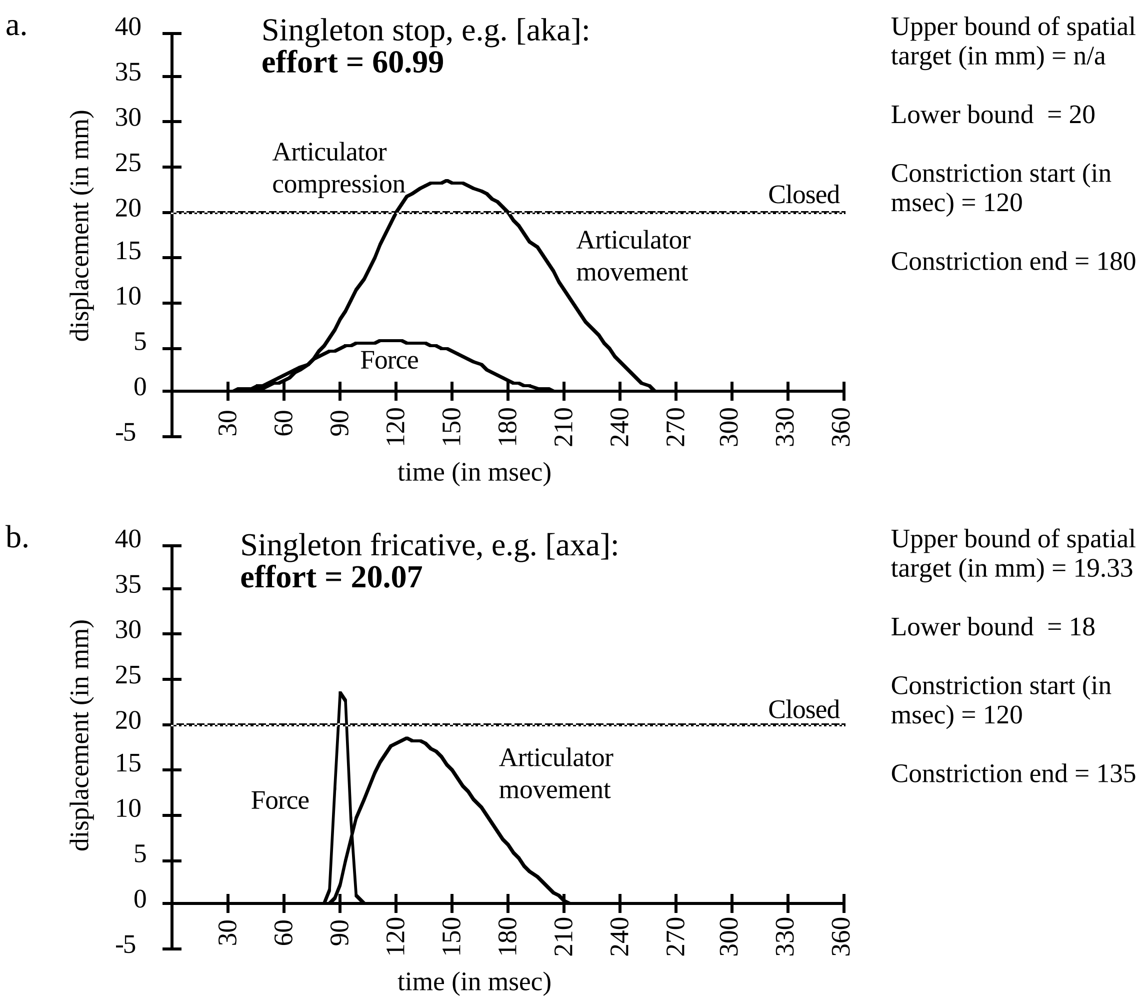


Figure 2-5. Force and displacement vs. time graphs for a singleton stop and fricative.

3. PRECISION

3.1. WHY ARE PRECISE GESTURES MORE DIFFICULT?

The notion of *gestural precision* has played an important role in Stevens' (1972) quantal theory of speech, i.e. that phonological systems tend to include sounds wherein relatively imprecise gestures yield a relatively invariant acoustic result. Precision can be evaluated by comparing the range of variation in some class of similar articulatory gestures to the variation in the acoustic results (see Mrayati et al. 1988).

(2-15) $$Precision = \frac{Acoustic\ variation}{Articulatory\ variation}$$

That is, when a small range of articulatory variation results in a large range of acoustic variation (i.e. the denominator in (2-15) is small, the numerator large), we have a high value for precision. Under these circumstances, in order to achieve a relatively stable acoustic target, only a minuscule amount of articulatory variation is permissible. For example, in order to achieve a bilabial whistle, the lips must form a narrow aperture, with relatively little variation; and the cheek walls, glottis, and lungs must maintain a fairly constant oral air pressure. Without these narrowly defined labial, cheek, glottal, and subglottal conditions, we get extreme acoustic variation from the intended whistle: perhaps silence, a flatulent trill, or fricated puffing. By comparison, a bilabial stop requires less precision: the lips simply have to be thrown against each other with sufficient force to form a seal. With variation in the magnitude of the closure gesture, we get variation in the amount of compression of the lower lip against the upper, and perhaps some variation in the duration of the closure; but modulo this durational variation, the acoustic result is extremely stable, namely the formant transitions, silence, and burst cues associated with a bilabial stop.

I will assume, however, that the difficulty associated with greater precision is ultimately reducible to the notion of effort developed above: presumably, precise gestures require the recruitment of more muscle groups, antagonistically counteracting one another to some extent, in order to control the approach to the target, temporally and spatially; whereas more ballistic gestures are less likely to achieve a narrow target reliably (either due to some inherent random variation in the physical result of some set of gestural commands, or due to context-induced variation). For example, it is frequently suggested that the general markedness of fricatives relative to stops is due to the greater precision required for partial constriction; but this observation can be recast in terms of the additional biomechanical force required to arrest the upward momentum of the articulator (as, for example, in sibilants: the sides of the tongue are stiffened and braced against the molar gumline, resulting in tongue-blade grooving) thus achieving a close constriction without going to full closure. For stops, however, a simple ballistic gesture gives the desired outcome, regardless of variation in the force applied, provided that the gesture is of sufficient force to achieve closure. That is, I understand precision not as a distinct *kind* of articulatory difficulty, but rather as an alternative strategy for getting at effort comparison. In principle, though, I assume that an increase in precision translates to an increase in total neuromuscular activation, and to an increase in the total force involved in the utterance.

The advantage of reducing precision to force is that it permits evaluation of tradeoffs between, say, increased precision and decreased displacement. Such evaluation proves to be crucial to the effort-based account of the non-stridency of spirantization outputs, and geminate resistance to spirantization, presented in Chapters 4 and 5. In contrast, the precision metric says nothing about displacement; evaluating effects of increased precision and decreased displacement using the precision metric amounts to a comparison of apples and oranges. Therefore, despite the usefulness of the precision metric in other areas of phonetic research, for this study of lenition we will persist in evaluating effort, whether due to precision or otherwise, in terms of biomechanical force.

3.2. PRECISION IN THE MASS-SPRING MODEL.

The computational mass-spring system, described in section 2.5. above, models directly *vertical* constriction and opening movement of an abstract active articulator. Consequently, this system cannot compare gestures involving different constriction location targets (e.g. velar vs. palatal place of articulation), which principally concern the horizontal dimension; nor can it reflect the greater precision involved in achieving a narrowly defined constriction location. Nevertheless, the model does reflect a kind of temporal precision, namely the precision involved in maintaining a partial constriction for an extended duration (crucial to our analyses of lenition patterns in Chapters 4 and 5). Recall that in this model, the targets must be achieved using at most one positive and one negative (bell-shaped) force "impulse." Ballistic gestures, such as the stop and the (very short) fricative in Figure 2-5, can be achieved using only a positive force impulse. The articulator is thrown upwards to form a full or partial constriction; but when the articulator reaches the peak of its (roughly parabolic) displacement curve, it immediately begins to return to rest position, by the passive forces of gravity and the contraction of the (virtual) springs. However, a more temporally controlled constriction, such as a long (geminate) fricative, cannot be so achieved under this model. Rather, such a gesture requires a negative force impulse, counteracting the positive force, to maintain the extended partial constriction while preventing the articulator from achieving full closure:

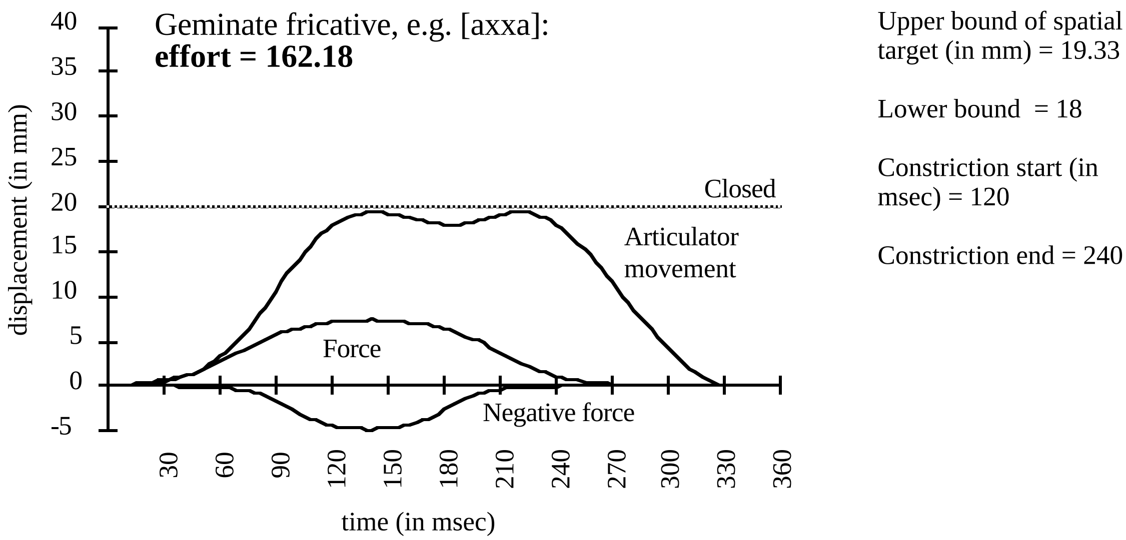


Figure 2-6. Force and displacement vs. time graph for a geminate fricative.

The higher effort cost required for this more temporally controlled constriction is attributable to this additional, negative force impulse, which counts as part of the total effort of the gesture:

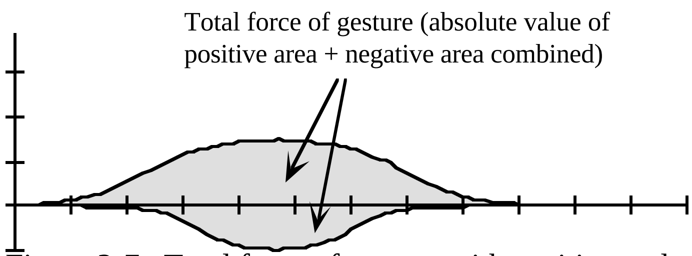


Figure 2-7. Total force of gesture with positive and negative force impulses.

4. Effort minimization and voicing lenition

4.1 Medial position

The discussion thus far, and the mass-spring model in particular, have focused upon the effort associated with consonantal constriction gestures. The connection between effort reduction and *voicing* lenition therefore requires some discussion.

Westbury & Keating 1986 demonstrate that, in utterance-medial position when preceded by a voiced sonorant, stops of normal duration (typically 50-80 msec) undergo *passive voicing*, unless they are devoiced by active abduction (or constriction) of the glottis, assuming an adducted rest position of the glottis. Voicing is able to occur because, under these conditions, oral air pressure is sufficiently lower than glottal air pressure (a difference of roughly 2,000 dyne/cm^2) to sustain vocal fold vibration. Thus, medial stop voicing processes can therefore be understood as a species of lenition, consistent with the effort-based approach proposed here, in that they plausibly involve the elimination of a glottal abduction (or constriction) gesture.[20]

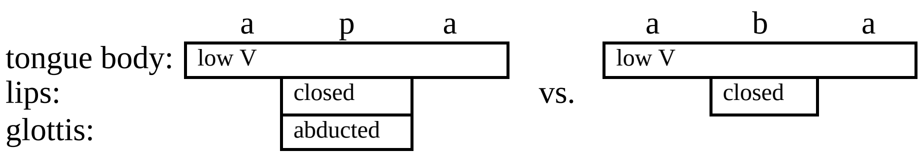


Figure 2-8. Gestural score for voiceless vs. voiced medial stops.

Voicing in this context thus affords an effort savings, because the devoicing (glottal abduction) gesture necessarily involves more effort than no gesture at all. Moreover, my own simulations, using the computational analog circuit model of vocal tract aerodynamics described in Westbury & Keating 1986, show that medial singleton fricatives likewise undergo passive voicing (see Chapter 5 section 3.1.4.2).

Moreover, in many languages, the constriction gesture in voiceless obstruents is of greater magnitude than in voiced obstruents , with a corresponding longer (acoustic)

[20]This analysis presumes that the "voicing" contrast in question indeed involves actual phonetic voicing, i.e. vocal fold vibration. It further presumes that the outputs of voicing lenition are not implemented with additional gestures, such as larynx lowering. Larynx lowering does frequently accompany voiced obstruents in initial position, being one of several strategies for expanding the oral cavity, hence facilitating the high transglottal pressure differential needed to initiate voicing; but Westbury & Keating's study suggests that such cavity expansion gestures are superflous in singleton medial stops.

closure duration (or interval of close constriction in fricatives), partially supporting the traditional notion that voiceless obstruents are "fortis," i.e. involving greater muscular force, and voiced obstruents "lenis" (see generally Ladefoged & Maddieson 1996, chapter 3), or [±tense], in the feature system of Chomsky & Halle 1968. Presumably, the more fortis gesture facilitates the maintenance of an occlusion (or partial constriction, in the case of fricatives) notwithstanding the greater oral pressure behind the constriction in voiceless obstruents. Moreover, longer constriction (i.e. beyond 50-80 msec, the point at which passive devoicing occurs) itself promotes devoicing, for the reasons discussed in the previous paragraph. Finally, to the extent that fortis constriction results in greater overall tensing of the vocal tract walls, this too promotes devoicing, since it reduces the capacity of the oral cavity to expand passively (see Rothenberg 1969). Thus, for languages in which fortis closure is a concomitant of voicelessness, the greater effort cost of the fortis gesture would provide an additional impetus for voicing lenition.

Thus, we can identify two phonetic scenarios under which medial voicing processes can plausibly be viewed as achieving a net reduction of effort cost (in accordance with the traditional classification of medial voicing as a species of lenition): passive voicing, by elimination of a glottal abduction gesture; and reduction to a briefer oral constriction. Nor are these two scenarios mutually inconsistent: plausibly, in many cases, the voiced candidate's lower effort cost, relative to its voiceless counterpart, is attributable both to the absence of glottal abduction, *and* a more lenis oral constriction.

4.2. Final and initial positions.

In utterance-final position, however, expiratory force, hence subglottal pressure, tends to decline sharply, as the respiratory system readies itself for post-utterance breathing. This drop-off typically brings subglottal pressure lower than the 2000 dyne/cm^2 differential needed to sustain voicing, typically resulting in passive devoicing within 35 msec of closure in a final stop (Westbury & Keating 1986). This devoicing can be avoided by deploying a variety of voicing-enabling gestures, for example intercostal contraction (raising subglottal pressure) or (more commonly) various oral cavity expansion gestures, such as larynx lowering and pharynx expansion (lowering oral pressure); but these voicing-enabling gestures carry some additional effort cost, just as active devoicing (through glottal abduction) does in medial position. Moreover, in an utterance-initial stop (in the absence of voicing-enabling gestures), subglottal and supraglottal pressure are roughly equal, and remain so until the stop is released; consequently the transglottal pressure differential required to initiate voicing (4000 dyne/cm^2) is frequently absent, resulting in passive devoicing of utterance-initial stops (again, in the absence of voicing-enabling gestures).

These observations suggest a solution to a problem raised in Chapter 1, section 2.7, in the discussion of John Harris' (1990) notion of voicing lenition as feature loss. Harris' idea is essentially correct: medial voicing can be characterized as loss of a glottal abduction component (in Harris' representational scheme, a h˚ element). Harris' problem is the assumption that voicelessness involves an active glottal abduction element *in all contexts*. Rather, voicing is the default result of vocal tract aerodynamics in medial position, while devoicing is the default result in final and initial position. We can thus reconcile the naturalness of final and initial devoicing processes with medial voicing processes. (See Chapter 7 for a formal effort-based analysis of final and initial devoicing in Tümpisa Shoshone.)

5. SUMMARY

Notwithstanding the skepticism of the Generative tradition toward functionalist principles such as effort minimization, it is possible to give an explicit characterization of effort in terms of the mechanical notion of *force*. Specifically, effort is characterized as a summation, over the gestures involved in an utterance, of the force involved in each gesture. This characterization permits intra-articulator (though not, as a practical matter, inter-articulator) comparison of gestures with respect to effort. This characterization, moreover, encompasses effects of *precision* on effort. Finally, the role of articulatory effort in grammar is formally expressed as a violable constraint, LAZY, which militates in favor of effort minimization.

Chapter 3:

Representational Issues

In this chapter, I demonstrate that it is possible to enrich phonological representations with gradient phonetic detail, such as effort cost, while maintaining a formal distinction between the contrastive vs. predictable (or freely varying) behavior of particular features, as well as their categorical or gradient variation, within particular sound systems.[21] This formal point is something of an excursus from the problems of lenition typology which are the focus of this dissertation; nevertheless, this issue is crucial to the general program of capturing substantive phonetic explanation within the formalism of phonological theory, of which the effort-based approach to lenition is an example.

A fundamental observation of phonological theory is that, out of the rich sound signal of speech, a small subset of phonetic properties is contrastive in any given language. Standardly, this observation is captured by excluding the non-distinctive phonetic properties from some levels of representation, yielding an abstract categorical representation of the contrasts among the speech sounds. Universally non-contrastive features have generally been assumed to be unspecified throughout the phonological component; thus, for example, Keating 1984 and Lombardi 1991 have argued against inclusion of particular laryngeal features in phonological representation on the grounds that they are never contrastive. Language-particular non-contrastive features have generally been assumed to be unspecified only in underlying representation and early

[21] An earlier version of this chapter has appeared as Kirchner 1997. Further note that many of the conclusions reached in this Chapter have been arrived at independently by Boersma (1997a,c).

stages of the derivation (e.g. Kiparsky 1982; Archangeli 1984, 1986; Archangeli and Pulleyblank 1986, 1994; Steriade 1987; Clements 1988). Moreover, since the discovery of extensive language-particular gradient alternations,[22] it has standardly been assumed that a subsequent phonetic component is needed, to fill in gradient and universally non-contrastive properties (e.g. Pierrehumbert 1980, Keating 1990; but see Pierrehumbert 1994 for an attack on this view of the phonetics-phonology "interface"). This model is schematized in Figure 3-1.

Underlying Representation	*(ideally) pure representation of contrast*
⇓	
Phonological component	*non-contrastive properties may be filled in, particularly if contrastive in other languages*
⇓	
Phonetic component	*remaining non-contrastive phonetic properties, including gradient values, filled in*
⇓	
Phonetic Representation	*representation of all speaker-controlled phonetic properties of the utterance*

Figure 3-1. The standard (representational) treatment of contrastiveness.

The crucial property of this model is its representational characterization of contrastiveness: an abstract representation of pure contrast must be assumed, i.e. underlying representation, with derivations from underlying to surface phonological representation (and perhaps extensive intermediate levels of representation), and from surface phonological representation to the sea of surface phonetic detail. As Steriade (1995b) has observed, the assumptions of this model have often been disregarded in practice. For example, the value of the feature [sonorant] is predictable for the classes of [+nasal], [-nasal, -continuant], and [-consonantal] segments; yet the feature defines such

---

[22]This includes alternations once thought to be categorical, e.g. final devoicing in German (Port and O'Dell 1985); and place assimilation in casual speech in English (Barry 1985, 1991; Nolan 1992).

an important natural class that phonologists have generally treated it as being present from the earliest stages of the derivation, even in these classes of segments. Similarly, syllable structure is not contrastive per se in any language, and so according to this reasoning ought to be consigned to the phonetic component; yet syllabification is generally considered to be the driving force behind much of phonology. However, I will not dwell on these inconsistencies, and whether they might be reconciled with the model's principal assumptions; rather, I argue that the principal assumptions are superfluous to an adequate treatment of contrastiveness, and therefore should be abandoned.

Further note the all-or-nothing characterization of contrastiveness in this model; whereas, what is a required is a more scalar expression of the potential contrastiveness of a particular feature (cf. Goldsmith's (1995: 9-13) observations along these lines). If a phonetic property is admitted to the pantheon of phonological features, it is formally equal to any other feature in its potential for signaling contrasts. Similarly, if a feature is contrastive in a given language, it must be present underlyingly, and there is no representational distinction between features which are contrastive in a broad array of contexts and contrasts which surface only in a particular environment. One may, of course, resort to context-sensitive underspecification (e.g. 'mid vowels are unspecified for [round] in unstressed syllables'), but if there is an independent rule [-high] - [-round] / $[C_0\_\_C_0]_{\sigma(\text{unstressed})}$, or an equivalent constraint, the assumption of underspecification is superfluous (leaving aside the feasibility of having underspecification conditioned by derived properties such as stress). A more accurate observation is that features fall along a continuum of potential contrastiveness. Some distinctions, such as vowel height, are contrastive in every language; others are contrastive in just a single language, e.g.

longitudinal vocal fold tension in Musey voiceless obstruents (Shryock 1995).[23] And of course many phonetic properties, such as the distinction between a 53 msec and 54 msec stop closure, are never contrastive. Similarly, while virtually every feature is subject to some distributional restriction, some features are typically contrastive in extremely narrow contexts. For example, contrastive consonant length is typically restricted to intervocalic position. In the standard approach, constraints or rules may be invoked to neutralize contrasts in particular environments; but if so, we now have two theoretical devices for explaining the absence of contrast: (a) representational restrictions (in UR or throughout the phonology), and (b) rules or constraints requiring the neutralization of a contrast. One may reasonably inquire whether the latter device is sufficient. Moreover, it is typically the contrasts which are banned outright in many languages which are severely restricted in their distribution in the languages that permit them. For example, Kaun (1994) observes that most languages do not permit a contrast in [round] independent of [back]; but in those languages that do, contrastive rounding is typically subject to vowel harmony, i.e. surface restrictions on the vowels that can occur with it (within some domain). The standard model fails to draw this connection.

I propose an alternative treatment, whereby the notion of contrastiveness emerges from Optimality Theoretic constraint interaction, specifically with regard to the notion of *faithfulness* (minimal divergence of output from input), without resorting to the representational and derivational assumptions of the standard model. This move permits phonological representations to include as much phonetic detail as may be necessary to

---

[23]See generally Ladefoged and Maddieson 1996, for a surprisingly large number of features which are contrastive in only one or two languages. The list of potentially contrastive features which they document, incidentally, is considerably larger than the feature inventory standardly contemplated by phonologists. Therefore, even if we take potential contrastiveness as a necessary condition for inclusion, a large expansion of the feature inventory beyond the standard 15 or so features is clearly in order.

adequately characterize a speaker's phonetic competence. A significant derivational residuum, namely the post-phonological phonetic component, is thereby eliminated from the theory, as is shown in Figure 3-2.

Underlying Representation	*no restrictions on this level of representation*
⇓	
Phonological component: Gen + H-eval	*contrastive, categorical behavior of particular features falls out from constraint system*
⇓	
Phonetic Representation	*representation of all speaker-controlled phonetic properties of the utterance*

Figure 3-2. A constraint-ranking treatment of contrastiveness.

Furthermore, this model captures the observations noted above: a more scalar notion of the potential contrastiveness of a given feature, and the connection between cross-linguistic markedness and restricted distribution of contrasts.

More importantly, this move permits direct reference to phonetic properties which, though never contrastive per se, may nevertheless play a role in conditioning phonological processes; we can thereby directly capture phonetic explanations for phonological generalizations which elude formal capture using standard, more abstract representations.

In principle, then, the phonological representation may (and, I will assume, does) contain all speaker-controlled articulatory properties, and all sound properties which the auditory system is capable of detecting: that is, an auditory representation in parallel with an articulatory representation, as in Flemming 1995. I further follow Flemming in representing the auditory properties as a sequence of matrices of auditory features, where the matrices may correspond to subsegmental units, such as stop closure, burst, and formant transitions.[24] The articulatory representation may be conceived of as a gestural score (e.g. Browman and Goldstein 1989, 1990).

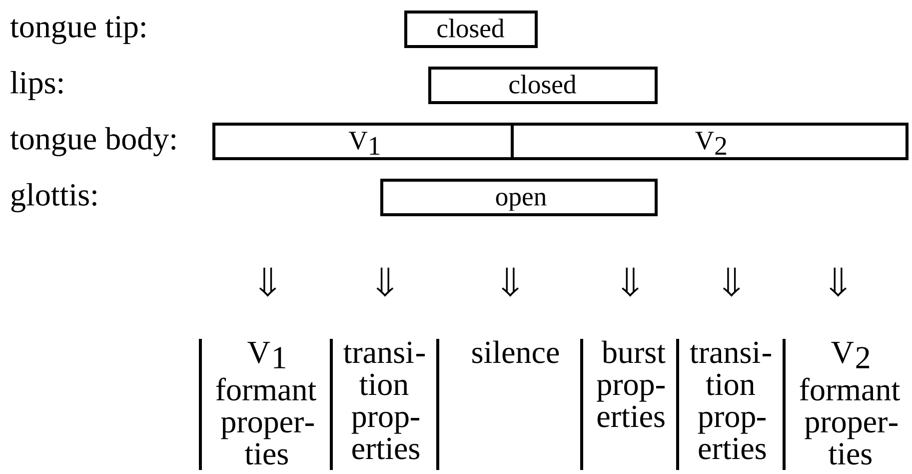


Figure 3-3. Relation between auditory and articulatory representations in this approach.

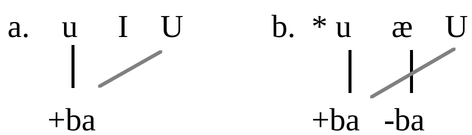
The mapping between the articulatory representation and auditory representation is, of course, determined by the laws of physics, and is not subject to cross-linguistic variation (see Flemming 1995). For reasons of familiarity, however, I will use traditional featural representations below, except in cases where phonetically richer representations are crucial.

[24]Alternatively, the matrices could represent fixed time slices, provided that the sampling rate is high enough to detect any changes over time in the auditory signal which speakers can actually perceive. This is an empirical question.

1. SMOLENSKY'S CHALLENGE TO THE STANDARD MODEL

Certain aspects of the standard model, particularly the assumptions surrounding underspecification theory, have been challenged since the development of Optimality Theory (Smolensky 1993, Inkelas 1994, Itô, Mester, and Padgett 1995, Kirchner 1995, Steriade 1995, Boersma 1997c). In particular, Smolensky has shown that the phonological "inactivity" of predictable features may be attributed to rankings of a particular class of constraints, rather than to the absence of such features from the representation at some stage of a derivation. For example, consider the standard underspecification-based treatment of the 'neutral' vowels [i,e] in Finnish backness harmony (Kiparsky 1981):

(3-1)



I = -low, -round, 0back vowel;
U = -low, +round, 0back vowel

That is, the non-low unrounded vowel in (3-1a) is unspecified for [back] at the point in the derivation where the spreading rule applies, and therefore is transparent to spreading of [back] onto the following vowel; whereas the [back] specification of non-neutral vowels such as [æ] (3-1b) makes them opaque. (The failure of [+back] to spread *onto* the neutral vowel is attributed to Structure Preservation, namely the ill-formedness of non-low back unrounded vowels in Finnish.) Assume, however, the Optimality Theoretic notion of faithfulness (minimal departure from identity of input and output), formalized, in part, in terms of feature-specific PARSE constraints (cf. Kirchner 1993, Orgun 1995, McCarthy and Prince 1995). Now the transparency of the Finnish neutral vowels can be analyzed in terms of the constraint hierarchy in (3-2) (modified slightly from Smolensky's presentation).

(3-2)	/u-I-U/	*[-low, -round,+back]	ALIGN(+back, right)	PARSE (back)	*[-low,+round, -back]	*EMBED
a. ☞	u-i-u					*
	u-ɨ-u	*!				
	u-i-ü		*!		*	
	u-ɨ-ü	*!	*		*	
	ü-i-ü			*!	**	
b.	/u-i-u/					
☞	u-i-u					*
	u-ɨ-u	*!		*		
	u-i-ü		*!	*	*	
	u-ɨ-ü	*!	*	**	*	
	ü-i-ü			*!*	**	
c.	/u-ɨ-ü/					
☞	u-i-u			**		*
	u-ɨ-u	*!		*		
	u-i-ü		*!	*	*	
	u-ɨ-ü	*!	*		*	
	ü-i-ü			**	*!*	

ALIGN(+back,right): requires that a [+back] specification be linked to a segment at the right edge of the word (thereby enforcing rightward spreading). *EMBED prohibits embedding of a [-back] domain inside a [+back] domain (i.e. in autosegmental terms, spreading which results in a line-crossing or gapped configuration). The correct outcome is obtained whether we assume an input in which the target vowels are unspecified w.r.t. [back] (3-2a), a fully specified input (b), or even an input whose specifications are *contrary* to the surface values (c). Since PARSE(back) is ranked below ALIGN(+back, right), it is better to spread [+back] than to preserve underlying values; and since *[-low,-round,+back] is ranked above PARSE(back), back unrounded high or mid vowels are ruled out in all contexts (even if this results in a *EMBED violation). That is, the formal expression of the unmarkedness of front unround vowels is shifted from the representations (i.e. the characterization of such sounds in terms of fewer underlying features) to the constraints, which directly state the preference for front unround and back round vowels. Indeed, this point is stated explicitly by Itô, Mester & Padgett (1995: 592): "[T]here is no need for a separate theory of feature minimization: the constraint hierarchy itself forces the correct output, irrespective of the specification of the input." The essential observation here is that low-ranking of a constraint on faithfulness to a particular feature results in phonologically inert behavior of that feature within the sound system. Consequently, restrictions on the presence of particular features, at underlying or intermediate levels of representation, are unmotivated in the OT framework.

2. A DEFINITION OF CONTRASTIVENESS

2.1. CONTRASTIVENESS OF PHONETIC REPRESENTATIONS

At this point, it is necessary to consider afresh precisely what we mean by 'contrastiveness.' Specifically, I wish to develop a characterization of contrastiveness which is independent of assumptions concerning the systematic presence or absence of particular features or feature values in underlying representation. Pretheoretically, speakers have intuitions that two distinct surface forms (traditionally 'phonetic representations', or 'PRs') may or may not count as 'significantly different,' depending on the phonological system, whether or not they are actual forms of the language. For example, [splɪk] and [splɛk] would presumably be considered 'different words' by most English speakers, whereas [splɪk] and [splɪk˺] (unreleased [k]) would not, though these are all nonce forms. This notion of significant difference, or contrastiveness, appears to require a level of representation distinct from the surface, traditionally 'underlying representation' or 'UR'; and the minimal theoretical assumption required to capture these intuitions of contrastiveness is that two contrastive PRs correspond to distinct URs. However, as illustrated in the previous section, there may be several possible URs for a given surface form under a grammar, and therefore it is necessary to speak of distinct *sets* of URs for some pair of PRs:

(3-3) **Dfn. Contrastiveness of PRs:** Two distinct PRs p and p' are contrastive under grammar G iff the set of URs for p under G is not identical to the set of URs for p' under G.

Thus, p and p' are contrastive in Figure 3-4a and b below, but not in 3-4c or d:

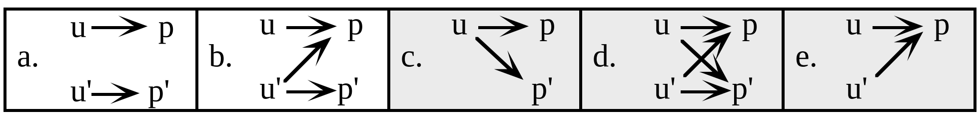


Figure 3-4. Constrastive and non-contrastive PRs.

In Figure 3-4e there is no contrast, trivially, since there is no distinct pair of PRs: this corresponds to obligatory neutralization to some surface value (e.g. final laryngeal neutralization in Korean, where a form such as [kap] could in principle derive from /kap/, /kap^h/ or /kap'/). Case 3-4b corresponds to an instance of optional neutralization, as in the [w]~[ʍ] distinction in certain dialects of English, where, e.g., <u>witch</u> is invariantly [wɪʧ], but <u>which</u> [ʍɪʧ] freely varies with [wɪʧ]: contrastiveness obtains, even though p and p' have a common UR (u'), since the *sets* of URs for each PR are not identical.

2.2. CONTRASTIVENESS OF FEATURES.

However, we want to be able to characterize the contrastive behavior of *features* within sound systems, as well as being able to identify contrastive pairs of forms. Of course, phonological representations consist not merely of features, but of temporal structure as well, which can be expressed in terms of relations between features, or between segments, which are themselves sets of features (cf. Bird & Klein 1992). For example, the relation precedes(dors,+strid) (or precedes(k,s)) contrastively distinguishes [tʰæks] ('tax' or 'tacks') from [tʰæsk] ('task'). Moreover, we presumably need to refer to the simulaneity (or overlap) of particular features within a segment (or some subsegmental temporal interval, as in Figure 3-3), i.e. the relation simultaneous(-voi,+strid) in an [s]; and the relation of shared affiliation to some prosodic constituent (e.g. tautosyllabic(-voi,+cont) (or tautosyllabic(t,ɹ)) is one of the things distinguishing nitrate from night rate). In this Chapter, however, I will use the term "feature" in an extremely broad sense, to encompass any property *or relation* contained in phonological representations, including the sorts of temporal relations discussed above, as well as the conventional sorts of phonological features.

Naively, a feature is contrastive just in case it is sufficient to distinguish a pair of contrastive PRs (analogous to the 'minimal pair' criterion of Structuralist phonemics).

2.2.1. MUTUALLY PREDICTABLE FEATURES. A complicating case, however, is presented when two or more features are mutually predictable. For example, if all sonorants in a sound system are voiced and all voiced sounds are sonorant, there will be no output pair which differs solely with respect to [sonorant], nor solely w.r.t. [voice]. By the naive definition, *neither* feature is contrastive; whereas we wish to say that one of the features is contrastive,

(3-4)	u ([-son,-voi])	-	p ([-son,-voi])
	u' ([+son,-voi])	-	p' ([+son,+voi])

namely the one which corresponds to an underlying distinction (i.e. in the feature [son]), and the other feature ([voi]) is predictable from it. Note that, under this treatment, the contrastive feature in (3-4) cannot be determined solely by inspection of the surface forms. Rather, identification of the contrastive feature is relative to the grammar in question: in particular, to the input-output mappings permitted by that grammar. I leave aside the distinct question of whether we can non-arbitrarily decide between alternative grammars which generate the same surface forms. As a practical matter, it may be necessary in such cases to speak of contrastive feature *sets*: that is, the set {[son],[voi]} is contrastive, without forcing a choice between [son] or [voi]. I leave aside the distinct question of whether we can non-arbitrarily decide between alternative grammars which generate the same surface forms. As a practical matter, it may be necessary in such cases to speak of contrastive feature *sets*: that is, the set {[son],[voi]} is contrastive, without forcing a choice between [son] or [voi].

2.2.2. Displaced contrasts. A further problem is posed by displaced contrasts. For example, in certain dialects of Basque, /e/ raises to [i], and /i/ raises to [i̝][25] when followed by another vowel (e.g. [seme] ('son'), [semie] ('the son') vs. [eri] ('village'), [eri̝e] ('the village') (Hualde 1991). That is, an underlying distinction in [high] is neutralized under hiatus; but there is a surface distinction in [+super-high].

(3-5) u ([+high,?super-high]) - p ([+high,+super-high])
u ([-high,?super-high]) - p ([+high,-super-high])

[25]Hualde transcribes this as [iy]; phonetically, the vowel in [iy] is presumably higher than plain [i], due to coarticulation with the glide. I assume that it is the extreme closeness of the vowel which gives rise to the stronger percept of a glide in the transition to the following vowel, rather than vice-versa.

We should not, however, treat [super-high] as a contrastive feature in Basque, although we can find surface minimal pairs, since it is predictable from the underlying value of [high]. On the other hand, we should not treat [high] as being contrastive *in this context*, since there is no surface [high] distinction. However, all cases of displaced contrast involve features which are directly contrastive in some other context in the sound system.[26] In Basque for example, [high] *is* contrastive by virtue of its behavior in non-hiatus contexts, where it surfaces directly; therefore the URs in (3-5) are underlyingly distinguished by a contrastive feature. Indeed, it is precisely the contrastiveness of [high] in pairs such as [seme] and [eri] that motivates the identification of [high] as the underlying distinction in the morphologically derived pair [semie] and [eri̩e]. More generally, we need not say that the contrastiveness of a feature F is established by F's behavior in contexts where an underlying F distinction is displaced: rather, if F is contrastive, this is so because of its behavior in other contexts. By excluding displaced contrasts from the definition, we are able to provide a formal characterization of contrastiveness in terms of constraint interaction, while keeping this task distinct from the non-trivial problem of opacity (i.e. the interactions which give rise to displaced contrasts) in non-serial Optimality Theory.

2.2.3. Free variation. To handle the complicating cases, then, we must require that a surface F distinction between p and p' corresponds to a minimal underlying F distinction w.r.t. u and u'. This condition brings us close to an adequate definition, but fails under a particular circumstance, namely neutralizing free variation, shown schematically in Figure 3-5:

---

[26]Modulo, of course, analyses using displaced underlying distinctions purely as diacritics marking exceptional forms: Kiparsky (1973) argues persuasively against such use of underlying features.

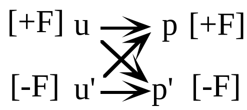


Figure 3-5. Free variation.

In this case, if we pick {<u,p>,<u',p'>} as the crucial pair of input/output mappings, we wrongly claim that F is contrastive, when plainly it is not. To exclude this scenario, we must attach a further condition, namely that one of the URs cannot map to one of the PRs.

2.2.4. The definition. We can now define contrastiveness of features as follows:

(3-6) **Dfn. Contrastiveness of a feature:** A feature F is contrastive under grammar G iff there is some pair of UR-to-PR mappings {<u,p>, <u',p'>} under G such that

(a) u and u' differ solely w.r.t. some specification for F, and

(b) p and p' differ w.r.t. a corresponding specification for F, *and*

(c) u' cannot map to p under G.

(Recall that the URs and PRs under discussion are possible, not necessarily actual forms.) I claim, then, that (3-6) corresponds to an adequate definition of contrastiveness of features. The fact that p and p' are contrastive PRs follows from condition (3-6c): u' can not map to p, therefore p and p' have distinct sets of URs, and so by definition (3-3) they are contrastive. To review the complicating cases:

(a) The treatment of mutually predictable features depends on the input/output mappings permitted by a particular grammar, thus the feature [sonorant] in (3-4) meets the definition (while [voice] does not), since the surface [αson, αvoi] distinction corresponds to an underlying distinction in [sonorant] alone.

(b) In cases of displaced contrasts, neither the underlying nor the surface distinction is treated as contrastive by virtue of its behavior in the displaced context, thus neither [super-high] nor [high] in (3-5) meet this definition, since the surface distinction corresponds to an underlying distinction in a different feature (though [high] is contrastive by virtue of its behavior in other contexts).

3. Faithfulness

Before demonstrating the connection between faithfulness to a particular feature and contrastiveness of that feature, we must make explicit our notion of featural faithfulness. McCarthy and Prince (1995) and Orgun (1995) present a generalized notion of faithfulness, defined in terms of correspondence between segments in an input and output, base and reduplicant, etc. That is, a class of faithfulness constraints requires each segment in an input to have a corresponding segment (indicated by coindexation) in the output. A further class of faithfulness constraints requires input and output correspondents to be identical w.r.t. specific features. I adopt this general conception of faithfulness here. However, the reference to segments is not crucial: we could instead speak of direct correspondence between the features (since a segment, if this notion is indeed motivated as an element of phonological representation, can be expressed simply as a set of features). Since we are concerned with the general behavior of features (broadly defined) here, this move simplifies the discussion considerably. Assuming a monostratal grammar, 'Input/Output' is equivalent to 'UR/PR', and we can define faithfulness between a UR and PR w.r.t. some feature F as follows:

(3-7) PRES(ERVE)(F,Input/Output): For all $\alpha \in \{+,-\}$, for each αF specification in the input there is exactly one corresponding αF specification in the output, and for each αF specification in the output there is exactly one corresponding αF specification in the input.

That is, all and only the following mappings w.r.t. F satisfy PRES(F):

(3-8) $+F_i$ - $+F_i$ $-F_i$ - $-F_i$ 0F - 0F

Privative F can be equated with +F above, assuming that no -F is possible. Given our broad use of "feature," the class of PRES(F,I/O) constraints therefore subsumes all notions of input/output faithfulness, whether featural, segmental, or prosodic.[27] I further assume that no constraint other than PRES(F,I/O) refers to the value of F in UR. As Orgun (1995) notes (extending an observation of Lakoff 1993), a parallel Optimality Theoretic model which permits context-sensitive constraints to refer to underlying properties has all the (presumably excessive) descriptive power of a serial framework, like *SPE*, with an unlimited number of intermediate levels of representation.

---

[27]Prince and Smolensky's (1993) arguments for independent ranking of segmental PARSE and FILL constraints concerned the distinction between languages that eliminate consonant clusters by inserting a vowel vs. deleting a consonant, and between languages that eliminate vowel sequences by deleting a vowel vs. inserting a consonant. These distinctions can be reanalyzed in terms of independent ranking of faithfulness constraints referring to vowel quality features vs. consonant manner or place features. Indeed, Prince and Smolensky (ch. 7) are forced to posit such a distinction in any case (in their terms, PARSE(C) vs. PARSE(V), and FILL(Ons) vs. FILL(Nuc)).

A final point concerns the treatment of phonological exchange, i.e. αF - -αF. Such mappings meet the definition of contrastive features (3-6), since the definition merely requires a surface distinction between p and p' corresponding to a underlying minimal distinction between u and u', even if the surface and underlying values are flipped. However, since we admit of no input/output 'anti-faithfulness' constraints w.r.t. F, such mappings cannot arise under any grammar. This prediction seems largely correct (see Janda 1987). Nevertheless, there are a few cases of morphologically conditioned exchanges, e.g. Dinka plural formation (Kenstowicz and Kisseberth 1979): if the base (singular) is [αlong], the derived (plural) form is [-αlong]. Such alternations can be handled in terms of morphologically conditioned constraints which require some feature in the surface form of the base to map to the opposite value in the morphologically derived form (i.e. morphologically conditioned output/output 'anti-faithfulness'), without referring to the value of the feature in UR. (Henceforth, I use PRES(F) as shorthand for PRES(F,I/O), since we will not discuss output/output faithfulness (or anti-faithfulness) constraints further in this Chapter.)

4. A LANGUAGE-SPECIFIC PREDICTABLE PROPERTY: ASPIRATION IN ENGLISH

To illustrate the connection between PRES(F) and the contrastiveness of F in a sound system, let us consider aspiration of stops in English. As a descriptive encapsulation, I posit the constraint in (3-9):[28]

---

[28]Phonetic studies (e.g. Pierrehumbert and Talkin 1992) indicate that the degree of aspiration in English actually varies gradiently, depending on the stress level and phrasal position of the relevant syllable. To handle this gradient variation, a more sophisticated, gradient version of the ASPIRATE constraint is required (see section 7 on gradient properties in the phonological representation); nevertheless, such elaboration does not change the essence of the present analysis, namely conflict between faithfulness and some constraint on the surface distribution of aspiration.

(3-9) ASPIRATE: A stop is [+spread glottis] iff it is [-voi], occurring in initial position in a stressed or word-initial syllable.

The English pattern is obtained under the ranking in the tableaux in (3-10):

(3-10)

			ASPIRATE	PRES(spread)
a.		pɪl - ˈpɪl	*!	
	☞	pɪl - ˈpʰɪl		*
b.		pʰɪl - ˈpɪl	*!	*
	☞	pʰɪl - ˈpʰɪl		
c.	☞	spɪl - sˈpɪl		
		spɪl - sˈpʰɪl	*!	*
d.	☞	spʰɪl - ˈspɪl		*
		spʰɪl - ˈspʰɪl	*!	

I assume that PRES(voi), the stress assignment constraints, etc. are all ranked above PRES(spread), therefore candidates [ˈbɪl], [pɪl] (unstressed), etc. are ruled out. Tableaux (3-10a) and (b) show that, regardless of underlying specification for [spread], a voiceless stop in initial position within a stressed syllable is aspirated on the surface. Tableaux (3-10c) and (d) show that, regardless of underlying specification for [spread], a voiceless stop in a non-syllable-initial position (more generally, a stop in any environment other than initial in stressed syllable) is realized as unaspirated. Moreover, we obtain the same result if the stop is unspecified for [spread]:

(3-11)

		ASPIRATE	PRES(spread)
	Pɪl - ˈpɪl	*!	*
☞	Pɪl - ˈpʰɪl		*
☞	sPɪl - ˈspɪl		*
	sPɪl - ˈspʰɪl	*!	*

Under this ranking then, for any pair of URs which differ solely w.r.t. [spread], e.g. /phɪl/, /pɪl/ (or /Pɪl/), the PR neutralizes to a particular value of [spread]: [+spread] in the aspiration environment, and [-spread] elsewhere. Therefore, [spread] is not contrastive under this grammar, by definition (3-6), at least for the class of stops.

If, however, ASPIRATE is ranked below PRES(spread) (and there is no other higher-ranking constraint on the distribution of [spread] in voiceless stops in this context), then [spread] is contrastive, as in Hindi:

(3-12)

		PRES(spread)	ASPIRATE
☞	pi - þi		*
	pi - þhi	*!	
	phi - þi	*!	*
☞	phi - þhi		
	Pi - ˈpi	*	*!
☞	Pi - ˈphi	*	

URs /pi/ and /phi/ are solely distinguished by [spread], and the value of [spread] in their respective PRs, [ˈpi] and [phi], matches the underlying value; therefore, by definition (8), [spread] is contrastive under this grammar.

Finally, consider a grammar where PRES(spread) and ASPIRATE are freely ranked.

(3-13)		PRES(spread)	ASPIRATE
☞	pi - þi		*
☞	pi - þhi	*	
	phi - þi	*!	*
☞	phi - þhi		
	Pi - ˈpi	*	*!
☞	Pi - ˈphi	*	

UR /pi/ freely varies between [ˈpi] and [ˈphi], while /phi/ is invariantly realized as [ˈphi], a pattern attested in Ao (Gurubasave-Gowda 1975). The PRs for /pi/ are [ˈpi] and [ˈphi]; and the PR for /phi/ is [ˈphi]. That is, a [+spread] UR is invariantly realized as surface [+spread], and a [-spread] specification may be realized as [-spread]; therefore, by definition (3-6), [spread] is contrastive under this grammar.

Observe that the contrastive or predictable status of aspiration in the foregoing tableaux depends on the satisfaction or violation of PRES(spread) in the winning candidates, which in turn depends on the ranking of PRES(spread) with respect to the constraint on its distribution. The predictable status of stop aspiration in English (3-10) and (3-11) in no way depends on the feature's absence from any level of phonological representation. We will now proceed to consider implications of this insight for inclusion of universally non-contrastive phonetic properties.

5. UNIVERSALLY NON-CONTRASTIVE PROPERTIES: VOWEL DURATION

To illustrate the explanatory potential of universally non-contrastive properties, we will consider the role of subphonemic vowel duration in a vowel centralization process in Nawuri, a Kwa language of Ghana (Casali 1995). Short non-back vowels (i,ɪ,e,ɛ) are centralized (ɨ,ɪ,#ə,ʌ), except in absolute word-initial or phrase-final position.[29]

(3-14) a. Non-back vowels centralize, except phrase-finally:

ləmbɨri	'black'
olɨŋ	'root'
təkpəri	(type of grass)
ɔ-kɪ#ŋ	'fish'
gɪ#baː (/gɪ-baː/)	'hand'
tʃʌmɪ#nɛː (/tʃɛ-mɪnɛː/)	'friend'
natɪ#ba (/natɪ ba/)	'walk and come'

b. Long vowels do not centralize:

bɪːla	'learn'

c. Word-initial vowels do not centralize:

ɪ sʌŋ ɪkɪ#ŋ	"The fish (pl.) remain'

Note that this centralization process cannot be relegated to the "phonetic" component: for, as Casali (1995) argues, the centralization process in turn conditions an unambiguously "phonological" process of rounding harmony. High vowels alternate, surfacing as [+round] when the following vowel is [+round] *and the target vowel is non-front, by application of the centralization process:* e.g. /gɪ-lɔ/ - [gʊlɔ] ('illness'); but /ɪ-

---

[29]Casali (1995) characterizes the centralization environment as 'interconsonantal'; however, VV sequences in Nawuri are subject to a set of hiatus-avoiding processes (coalescence and glide formation), which yield long vowels (Casali 1996). Therefore, the interconsonantal condition reduces to the (surface) vowel length condition.

kɔ/ - [ɪkɔ] ('wars'), not *[ʊkɔ] (the prefix vowel in the latter form is not in the centralization environment, because it is word-initial). The centralization condition on the rounding process can be readily understood in terms of the markedness of front, round vowels. But in order to capture this explanation, and thus account for the differing rounding behavior of consonant-initial and vowel-initial prefixes, the centralization process must apply no later than the rounding harmony process. And the rounding harmony process cannot be relegated to the "phonetics," as it fails to apply across word boundaries, and results in neutralization with underlyingly [+round] vowels.

Rod Casali (p.c.) observes, from both auditory impressions and instrumental measurements, that the 'short' vowels in the contexts in which centralization is blocked are phonetically longer than the vowels in the target contexts, albeit shorter than a truly (i.e. contrastively) 'long' vowel. The extra vowel duration is attributable to cross-linguistically common phenomena of word-initial and phrase-final lengthening (Oller 1973, Klatt 1975). The centralization process can now be readily understood as articulatory 'undershoot' (Lindblom 1963; Moon & Lindblom 1994): the tongue body achieves the peripheral 'front' target only in long *or phonetically lengthened* vowels, when it has enough time to do so without deploying a high velocity fronting gesture.[30] The alternative, stipulation of a list of contexts where centralization occurs, is formally inelegant as well as phonetically uninsightful.

Assume a feature, [partially long], which distinguishes the truly short vowels from the long or phonetically lengthened vowels. By allowing the phonology to refer directly

---

[30]Further supporting the undershoot account, Rod Casali (p.c.) reports that centralization occurs to a lesser degree when the target vowel follows a coronal consonant, i.e. when the tongue body is already relatively 'front' due to the advancement of the tongue blade for the coronal articulation.

to such a feature, despite its universal non-contrastiveness, the undershoot explanation can be captured, e.g. in terms of a constraint, *[+front, -partially long].[31] To complete the analysis, we merely need to rank *[+front,-partially long] above PRES(front).

(3-15)

		*[+front,-pl]	PRES(front)
	o-liŋ - oliŋ	*!	
☞	o-liŋ - olɨŋ		*
	natɪ ba - natɪ ba	*!	
☞	natɪ ba - natɪ#ba		*
☞	lembiri - ləmbɨri]phrase		
	lembiri - ləmbɨrɨ]phrase		*!
☞	bɪːla - bɪːla		
	bɪːla - bɨːla		*!

Let us assume that the corresponding faithfulness constraint, PRES(partially long), is universally so low-ranked that it is never active[32] (presumably reflecting the lack of perceptual salience of this acoustic cue), or equivalently, that there is no faithfulness constraint for partially long in the constraint set. (This stipulation is analogous to the standard assumption that [partially long] is unspecified in the phonology.) The distribution of this feature is therefore determined by the following constraints:

[31] More elegantly and directly, the undershoot explanation can be captured in terms of the same LAZY-vs.-faithfulness conflict deployed elsewhere in this dissertation. This analysis involves the decomposition of LAZY into a series of effort thresholds, a move motivated in Chapter 7. To sketch such an analysis, let us assume that realization of a front vowel generally requires no more than *x* amount of effort. But when the vowel is extra-short, some greater amount of effort is required (x+1), due to the high velocity of the fronting gesture. Thus, under the ranking Lazy_{x+1} » PRES(front) » Lazy_x, centralization occurs in extra-short vowels, and is blocked elsewhere.

[32] A constraint is 'active' on an input if it is satisfied by some candidates and violated by others, and no higher ranking constraint has already ruled out all of the satisfiers or all of the violators (Prince and Smolensky 1993:82).

(3-16) PHRASE-FINAL LENGTHENING: V]$_{Phrase}$ - [+partially long].

WORD-INITIAL LENGTHENING: [$_{Word}$V - [+partially long].

*HALF-LONG: [-long] - [-partially long].[33]

As shown in (3-17), for any pair of URs which differ solely w.r.t. [partially long], the distinction neutralizes in PR; the surface value of [partially long] is conditioned by the position of the vowel within the word or phrase.

(3-17)

		PHRASE-FINAL	WORD-INITIAL	*HALF-LONG
☞	[...V...]$_{Wd}$ -[...V...]$_{Wd}$			
	[...V...]$_{Wd}$ -[...V·...]$_{Wd}$			*!
	[$_{Wd}$V -[$_{Wd}$V		*!	
☞	[$_{Wd}$V -[$_{Wd}$V·			*
	V]$_{Phr}$ -V]$_{Phr}$	*!		
☞	V]$_{Phr}$ -V·]$_{Phr}$			*
☞	[...V· ...]$_{Wd}$ -[...V...]$_{Wd}$			
	[...V· ...]$_{Wd}$ -[...V·...]$_{Wd}$			*!
	[$_{Wd}$V· -[$_{Wd}$V		*!	
☞	[$_{Wd}$V· -[$_{Wd}$V·			*
	V·]$_{Phr}$ -V]$_{Phr}$	*!		
☞	V·]$_{Phr}$ -V·]$_{Phr}$			*

Therefore by (3-6), [partially long] is not contrastive. Under the opposite ranking, with *HALF-LONG dominating the lengthening constraints, all non-long vowels would be realized as [-partially long] in all contexts; but in no case can the feature behave contrastively.

Once again, the non-contrastive behavior of the feature in question emerges from the constraint system; we require no assumption that such properties are excluded from

[33][+long] entails [+partially long] by definition.

phonological representation, or from any level of representation within the phonological component. The sole distinction between this case and the discussion of aspiration in the previous section is that [partially long] is non-contrastive under any ranking, due to the lack of a PRES(partially long) constraint; and since rankings are all that distinguish phonological systems in Optimality Theory, this amounts to showing that the feature is non-contrastive universally. We are able to include this phonetic property in the phonological representation, thereby capturing insights into Nawuri vowel centralization; but we do not generate spurious systems in which [partially long] is contrastive per se.

This result does not translate elegantly into a rule-based framework. We would have to stipulate that *every* language has a rule or set of rules that neutralize this feature in all contexts, contrary to the general view of rules as language-specific. Nor is it a simple matter to stipulate these universal neutralization rules. For example, the distinction between released and unreleased stops is universally non-contrastive, though stop release is phonologically relevant in licensing contour segments (partially nasalized or affricated) (Steriade 1993). Nevertheless, this feature neutralizes (pre-pausally) to [+released] in French, [-released] in Korean, and is in free variation in English. In a rule-based framework, we must then posit three distinct neutralizing rules for the three languages; and there is no unified formal expression of the feature's non-contrastiveness. In the OT formalism, however, what unifies the three cases is the lack of a PRES(released) constraint; and what distinguishes them is the ranking (or non-ranking) of conflicting constraints on the surface value of [released], such as 'Stops must be released' vs. 'Coda stops must be unreleased.'

6. GENERALIZING THE RESULT: THE CONTRASTIVENESS THEOREM

The ability to characterize the predictable or contrastive status of features in terms of the interaction of PRES(F) constraints and the rest of the constraint system (either under particular rankings or universally), rather than in terms of representational restrictions, is not limited to the cases just considered, but rather is fully general.

(3-18) **<u>The Contrastiveness Theorem</u>**

For all features F, F is contrastive under grammar G iff

(a) there is a constraint PRES(F), *and*

(b) there is some PR p such that for any UR u, if the mapping u - p is allowed under G, the mapping satisfies PRES(F) w.r.t. some F specification in u or p, if any.

To prove (3-18), it is sufficient to show that (I) if both the conditions in (3-18) hold, F is contrastive, and (II), if either of the conditions in (3-18) do not hold, F is not contrastive. Recall that under our definition, F is contrastive iff there is some pair of UR-to-PR mappings {<u,p>, <u',p'>} such that (a) u and u' differ solely w.r.t. some specification for F, and (b) p and p' differ w.r.t. a corresponding specification for F, and (c) u' cannot map to p. (Again, the URs and PRs under discussion are possible, not necessarily actual, forms.)

I. Assume that both the conditions in (3-18) hold: there is a constraint PRES(F), and it is unviolated in the mapping between p and all its possible URs w.r.t. some F specification in p, if any. Therefore, by the definition of PRES(F) (3-7), for some F specification in p, u must have a corresponding identical F specification, or if p has no F specification, then neither does u. Let u' be a UR which differs from u solely w.r.t. F. That is,

Table 3-1. Values for feature F, in URs and PRs.

if p contains:	then u contains:	and u' contains:
$+F_i$	$+F_i$	$-F_i$ or 0F
$-F_i$	$-F_i$	$+F_i$ or 0F
0F	0F	$+F_i$ or $-F_i$

Since u' - p would violate PRES(F), u' cannot map to p. We next show that the PR for u' must differ from p w.r.t. F. For some $\alpha \in \{+,-,0\}$,

(3-19)

			(Other higher- or equally ranked constraints)	PRES(F)
a. ☞	i.	UαFV (= u) - WαFX (= p)	<a>	
	ii.	UαFV (=u) - W¬αFX (= p'_1)	<b>	*
	iii.	UαFV (= u) - YαFZ	<c>	
	iv.	UαFV (= u) - Y¬αFZ (= p'_2)	<d>	*
b. ✘	i.	U¬αFV (= u') - WαFX (= p)	<a>	*
	ii.	U¬αFV (= u') - W¬αFX (= p'_1)	<b>	
✘	iii.	U¬αFV (= u') - YαFZ	<c>	*
	iv.	U¬αFV (= u') - Y¬αFZ (= p'_2)	<d>	

U, V, W, X, Y, Z denote any phonological material. ✘ indicates a necessarily losing candidate. <a>, <b>, <c>, <d> show dependencies between the presence of a violation mark in tableaux (a) and (b), depending on assumptions concerning the other constraints: that is, whatever violations of higher- or equally ranked constraints are incurred by candidate (a-i), these violations are likewise incurred by candidate (b-i). These dependencies follow from the assumption that no constraint other than PRES(F) refers to the underlying value of F: since the URs in (a) and (b) are otherwise identical, all constraints violated by a given PR in (a) must be violated by the same PR in (b), and vice-versa. <a> corresponds to no worse violation marks than are present in <c>, otherwise candidate (a)(i) would lose to (a)(iii), contrary to our assumption that p is an output for u. We have already shown that u' cannot map to p, i.e. candidate (b)(i) necessarily loses. Moreover, since the violations in <a> are no worse than those in <c>, (b)(iii) must lose as well. Therefore the PR for u' (either p'_1 or p'_2) differs from p w.r.t. F. There is therefore a pair of PRs, p and p' which differ w.r.t. F; among the inputs to p and p' are u and u' respectively, which differ solely w.r.t. F; and u' does not map to p. By definition (3-6), F is contrastive under G. This result is exemplified by the rankings of PRES(spread) and ASPIRATE for Hindi (3-12) and Ao (3-13).

II. Assume that either of the conditions in (3-18) do not hold. There is no PRES(F) constraint, or there is no PR p such that all URs for p satisfy PRES(F).

(3-20)

		Other constraints	PRES(F)
☞	UαFV (= u) - WαFX (= p)	<a>	
	UαFV (= u) - W¬αFX (= p')	<b>	*
☞	U¬αFV (= u) - WαFX (= p)	<a>	*
	U¬αFV (= u') - W¬αFX (= p')	<b>	

p' may also be a winner, in free variation with p, if the violations in <b> are equal to <a>. In either case, since PRES(F) can be violated, no other constraint prevents u and u' from both mapping to p. Since u' can map to p, by definition (3-6), F is not contrastive under G. This result is exemplified by the status of [partially long] universally (3-17), and by the rankings of PRES(spread) and ASPIRATE for English (3-10, 3-11).

Therefore, the conditions in (3-18) are both necessary and sufficient to show that a feature is contrastive.

The assumption of strongly parallel, one-step UR-PR mapping is not crucial to this result. In a multi-stratal grammar, F is contrastive just in case (3-18) (substituting 'input/output' for 'UR/PR') holds true for F at each stratum. We have shown that, on a single round of evaluation, an underlying distinction in F maps to an output F distinction just in case (3-18) holds w.r.t. F. In a multi-stratal OT grammar, this output is then taken as the input for another round of candidate generation and evaluation. But if (3-18) characterizes the behavior of F on the second stratum as well, the same result obtains, and so on, regardless of the number of strata.

7. Categorical Effects with Continuous Representations

It is standardly claimed that 'phonological' representations are categorical, whereas 'phonetic' representations are gradient. Consider a phonetic dimension such as vowel height. For purposes of phonological analysis, this phonetic continuum is standardly divided into three regions, low, mid and high, which are formally represented in terms of two binary features: [low] and [high], as shown in Figure 3-6.

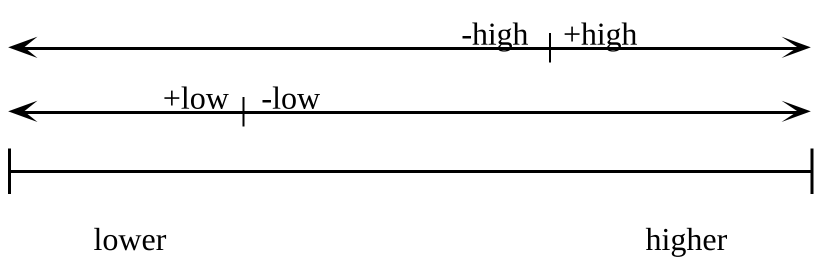


Figure 3-6. Vowel height continuum subdivided using two binary features.

Clearly, the claim of phonological categoricalness cannot mean that there is at most a binary distinction for any phonetic dimension; for at least a ternary distinction in vowel height is required.[34] If, however, the claim of phonological categoricalness is that the phonology represents phonetic dimensions in terms of *some* number of discrete, binary features, then the distinction between categorical and gradient representations is empirically vacuous, since this technique can be applied recursively to yield a quasi-continuum. Let the vowel height dimension be subdivided into, say, 100 features of the form [±vowel height (Vht) > n] as shown in Figure 3-7.

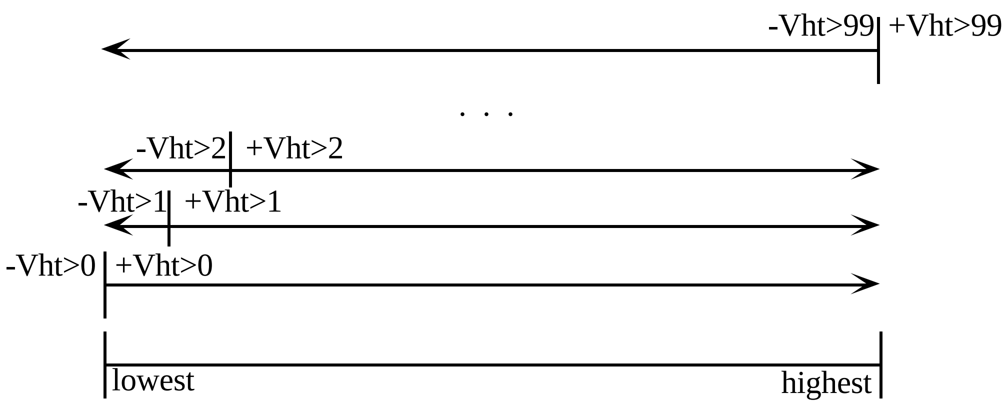


Figure 3-7. Vowel height continuum subdivided using 100 binary features.

This scale is certainly fine enough for any linguistic phonetic analysis, but if a closer approximation to a true continuum were required, we could simply subdivide the dimension into an even greater number of binary distinctions, such that the increments approach infinitesimality.[35]

If there is any truth to the claim of phonological categoricalness, it lies in the observation that a small number of points along the phonetic dimension will be

---

[34]Nor is non-binarity unique to vowel height: ternary (or greater) distinctions are also required for the F0 frequency dimension (i.e. tone), and for the 'sonority' dimension (e.g. Clements 1990).
[35]Similarly, the categorical treatments of aspiration and vowel duration in sections 4 and 5 can be recast in gradient terms by adopting more continuous representations of voice onset time and vowel duration, respectively.

contrastive in any given language. For example, we do not find languages which have anything approaching 100 contrastive degrees of vowel height. But this is simply a special case of distinguishing between contrastive and non-contrastive properties, which can be adequately handled in terms of the constraint system, without representational restrictions, as shown in the preceding sections.[36] That is, the categorical behavior of phonological objects emerges from the constraint system, rather than by reifying the categories in an abstract representation.

To make this clear, let us adopt a quasi-continuous representation of vowel height, as in Figure 3-7, to describe a language with a three-way distinction of vowel height. We simply need to identify the two features which serve as the boundaries between mid and non-mid vowels in the relevant language: for the sake of concreteness, say [Vht>33] and [Vht>67]:

(3-21)	[-Vht>33]	=	[+high]	=	high
	[+Vht>33, -Vht>67]	=	[-low, -high]	=	mid
	[+Vht>67]	=	[+low]	=	low

To capture the three-way contrast, it suffices to assume that PRES(Vht>33) and PRES(Vht>67) are ranked such that the conditions of (3-18) hold, i.e. for some set of mappings between a PR and all of its URs, the mappings satisfy PRES(Vht>33) and PRES(Vht>67); and the conditions of (3-18) do not hold for the remaining possible

[36]This point has also been made by Steriade (1995) and Flemming (1995), though Flemming's approach relies on a family of constraints which refer explicitly to the maintainance of contrasts (by globally comparing inventories of possible words), rather than relying on faithfulness constraints.

distinctions in F1 frequency, e.g. [Vht>60], either because there is no PRES(Vht>60) constraint, or this constraint is inactive due to low ranking.

A class of constraints which might render certain faithfulness constraints inactive in all contexts are the fortition constraints, which favor maximal dispersion of the values within the relevant perceptual dimension (cf. Liljencrants and Lindblom 1990, Flemming 1995), thereby enforcing strict categoricalness.

(3-22) DISP(Vht,binary): Vht = 0 or 100

DISP(Vht,ternary): Vht = 0, 50 or 100

DISP(Vht,quaternary): Vht = 0, 33, 67, or 100 etc.

Under the ranking shown in tableau (3-23), only three possible surface vowel heights are possible (0, 50 or 100), regardless of underlying values of vowel height, notwithstanding the presence of a constraint PRES(Vht>60) within the hierarchy.

(3-23)

	a. V height = 84	DISP (tern)	PRES (Vht>33)	PRES (Vht>67)	PRES (Vht>60)	DISP(bin)
	Vht = 0		*!	*	*	
	Vht = 50			*!	*	*
	Vht = 60	*!				*
☞	Vht = 100					
	b. Vht = 61					
	Vht = 0		*!		*	
☞	Vht = 50				*	
	Vht = 60	*!			*	*
	Vht = 100			*!		*
	c. Vht = 29					
☞	Vht = 0					
	Vht = 50		*!			
	Vht = 60	*!	*			*
	Vht = 100			*!	*	*

Of the three surface values permitted by DISP(Vht,ternary), the one assigned to a particular UR value is the one which crosses no 'boundary,' i.e. violates no active PRES constraint. For example, in (3-23a), mapping from an underlying value of 84 to a surface value of 100 does not cross the 33 or 67 boundaries enforced by the active constraints PRES(Vht>33) and PRES(Vht>67); if, however, the UR is lower than 67, as in (3-23b), the surface value which crosses no boundary is 50 (though it violates the inactive PRES(Vht>60)). Again, the ternary surface distinction is attributable entirely to the constraint system; it does not depend on the exclusion of the non-contrastive vowel height distinctions from phonological representation.

Ironically, this approach now appears to be *too* categorical, i.e. predicting a narrow set of precise and invariant surface values for vowel height, when in fact we observe a great deal of surface variation; indeed, one of the points of having a phonetically rich representation is to capture gradient as well as categorical variation within a unified phonological formalism. However, we can reintroduce variation into the picture in a number of ways.

First, to the extent that the variation is completely *free*, the dispersion fortition constraints, rather than identifying precise values, might specify permissible *ranges* of values:

(3-24) DISP(Vht, ternary) (revised): Vht = 0-12, 45-53, or 89-100

Second, *context-sensitive* variation can be captured in terms of conflict between the fortition constraints and other, context-specific, constraints on vowel height, e.g. articulatory constraints favoring reduced vowels in fast speech:

(3-25)	Vht = 15 (fast speech)	FAST-SPEECH REDUCTION[37]	PRES(Vht>33)	POLAR(tern)
	Vht = 0	*!		
☞	Vht = 21			*
	Vht = 50		*!	
	Vht = 15 (normal speech)			
☞	Vht = 0			
	Vht = 21			*!
	Vht = 50		*!	

The result here is a (somewhat simplified) case of context-sensitive variation in the value of Vht, depending on speech rate.

Finally, I do not attempt to spell out here the relation between this gradient but still abstract scale of vowel height and actual physical measures thereof: i.e. frequency of

[37]A more sophisticated, effort-based approach to fast-speech reduction is presented in Chapter 6.

the first formant in Hz (acoustically), or Mels (perceptually); or (roughly speaking) tongue body height in millimeters (articulatorily). Consequently, there is room for *inter-speaker variation* in this relation. If we are simply concerned with modeling the production of a single speaker, it is possible (if non-trivial) to substitute, e.g., Hz values for the abstract values, as we shall see in the following section. However, inter-speaker variation remains an unsolved problem in this as in all existing theories of phonology and phonetics.

8. Lenition or Fortition with Indeterminate UR's

Lozano (1979) raises the question of why alternations or allophonic relations between stronger and weaker consonants should be analyzed as lenition of underlying stronger consonants. Specifically, Lozano argues that so-called spirantization patterns in a number of Spanish dialects could be analyzed as fortition of underlying fricatives to stops, with equal plausibility to the traditional spirantization analysis.

In the OT approach sketched above, the choice of underlying representations turns out to be much less crucial to the analysis: the burden of characterizing the surface patterns falls on the constraint system, rather than assumptions about the inventory of possible segments in UR. It is therefore worthwhile to appraise Lozano's question afresh, in light of these new assumptions. Following Lozano, I will take spirantization, or patterns relating stops and continuants, as emblematic of the general lenition vs. fortition question.

Patterns which involve neutralization of a stop/continuant contrast to a surface continuant I will refer to as *lenition*, while neutralization to a stop I will refer to as *fortition.* Use of these terms does not, however, necessarily imply that the relevant consonants have an underlying value for continuancy which differs from their surface value. Consider, for example, a sound system in which there is a word-initial stop/fricative distinction; while elsewhere, only fricatives occur:

(3-26)

		PRES(cont/#__)	LAZY	PRES(cont)
	aba - aba		**!	
☞	aba - aβa		*	*
	aβa - aba		**!	*
☞	aβa - aβa		*	
☞	ba - ba		**	
	ba - βa	*!	*	*
	βa - ba	*!	**	*
☞	βa - βa		*	

As shown in (3-26), the constraint hierarchy characterizes the surface pattern without regard to the underlying value of continuancy in [aβa]. On the other hand, in the case of a distinction between, say, a lexical item which invariantly surfaces as a fricative [oβ], and a lexical item [ab/aβ], whose realization varies between stop and fricative, depending on context, we must assume that the latter is underlyingly [-cont] (if it were [+cont], it would behave identically to [oβ]).

A case of word-initial fortition, which neutralizes a medial stop/fricative contrast, corresponds to the ranking *+cont/#__ » PRES(cont):

(3-27)		*+cont/#__	PRES(cont)
☞	aba - aba		
	aba - aβa		*!
	aβa - aba		*!
☞	aβa - aβa		
☞	ba - ba		
	ba - βa	*!	*
☞	βa - ba		*
	βa - βa	*!	

Again, in the neutralizing case, [ba], the value of [cont] in UR is not crucial to the analysis, so long as the lexical item in question is invariantly [-cont] on the surface.

Finally, cases of perfect complementary distribution between a stop and a fricative must be analyzed as both lenition (in the spirantizing contexts) *and* fortition (in the fortifying contexts). That is, it is not sufficient for the constraint system to cause an underlying stop to be realized as a fricative in the correct context: since we make no assumptions about inputs, the constraint hierarchy must also ensure that an underlying fricative will fortify to a stop, in the contexts where the fricatives do not occur. I will refer to such patterns as (for want of less cumbersome term) "complementary lenition/fortition." Thus, for example, a pattern of word-initial stops in complementary distribution with fricatives corresponds to the following ranking:

(3-28)		*+cont/#__	LAZY	Pres(cont)
	aba - aba		**!	
☞	aba - aβa		*	*
	aβa - aba		*!	
☞	aβa - aβa			*
☞	ba - ba		**	
	ba - βa	*!	*	*
☞	βa - ba		**	
	βa - βa	*!	*	*

Again, the underlying value of continuancy is not crucial to the analysis. Note that Pres(cont) in (3-28) is completely inactive, reflecting the non-contrastive status of [cont] under this hierarchy.

Our response to Lozano's question, then, is: (a) neutralization to a fricative is lenition; (b) neutralization to a stop is fortition; (c) complementary distribution between stop and fricative is both lenition and fortition; and (d) (except in the case of a difference between lexical items with consistent continuancy values and items with varying values) the choice between underlying stops and fricatives is not crucial to the analysis of the surface patterns, and so becomes a non-issue.[38]

9. SUMMARY

Let us return to the observation that, contrary to the standard approach, contrastiveness is not an all-or-nothing property; that features fall along a continuum of potential contrastiveness. In the approach developed herein, the contrastiveness of a feature follows from the satisfaction of PRES(F), which in the OT framework in turn depends on the ranking of PRES(F) relative to potentially conflicting constraints, namely constraints on the surface distribution of F. The higher the ranking of PRES(F), the more distributional constraints PRES(F) outranks, hence the broader the contexts in which F is contrastive.

Intuitively, the position of PRES(F) in the constraint hierarchy of a grammar for a given speaker corresponds to the degree to which the speaker attends to feature F in the

[38]This position is modified somewhat in Chapter 9, in response to the problem of capturing stabilized context-sensitive lenition patterns.

mapping between input and output. Thus, for example, speakers of Hindi attend to stop aspiration distinctions in a way that English speakers do not. Although constraint ranking is generally a language-specific matter, it must be recognized that certain featural distinctions are inherently more salient than others, e.g. [consonantal] (characterized by abrupt, large-scale changes in amplitude) vs. [longitudinal vocal fold tension] (principally cued by subtle F0 distinctions in the beginning of a following vowel) mentioned in the Introduction. The notion of inherent salience can be formalized in terms of a set of universal ranking conditions such that for a certain feature F, PRES(F) outranks certain constraints on the distribution of F; or PRES(F) outranks PRES(G) (for some other feature G), reflecting the claim that F is inherently more salient than G (cf. Jun 1995). Features which are inherently highly salient have corresponding faithfulness constraints which are universally highly ranked; while inherently subtler features have lower-ranked faithfulness constraints. And, as discussed in Sections 5 and 6, universally non-contrastive features lack a faithfulness constraint altogether. (The determination of the relative salience of particular auditory cues constitutes a set of empirical questions, to be resolved through perceptual phonetic experimentation.)

Moreover, this approach captures the connection between the frequent non-contrastiveness of some feature and its restriction to narrow environments when it is contrastive. The lower the ranking of PRES(F), the more constraints on the distribution of F dominate PRES(F), hence the narrower the contexts in which F is contrastive, and the greater the likelihood that F is not contrastive in any context at all. Central to this treatment of contrastiveness is the Optimality Theoretic notion of faithfulness. The analog of faithfulness in rule-based frameworks is a mere default state, the absence of neutralization rules w.r.t. a particular feature in a given language; thus there is no corresponding formal expression in such a framework of a feature's cross-linguistic tendency to resist neutralization.

In sum, I have shown that some standard assumptions concerning phonological representation, and the 'phonetics-phonology interface,' warrant reevaluation in light of Optimality Theory. Specifically, the motivation for excluding non-contrastive properties from the phonological representation, or any derivational stage therein, evaporates under Optimality Theoretic analyses which include feature-specific faithfulness constraints. By the Contrastiveness Theorem, the contrastiveness of a particular feature depends entirely on whether there is a corresponding faithfulness constraint which is satisfied under some set of mappings, which in turn depends on the position of the constraint within the constraint hierarchy. This result extends even to properties which are not contrastive in any language, if we simply assume that such properties lack a corresponding PRES constraint. Finally, the distinction between categorical and gradient properties, standardly assumed to characterize the difference between phonological and phonetic representation, proves to be a special case of the previous result. Consequently, we may capture the categorical and contrastive *behavior* of particular phonetic properties (and the predictable or gradient behavior of the remainder) in terms of constraint interaction, while using *representations* which in principle may contain complete phonetic detail, including gradient properties such as effort cost. Thus, Optimality Theory permits the removal of the representational blinders imposed by the standard treatment of contrastiveness: a move which promises to lead to new insights into the phonetic bases of many phonological phenomena, including lenition processes.

Chapter 4:

Spirantization and Stridency

The lenition surveys (see Chapter 1 and Appendix) support a generalization concerning spirantization: unaffricated stops never lenite to strident fricatives, such as [s] or [f]. Rather, stops typically spirantize to weakly fricated or approximant continuants such as [β,ð,ɣ] or [β̞,ð̞,ɰ]. Although assibilatory spirantization processes (e.g. t - s,ʃ) are attested, these are restricted to contexts in which the stop is inherently somewhat affricated. In this chapter, I document these generalizations, and demonstrate that they fall out from the effort-based approach, coupled with certain plausible assumptions concerning the effort cost of strident relative to nonstrident continuants. Moreover, this generalization is reconciled with the apparently conflicting observation, from studies of segment inventories (e.g. Maddieson 1984, Ladefoged & Maddieson 1996) that these nonstrident continuant consonants, which are favored as outputs of lenition, are highly *dis*favored (i.e. "marked") relative to strident fricatives in the general case.

1. Documenting the Generalization.

1.1. Stridency.

Stridency, as used herein, refers to high noise intensity of fricatives and affricates. However, in the absence of explicit description of sounds as strident or nonstrident in the survey grammars, I must rely upon inferences from transcriptions. Sibilants (s, ʃ, tʃ, z, etc.) and labiovelar fricatives (f, v) are generally classified as [+strident], whereas (θ, ð, ɸ, β) are [-strident] (Chomsky and Halle 1968).

Strident fricatives (particularly sibilants) are generally considered to be unmarked relative to nonstrident continuant consonants (excluding glides and liquids) (Maddieson 1984, Ladefoged and Maddieson 1996). The relative predominance of strident fricatives in segment inventories is clear from the following table, adapted from Maddieson 1984: 45:

Table 4-1. Frequency of fricatives in Maddieson's (1984) segment inventory database.

s	266	z	86	ɬ	30	ɮ	7
ʃ	146	ʒ	51	θ	18	ð	21
f	135	v	67	ʂ	17	ʐ	3
x	75	ɣ	40	ç	16	ʝ	7
χ	29	ʁ	13	ħ	13	ʕ	9
ɸ	21	β	32				

This unmarkedness presumably reflects the perceptual salience of the stridents, precisely because of their noisiness relative to the nonstrident continuants.

1.2. Coronals

1.2.1. No spirantization to a sibilant. Despite this general relative unmarkedness, I have encountered no genuine cases of a non-affricated coronal stop synchronically leniting to a sibilant fricative. Rather coronal stops commonly lenite to non-sibilant fricatives. Thus in Mexican Spanish (Harris 1969), for example, [d] alternates with [ð], e.g. [dewðas] ('debts') vs. [aj ðewðas] ('there are debts'). Similarly, in Florentine Italian, [t] and [d] alternate with [θ] and [ð] (or corresponding approximants, in fast or informal speech) (4-1a); only the sibilant palato-alveolar affricates lenite to sibilant fricatives (4-1b).

(4-1) Florentine Italian (Giannelli & Savoia 1979)

a.	tattsa	la {θ/θ̩}attsa	domani	e {ð/ð̩}omani
	'cup'	'the cup'	'tomorrow'	'it is tomorrow'
b.	tʃena	la ʃena	dʒorni	i ʒorni
	'supper'	'the supper'	'days'	'the days'

Additional examples of coronal spirantization outputs are presented in the following table:

Table 4-2. Spirantization of coronal stops to nonstrident fricatives or approximants.

Language	Reference	Description
Early Modern Greek	Bubeník 1983	t^h > θ context-free
Badimaya	Dunn 1988	d -ð /V__V
Cardiff English	Collins & Mees 1990	d -ð medially (optional)
Catalan	Hualde 1992	d -ð̞ /non-initially, except after n or l
Dahalo	Tosco 1991	d -ð /V__V
Florentine Italian	Giannelli & Savoia 1979	t -θ /V__({r,l,j,w})V, and elsewhere in casual speech; d -ð /V__({r,l,j,w})Vin natural speech
Germanic (Gothic)	Bennett 1980	d -ð/V__V
Gujarati	Cardona 1965	d̪ -ð /murmured V__V
Kabylie Berber	Chaker 1983	d > ð context-free
Karao	Brainerd 1994	t,dʒ - θ,j (context unclear)
Ladakhi	Koshal 1976	d - ð non-initially (optional)
Liverpool English	Wells 1982	t,d - θ̠,ð̠ word-finally and /V__V
Manobo	Reid 1971	d - ð (context unclear)
Proto-Germanic	Meillet 1970	d^h,t^h > ð,θ
Purki	Rangan 1979	d̪ - ð /V__V
Shina	Rajapurohit 1983	d - ð /V__V
Somali	Armstrong 1964	d - ð̞ /stressedV__V
Mexico City Spanish	Harris 1969	d - ð non-initially, except after /n/ or /l/
Taiwanese	Hsu 1995	t - ð non-initially
Tiberian Hebrew	Malone 1993	d,t - ð,θ/V__
Tümpisa Shoshone	Dayley 1989	t - ð/{i,e}__
Tzeltal	Kaufman 1971	d - ð/V__{V,+}
Uradhi	Dixon 1979	t - ð (context unclear)
Warndarang	Heath 1980	d^j - j/V__V
Yindjibarndi	Wordick 1982	d̪,d^j - ð̞,j/V__V

Coronal stops also frequently lenite to flaps (particularly if contrastively retroflex or apical), as in the following examples:

Table 4-3. Lenition of coronal stops to flaps.

Language	Reference	Description
American English	Kahn 1976	t,d - ɾ/stressedV__V and /#__
Chitwan Tharu	Leal 1972	ʈ - ɾ/V__V, d,d^h - ɾ /__#
Djabugay	Patz 1991	d - ɾ/stressedV__V
Gujarati	Cardona 1965	ɖ - ɽ/utterance-medially, except after ɳ or another ɾ, and /__# (optional)
Halabi	Singh 1977	ɖ,$ɖ^h$ - ɾ/non-initially, except after ɳ
Kanakuru	Newman 1974	T - ɾ/V__V
Kashmiri	Kachru 1969	ɖ - ɽ/V__V and /__#
Kupia	Christmas & Christmas 1975	ɖ - ɽ/V__Vunstressed (obligatory), ʈ - ɾ/V__Vunstressed (optional)
Lamani	Trail 1970	ɖ - ɽ/non-initially, except after ɳ or ɭ
Lowland Murut	Prentice 1971	d - ɾ/-cons__
Malayalam	Mohanan 1986	t,ɖ - ɾ/+son,-nas__V
Moghamo	Stallcup 1978	t - ɾ (context unclear)
Nepali	Bandhu 1971	$ɖ^h$ - $ɾ^h$ non-initially
Panyjima	Dench 1991	ʈ - ɾ/V__V
Purki	Rangan 1979	d - ɾ/V__V
Sawai	Whistler 1992	d - ɾ/V__
Shina	Rajapurohit 1983	ɖ - ɽ/V__V
Tauya	MacDonald 1990	t - ɾ/V__
Tsou	Wright 1996	ɗ - ɺ (lateral flap)/__a
Tümpisa Shoshone	Dayley 1989	T - ɾ/V+back__

The typology of flapping processes is discussed more fully in Banner-Inouye 1995. Coronal stops can also lenite to lateral liquids (e.g. Sotho, Doke 1957; Proto-Bantu, Greenberg 1948; Limbu, van Driem 1987).

1.2.2. ASSIBILATION. Assibilatory spirantization processes (e.g. t - s,ʃ) are attested, but these prove to be restricted to contexts in which the stop is inherently somewhat affricated, thus conforming to the generalization that *unaffricated* stops never spirantize to sibilants. The lenition survey of Lavoie (1996: 294) presents the following synchronic cases in which an unaffricated stop changes to a sibilant fricative:

Table 4-4. Synchronic assibilatory spirantizations, from Lavoie 1996.

Language	Reference	Description
Ancient Greek	Sommerstein 1973	t - s / V__ i,y
Nez Perce	Aoki 1970	c - s / __ {n,w}
Turkana	Dimmendaal 1983	t - s / __ [-low, -back]

Fricated release. The Ancient Greek and Turkana cases exemplify conditioning of assibilation by a following high (or non-low) front vocoid, also seen in English morphophonemic alternations such as [pɹɛzɪɾɛnt] ~ [pɹɛzɪɾɛnsi] ('presidency'), /əlɛkt + jən/ - [əlɛkʃən] ('election'). In fact, this is also a common environment for assibilatory affrication, e.g. Québecois French t,d - ts,dz /__{i,y,ɪ,ʏ}; Japanese t,d - ts,dz /__ɨ[39]; often combined with palatalization, e.g. Japanese t,d - tʃ,dʒ/__i. In this environment, the stop's release is inherently somewhat fricated, due to the closeness of the tongue blade to the hard palate as it is released into the following vowel: the affrication is stronger the higher and fronter the vowel, and the greater the coarticulation between the stop and the following vocoid (e.g. in English casual speech /mit ju/ - [mitʃjə] ('meet you')); it is also stronger in voiceless than voiced stops, due to the typically louder release burst in the former (Ohala 1983, Jaeger 1978). The assibilation of palatal stops in Nez Perce can be understood in similar terms: the close position of the tongue blade during the release is inherent in the palatal place of articulation. Lavoie further notes a diachronic case of assibilatory spirantization of a retroflex stop in the Dravidian language Pengo (Burrow & Bhattacharya 1970), which admits of the same explanation. Like palatals, retroflex stops often have a fricated release, e.g. Iaai (Maddieson & Anderson 1994: 178-179), Ndumbea (Gordon & Maddieson 1996: 37), due to the forward uncurling of the tongue blade as the retroflex closure is released: this keeps the tongue tip in close constriction relative to the

[39]This vowel is commonly transcribed as [u], presumably because it is standardly so used for purposes of transliteration into Western orthographies; but the vowel is actually a unrounded centralized high vowel, with some degree of bilabial constriction.

postalveolar-to-alveolar region (modulo dramatic jaw lowering), hence friction (Victoria Anderson, p.c.); cf. assibilation of retroflex trills, e.g. in Czech, which also occurred in Pengo apparently as part of the same sound change as the assibilation of the retroflex stop.[40]

Fortition to a sibilant, lenition to a fricative. These stops with a relatively noisy release may then be fortified to a true sibilant affricate, i.e. with tongue-blade grooving (as has occurred, for example, in English, with lexicalized forms such as gotcha and betcha): this is the step which we may properly call "assibilation." The pre-[i], palatal, and retroflex assibilations therefore can be understood as fortitions, presumably as a strategy of avoiding distinctions between stops with a weakly fricated release and true strident affricates (see Flemming's (1995: 91) similar analysis of assibilatory affrication in Eastern Arrernte). Moreover, once this fortitional assibilation process has applied, the resulting sibilant affricate can spirantize to a fricative, due to effort minimization, just like underlyingly sibilant affricates (cf. the discussion of Florentine /tʃ,dʒ/ - [ʃ,ʒ] above). A formal analysis of these interactions is presented in Section 5 below. In sum, lenition of a coronal stop to a sibilant fricative is by no means attested in the general case; it occurs only in contexts where the release of the stop inherently has significant friction.

A series of sound changes. Furthermore, synchronic assibilatory spirantization must be distinguished from a series of sound changes which ultimately result in a sibilant fricative. Such a chain of events is known to have occurred, for example, in Ashkenazi Hebrew. In Ancient Hebrew, non-emphatic stops (p,t,k,b,d,g) post-vocalically

---

[40]Cf. the considerable affrication of /tɹ/ clusters in English (which typically involve coarticulatory retraction of the /t/), as reflected, for example, in Read's (1975: 79-104) finding that English-speaking children, when learning to spell, frequently represent /tɹ/ and /dɹ/ clusters as chr and jr respectively.

spirantized to nonstrident fricatives (ɸ,θ,x,β,ð,ɣ); much later, in Ashkenazi dialects, [θ] fortified to [s]. This series of changes has given rise to a situation in Ashkenazi Hebrew of lexical items containing [s] as the reflex of Ancient Hebrew *t; however, there has never been a stage of Hebrew in which [t] productively and regularly alternated with [s].[41]

A similar account can be given of a putative t - s lenition in Liverpool English.[42] Wells (1982) notes that in this dialect, word-final and intervocalic /t/ spirantizes to a non-sibilant alveolar fricative [ð̠]. This fact is also mentioned by John Harris (1990); but Harris further claims that, for some speakers, this [ð̠] neutralizes to [s], rendering words such as <u>letter</u> and <u>lesser</u> homophonous. This was apparently not the case some years earlier, when Wells collected his data, for he expressly denies that such neutralization occurs. Thus, there is clear evidence of an intermediate ð̠ stage in the Liverpool t > s change. Moreover, Harris gives no evidence suggesting that the speakers with [s] rather than [ð̠] exhibit productive t - s alternations. Therefore I conclude that Liverpool English has never had a synchronic process converting /t/ to [s].

1.3. Labials

Furthermore, bilabial stops overwhelmingly lenite to nonstrident bilabial or labiovelar fricatives or approximants, notwithstanding the general unmarkedness of the labiodentals, [f] and [v], among the class of labial fricatives.

---

[41]The t>θ>s development in Ashkenazi is further complicated by the fact that Tiberian Hebrew [tˤ] (i.e. emphatic t), has neutralized to [t], and [sˤ] has neutralized to [s]. Thus, although in Modern Ashkenazi Hebrew the t>θ>s changes have given rise to some morphophonemic alterations between [t] and [s], these are not part of any systematic phonological pattern, as there are also [t]s and [s]s (reflexes of the emphatics), occurring in the same contexts, which never alternate.
[42]Also referred to as the "Merseyside" or "Skouse" dialect.

Table 4-5. Spirantization of labials to nonstrident fricatives or approximants.

Language	Reference	Spirantization pattern
Apatani	Abraham 1985	b -β /V_V+back
Bashkir	Poppe 1964	b -β /V__V
Basque	Hualde 1993	b -β /V__V
Catalan	Hualde 1992	b -β̞ /non-initially, except after m
Dahalo	Tosco 1991	b -β /V__V
Danish	Bauer et al. 1980	b -β̞ medially
Germanic (Gothic)	Bennett 1991	b -β /V__V
Gujarati	Cardona 1965	b^h -β /V__V
Kanakuru	Newman 1974	p -w /V__V
Kupia	Christmas and Christmas 1975	p -ɸ /V_Vunstressed
Ladakhi	Koshal 1976	b -β non-initially
Lama	Ourso & Ulrich 1990	p -w /__#
Limbu	van Driem 1987	b -w /V__V
Lowland Murut	Prentice 1971	b -β /-cons__
Malayalam	Mohanan 1986	b -β̞ /+son,-nas__V
Manobo	Reid 1971	b -β or β̞ (context unclear)
Middle Korean	Ramsey 1991	b > β /V__V
Moghamo	Stallcup 1978	p - β (context unclear)
Nepali	Bandhu 1971	p^h -ɸ/V__, b^h -β/__#
Pennsylvania German	Kelz 1971	b -β/V__V
Proto-Bantu	Greenberg 1948	b -β except after m
Senoufo	Mills 1984	b -β medially
Shina	Rajapurohit 1983	b -β /V(r)__V
Somali	Armstrong 1964	b -β̞ /stressedV__V
Spanish	Harris 1969	b -β non-initially, except after m
Tamazight Berber	Abdel-Massih 1971	b -β context - free
Tatar	Poppe 1963	b-β context-free
Tiberian Hebrew	Malone 1993	p,b -ɸ,β /V__
Tzeltal	Kaufman 1971	b -β /V__{V,+}
Tümpisa Shoshone	Dayley 1989	p -ɸ,β non-initially except after m
Uradhi	Dixon & Blake 1979	p -β (context unclear)
Uzbek	Sjoberg 1963	p,b -ɸ,w medially
Warndarang	Heath 1980	p -w /V__V
Yana	Sapir & Swadesh 1960	b -w /V__V
Yindjibarndi	Wordick 1982	p -w /V__V

Moreover, in Nkore-Kiga (Taylor 1985), /b/ spirantizes to a labiodental approximant, [ʋ], in intervocalic position.

The only ostensible synchronic cases of spirantization to a labiodental in the lenition surveys are Amele (Roberts 1987), Kanuri (Lukas 1967), and Irish (Ní Chiosáin 1991). Lavoie (1996) summarizes Amele as having intervocalic lenition of /p/ to [f]. In fact, Roberts describes a phoneme /f/ which freely varies with [p] in all contexts in which /f/ occurs, e.g. [pupu] ~ [fufu] ('wind') (pp. 333, 337). This /f/ is in contrast with a bilabial stop, realized as [p] word-finally, and [b] elsewhere. The Amele [p/f] variation thus represents a case of optional fortition to [p], not lenition to [f]. As for Kanuri, though Lukas uses the symbols [f,v], the text expressly describes this "[f]" as bilabial, not labiodental; and the "[v]" presumably is bilabial as well. Finally, Irish has a consonant mutation process which, inter alia, involves alternations between [p] and [f] in certain morphosyntactic contexts (typically as the expression of a preposition) (see Chapter 1, section 2.6). Having become morphology-driven rather than effort-driven in the synchronic grammar, this process is no longer bound by considerations of effort minimization in the choice of output; the output can thus fortify [ɸ] (< p) to [f] while still participating in productive stop ~ fricative alternations.

1.4. DORSALS.

As for dorsal consonants, their non-stridency under spirantization is difficult to confirm (or falsify) from a survey of descriptive grammars, since there are no standardly used phonetic symbols distinguishing noisy from quiet velar or uvular fricatives, and descriptive grammars rarely comment on the noise intensity of the fricatives. However, Harris' (1969) careful phonetic description of Mexican Spanish makes clear that the outputs of spirantization, including [ɣ], are nonstrident, as does Giannelli and Savoia's (1979) description of Florentine Italian. But I am aware of no grammars which explicitly characterize any velar spirantization output as [+strident]. We may further observe that velar stops can spirantize to approximants, e.g. [ɰ], (e.g. Somali, Armstrong 1964; Catalan, Hualde 1992), or [w], e.g. Warndarang (Heath 1980), neutralizing with labials; or they can debuccalize to [h] (Florentine, Giannelli & Savoia 1979), or to [ʔ] (West Tarangan, Nivens 1992), losing their place of articulation. But there are no synchronic lenition processes whereby /k/ goes to a noisy fricative such as [s] or [f].[43] Therefore, the generalization of nonstrident spirantization outputs does not reduce to preservation of place of articulation.

In sum, in those languages for which we have careful phonetic descriptions, spirantization outputs are described as having weak friction, no friction, or some kind of phonetically or pragmatically conditioned variation between the two. In less explicit descriptions, the nonstridency of the spirantization outputs can be inferred from the absence of sibilants and labiodentals. And though a series of sound changes may result in strident fricatives as reflexes of stops, the claim is that, modulo cases of consonant mutation, there are no productive phonological processes converting unaffricated stops to strident fricatives.

---

[43]An apparent exception is English velar softening (k - s). However, if this is regarded as a synchronic process at all, it falls under the assibiliation rubric, being conditioned (at least historically) by a following high front vocoid, as it was in Late Latin.

2. AN EFFORT-BASED EXPLANATION.

2.1. ISOMETRIC TENSION

Recall that spirantization, under the effort-based approach, is viewed as reduction of the magnitude of a stop gesture (Figure 4-1a), for reasons of effort minimization, to the point where closure is lost. The interval of close constriction in such a reduced gesture is brief (Figure 4-1b), hence the duration of friction will be correspondingly brief; and because there is no opportunity for significant build-up of air pressure behind the constriction, the friction will be relatively weak, or completely absent.[44]

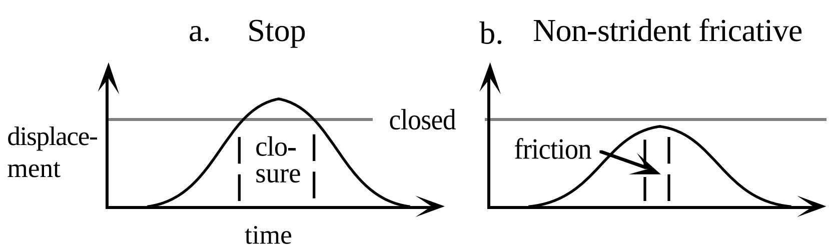


Figure 4-1. Schematic displacement-vs.-time graphs for a stop and nonstrident fricative.

In contrast, strident fricatives involve a more controlled constriction: they are typically longer than nonstridents (Nartey 1982); and sibilants in particular involve bracing of the sides of the tongue blade against the molar gumline, to produce a grooved midsagittal channel for the airflow, which, directed against the teeth, results in intense fricative

[44]Cf. Romero's (1996) observation that in Andalusian Spanish, duration of constriction appears to be the main articulatory distinction between (non-strident) fricatives and approximants. Romero's finding accords with the assumption above: a shorter interval of close constriction weakens (or, in the case of Andalusian, eliminates) the friction. However, Romero's assumption that duration, rather than constriction degree, distinguishes fricatives from approximants cross-linguistically, seems suspect. It should be noted that the Andalusian "fricatives" derive from /s/ + stop clusters, whereas the approximants derive from single stops. The greater duration of the former may therefore be due to preservation of the greater cluster duration (with greater friction as a by-product). It may well be the case that in other languages, where fricatives do not have this correspondence to underlying clusters, the distinction between fricatives and approximants would be one of constriction degree rather than, or in addition to, duration.

energy. More generally, I assume that strident fricatives (Figure 4-2) require a relatively precise, sustained close constriction, in order to generate highly turbulent airflow.[45]

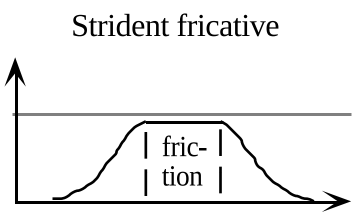


Figure 4-2. Schematic displacement-vs.-time graphs for a strident fricative.

An account of the non-stridency of spirantization outputs can now be constructed, applying the idea that gestural precision results in increased effort cost, as discussed in Chapter 2, section 3. Specifically, I assume that for the strident fricative, in order to achieve the delicate balance of holding the articulator in closely constricted position, but preventing it from going all the way to closure, *isometric tension*, i.e. exertion of force in opposition to the main constriction gesture, is required. The total effort cost of the constriction gesture plus the opposing force is greater than the effort cost of the corresponding stop.

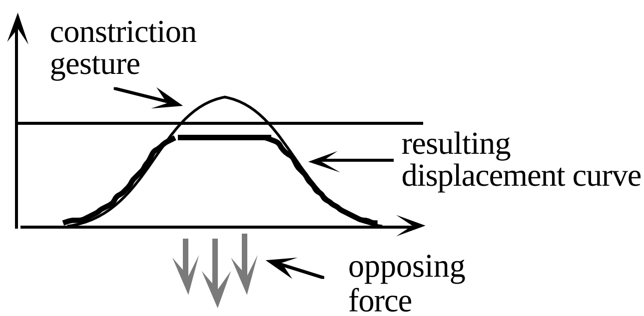


Figure 4-3. Schematic: sustained constriction achieved by isometric tension.

[45]Alternatively, strident friction might be generated by dramatically increasing airflow during the close constriction interval, e.g. by contracting the intercostal muscles to boost subglottal pressure. But the alternative strategy presumably carries its some additional effort cost as well.

The figure above, however, should not be interpreted as claiming that this opposing force necessarily takes the form of a directly opposing gesture. For example, as noted above, sibilants characteristically involve transverse stiffening of the tongue blade, which is braced against the molar gumline: in this case, the transverse bracing gesture is oblique to the direction of the main constriction gesture; nevertheless, we can speak of the extent to which the force exerted by the the transverse bracing gesture opposes the force of the main constriction. This is the sense in which we use "opposing" or "negative force" herein.

2.2. Results of the mass-spring model

Recall from Chapter 2 that the mass-spring model of consonant constriction permits the user to specify temporal and spatial targets for consonant constriction, and finds the optimal function of force against time such that the virtual articulator achieves those targets. Effort is computed by summing over the absolute values of the positive and negative forces for each timestep of the simulation. Under this model, then, the strident fricative (i.e. with a sustained partial constriction) emerges as more effortful (65.98) than the corresponding stop (60.99), as well as the nonstrident fricative (i.e. with a brief interval of close constriction) (25.64):

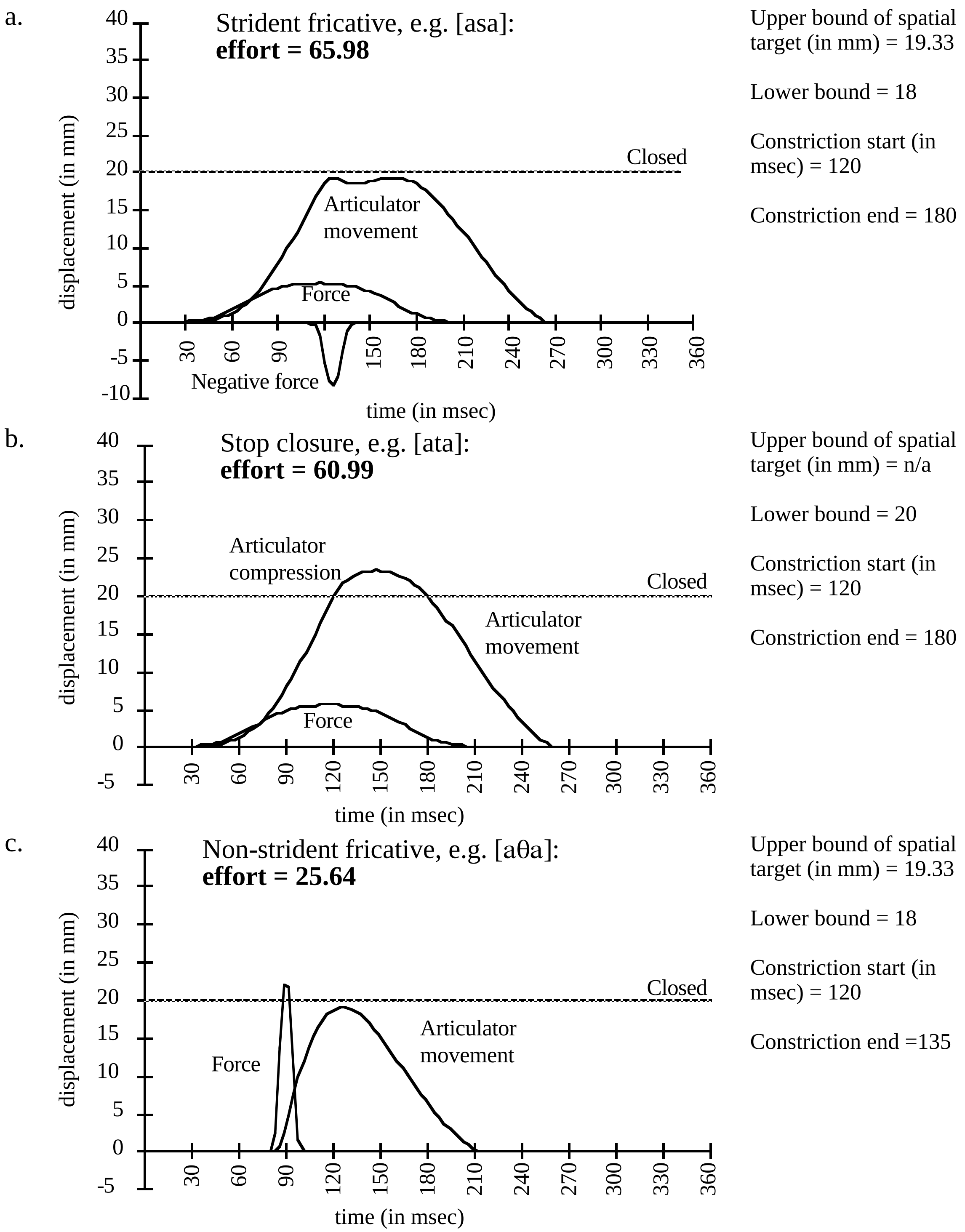


Figure 4-4. Outputs of the mass-spring model, for a stop, and strident and nonstrident fricative.

In particular, the source of the higher effort in the strident fricative (Figure 4-4a) is the extra, negative force impulse which is required to achieve a sustained partial constriction. Without this antagonist impulse, the articulator goes all the way to full closure (b). In contrast, the partial constriction in the nonstrident fricative (c) is achieved simply by throwing the articulator towards the target, and then immediately letting it return to rest position, due to the passive force of gravity and the movement spring. (Presumably, the unusually spiky force function in the nonstrident fricative is due the fact that the constriction does not need to be maintained for any appreciable length of time.) The mass-spring model thus lends support to the isometric tension hypothesis for strident fricatives, discussed above.

3. FORMAL CAPTURE OF THE PHONETIC EXPLANATION

Given the relative effort levels required for stops, strident fricatives, and nonstrident fricatives, as discussed above, if LAZY » PRES(continuant), we obtain spirantization to a nonstrident continuant [ð]:[46]

(4-2)

	Input: d	LAZY	PRES(cont)
	d	**!	
☞	ð	*	*
	z	**!*	*

Crucially, no ranking of LAZY with other constraints permits an input stop to map to a strident fricative, because the strident incurs a worse violation of LAZY than the input stop.

---

[46]Note that the assignment of multiple violation marks to particular candidates in this and subsequent tableaux does not reflect absolute effort quantification -- it merely shows how the candidates fare relative to one another with respect to violation of LAZY.

This result holds true, even if we introduce an active fortition constraint, which militates against nonstrident fricatives (motivated by the cross-linguistic markedness of these continuants, as noted in section 1). While such a constraint, if ranked above LAZY, is capable of blocking spirantization, it cannot cause the strident candidate to emerge as the winner, again because the strident incurs a worse violation of LAZY than the input stop:

(4-3)

	Input: d	*[+cont,-son,-strid]	LAZY	PRES(cont)
☞	d		**	
	ð	*!	*	*
	z		**!*	*

The unattested result could arise only if there were a constraint favoring (all manner of) fricatives over stops: in combination with the *[+cont,-son,-strid] constraint, such a constraint could induce a stop to sibilant process. But since, by hypothesis, lenition is driven by effort minimization, rather than being a sort of increase in aperture for its own sake, there would appear to be no motivation for such a constraint.

In sum, the generalization concerning the nonstridency of spirantization outputs falls out from a notion of effort minimization, formalized as the constraint LAZY, ranked relative to constraints on faithfulness to manner features, e.g. PRES(cont), or fortition constraints, such as *[+cont,-son,-strid].

4. Lenition of Affricates

Recall, however, that there is an exception to the stridency generalization, if the input is an affricate with a strident release, e.g. /tʃ/ - [ʃ], as in Florentine. Again, the mass-spring model supports the assumption that such a reduction is effort minimizing:

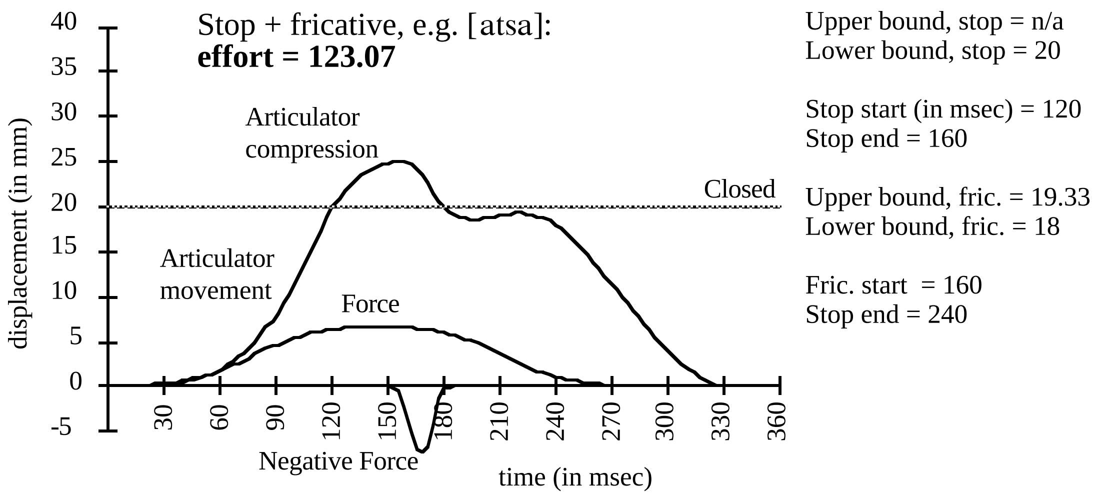


Figure 4-5. Output of mass-spring model for affricate.

Comparing the "affricate" in Figure 4-5 with the strident fricative in Figure 4-4a, it appears that the affricate is more effortful (123.07 vs. 65.98).

Now, if preservation of the underlying stridency (i.e. PRES(strid)) is ranked high, we will obtain a strident fricative as the output, hence /tʃ/ - [ʃ], as in Florentine.

(4-4)

	Input: la tʃena	PRES(strid)	LAZY	PRES(cont)
	la tʃena		****!	
☞	la ʃena		***	*
	la tena	*!	**	
	la jena	*!	*	*

Therefore, the effort-based approach correctly predicts this exception to the non-stridency generalization. This approach also predicts that affricates can lenite to approximants, or to unaffricated stops (depending on the ranking of PRES(strid) and PRES(son) relative to LAZY), since these outputs all involve less effort than the input affricate. The approximant outcome occurs in Florentine, in lower speech registers: /tʃ,dʒ/ further reduce to (unfricated) [ʃ̞,ʒ̞] (see Chapter 8). The deaffrication outcome is attested in Pennsylvania German, in which /pf/ reduces to [b] intervocalically and word-finally (Kelz 1971).

5. ANALYSIS OF ASSIBILATION

Recall from section 1.2.2 above that assibilation is restricted to contexts in which the release of the stop is inherently fricated; and that assibilation can viewed as a neutralization of this inherently fricated stop with a true sibilant affricate. The prohibition of the intermediate category, i.e. the weakly fricated stop, can be expressed as the constraint *[+fricated release, -strident]. Now, consider an input /ti/ sequence:

(4-5)

	/ti/	*[+fricated release, -strident]	PRES(strid)	LAZY
	t^si (w/ weakly fricated release)	*!		*
☞	tsi		*	**

Since this sequence cannot be realized without some release friction, either weak or strident, if the fortition constraint dominates PRES(strid) and LAZY, the output will be a strident affricate. Moreover, in section 4 we determined that a strident affricate is more effortful than a strident fricative (Figures 4-4a, 4-5). Therefore, if *[+fricated release, -strident] and LAZY both dominate PRES(cont), the output is a strident fricative:

(4-6)

/ti/	*[+fricated release, -strident]	LAZY	PRES(cont)	PRES(strid)
t^si (w/ weakly fricated release)	*!	**		
tsi		****!		*
☞ si		***	*	*
θi	*!	*	*	

Thus, an input stop can assibilate either to a strident affricate or a fricative; but such assibilation is limited to environments where the stop would inevitably be realized with some release friction.

6. SUMMARY AND DISCUSSION

In the foregoing discussion, the following effort relations were established:

(4-7) Strident Affricate > Strident Fricative > Stop > Nonstrident Fricative

From these relations, the generalization of non-stridency of spirantization outputs was obtained, using the device of conflict between LAZY and lenition-blocking constraints.

6.1. COMPARISON: STANDARD MARKEDNESS THEORY

In contrast, standard approaches to markedness cannot readily reconcile the prohibition on strident fricatives as spirantization outputs with the *general* unmarkedness of this class of fricatives.

(4-8) Non-effort-based approach:

	Input: d	*[+strident]	SPIRANTIZATION-INDUCING CONSTRAINT	PRES(cont)
	d		*!	
	z	*!		*
☞	ð			*

That is, if we posit a general markedness constraint disfavoring stridents, in order to obtain the correct output in (4-8), it is predicted that nonstrident fricatives should be preferred under all circumstances, when in fact languages display the very opposite preference, for all purposes except as spirantization outputs.

Therefore, there seems to be no recourse in such approaches except to stipulate the generalization as a condition on spirantization operations.[47] In the effort-based approach, however, we can distinguish between *articulatory markedness* (i.e. effort minimization), which favors the nonstrident continuants, and *perceptual markedness*, which favors the stridents (cf. Steriade 1995). The preferred sound in a given case will depend on whether articulatory or perceptual factors are driving the sound pattern in the relevant context. Since articulatory concerns drive lenition patterns, nonstrident continuants are correctly predicted to be only possible outcome of spirantization.

6.2. COMPARISON: FAITHFULNESS TO NONSTRIDENCY, ARTICULATORY PHONOLOGY

Alternatively, one might attribute the nonstridency generalization to faithfulness to the nonstridency of the input stop. Such an account is implicit in the framework of Articulatory Phonology (e.g. Browman & Goldstein 1990, 1992), which represents

---

[47]In OT, since operations cannot be referred to per se, this stipulation would have to take the form of a condition on mappings from input stops to output continuants.

lexical items as "scores" of articulatory gestures (I also adopt such representations, in Chapter 2, but Articulatory Phonology does not posit parallel auditory representations as I do). In particular, Articulatory Phonology attempts to analyze highly productive, phrasal, casual-speech processes in terms of a relatively restricted set of operations on these otherwise invariant gestural scores, principally modification of intergestural timing, and, more importantly for our purposes, gestural reduction. If one begins with a simple nonstrident coronal closure gesture, and reduces the magnitude of the gesture, it is not surprising that one ends up with a simple nonstrident continuant coronal gesture, rather than the more complex grooved tongue-blade configuration of a sibilant. For the same reason, one would not expect to derive a strident labiodental fricative from reduction of a bilabial closure gesture.

In order for this Articulatory Phonology-inspired account to go through, however, it must be the case that lenition processes are strictly operations of reduction. Further modifications of the underlying set of gestures must be universally prohibited, otherwise there is no reason why reduction of a coronal stop could not be accompanied by tongue-blade grooving, to enhance the audibility of the output continuant. But lenition processes can in fact involve modifications of the original gestures, beyond mere reduction. For example, Romero (1996) observes, from electromagnetic articulometry data on Andalusian Spanish lenition, that gradient reduction of constriction *degree* is accompanied by modification of the constriction *location* target: the reduced labials tend to be more retracted (closer to labiodental) and the reduced dentals tend to be more advanced (closer to interdental) than the corresponding stops (p. 62).

Giannelli & Savoia's (1979) description of Florentine Italian lenition likewise casts doubt on the characterization of lenition as pure gestural reduction. Specifically, they report that in careless (*trascurato*) speech, /ɡ/ debuccalizes to [ɦ] (i.e. "voiced h").

(17) gamba / la ɦamba 'leg / the leg'
grattare / e si ɦratta 'to scratch / (s)he scratches'

Note that Giannelli & Savoia's transcription of the consonants is extremely narrow (distinguishing, for example, three different degrees of constriction in continuant consonants; and distinctions between ordinary [h], a somewhat weaker voiceless variant [ḥ], and this voiced [ɦ]); moreover, these transcriptions are reported to be based on spectrographic analysis. Assuming, then, that Giannelli and Savoia's transcriptions are accurate, the dorsal closure gesture of the /ɡ/ is being replaced, in this debuccalization process, by a (weak) glottal abduction gesture. However, there is no reason to suppose that this glottal abduction gesture is present in the original gestural plan for the voiced stop; indeed, glottal abduction would tend to result in devoicing of the stop. Simple gestural reduction, therefore, cannot derive [ɦ] from /ɡ/. Rather, it would appear that the [ɦ] serves as a *perceptual* vestige of the input voiced stop (i.e. satisfying perceptually based faithfulness constraints, as posited under the effort-based approach to lenition). Moreover, if any process is to be viewed as part of the synchronic speech production system which Articulatory Phonology seeks to model, it is this Florentine lenition process: the variation is sensitive to speech rate and register; it applies without lexical exceptions, in phrasal domains as well as within lexical items; and it constitutes part of a more general pattern of gradient lenition (see Chapter 8). For /ɡ/ in particular, the [ɦ] allophone is but one point on a continuum of increasingly lenited allophones, from [g] to

Ø. (An additional, diachronic, case of g > h debuccalization (apparently in intervocalic position) is found in Ainu (Vavin 1993)).

Similarly, in the London and Fife dialects of English (John Harris 1990) /t/ debuccalizes to [ʔ], in intervocalic and final position. That is, loss of the coronal closure gesture (and glottal abduction gesture) is compensated for by insertion of a glottal *constriction* gesture. Though it might be suggested that the final /t/'s are already glottalized in the input in this case (since coda glottalization of /t/ is well attested, for example in American English), there is no reason to suppose that the input /t/'s have any glottal constriction component in *intervocalic* position. A further case of intervocalic stop debuccalization, in this case /k/, is found in West Tarangan (Nivens 1992). It appears from Lavoie's (1996) survey (p. 291), however, that fricatives (including voiced fricatives, which should have no significant glottal abduction) debuccalize exclusively to [h].

Table 4-6. Fricative debuccalization outcomes.

Language	Reference	Debuccalization pattern
Miami (Illinois)	Costa 1991	s,x,θ,ʃ,tʃ,ç > h /__-voi stop
Latin American Spanish	Lipski 1984	s - h /V__V and /__# in polysyllabic words
Proto-Greek	Sommerstein 1973	s >h /__V
Middle Chinese	Pulleyblank 1984	x > h(context unclear)
Páez	Gerdel 1985	x - h/V__V
Navaho	Kari 1976	x - h non-initially
Babine	Story 1984	x > h stem-finally
Canelakraho	Popjes & Popjes 1986	j,x > h initially
Pipil	Campbell 1985	w - h word-finally and /__C

Plainly, what is being preserved in these debuccalization cases is not any *gestural* component of the input stop or fricative, but one of its acoustic/auditory properties, i.e. an interval of silence or noise.

Moreover, in Kanakuru (Newman 1974), Lama (Ourso & Ulrich 1990), Limbu (van Driem 1987), Uzbek (Sjoberg 1963), and Warndarang (Heath 1980), bilabial stops reduce to the labiovelar glide, [w]: that is, as the bilabial gesture reduces, the resulting glide is perceptually enhanced by insertion of a dorsal raising/backing gesture.[48] In Nkore-Kiga (Taylor 1985), /b/ reduces to a labiodental approximant, [ʋ], a modification of the constriction location target of the underlying gesture (cf. Romero's similar finding for Andalusian Spanish, above); but, crucially, this modification does not result in a strident labiodental fricative. Finally, in Warndarang (Heath 1980), /k/ reduces to [w], neutralizing with the reduced labial: in this case, the velar glide resulting from reduction of the underlying /k/ is enhanced by insertion of a labial rounding gesture.

Thus, it is not correct to assume that lenition processes are strictly characterized as gestural reduction; rather, such reductions are commonly accompanied by gestural insertion or modification, presumably for purposes of perceptual enhancement of the lenited output. Lenition is thus more accurately characterized as *substitution of a less effortful set of gestures*, the selection of which is constrained by the hierarchy of active faithfulness and fortition constraints under a given grammar.

But perhaps the alternative account can be maintained, if we stipulate that lenition processes may modify underlying gestures beyond reduction, provided that faithfulness to

[48]See Stevens, Keyser & Kawasaki (1986) on the acoustic basis for viewing the combination of tongue-body raising/backing and labial rounding as an enhancement.

the feature [strident] is satisfied. That is, the constraint PRES(strident) is assumed to be inviolable. However, even this condition is empirically inadequate. For example, in Florentine Italian, the strident fricatives (s,z,ʃ,ʒ,f,v) reduce to corresponding approximants in casual speech, violating PRES(strident) (see Chapter 8). Furthermore, we have observed, in sections 1.5 and 5, that PRES(strident) is violable in cases of assibilation. In sum, PRES(strident) is indeed violable, provided that the output is less effortful than the input, in the case of lenition (as in Florentine), or in case the selection of the strident candidate is attributable to some constraint other than LAZY, as in assibilation. Therefore, the notion of effort minimization is indeed crucial to an adequate account of the non-stridency generalization.

Chapter 5:

Geminates

It is a well known observation that phonological processes which apply to short segments frequently fail to apply to corresponding long ("geminate") segments. For example, post-vocalic spirantization of velar stops in Tigrinya yields [ʔa-xalɨb] 'dogs' (cf. [kəlbi] 'dog'), but [fəkkərə] 'boasts', not [fəxkərə] nor [fəxxərə] (Kenstowicz 1982).[49] This phenomenon of geminate "inalterability" or "blockage" has been the subject of a number of proposals within the framework of Autosegmental Phonology, most influentially Hayes 1986 and Schein & Steriade 1986.[50] Subsequent research, however, has revealed that these proposals make incorrect predictions as to the class of processes which display inalterability (see Inkelas & Cho 1993). As Churma 1988 observes, geminate inalterability holds true as a universally inviolable condition only in the domain of *lenition* phenomena, a generalization which the classic inalterability approaches fail to capture. Moreover, as Elmedlaoui 1993 notes, within the domain of lenition phenomena, the classic approaches are insufficiently restrictive: they fail to rule out processes which specifically target geminates for lenition, e.g. /kk/ - *[xx], or which convert an underlying singleton to a lenited geminate, e.g. /k/ - *[xx]; and they fail to draw a connection between inalterability and the general markedness of "weaker" (i.e. continuant and voiced (obstruent)) geminates, whether derived via some lenition process or present underlyingly.

---

[49]Here and throughout, transcriptions have been modified to conform with IPA. For consistency, I transcribe geminates with doubling (e.g. kk, aa) rather than the length diacritic (kː, aː); this practice is without theoretical significance.

[50]See also Guerssel 1977 for a pre-autosegmental treatment.

This chapter begins with a review of previous approaches to geminate inalterability. Next, building on the work of Churma and Elmedlaoui, and the lenition surveys (see Appendix), I identify and document the following specific generalizations concerning geminates and lenition:

(5-1) a. No process converts a stop (geminate or otherwise) to a geminate with reduced oral constriction (section 2.2.1).

b. No process converts a (tautomorphemic) geminate stop to a "half-spirantized" cluster, e.g. /kk/ - *[xk] (section 2.2.2).

c. No process converts a voiceless segment (geminate or otherwise) to a voiced geminate obstruent (section 2.2.3).

d. "Partial geminates" (i.e. homorganic nasal + stop or lateral + stop clusters) behave identically to full geminates with respect to reduction of oral constriction; but, unlike full geminates, they readily undergo voicing (section 2.2.4).

e. No occlusivization nor obstruent devoicing process targets singletons to the exclusion of geminates (section 2.2.5).

f. The presence of a geminate continuant consonant, or voiced geminate obstruent, in the segment inventory of a language (whether derived or underlying) implies the presence of a corresponding non-continuant or voiceless geminate, respectively (section 2.2.6).

The generalizations in (5-1) are shown to follow from the effort-based approach to lenition, outlined in Chapter 1, coupled with plausible phonetic assumptions concerning the effort required to produce geminates. Specifically:

A. More effort is required to produce a geminate continuant consonant than a geminate stop (the inverse of the situation in singletons), due to the isometric tension required to maintain a prolonged steady-state partial constriction, cf. the discussion of strident fricatives in Chapter 4. Therefore it is never optimal to reduce the oral constriction of a geminate stop (except by shortening it) (section 3.1.2).

B. More effort is required to produce a voiced geminate obstruent than a voiceless geminate (the inverse of the situation in singleton obstruents in medial position), due to the aerodynamic conditions required to sustain voicing, cf. Ohala 1983. Therefore it is never optimal to voice a geminate obstruent (section 3.1.4).

In section 4, I propose an account of the (often distinct) behavior of heteromorphemic geminates, relying upon paradigmatic (output-output) faithfulness constraints (e.g. Benua 1995, 1997; Flemming 1995). Finally, in section 5, the effort-based approach is compared with a non-effort-based approach to these generalizations.

1. Previous Treatments of Geminate Inalterability

1.1. The classic autosegmental proposals. Steriade 1982, Hayes 1986, and Schein & Steriade 1986 attempt to derive geminate inalterability effects largely from general assumptions of the theory of representations of Autosegmental Phonology (e.g. Goldsmith 1976, Clements and Keyser 1983).

```
(5-2) a. True geminate:  X   X    b.  Singleton:  X     c.   Heteromorph. geminate:   X + X
                          \ /                     |                                   |   |
                        Melody                  Melody                            Mel. Mel.
```

In particular, "true" geminates (5-2a) can be distinguished from both singletons (b) and heteromorphemic ("fake") geminates (c), in that true geminates involve *multiple association* of the melody (featural content of the segment) to the segment-timing or prosodic units which dominate it. Hayes attributes inalterability to the following notational convention:

(5-3) *Linking Constraint*. Association lines in structural descriptions are interpreted as exhaustive.

Thus, a rule such as Tigrinya spirantization (5-4a) cannot apply to a geminate, because the structural description of the rule refers to a single association line between the target dorsal consonant and its timing unit, whereas a geminate is associated with two timing units (5-4b).

```
(5-4)  Tigrinya            a.  V      C            b.  C   C
       spirantization:         | . - ' |                \  /
                            +cont  dors                 dors
```

Schein & Steriade propose a somewhat more narrowly drawn convention:

(5-5) *Uniform Applicability Condition ("UAC")*. Given a node *n*, a set S consisting of all nodes linked to *n* on some tier T, and a rule R that alters the contents of *n*: a condition in the structural description of R on any member of S is a condition on every member of S.

The principal difference between the two conventions is that the Linking Constraint blocks rule application when the target or trigger is a geminate; whereas the UAC blocks only when the target is a geminate, by virtue of the "alters the contents" clause. However, both approaches focus upon the *representational* distinction between single and multiple autosegmental association to block certain rules from applying to geminates. Furthermore, both approaches elegantly handle the distinct behavior of tauto- and heteromorphemic geminates, e.g. in Tigrinya, where the first half of heteromorphemic geminates undergo spirantization, just like singletons (see sections 2.2.2, 4): heteromorphemic geminates are singly linked (5-2c), just like singletons (b).

Neither approach, however, draws a connection between inalterability and lenition phenomena. Neither approach prohibits rules which specifically target geminates for full or partial spirantization or voicing. And neither approach draws a connection between inalterability effects and the general markedness of geminate continuant consonants and geminate voiced obstruents, as reflected in segment inventories. Rather, these approaches predict that inalterability effects are tied to what Schein & Steriade call "structure-dependent" rules, which refer to information on both melodic and timing-unit tiers. Such rules necessarily refer to the linkages between these tiers, thus invoking blocking by the Linking Constraint or UAC. Inkelas & Cho (1993), however, demonstrate that this prediction is false. Syllabification processes (whether formalized in terms of rules or constraints, cf. Itô 1986) refer to prosodic and melodic information, and thus should invariably display inalterability effects due to the Linking Constraint.[51] Yet in Korean, for example, the rule or constraint which eliminates velar nasals in onset position applies

[51]Indeed Itô (1986) relies upon the Linking Constraint to account for geminates' immunity to a phonotactic coda place constraint in Japanese and other languages. Note that syllabification rules or constraints do not invoke inalterability under Schein & Steriade's UAC, however, since the rules in question, as standardly conceived, do not alter the contents of the multiply linked node.

to singletons and geminates alike (e.g. [kaŋ] ('river'), but *[aŋa], *[aŋŋa]). Similarly, geminates are never immune to rules or constraints requiring sonority sequencing within coda and onset clusters: thus in Latin, a coda [kl] cluster is ill-formed, whether the [l] is a singleton (e.g. *[akl.ta]) or the first half of a geminate (e.g. *[akl.la]). Inkelas & Cho further note that the Linking Constraint and UAC do not hold true for "long-distance" (i.e. segmentally non-adjacent) multiple linking, as Hayes acknowledges. For example, tones which are associated to multiple syllables are not typically immune to processes affecting singly-linked tones.

Finally, Inkelas & Cho observe that it is frequently possible to formulate rules either as structure-dependent or segmental, so as to place the rule within or outside the purview of the Linking Constraint or UAC; thus, the predictions these approaches make, as to which processes will or will not exhibit geminate inalterability effects, are not as strong as initially meets the eye (as Hayes (p. 344) acknowledges).[52] Indeed, this criticism can be taken considerably further: to the extent that these approaches attempt to constrain possible individual *rules*, without thereby constraining *sound systems*, they are empirically vacuous. Thus, for example, nothing in these approaches rules out "Zigrinya," a hypothetical language with a general post-vocalic spirantization rule, as in Tigrinya, *plus* a rule specifically spirantizing post-vocalic geminates. Zigrinya thus achieves by a combination of licit rules the same unattested sound pattern which the Linking Constraint and UAC purport to rule out.

---

[52]For example, Schein & Steriade analyze Turkish depalatalization as changing a coda velar to [+back], thereby invoking the UAC, and correctly accounting for the fact that this depalatalization does not affect a multiply-linked [-back] specification. However, it is equally possible (and, as Inkelas & Cho argue, more elegant) to express this as a rule delinking a [-back] specification from a coda velar; this delinking rule does not invoke the UAC, since it does not "alter the contents" of the multiply-linked node.

1.2. GENERALIZED INALTERABILITY. In contrast to the foregoing approaches, Inkelas & Cho 1993 challenge the basic assumption that geminate inalterability is a discrete phenomenon. Inkelas & Cho observe that the blocking of phonological rules is not a phenomenon confined to geminates. For example, the "opaque" behavior of certain vowels in harmony processes, and lexical exceptionality, are also examples of rule blocking. They further identify *prespecification* as the generalized blocking mechanism, and predict that it is the class of feature-filling rules which systematically display inalterability effects (whether geminate inalterability, or other blocking phenomena). For example, under their analysis, Latin coda [l] velarization involves a rule assigning onset [l] a [-back] specification. This rule applies to the geminates, because they are in onset position (it does not matter that they are also in coda position). Other (i.e. coda singleton) laterals undergo a context-free feature-filling rule making laterals [+back]. The onset rule is ordered before the context-free rule, by virtue of the Elsewhere Condition (Anderson 1969, Kiparsky 1973). But the context-free rule is blocked from applying to the geminates (or other onset [l]'s), because they are already specified for [back]. Other prespecifying rules may specifically target geminates: e.g. in Berber, a rule specifies geminate consonants as [-cont], which bleeds an "elsewhere" rule assigning [+cont].

Despite their heavy reliance on rule ordering and underspecification, Inkelas & Cho's notion of blocking through prespecification translates rather neatly into OT, as blocking through higher-ranked constraints (cf. the discussion of blocking and triggering in Prince & Smolensky 1993, chs. 3-4). For example, Inkelas & Cho's analysis of Latin [l] velarization can be restated as follows:

(5-6)

		ONSET L: *[+back, +lateral] in onset	ELSEWHERE L: *[-back,+lateral]
☞	al.la		*
	ał.ła	*!	
	al.ta		*!
☞	ał.ta		

But while Inkelas & Cho's approach, particularly in its OT reformulation, gives us a general mechanism for the blocking of phonological processes, it does not account for the generalizations identified in (5-1), which specifically concern geminates and lenition. Inkelas & Cho, acknowledging Churma's (1988) observations along these lines, attempt to draw a connection between geminate inalterability and lenition, as follows. They assume, following Hyman 1985 and Hayes 1989, that (underlying) geminates are linked to moras in underlying representation; whereas other segments must be assigned moras by rule. Moraification rules often impose minimum sonority requirements on coda consonants, e.g. Hausa, which requires codas to be [+sonorant]. The geminates escape this condition of the moraification rule, however, because they are already moraified. However, this analysis only extends to cases of *coda* lenition. In Tigrinya and Hebrew, geminate inalterability effects are observed, although lenition occurs in intervocalic onset, as well as coda, position (i.e. post-vocalically). For these sorts of cases, Inkelas & Cho must stipulate that a rule specifying geminates as [-cont] has priority over a rule specifying post-vocalic obstruents as [+cont] (Elsewhere Condition ordering does not obtain, as there is no subset relation between the structural descriptions). Inkelas & Cho predict that the two cases are typologically distinct: that coda lenition processes are systematically blocked in geminates, whereas other processes vary in this regard. But no such distinction emerges from the lenition survey (section 2 below). Rather, the data support the stronger generalization that spirantization and voicing processes *never* apply to geminates, regardless of the conditioning environment. Furthermore, Inkelas & Cho's analysis rests upon the problematic assumption that lenition can be equated with sonority promotion, cf. Chapter 1.

In a footnote to their conclusion, however, Inkelas & Cho (1993: 569) observe:

> Although they are arbitrary under our analysis, certain of the allophonic alternations involving geminates have a plausible phonetic basis. For example, the fact that voicing is harder to maintain over longer durations might motivate the distribution of [voice] in Berber ... in which singletons but not geminates are voiced.

This is precisely the sort of analysis provided under the effort-based approach, in section 3 below.

2. GENERALIZATIONS

2.1. THE NON-UNITY OF INALTERABILITY EFFECTS. As a preliminary matter, note that I am *not* claiming that geminate inalterability effects are to be found only in the domain of lenition processes; nor is it my goal to develop a unified account of all inalterability effects. For example, rounding harmony, i.e. unbounded extension of a lip rounding gesture, is not plausibly regarded as a species of lenition. Yet, in Maltese, rounding harmony fails to apply to long vowels, e.g. /kitbuulik/ - [kitbuuluk] ('he wrote it to you'), but /ʃurbitiilim/ - [ʃurbutiilim] ('she drank it (fem.) from them'); and virtually the same pattern obtains in Tigre (McCarthy 1979, Schein & Steriade 1986). Such resistance to rounding neutralization is plausibly analyzed in terms of interaction between a constraint which induces rightward spreading of [round] (e.g. ALIGN(rnd,R), cf.

McCarthy & Prince 1993) and a *positional faithfulness* constraint (see Jun 1995, Steriade 1995, Beckman 1997), specifically referring to vowel features in long-vowel position:

(5-7)

		PRES(rnd/long V)	ALIGN(rnd,R)	PRES(rnd)
	ʃurbitiilim - ʃurbitiilim		***!	
☞	ʃurbitiilim - ʃurbutiilim		**	*
	ʃurbitiilim - ʃurbutuulum	*!		***
	kitbuuliik - kitbuulik		*!	
☞	kitbuuliik - kitbuuluk			*

Presumably this positional faithfulness constraint reflects the greater perceptibility of vowel quality distinctions in long vowels, cf. Kaun 1994, Jun 1995, Flemming 1995. For our purposes, the important observation is that the blocking effect under this analysis is *violable*: for under the opposite ranking of PRES(rnd/long V) and ALIGN(rnd,R), no geminate blocking obtains. This prediction is confirmed by Khalkha Mongolian (Street 1962, Schein & Steriade 1986), in which [round] (and [back]) harmony targets long and short vowels alike: [aabaas] ('father-abl.'), [odoogoos] ('now-abl.'), [gerees] ('house-abl.'), [tœrœœs] ('state-abl.').

In contrast, an examination of the behavior of geminates under lenition reveals cross-linguistically robust generalizations, namely the geminate lenition generalizations in (5-1), documented below. I take this as motivation for a distinct account of geminate inalterability under lenition: it is the goal of this chapter to develop such an account. I shall not address the question of whether the remaining non-lenitional geminate inalterability effects can be handled exclusively in terms of the positional faithfulness approach sketched in (5-7), or whether there may be yet further sources of geminate inalterability effects.

Finally, note that, in assuming that geminate inalterability is not a unified phenomenon, I am not diverging from the consensus of previous approaches. For example, Schein & Steriade attribute the Maltese and Tigre blocking not to their general principle of geminate blockage, but to a language-specific metrical condition on the harmony rule: [round] can only spread rightward within a foot; and the long vowel serves as the head of a new foot. Churma 1988 appears to concur in this non-unified treatment; for he claims that harmony processes never display ("genuine") inalterability effects, despite his awareness of the Maltese and Tigre cases. More explicitly, Inkelas & Cho (1993: 557) take the position that "the mere survival of a geminate in a language which has a rule that in principle could affect that geminate does not necessarily mean that a genuine case of geminate blocking has occurred." They identify a large class of "pseudo-inalterability" effects, which they attribute to counterfeeding rule ordering, or to the fact that, for various reasons, geminates fail to meet the structural description of the rule in question.[53]

2.2. DOCUMENTING THE GENERALIZATIONS

2.2.1 NO ORALLY REDUCED GEMINATE STOPS. *No process converts a stop (geminate or otherwise) to a geminate with reduced oral constriction.* This generalization, together with several of the generalizations below, is a somewhat more specific restatement of Churma's (1988) original claim that "aside from degemination, no weakening process may affect a geminate consonant."[54]

---

[53]Only Hayes 1986 appears to assume that all instances of inalterability require a unified treatment. The price Hayes pays for this broad-scope theory is a weakening of its predictive power: for he acknowledges (p. 344) that under his approach it is impossible to identify the necessary conditions for a process to display geminate inalterability.

[54]That some "weakening" (i.e. effort-reducing) processes, other than spirantization and obstruent voicing, do apply to geminates is documented in section 2.2.1.4 below. As Elmedlaoui 1993 notes, Churma's claim bears some resemblance to the earlier "Inertial Development Principle" of Foley 1977, which states, in essence, that "weak" segments are preferentially targeted by weakening processes, and "strong" segments

2.2.1.1. SPIRANTIZATION. A classic example of such geminate resistance to oral reduction is the Tigrinya spirantization pattern, alluded to in the introduction to this chapter, and more fully exemplified below (data from Kenstowicz 1982):

(5-8)	a.	kəlbi	'dog'
		ʕarat-ka	'bed-2sg.m.'
		qətəl-ki	'kill-2sg.f. perfect'
	b.	kətəma-xa	'town-2sg.m.'
		mɨrax-na	'calf-3sg.f.'
		ʔa-xalɨb	'dogs'
		ʔɨti xalbi	'the dog'
	c.	qətəl-a	'kill-3pl.f. perfect'
		tɨ-χətl-i	'kill-2sg.f. imperfect'
	d.	fəkkərə	'boasts'
		qətəl-na-kka	'we have killed you (masc.)'

That is, post-vocalic velars (5-8b) and uvulars (c) spirantize, but geminates (d) remain stops. Tiberian Hebrew, with a similar pattern of post-vocalic spirantization of labials, velars, and (non-emphatic) coronals, does indeed display alternations between geminate stops and fricatives (data from Elmedlaoui 1993).[55]

---

for strengthening processes. But since Foley explicitly refuses to attribute any consistent phonetic content to his notions of weakening or strengthening, it is difficult to evaluate the empirical predictions which follow from this principle.

[55]Previous treatments of blocking of lenition in the emphatic (pharyngealized) stops have assumed some property, shared with geminates, that makes them immune to spirantization namely [+tense] (Prince 1975) or [-released] (McCarthy 1981), but without presenting any argument that the emphatics actually had these phonetic properties. Moreover, these previous treatments conflate a cross-linguistic generalization

(5-9)

Causative Perfect	Basic Perfect	Gloss
zikkeer	zaaxar	remember
kippeer	kaaɸar	cover
biddeel	baaðal	separate
pitteeah	paaθah	open
piggeeʃ	paaɣaʃ	meet

But, crucially, the spirantized class is limited to surface singletons, thus illustrating a corollary generalization: geminate stops can undergo oral reduction, but only if they surface as singletons.[56] As Elmedlaoui (1993) observes, the inalterability generalization thus properly focuses not on whether geminates are licit *inputs* to spirantization processes, but whether spirantization processes may yield *output* geminates. Further examples of geminate blocking of reduction of stops to continuants appear in the following table.

(geminate inalterability under lenition) with a language-specific blocking effect (compare lenition of /q/ in Tigrinya vs. blocking of /q/ lenition in Hebrew (cf. McCarthy 1988, arguing that the uvular [q] is a pharyngealized dorsal)). Instead, I suggest that this blocking is attributable to the need to avoid neutralization with the true pharyngeals ([ħ] and [ʕ]), which are distinct phonemes in Tiberian Hebrew. Spirantization of a pharyngealized stop would yield a pharyngealized continuant with weak (or no) coronal or dorsal friction, bringing the output too close to a true pharyngeal continuant. (Indeed, in Biblical Aramaic, /dˤ/ neutralized to [ʕ], see Elmedlaoui 1993: 143.) Formally, then, I posit a fortition constraint that rules out the intermediate category:
*[+cont,-strid,+cons,phar] = "no non-strident secondarily pharyngealized continuants"
under the assumption that the true pharyngeals are [-cons]). Ranked above LAZY, this constraint blocks spirantization of the pharyngealized consonants.

[56]Under traditional phonemic or generative analyses, the target consonant is underlying a singleton, and the gemination in the 'causative' column is derived, in which case Tiberian Hebrew does not exemplify degemination-cum-spirantization of underlying geminates. In the OT framework, however, since all inputs are admitted by GEN ("Richness of the Base," Prince & Smolensky 1993, ch. 9), the systematic absence of geminates in the 'basic perfect' column cannot be attributed to the absence of geminate inputs; rather, there must be an active constraint prohibiting geminates in perfect forms, just as there is an active constraint requiring medial geminates in the 'causative' forms. That is, if an input were to contain a geminate stop, it would degeminate and spirantize on the surface.

Table 5-1. Blocking of spirantization in geminates.

Language	Reference	Description of process
Florentine Italian	Giannelli & Savoia 1979	Lenition (ranging from spirantization to complete elision, depending on rate and register) blocked in geminate obstruents and non-continuants
Hausa	Klingenheben 1928	b,d,g - w,r,w in coda, blocked in geminates
Malayalam	Mohanan 1986	Stops - approximants (or apical tap) in the context /[+son,-nas]__V, blocked in geminates.
(Proto-) Berber	Saib 1977	Stops - fricatives (context-free), blocked in geminates[57]
Tamil	Christdas 1988	Voicing and spirantization in medial position, blocked in geminates
Tiberian Hebrew	Malone 1993	Post-vocalic non-emphatic stops spirantize, blocked in geminates
Tigrinya	Kenstowicz 1982	Post-vocalic velars and uvulars spirantize, blocked in geminates
Tümpisa Shoshone	Dayley 1989	Spirantization, flapping blocked after a homorganic nasal and in geminates; voicing blocked in geminates

Note that inalterability under spirantization holds true for geminate nasal as well as oral stops, as seen (non-vacuously) in Tümpisa Shoshone:

[57]At some point prior to Modern Berber, this spirantization pattern came to be reanalyzed as a process occlusivizing geminate fricatives, see Saib 1977.

(5-10)	a.	siimɔɔtɨ	-	siiw̃ɔɔɾɨ	'ten'
		senu	-	sej̃u	'therefore'
	b.	tapettʃi	-	taβettʃi	'sun'
		tsitoohi	-	tsiðoohi	'push'
		puhakantɨ	-	puhaɣandɨ	'shaman'
	c.	uttunna	-	uttunna	'to give'
		kimmakinna	-	kimmaɣinna	'to come here'
	d.	patɨasɨppɨ	-	paɾɨasɨppɨ	'ice'
		uttunna	-	uttunna	'to give'
		punikka	-	punikka	'see, look at'
	e.	taziumbi	-	taziumbi	'star'
		ɨntamɨʔi	-	ɨndaw̃ɨʔi	'your little brother'
		tɨppisihpuŋki	-	tɨppiʃiɸuŋgi	'stinkbug'

Non-initial singleton nasals spirantize (5-10a), as do oral stops (b); but this lenition is blocked in geminate nasal (c) and oral stops (d), as well as partial geminate clusters (e).[58] (The Tümpisa Shoshone lenition pattern is discussed more fully in Chapter 7.)

2.2.1.2. FLAPPING. Tümpisa Shoshone (and Hausa) further demonstrate that the geminate inalterability effect is not limited to spirantization per se: geminate inalterability also obtains under *flapping* (i.e. reduction of closure duration in coronal stops, see Inouye 1995). It might be objected that this observation is trivial, true by definition; that the closure duration of stop cannot be radically temporally reduced, as in a flap, and still remain a geminate. However, it is logically possible that a flapping process applying to a

[58]Coronals lenite by flapping rather than spirantizing when following a back vowel, see Chapter 7.

geminate stop would yield a long trill; but such a result appears to be unattested. Moreover, the same effect shows up in partial geminates in Lamani (Trail 1970): flapping is blocked after a homorganic nasal or lateral.

2.2.1.3. GENERALIZATION. In contrast to this wealth of cases showing geminate inalterability under processes of spirantization, flapping, and reduction to approximants, cases where these processes apply to geminates (without concomitant degemination) appear to be completely unattested, based on the previous inalterability literature, and the lenition surveys. Indeed, far from reducing their oral constriction, there is a positive tendency for geminate consonants to occlusivize, see section 2.2.6.

To my knowledge, the only ostensible (partial) counter-example to geminate inalterability under spirantization involves a detail of the Florentine Italian facts (Giannelli & Savoia 1979). Singleton intervocalic stops lenite, from fricatives all the way to Ø, depending on speech rate and register, particularly in intervocalic position:

(5-11)

	Slow/ Careful	*Moderate/ Natural*	*Fast/ Careless*	*Extremely Fast/Careless*	
/la tavola/	la θavola	la θ̞avola	la (ð̞)aol̞a	la aol̞a	'the table'
/e dɔrme/	e dɔrme	e ð̞ɔrme	e ð̞ɔrm̞e	e ɔɹm̞e	's/he sleeps'

This spirantization (and further reduction) is generally blocked in geminate stops; however, at the fastest rate and lowest register of speech, in intervocalic position, even geminates can spirantize, to very close fricatives.[59]

---

[59]Such a counterexample might be dismissed as mere "phonetics," beyond the purview of phonological theory. However, since my approach recognizes no modular distinction between phonological and phonetic processes, I cannot avail myself of this traditional "out."

(5-12)		*Slow/ Careful*	*Moderate/ Natural*	*Fast/ Careless*	*Extremely Fast/Careless*	
	/bɾutto/	[bɾutto]	[bɾutto]	[bɾutto]	[bruθ̞θ̞o]	'ugly'
	/fɾeddo/	[fɾeddo]	[fɾeddo]	[fɾeddo]	[fɾeð̞ð̞o]	'cold'

However, it is not clear, despite the transcription, that these spirantized segments are in fact phonetically geminates. Giannelli & Savoia give no data on the actual duration of these consonants; but as this is essentially a very-fast-speech phenomenon, it is unlikely that the duration of the fast-speech “geminate" approaches the typical duration of a geminate in slow or normal speech.[60] Assuming that the lenited geminates have in fact degeminated, in this phonetic sense, they do not constitute a counterexample to the generalization. On the contrary, these facts provide a striking example of geminate resistance to lenition. In this dialect, in which all singleton stops reduce to weak approximants or Ø, even in normal speech, geminates do not spirantize at all, except in the fastest speech style, when they are presumably no longer realized with typical geminate duration.

2.2.1.4. LENITION OF GEMINATES OTHER THAN REDUCTION OF ORAL CONSTRICTION IN STOPS. Finally, note that the generalization distinguishes between reduction of oral constriction in geminate stops and other forms of lenition. It has already been noted that geminates can lenite by degeminating, and that degemination potentiates

[60]Giannelli & Savoia could reasonably transcribe these spirantized segments as "geminates," notwithstanding their *phonetic* degemination, i.e. a substantial reduction in their duration, because they do not neutralize with the category of short consonants. In Florentine the consonant “length” contrast is supported by at least three cues in addition to consonant duration itself: (a) the shortened duration of the vowel that precedes the geminate (Smith 1992); (b) in the case of the voiceless geminate, an aspirated release, as Gianelli & Savoia note; and (c) reduced acoustic energy compared to corresponding singletons, due to the more fortis constriction in the geminates (i.e. the lenited geminates are near-stops whereas the lenited singletons are weak approximants or Ø).

further lenition in Tiberian Hebrew (5-9) and Florentine (5-12). Hebrew also contains a case of degemination tout court: "guttural" (pharyngealized) consonants degeminate, context-free (Hayes 1986). Although such degemination does constitute (temporal) reduction, the output ceases to be a geminate, and thus the NO ORALLY REDUCED GEMINATE STOPS generalization is maintained.

Geminates also commonly undergo loss of a distinct release of the first half of a geminate, e.g. in English /bʊk/ (with optionally released [k]) + /keis/ - [bʊk̚kʰeis] ('book-case'). This elimination of the opening gesture is clearly a species of articulatory reduction, although this occurs so ubiquitously in geminates that its status as a lenition process is easily overlooked. This loss of release presumably lies behind Ancient Greek “deaspiration" of the first half of a geminate stop (Hayes 1986). For aspiration (in the typical sense of *post*-aspiration, i.e. long lag voice onset time) is a property of the stop's release; thus an unreleased stop cannot bear (post-)aspiration.

The distinction between oral reduction of stops and other forms of lenition also appears in Florentine rhotic reduction (Giannelli & Savoia 1979). Florentine, like many Romance dialects, has a contrast between a long alveolar trill (e.g. [korriðojo] 'corridor') and a short trill or tap (e.g. [la ʃe{ɾ/r}a] 'the wax'). In natural speech styles, both the long trill (e.g. [korriðojo] 'corridor') and the short trill or tap (e.g. [la ʃe{ɾ/r}a] 'the wax') optionally lenite to approximants, without any temporal reduction of the longer rhotic: [koɹɹiðojo] vs. [ʃeɹa].[61]

[61]This reduction is not restricted to extremely fast speech; moreover, the two lenited rhotics appear to be distinguished solely by duration; thus it does not seem plausible to claim here, as I did with regard to spirantization, that the lenited geminate is in fact phonetically degeminated.

2.2.2. NO HALF-SPIRANTIZATION. *No process converts a (tautomorphemic) geminate non-continuant to a "half-spirantized" cluster, e.g. /kk/ - *[xk].* We have already seen in the previous section that spirantization processes are no more able to yield half-spirantized clusters than they are able to yield fully spirantized geminates. Thus, in Tigrinya, [fəxkərə] and [fəxxərə] are both equally impossible outputs for /fəkkərə/. More generally, cases of spirantization of the first half of a (tautomorphemic) geminate appear to be unattested, based on the previous inalterability literature and the lenition surveys. In heteromorphemic geminates, however, half-spirantization is attested, to wit, in Tigrinya: /mɨrak-ka/ - [mɨraxka] ('calf-2sg.m.'). On the other hand, this distinct behavior of heteromorphemic geminates under spirantization in Tigrinya is not universal: in Tiberian Hebrew, heteromorphemic geminates resist spirantization just as the tautomorphemic geminates do: e.g. [kaarattii] ('I cut'), cf. [kaaraθ] ('he cut').

Furthermore, it is necessary to distinguish between half-*spirantization* (in the narrow sense of reduction to a fricative), and half-*gliding* of geminates, which is attested in Maxakalí (Gudschinsky, Popovich & Popovich 1970): /mattɨk/ - [m^{b}aətɨx] ('happy'), /kaktʃoppit/ - [kakʃoəpiyə] ('boy'), /kɨtʃakkɨk/ - [kaʃaikɨx] ('capybara (type of rodent)'). For our purposes, the crucial observation is that the Maxakalí vocoid corresponding to the first half of the geminate is not a *steady-state constriction*, but a (somewhat attenuated) transition from the vowel into the following (singleton) obstruent.[62] It is also necessary to distinguish half-spirantization from half-debuccalization, attested in the Icelandic process of "pre-aspiration," Thráinsson 1979, whereby voiceless geminate stops reduce to h + stop clusters (e.g. /kappi/ - [kahpi] ('hero'). As Thráinsson observes, this is simply

[62] Gudschinsky et al., p. 77, explicitly describe these vocoids as "phonetic transition phenomena."

degemination of the oral constriction gesture, leaving the long glottal abduction gesture unchanged.

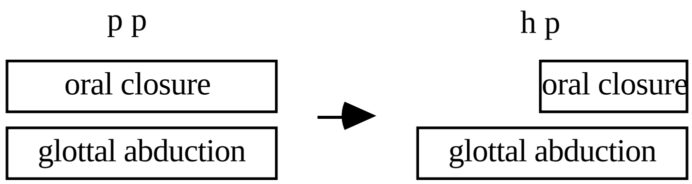


Figure 5-1. Icelandic geminate preaspiration as shortening of oral closure gesture.

Since the oral constriction degeminates, this process (vacuously) conforms to the NO HALF-SPIRANTIZATION generalization, as well as the NO ORALLY REDUCED GEMINATE STOPS generalization.

2.2.3. NO VOICING OF GEMINATES. *No process converts a voiceless segment (geminate or otherwise) to a voiced geminate obstruent.* Blocking of voicing in geminate obstruents has already been exemplified in the Tümpisa Shoshone data (5-10). That is, all obstruents undergo voicing, except in utterance-initial position, and in (full) geminates (Dayley 1989).[63] Additional examples of geminate inalterability under voicing include:

[63] In addition, utterance-final syllables optionally devoice in their entirety, e.g. $[\ldots$tɨppiʃiɸuŋki̥$]_{\text{Utterance}}$ ('stinkbug').

Table 5-2. Blocking of voicing in geminates.

Language	Reference	Description of process
Berber	Saib 1977	Pharyngealized obstruents - voiced (context-free), blocked in geminates
Cuna	Sherzer 1975	Voicing in medial position, blocked in geminates
Florentine Italian	Giannelli & Savoia 1979	In fast/casual speech styles, voiceless stops, which otherwise spirantize to voiceless fricatives or approximants, further reduce to voiced approximants; this is blocked in geminates.
Gallo-Romance	Bourciez & Bourciez 1967	Sound change: intervocalic /t/ underwent voicing, while /tt/ degeminated without voicing
Malayalam	Mohanan 1986	Stops become voiced in the context /+son__V or /+nas__; blocked in geminates
Somali	Armstrong 1964	Intervocalic voicing, blocked in geminates
Tamil	Christdas 1988	Voicing (and spirantization) in medial position, blocked in geminates
Tümpisa Shoshone	Dayley 1989	Non-initial obstruents are voiced, blocked in geminates

In contrast, voicing processes which do apply to full geminates appear to be unattested, based on the previous inalterability literature, and the lenition surveys. See also Hock 1991, who concurs that such processes are unattested, but views this as an accidental gap. Indeed, far from undergoing voicing processes, geminate stops show a positive tendency to devoice, as discussed in section 2.2.6 below.

An interesting question is whether there is a “no half-voicing" generalization, paralleling the NO HALF-SPIRANTIZATION generalization above. Dayley's (1989) description of Tümpisa Shoshone indicates that geminates can indeed be "split" with respect to voicing (contrary to the predictions of the autosegmental geminate inalterablity accounts).

(5-13)	a.	tahaɸi̥	'snow'
		huβiaɾɨxi̥	'sing'
		peθɨ̥	'daughter'
		mõz̥o̥	'whiskers'
		puhaɣãnt̥ɨ	'shaman'
	b.	pũñĩkkḁ	'see, look at'
	c.	kĩmmãɣĩnn̥ḁ	'to come here'
		uttũnn̥ḁ	'to give'
		suɾɨ̃mm̥ɪ̥	'those'
		pɨe ðuɣʷãnni j̃ããɣĩnn̥ḁ	'it's already getting dark'

Specifically, in utterance-final position, the final vowel and the preceding consonant, are optionally realized as voiceless (5-13a), whereas the consonants would otherwise be voiced in non-initial position (see Chapter 7). Geminate obstruents (b) are predictably voiceless in all contexts. The important point is that the geminate nasals (c) are split, by application of this devoicing process, into voiced and voiceless components. Thus, geminate inalterability does not appear to be a generalization about multiply linked nodes, but rather about oral reduction and voicing of consonants with prolonged oral constriction.

Armstrong's (1964) description of Somali further suggests that partial voicing of geminate obstruents can occur. Armstrong observes that the Somali geminate stops, which she transcribes as voiced, in fact "do not sound fully voiced," and in some cases are completely voiceless. That is, the geminates in question are (somewhat variably) realized as partially voiced/partially devoiced. This phonetic description comports well with a characterization of these Somali geminates as *passively devoiced;* that is, voicing ceases roughly 60 msec into the closure, not due to any active adjustment of the glottis, but due to a build-up of oral air pressure behind the closure, which makes continued transglottal airflow, hence voicing, impossible. An analysis of passive devoicing in geminates is presented in section 3.1.4.

2.2.4. NO REDUCTION OF PARTIAL GEMINATES. *"Partial geminates" (i.e. homorganic nasal + stop or lateral + stop clusters) behave identically to full geminates with respect to reduction of oral constriction; but, unlike full geminates, they readily undergo voicing.* By "partial geminates," I mean adjacent consonants which share an oral constriction gesture, where the first consonant does not have a distinct release: typically, this includes homorganic nasal + stop or lateral + stop clusters (but it excludes flap + stop clusters, because the flap is released).[64] Blocking of spirantization in partial geminates has already been discussed in connection with Tümpisa Shoshone (5-10e). This inalterability effect is further exemplified in Spanish (Harris 1969):

---

[64]This view may be contrasted with the autosegmental idea that a cluster with multiple linking of any feature-geometric node is a partial geminate, e.g. [mn] (nasality), [sk] (voicelessness), or (ws] (continuancy).

(5-14)

a.	aβa 'bean'	aða 'fairy'	aɣa 'make'
	kalβo 'bald'	--	alɣo 'something'
	aβla 'speak'	aðlateɾes 'lackies'	aɣlomeɾaɾ 'to cluster'
	aɾβol 'tree'	aɾðe 'burn'	aɾɣamasa 'mortar'
	aβɾa 'will have'	paðɾe 'father'	aɣɾio 'sour'
	xajβo (no gloss)	najðen 'nobody'	kajɣa 'fall'
	aβjeɾto 'open'	aðjestɾaɾ 'to guide'	siɣjendo 'following'
	ewβolja (no gloss)	dewða 'debt'	sewɣma 'zeugma'
	aβwelo 'grandfather'	aðwana 'customhouse'	aɣweɾo 'fortune-teller'
	aðβeɾso 'unfavorable'	aβðomen 'abdomen'	suβɣlotal 'subglottal'
	suβmaɾino 'submarine'	aðmiɾasjon 'admiration'	diaɣnostiko 'diagnostic'
b.	bomba 'bomb'	donde 'where'	gaŋga 'bargain'
		kaldo 'hot'	

That is, non-initial voiced stops spirantize (5-14a), except when following a homorganic nasal or lateral (b).[65] Additional cases appear in the following table:

[65]Note however that in [aðlateɾes], spirantization does occur. The effort-based approach predicts blocking in this context, just as in its mirror image, as /ld/ and /dl/ are both partial geminates. However, /dl/ clusters are marginal in Spanish, being restricted, so far as I am aware, to heteromorphemic contexts involving rather learned words. Thus the spirantization in [aðlateɾes] may be an effect of paradigmatic faithfulness with respect to the prefix [að-] (see section 4.1), rather than natural conditioning of /d/ spirantization in pre-[l] position.

Table 5-3. Blocking of spirantization, flapping in partial geminates.

Language	Reference	Description of process
Lamani	Trail 1970	Flapping blocked after a homorganic nasal or lateral
Malayalam	Mohanan 1986	Spirantization, flapping do not apply following a homorganic nasal
Proto-Bantu	Greenberg 1948	Spirantization (context-free), blocked after homorganic nasal
Tümpisa Shoshone	Dayley 1989	Spirantization blocked in homorganic nasal-stop clusters

More generally, oral reduction of all or part of a tautomorphemic homorganic nasal stop or lateral-stop cluster is unattested in the inalterability literature and the lenition surveys. Far from spirantizing, consonants show a positive tendency to occlusivize when adjacent to a homorganic nasal, as in the following Kikuyu post-nasal alternations (Armstrong 1967):

(5-15)	mbureetɛ	'lop off'	cf. βura
	mbaareetɛ	'look at'	cf. βaara
	ndeheetɛ	'pay'	cf. reha
	nduɣeetɛ	'cook'	cf. ruɣa
	ŋgoreettɛ	'buy'	cf. ɣora
	ŋgaeetɛ	'divide'	cf. ɣaja

Similarly, pre-nasal occlusivization is seen in certain dialects of American English, e.g. [bɪdnəs] ('business'), ɪdnɪt ('isn't it').

However, nasal + stop clusters show no parallel blocking of voicing. This is seen in Tümpisa Shoshone (5-10), where voicing applies to post-nasal stops (though spirantization is blocked), e.g. /ɨntamɨʔi/ - [ɨndaw̃ɨʔi] ('your little brother'). A virtually identical pattern of post-nasal voicing, but blocking of spirantization, is observed in Malayalam (Mohanan 1986). Indeed, Hayes & Stivers (in progress) observe that stops very commonly undergo voicing in post-nasal position (regardless of homorganicity), as seen in the following alternations (5-16b) from Wembawemba:

(5-16)	a.	/taka/	takə	'to hit'
		/milpa/	mɪlpə	'to twist'
	b.	/jantin/	jandɪn	'me'
		/panpar/	panbər	'shovel'

See also Pater 1995, 1996.

2.2.5. NO EXCLUSIVE OCCLUSIVIZATION OR DEVOICING OF SINGLETONS. *No occlusivization or obstruent devoicing process targets singletons to the exclusion of geminates.* This claim, the flip side of geminate resistance to oral reduction and voicing, is originally due to Churma (1988), who refers to these processes more loosely as "strengthening." Thus, one may find languages in which both geminate and singleton obstruents are uniformly realized as stops (that is, all obstruents occlusivize), e.g. Warray, Mayali, and numerous other Australian languages (see Evans 1995); and there are languages in which only geminates occlusivize, e.g. Modern Berber (Saib 1977), Luganda (/jj, ww/ - [ɟɟ, gg^w], Clements 1986, Churma 1988), and Malayalam (/rr/ - [tt], Mohanan 1986). But there appears to be no language in which singletons occlusivize to the exclusion of geminates.

Similarly, there are languages in which both geminate and singleton obstruents uniformly surface as voiceless (that is, all obstruents devoice), e.g. Delaware (Maddieson 1984); and Ohala 1983 cites Nubian as a case of geminate devoicing:

(5-17)

Noun stem	Stem + 'and'	Gloss
sɛgɛd	sɛgɛttɔn	father
kadʒ	kattʃɔn	scorpion
mʊg	mʊkkɔn	dog

(See also the devoicing of the geminate rhotic trill in Malayalam, noted above). But there appears to be no language in which singletons devoice to the exclusion of geminates.

2.2.6. INVENTORY ASYMMETRIES. *The presence of a geminate continuant consonant, or voiced geminate obstruent, in the segment inventory of a language (whether derived or underlying) implies the presence of a corresponding non-continuant or voiceless geminate, respectively.* This generalization is a narrower restatement of Elmedlaoui's original claim that segment inventories never have "weaker" (i.e. higher sonority) geminates without also having "stronger" geminates.[66] We have already seen that geminates resist spirantization and (obstruent) voicing processes. However, this is not exclusively a claim about lenition *processes*; for the same asymmetry is seen in segment inventories, regardless of the phonemic status of the "weaker" geminate. This generalization is borne out by the inventories discussed in Maddieson 1984.

---

[66]To the extent that Elmedlaoui generalizes this claim in terms of the sonority hierarchy, it is false. Elmedlaoui's claim predicts, for example, that all languages with long vowels also have geminate consonants (falsified by a significant share of the world's languages, including Yidiɲ, Dixon 1977); and that all languages with geminate sonorant consonants also have geminate obstruents (falsified by Ponapean, which has geminate nasals and liquids, but no geminate obstruents, Rehg & Sohl 1981).

First consider the stop/fricative asymmetry:

Table 5-4. Segment inventories: geminate stops and fricatives.

Labial		**Coronal**		**Dorsal**	
geminate stop/affricate	*geminate fricative*	*geminate stop/affricate*	*geminate fricative*	*geminate stop/affricate*	*geminate fricative*
Punjabi		Punjabi	Punjabi	Punjabi	
Finnish		Finnish	Finnish	Finnish	
Yakut		Yakut	Yakut	Yakut	
Japanese		Japanese	Japanese	Japanese	
Maranungku		Maranungku		Maranungku	
Delaware		Delaware		Delaware	
Lak		Lak	Lak	Lak	Lak
Wolof		Arabic	Arabic	Arabic	Arabic
Arabic	Arabic	Shilha	Shilha	Shilha	Shilha
Shilha	Shilha	Somali		Somali	
Somali		Kaliai		Ngizim	
	Greenlandic	Wichita	Wichita		Greenlandic
			Greenlandic		
			Iraqw		

It can be seen from this table that languages with geminate obstruents overwhelmingly have geminate stops, and may also have fricatives. But the only reported cases of inventories with a geminate fricative but no corresponding stop are Greenlandic and Iraqw. It is clear, however, from Rischel's (1974) description (Maddieson's principal source) that Greenlandic does in fact have surface geminate stops as well as fricatives :

(5-18) suraajuwippuq 'is incessant'

kamittaq 'new boot'

puwijuwikkijuppaa 'never forgets it'

qaqqaq 'mountain'

As for Iraqw, Maddieson characterizes the inventory as having a distinction between [s] and [ss], with no length distinction in the stops. I have not been able to obtain Whiteley's (1958) description, on which Maddieson's Iraqw inventory is based; however, Nordbustad's (1985) grammar of Iraqw makes clear that geminate stops are present in the inventory:

(5-19)	a daqqáw	'I am in the act of going'
	a tsattá ɬaáʔ	'I want to cut'
	gwa tuntukká	'she has not covered it'

Curiously, Nordbustad gives no examples with [ss]. Other descriptions of the Iraqw sound system available to me (Maghway 1995, Mous 1992) do not mention any length contrasts at all. Nordbustad observes that "double consonants are rare." Apparently, Iraqw presents a case of a rather marginal (perhaps diachronically waning) consonant length contrast, in stops as well as fricatives; hence the variation in the descriptions. In any event, there appears to be no grounds for viewing Greenlandic as a counterexample to the INVENTORY ASYMMETRIES generalization.

Finally, consider the voicing asymmetry:

Table 5-5. Segment inventories: geminate voiced and voiceless obstruents.

Labial		**Coronal**		**Dorsal**	
geminate voiceless obstruent	*geminate voiced obstruent*	*geminate voiceless obstruent*	*geminate voiced obstruent*	*geminate voiceless obstruent*	*geminate voiced obstruent*
Punjabi	Punjabi	Iraqw		Punjabi	Punjabi
Finnish		Punjabi	Punjabi	Finnish	
Yakut		Finnish		Yakut	
Japanese		Yakut	Yakut	Japanese	
Maranungku		Japanese		Maranungku	
Delaware		Maranungku		Delaware	
Lak		Delaware		Lak	Lak
Arabic	Arabic	Lak		Arabic	Arabic
Shilha	Shilha	Arabic	Arabic	Shilha	Shilha
Greenlandic		Shilha	Shilha	Greenlandic	Greenlandic
	Somali	Wichita			Somali
	Wolof	Lak			
		Japanese			
			Somali		

Again, the table reveals an overwhelming pattern of voiceless geminates, or voiceless and voiced geminates. But the only reported cases of inventories with voiced geminate obstruents without voiceless counterparts are Somali and Wolof. However, Armstrong 1964 (Maddieson's source for the Somali inventory) states that the Somali "voiced" geminates in question "do not sound fully voiced," and in some cases are in fact completely voiceless.[67] As discussed in section 2.2.3, I assume that these geminates are passively devoiced. As for Wolof, Sauvageot's (1965) grammar of the Dyolof (or Jolof) dialect (Maddieson's source for the Wolof inventory) in fact lists a number of forms with voiceless geminates:[68]

---

[67] An exception are the post-alveolar stops, transcribed as [ɖ]/[ɖɖ] which Armstrong describes as fully voiced, in both the singleton and the geminate. Armstrong notes, however, that this sound is not a simple voiced stop. Unlike the rest of the stop series, it involves pharyngeal constriction, as well as being somewhat implosive. Moreover, there is no voiceless correspondent to [ɖ] in the singleton series, as there are for the other voiced singleton stops.

[68] Maddieson's characterization appears to be based on Sauvageot's statement (p. 17) that "Le parler possède une corrélation de gémination consonantique. Celle-ci ... est, semble-t-il, limitée aux occlusives sonores; ... aux nasales; ... [et] à la latérale." Unfortunately, Sauvageot does not explain how this characterization is to be reconciled with the voiceless geminate forms cited above, which appear in later sections of the grammar.

(5-20) tappu 'needle'

atte 'to judge'

fɛtt^{j}i 'undress, untie'

tɛkki (no gloss)

Moreover, more recent grammars and dictionaries of (standard) Wolof (e.g. Ka 1994, Munro & Gaye 1991) make clear that voiceless geminates are indeed part of the inventory; and Omar Ka (p.c.) denies that any dialect of Wolof lacks voiceless geminates.

In sum, the segment inventories listed by Maddieson 1984 contain, upon closer examination, no actual counterexamples to the INVENTORY ASYMMETRIES generalization.

3. AN EFFORT-BASED ACCOUNT OF THE GENERALIZATIONS

3.1. EFFORT IN GEMINATES.

In this section, I motivate a set of effort relations between geminates and possible lenited correspondents, before proceeding to a formal analysis of the geminate lenition generalizations.

3.1.1. DEGEMINATION. Under the computational mass-spring model of consonant articulation presented in Chapter 2, the result emerges that geminate stops are more effortful than singleton stops (as one would expect):

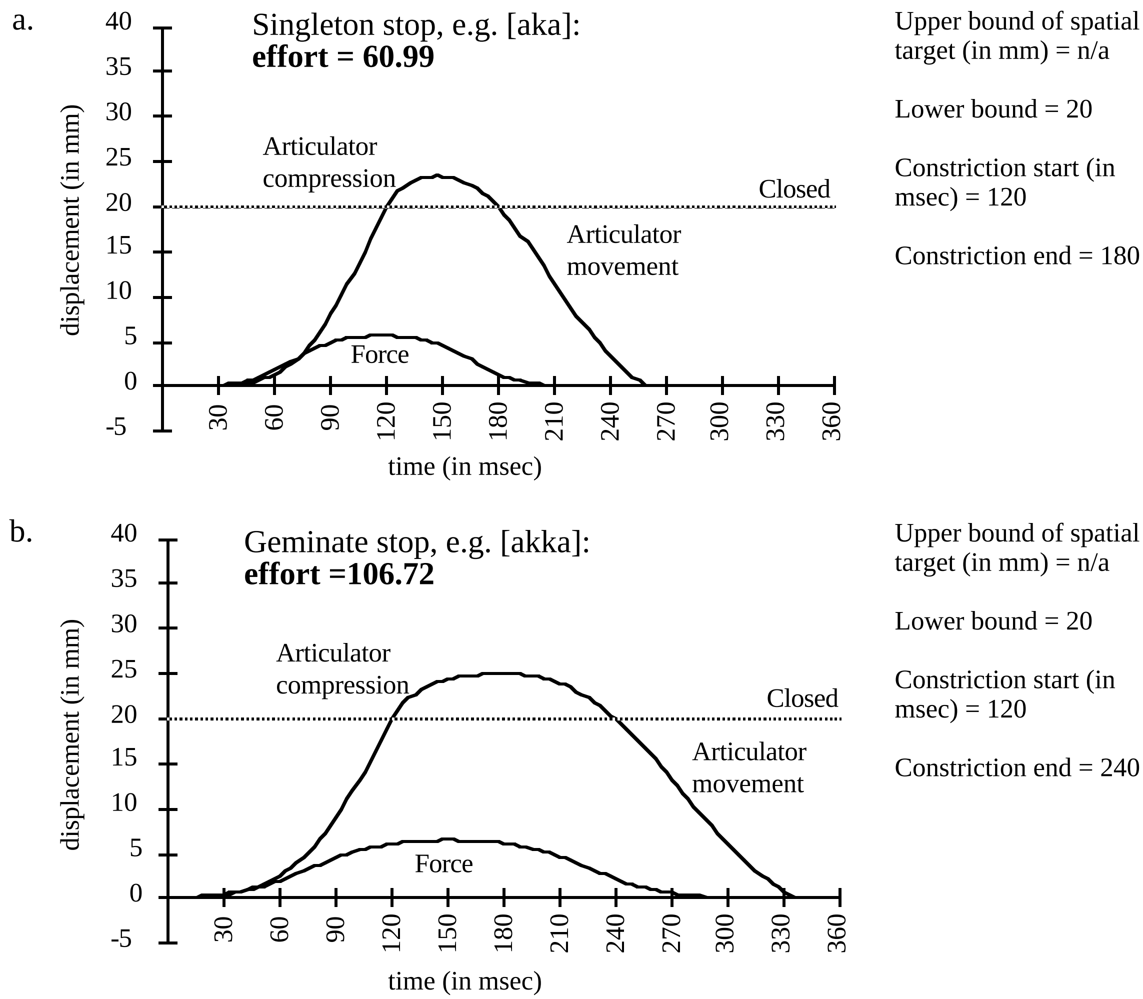


Figure 5-2. Outputs of mass-spring model for singleton and geminate stops.

Let [length] informally denote the property that distinguishes geminates from singletons, however this is to be formally represented (featurally, segmentally, or moraically). Now, degemination obtains under the ranking LAZY » PRES(length).

(5-21)

	Input: akka	LAZY	PRES(length)
	akka	**!	
☞	aka	*	*

The treatment of degemination is thus unified with the general effort-based approach to lenition outlined in Chapter 1.

3.1.2. REDUCTION OF ORAL CONSTRICTION. Recall that under this effort-based approach (see Chapters 1, 4), spirantization is treated as reduction of a stop gesture, for reasons of effort minimization, to the point that closure is lost. Unlike the sort of brief fricative constriction shown in (Figure 5-3a), however, geminate fricatives, and more broadly, geminate continuant consonants, involve a prolonged steady-state constriction, by definition (Figure 5-3b):

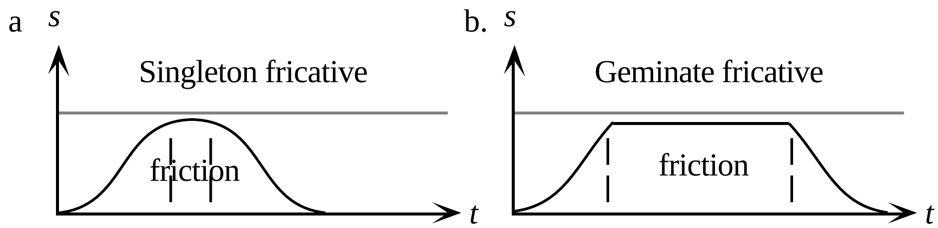


Figure 5-3. Schemata of displacement-vs.-time, for singleton and geminate fricatives.

That is, if the period of friction is not maintained for an extended duration, the consonant is not a geminate. As with the analysis of strident fricatives in Chapter 4, to achieve this prolonged steady-state constriction, I assume that the upward movement of the active articulator must be arrested by an active antagonistic force applied to the same articulator, i.e. isometric tension. This assumption is supported by the mass-spring model, which is able to achieve a geminate fricative only by opposing the positive force impulse with a negative force impulse:

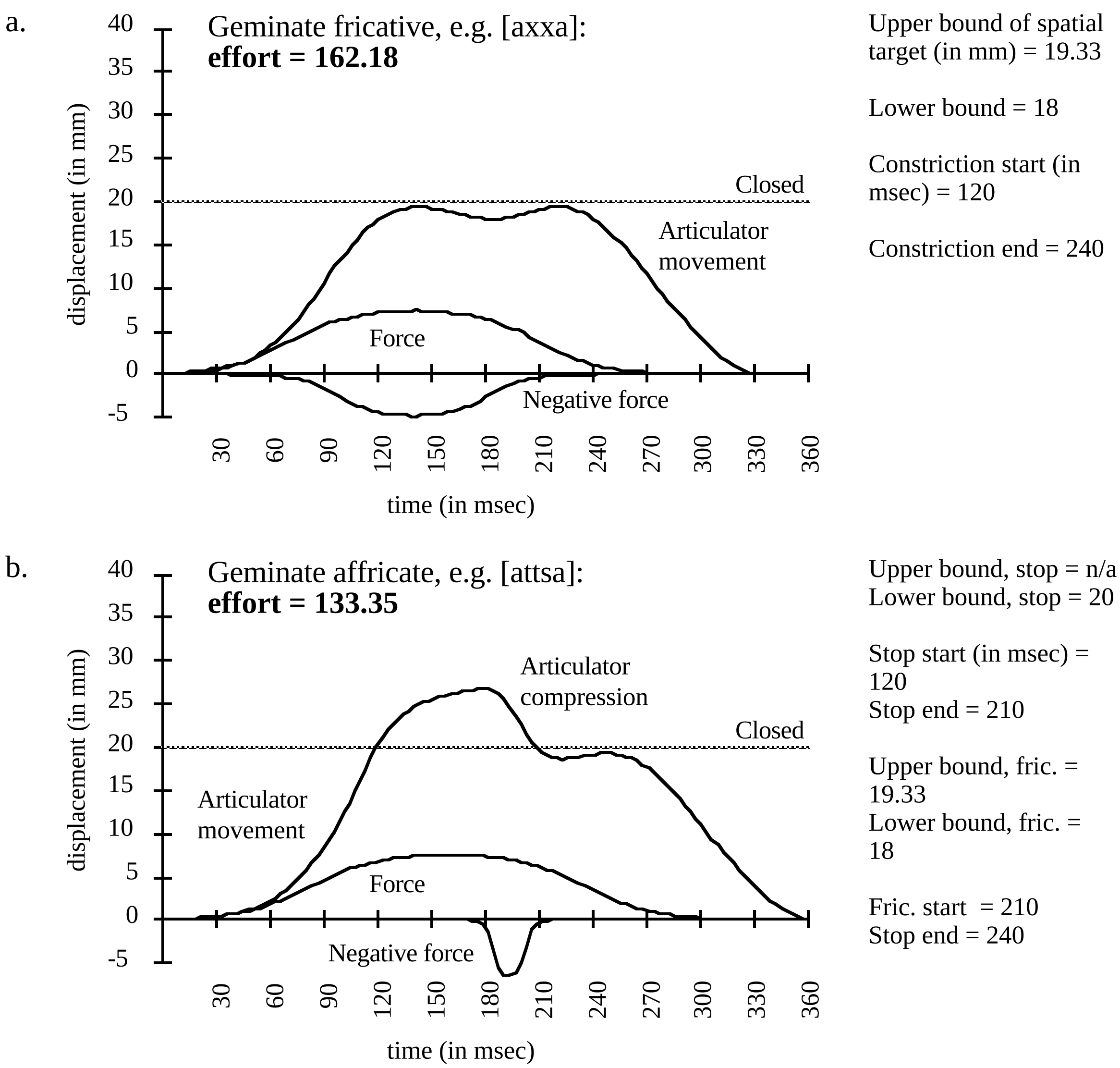


Figure 5-4. Output of mass-spring model for geminate fricative and affricate.

By comparison with the singleton and geminate stops (Figure 5-2a,b, effort = 60.99, 107.72 respectively), and the singleton strident and nonstrident fricatives (Figure 4-4a,c, effort = 65.98, 20.07 respectively), the geminate fricative in Figure 5-4a emerges as more effortful than the geminate stop, the geminate affricate (Figure 5-4b), or any of the singletons.

3.1.3. HALF-SPIRANTIZATION AND PARTIAL GEMINATES. For similar reasons, half-spirantization of a geminate stop also increases its effort cost. Again, this assumption is supported by the mass-spring model:

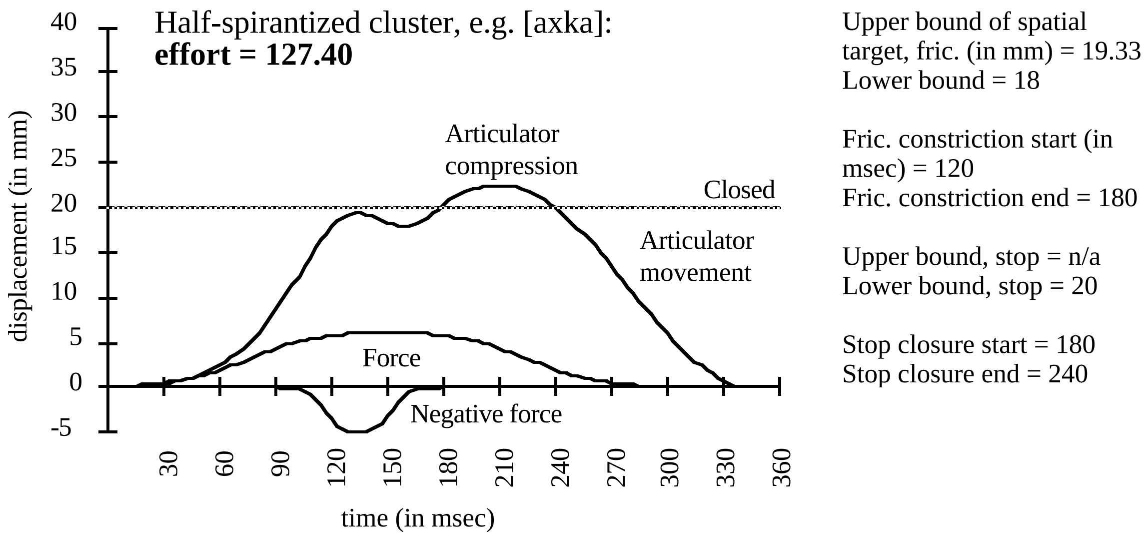


Figure 5-5. Output of mass-spring model for half-spirantized geminate.

The half-spirantized geminate (effort = 127.40) emerges as more effortful than the geminate stop (Figure 5-2b, effort = 107.72), again due to the negative force impulse required for the former.

Note, however, that the "problem" in the half-spirantized gesture necessitating substantial isometric tension is the sustained constriction of the fricative. If the lenited portion of the geminate does not involve a steady-state constriction (e.g. the glide + homorganic stop clusters of Maxakalí, see section 2.2.5), we can obtain this result simply by slowing down the transition into the closure, with a minute negative force impulse, as shown below.

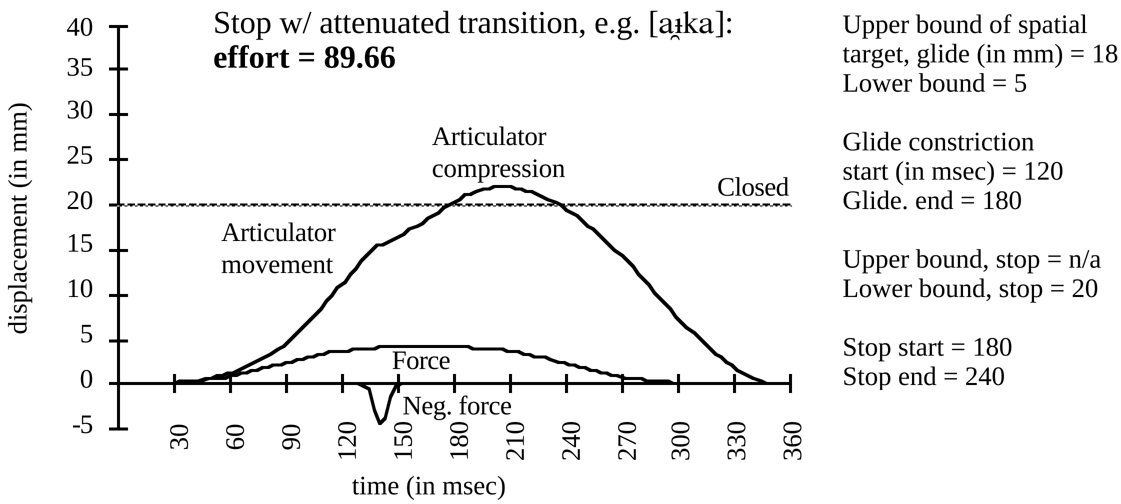


Figure 5-6. Output of mass-spring model for stop with attenuated transition.

This stop with attenuated transition has a lower effort cost (89.66) than the geminate stop (Figure 5-2b, effort = 107.72).

Note that the foregoing conclusions concerning spirantization and half-spirantization of geminate stops apply equally to partial geminates, assuming that the partial geminates in question are of comparable duration to full geminates (as is documented, for example, in Sinhalese, Ladefoged & Maddieson 1996). For the effort relations above refer to oral constriction gestures; and, as schematized in Figure 5-7, the oral constriction of a partial geminate is equivalent to that of a full geminate.

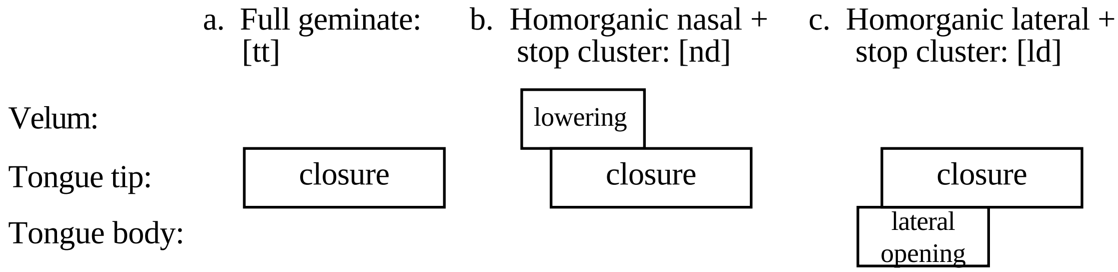


Figure 5-7. Gestural scores for full and partial geminates.

There are, however, cases of homorganic nasal + stop and lateral + stop clusters whose duration is closer to that of a singleton (as documented, for example, in English, Browman & Goldstein 1990: 367). For these short clusters, a distinct explanation is available for their conformity to the partial geminate generalization. Specifically, these clusters are perceivable as clusters (or, alternatively, but non-crucially to the present analysis, as contour segments), precisely because they have a release burst, an abruptly discontinuous acoustic event from the nasal or lateral portion which precedes it (cf. Steriade's (1993) observation that closure and release are essential to a contour segment). Spirantization or other oral reduction of such a cluster (or contour segment) would result in loss of the burst, without which the orality (or centrality) of the release is presumably imperceptible. Thus, either spirantization is blocked (e.g. by faithfulness to the [-nas] (or [-lat]) specification of the release portion), or spirantization occurs, resulting in apparent neutralization with a plain nasal or lateral singleton. (For example, in American English, in casual speech productions of /nt/ clusters, e.g. in <u>winter</u>, the cluster degeminates to a singleton nasalized flap: [wĩɾ̃ɚ].) But in either case, the generalization is maintained: partial geminates, like full geminates, are immune to spirantization, unless they concomitantly degeminate.

3.1.4. VOICING

3.1.4.1. STOPS. Ohala 1983 identifies a straightforward aerodynamic explanation for the markedness of geminate voiced stops, namely their tendency to passively devoice (unlike medial singletons, which passively *voice*, see Chapter 2 section 4). As air pressure builds up in the oral cavity during a stop, the trans-glottal air pressure differential drops below what is required to keep the vocal folds vibrating (roughly 2,000 dyne/cm^2), and voicing ceases, typically 60 msec into the closure for an alveolar stop (slightly earlier for a velar, and later for a labial) (Westbury and Keating 1986). Voicing can be extended during an oral stop by various cavity expansion gestures, e.g. pharynx expansion and larynx lowering (Rothenberg 1969).[69] However, to sustain voicing for the duration of a geminate, typically over 150 msec., "heroic" cavity expansion is required, which necessarily involves additional effort. Consequently, a voiced geminate stop $>_{\text{effort}}$ a (substantially) voiceless geminate stop (the reverse of the situation in medial singletons).

Note, however, that *partial* geminates present none of these devoicing problems. For the air is vented during the nasal or lateral portion of a partial fricative, preventing significant build-up of oral pressure. (This is a fortiori the case in short partial geminate clusters.) Indeed, Hayes & Stivers (in progress) suggest that the velic raising that occurs toward the end of a nasal + stop cluster actually facilitates voicing, by expanding the oral cavity during the oral portion of the cluster.

3.1.4.2. FRICATIVES. Geminate fricatives likewise tend to passively devoice. My own simulations, using the analog circuit model of vocal tract aerodynamics

---

[69]The other principal strategy of avoiding passive devoicing, "nasal leak" (allowing air to leak out the nasal passages), carries a perceptual cost: risking confusion of the stop with a nasal consonant.

described in Westbury & Keating 1986, show that with a close fricative (oral aperture = 20 mm^2), passive devoicing occurs (i.e. the transglottal pressure differential falls below 2,000 $dyne/cm^2$) at 95 msec in an alveolar. With a more open fricative (30 mm^2),[70] the point of passive devoicing is postponed to 166 msec for an alveolar, too late to account for devoicing in a geminate fricative; and indeed labials appear never to reach the point of passive devoicing. However, this assumes that the glottal aperture of a voiced fricative is equivalent to that of a voiced stop (estimated at 4 mm^2 by Westbury & Keating, averaging over vibratory cycles). In fact, Ohala 1983 observes that the glottis is typically somewhat more abducted in a voiced fricative than it is in a voiced stop. This is because fricatives crucially involve a pressure differential at the place of oral constriction (as well as at the glottis, if voiced):

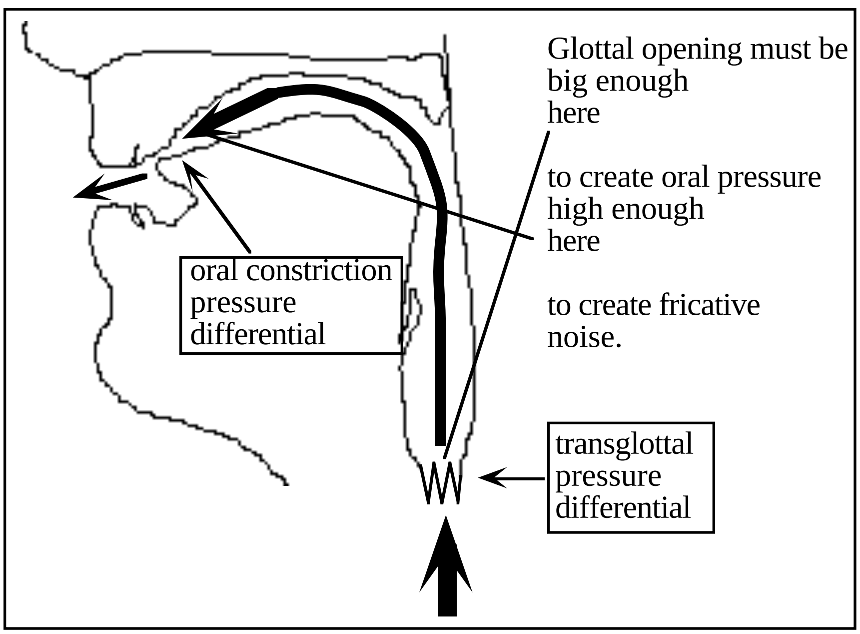


Figure 5-8. Aerodynamic considerations in fricative voicing and glottal aperture.

[70]Clark & Yallop 1990:84 indicate that 30 mm^2 is a plausible upper bound on oral aperture in fricatives.

With an increase in glottal aperture of just 1.5 mm^2, the build-up of oral pressure, hence passive devoicing, occurs much earlier.[71] The results are summarized in Table 6.

Table 5-6. Point of passive devoicing, medial geminate fricative, in msec.

	Oral aperture = 20 mm^2			Oral aperture = 30 mm^2		
Glottal aperture = 4 mm^2	125 (lab)	95 (alv)	85 (vel)	-- (lab)	166 (alv)	127 (vel)
Glottal aperture = 5.5 mm^2	95 (lab)	89 (alv)	80 (vel)	117 (lab)	99 (alv)	85 (vel)

Thus, a substantial portion of the geminate fricative is devoiced, even if the fricative is quite open, unless cavity expansion gestures or other heroic voicing strategies are employed, just as in geminate stops. We may therefore conclude that a voiced geminate fricative $>_{effort}$ a (substantially) voiceless geminate fricative.

2.1.5. SUMMARY. The effort relations motivated above can be presented in the form of a Hasse diagram,[72] which conveys the additional relations which follow from transitivity:

[71] The slight abduction assumed here is still a long way from a truly spread glottis (e.g. in actively devoiced stops), estimated at 32.5 mm^2 by Westbury & Keating.

[72] To be read as follows: a consonant of type A is of greater effort than a corresponding consonant of type B if A is connected to B by a downward path within the lattice.

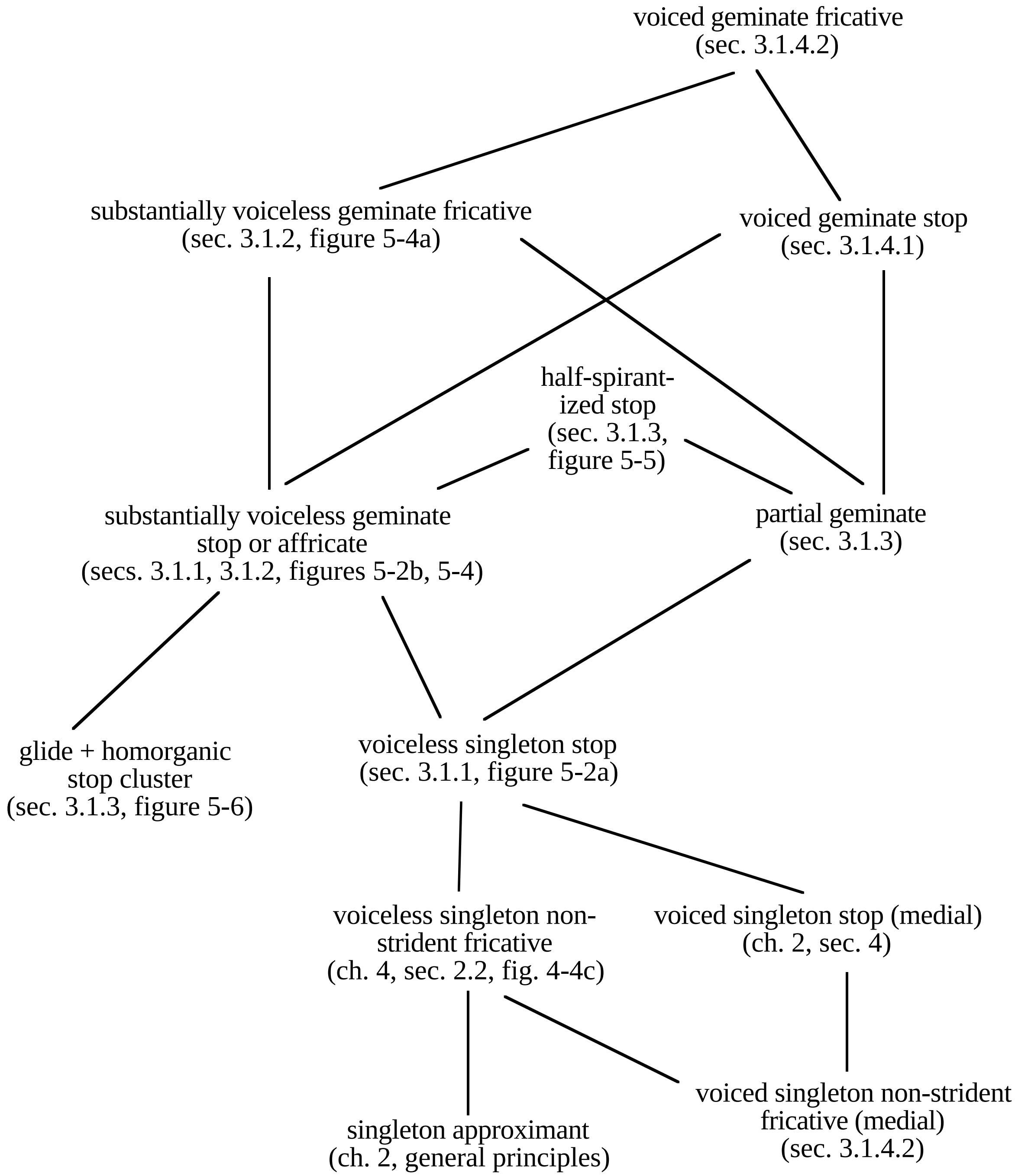


Figure 5-9. Hasse diagram of previously inferred effort relations among geminate and singleton consonant types.

3.2. FORMALLY CAPTURING THE GEMINATE LENITION GENERALIZATIONS

3.2.1. NO ORALLY REDUCED GEMINATES, ETC. Recall that geminate continuants are more effortful than geminate stops (section 3.1.2). Given this effort relation, under the constraint system posited under the effort-based approach to lenition, no ranking of the relevant constraints (i.e. PRES(cont) and LAZY) allows an input stop (geminate or otherwise) to map to an output geminate continuant, since the latter candidate fares worse than a geminate stop with respect to both constraints.

(5-22)

	ap(p)a	LAZY	PRES(cont)
☞	appa	*	
	aϕϕa	**	*

(I further assume that the constraint set contains no perceptually based constraint favoring the fricative candidate; since, by hypothesis, lenition is characterized as effort reduction, rather than increase in aperture for its own sake, there would appear to be no motivation for such a constraint.) An input geminate stop can only yield a spirantized output if the output degeminates as well.

(5-23)

	appa	LAZY	PRES(cont)	PRES(length)
	appa	**!		
	aϕϕa	***!	*	
☞	aϕa	*	*	*

Likewise, because of the effort relations *{Geminate fricative, Half-spirantized stop}* > *{Geminate stop, Partial geminate}* (section 3.1.3), neither full nor partial geminates can undergo spirantization nor half-spirantization (modulo the heteromorphemic geminate case considered in section 4).

(5-24)		LAZY	PRES(cont)
☞	appa - appa	*	
	appa - aɸpa	**	*
☞	ampa - ampa	*	
	ampa - amɸa	**	*
	ampa - aw̃ɸa	**	*

In sum, since reduction of oral constriction in geminate stops increases the effort cost, due to the increased isometric tension involved (see sections 3.1.2, 3.1.3), oral reduction of geminates is ruled out universally, and the NO ORALLY REDUCED GEMINATES, NO HALF-SPIRANTIZATION, and NO REDUCTION OF PARTIAL GEMINATES generalizations are captured. That is, no process converts a consonant (geminate or otherwise) to a geminate with reduced oral constriction; no process converts a (tautomorphemic) geminate stop to a "half-spirantized" cluster, e.g. /kk/ - *[xk]; and partial geminates behave identically to full geminates with respect to reduction of oral constriction.

3.2.2. VOICING. By precisely the same reasoning from relative effort costs (see section 2.1.4), voicing of geminate obstruents is prohibited, and the NO VOICING OF GEMINATES generalization is captured.

(5-25)		LAZY	PRES(voi)
☞	appa - appa	*	
	appa - abba	**	*

That is, no process converts a voiceless segment (geminate or otherwise) to a voiced geminate obstruent.

3.2.3. OCCLUSIVIZATION, DEVOICING. In singletons, occlusivization processes must be attributed to fortition constraints. Let us denote such a constraint as *[+cont] / K, where K refers to the context in which occlusivization occurs. If this constraint is active in some grammar, it must outrank both PRES(cont) and LAZY.

(5-26)

	Input: ϕ	*[+cont] / K	PRES(cont)	LAZY
☞	p in context K		*	**
	ϕ in K	*!		*

By section 3.1.2, *Geminate continuant $>_{effort}$ Geminate stop*, therefore LAZY disfavors the fricative geminate. And since the fortition constraint must outrank PRES(cont), the only constraint which potentially blocks occlusivization,[73] it follows that the geminate must occlusivize as well.

(5-27)

	Input: ϕϕ	*[+cont] / K	PRES(cont)		LAZY
☞	pp in context K		*		*
	ϕϕ in K	*!			**

(LAZY is split off from the rest of the tableau above to indicative that its ranking relative to the other constraints does not affect the result here.) By the same reasoning from relative effort costs (see section 3.1.4), the same result obtains for geminate devoicing.

(5-28)

	Input: bb	*[+voi] / K	PRES(voi)		LAZY
☞	pp in context K		*		*
	bb in K	*!			**

[73]This presumes that the constraint set does not include a positional faithfulness constraint specifically referring to continuancy in long consonants.

Consequently, the NO EXCLUSIVE OCCLUSIVIZATION OR DEVOICING OF SINGLETONS generalization is captured. That is, no occlusivization nor obstruent devoicing process targets singletons to the exclusion of geminates. In contrast, occlusivization or devoicing of *geminates* (e.g. Berber, Schein & Steriade 1986) obtains under any ranking in which LAZY or the relevant fortition constraint, dominates PRES(cont) or PRES(voi), respectively.

3.2.5. INVENTORY ASYMMETRIES. As shown in the previous section, to obtain surface geminate continuants or voiced geminate obstruents, PRES(cont) or PRES(voi) must dominate LAZY, and any applicable fortition constraints (otherwise occlusivization or devoicing will occur):

(5-29)

		PRES(cont)	PRES(voi)	LAZY	*+cont	*+voi
	ɸɸ - pp	*!		*		
☞	ɸɸ - ɸɸ			**	*	
	bb - pp		*!	*		
☞	bb - bb			**		*

But under this ranking, an input geminate stop, or voiceless geminate obstruent, surfaces unchanged (and such inputs must be allowed, by the OT tenet of Richness of the Base):

(5-30)

		PRES(cont)	PRES(voi)	LAZY	*+cont	*+voi
☞	pp - pp			*		
	pp - ɸɸ	*!		**	*	
	pp - bb		*!	**		*

Consequently, the INVENTORY ASYMMETRIES generalization is captured. That is, the presence of a geminate continuant consonant, or voiced geminate obstruent, in the segment inventory of a language (whether derived or underlying) implies the presence of a corresponding non-continuant or voiceless geminate, respectively.

4. HETEROMORPHEMIC GEMINATES

4.1. THE DISTINCT BEHAVIOR OF HETEROMORPHEMIC GEMINATES.

In section 3.2.3 the effort-based approach appears to rule out half-spirantization of geminates. Yet in the discussion of Tigrinya in section 2.2.2, it is conceded that half-spirantization is indeed possible, provided that the geminates are heteromorphemic. The Tigrinya facts were taken by Hayes 1986 as precluding any sort of phonetically-based account of geminate inalterability. If inalterability is attributed to phonetic considerations, how, then, could heteromorphemic and tautomorphemic geminates (which are typically phonetically indistinguishable) behave differently from one another? The answer lies in OT's capacity for interaction between purely articulatory constraints such as LAZY and constraints which do refer to morphological affiliation.

Specifically, a class of paradigmatic faithfulness constraints (also called output-output faithfulness, uniform exponence, paradigm uniformity, and allomorphy minimization constraints) has been motivated by such phenomena as base-reduplicant correspondence, base-derivative correspondence in truncation patterns, and cyclicity effects, see Benua 1995, 1997; Flemming 1995; Kenstowicz 1995; McCarthy and Prince 1995; Steriade 1996; Burzio 1997. These constraints are formally similar to the input-output faithfulness constraints employed in Chapter 3 above, but the comparison is between two morphologically related surface forms, typically a base and its derivative. Unlike input-output faithfulness, these paradigmatic constraints can enforce identity between an output and its morphological base with respect to phonologically derived surface properties of the base, including lenition.

The Tigrinya pattern of half-spirantization of heteromorphemic geminates now follows from undominated ranking of PRES(BASE/DERIVATIVE, cont):

(5-31)

	Input: /mɨrak-ka/ (base = [mɨrax])	PRES(B/D, cont)	LAZY	PRES(I/O, cont)
	mɨrakka	*!	*	
☞	mɨraxka		**	*
	mɨraxxa		***!	**

The final candidate, [mɨraxxa], loses purely on grounds of effort minimization. The interesting result in this tableau is the selection of [mɨraxka] over [mɨrakka]. Spirantization occurs in [mɨraxka] not because it serves the goal of effort minimization (in fact, it fares worse on this score than the competing candidate [mɨrakka]), but because it promotes similarity between the output and its base, [mɨrax], in which spirantization *is* motivated by LAZY. If, however, PRES(O/O, cont) is subordinated to LAZY, heteromorphemic geminates will be inalterable under spirantization, just like tautomorphemic geminates, as we find in Tiberian Hebrew. Finally, in tautomorphemic geminates, paradigmatic concerns do not enter the picture (there can be no separate base containing a spirantized singleton), and so half-spirantization is ruled out under any ranking.

4.2. HETEROMORPHEMIC DERIVED GEMINATES

A remaining issue concerns the behavior of heteromorphemic geminates which are derived by assimilation. These derived geminates never undergo half-spirantization

(Guerssel 1977); that is, they do not pattern with other heteromorphemic geminates, but with the "true" (monomorphemic) geminates, in being systematically inalterable under spirantization. For example, in Tigrinya, /jɨ-t-kəfət/ ('open-passive-jussive') surfaces as [jɨkkəfət], not [jɨxkəfət] (Kenstowicz 1982). This generalization was adduced by Steriade (1982), Hayes (1986), and Schein & Steriade (1986) as confirmation of the autosegmental true/fake geminate distinction. The reasoning is that the assimilation process, expressed in the autosegmental framework as a feature-spreading operation, gives rise to a multiply linked structure (5-32a), identical to an underlying true geminate (b), rather than the fake geminate (c).

```
(5-32)  a. C     C          b. C   C          c. C   C
           ╪ ⟍   |              \ /              |   |
           t   ⟍ k               k               k   k
```

Therefore, the derived geminate is correctly predicted to be inalterable under spirantization, due to the UAC or the Linking Constraint, as is the underlying true geminate.

However, this subgeneralization, at least as it relates to the behavior of geminates under lenition, falls out from the effort-based approach as well, without resorting to an abstract representational distinction between true and fake geminates. The answer to Guerssel's observation lies in the insight that place assimilation in consonant clusters is driven by effort minimization, just as ordinary lenition is (see, e.g., Stampe & Donegan 1979, Jun 1995).[74] In theory-neutral terms, the question is why a heteromorphemic,

[74]In the case of Tigrinya, a language-specific solution happens to be available. Only dorsal consonants spirantize; thus, the 'passive' prefix /-t-/ surfaces as [-cont] in all output forms. Therefore, /t+k/ - [xk] violates the undominated constraint PRES(O/O, cont), losing to [kk]. However, this result is dependent on the absence of coronal spirantization in Tigrinya. In this discussion, in the interest of obtaining a more

heterorganic cluster such as /t + k/ can partially spirantize to [θk], or undergo place assimilation [kk], but not both [xk]. First, we can infer that a half-spirantized geminate is more effortful than a non-homorganic fricative + stop cluster. This inference is supported by the mass-spring model: the effort costs of a singleton (strident) fricative plus the singleton stop, i.e. [sk] (Figure 4-2a+b, total effort = 126.97) is lower than that of the partially spirantized geminate, [xk] (Figure 5-5, effort = 127.40) (a fortiori if the friction is nonstrident, i.e. [θk]); while the full geminate stop, [kk] has an effort cost of 106.72 (Figure 5-2b).

Now, assuming there are paradigmatically related forms with surface [θ], the possible outputs are either [θk] (5-33a) or [kk] (5-33b, the correct result for Tigrinya), depending on the ranking of output-output faithfulness to continuancy (specifically, comparing the derived form with other members of the affixal paradigm) relative to LAZY:[75]

---

general account of Guerssel's generalization, I instead analyze a hypothetical variant of Tigrinya, with coronal spirantization.

[75] PRES(place features) » LAZY is also sufficient to block an assimilated output.

(5-33) a.

	Input: /jɨ-t-kəfət/ (/-t-/ related to surface [θ] in other outputs)	PRES(AFFIXAL PARADIGM/D, cont)	LAZY
☞	jɨθkəfəθ		**
	jɨkkəfəθ	*!	*
	jɨxkəfəθ		***!

b.

	Input: /jɨ-t-kəfət/ (/-t-/ related to surface [θ] in other outputs)	LAZY	PRES(AFFIXAL PARADIGM/D, cont)
	jɨθkəfəθ	**!	
☞	jɨkkəfəθ	*	*
	jɨxkəfəθ	**!*	

But no ranking of the relevant constraints permits the half-spirantized geminate, [xk], to emerge as the winner. More generally, since the place-assimilated partially-spirantized candidate incurs a higher effort cost than the unassimilated or unspirantized candidate, the partially-spirantized derived geminate cannot as emerge as the winner. This result thus falls out from Jun's (1995) treatment of place assimilation in consonant clusters: such assimilation is simply (effort-driven) lenition of C_1 coupled with compensatory extension of the gesture of C_2 (see the discussion of Jun's treatment as it relates to the effort-based approach to lenition in Chapter 1 section 3.3.4).[76]

It is thus possible to account for the distinct inalterability behavior under lenition of heteromorphemic geminates, underlying and derived, within a phonetically-based

[76]A further alternation in Tigrinya, previously adduced in support of the true/fake geminate distinction, involves the 3d. sg. pronominal suffixes -o and -a (masc. and fem. respectively), which induce gemination of the stem-final consonant, e.g. [jɨbarɨx] ('bless-jussive'), but [jɨbarɨkko]. I suggest, however, that the distinct behavior of -ka vs. geminating -o and -a can equally be handled in terms of a Class II vs. Class I affix distinction. As Benua (1997) has proposed, Class I affixes correspond to lower-ranked versions of B/D faithfulness constraints. In Tigrinya, the selection of the geminate stop candidate [jɨbarɨkko] over the partially spirantized candidate [jɨbarɨxko] now follows from ranking LAZY above PRES(B/D(CLASS I)/cont). While [mɨraxka]still defeats [mɨrakka] (see tableau 5-31) because PRES(B/D(CLASS II)/cont) » LAZY.

approach, and without resorting to a representational distinction between true and fake geminates.

5. SUMMARY AND ALTERNATIVE APPROACHES

By reducing geminate inalterability to considerations of effort minimization, the effort-based approach to lenition achieves a greater depth of explanation, as well as better empirical coverage, than previous accounts of this class of phenomena. Since I am presenting this account of the geminate lenition generalizations as a significant part of the motivation for an effort-based approach to lenition, an important remaining question is whether these generalizations could be as elegantly captured without explicit reference to effort in the formalism.

5.1. MARKEDNESS CONSTRAINTS

Let us consider a theory in which lenition is attributed to a scalar REDUCE constraint, favoring reduction of constriction degree (e.g. approximant < fricative < stop), homologous to LAZY, but without explicitly referring to effort. The blocking of lenition in geminates could be attributed to a miscellany of markedness constraints, such as *[+cont,-cons,+length], *[+voi,-son,+length], *NASAL + FRICATIVE CLUSTERS, etc.

However, it would be a mistake to suppose that the cross-linguistic geminate lenition generalizations are captured in such a theory. For under rankings in which REDUCE dominates one of the markedness constraints, the generalization embodied by the markedness constraint evaporates:

(5-34)

	REDUCE	*[+cont,-cons,+length]
pp	**!	
☞ ϕϕ	*	*

The markedness constraints must therefore be stipulated to be inviolable; however, this ranking condition holds true only under lenition: for the faithfulness constraints, e.g. PRES(cont) and PRES(voi), must be able to dominate the markedness constraints, otherwise *contrastively* voiced or continuant geminates are incorrectly ruled out. By comparison, under the effort-based approach, the geminate lenition generalizations follow from the constraint set, without any ranking stipulations (process-specific or general). It therefore appears that direct reference to effort in the formalism does afford a substantially more elegant and insightful treatment of the geminate lenition generalizations. More generally, it is unclear how a non-effort-based approach, such as this REDUCE + markedness constraint system, could be extended to capture such aspects of lenition typology as the naturalness of lenition in intervocalic position, the increasing prevalence of lenition in faster speech rates, and the relation between spatial reduction of consonant constriction (in spirantization), temporal reduction (in degemination and flapping), and laryngeal adjustments (in voicing). In contrast, all these phenomena receive natural treatments under the effort-based approach, as outlined in Chapters 1 and 6.

5.2. ARTICULATORY PHONOLOGY

Alternatively, it might be claimed that the geminate lenition generalization (at least with respect to oral constriction) follows from simple reduction of otherwise invariant gestures. That is, if we take a position-vs.-time curve for a long oral closure gesture (Figure 5-10a), and simply reduce the magnitude of the gesture, without otherwise modifying the "shape" of the curve, the immediate result is shortening of the closure duration (b); and further reduction, to the point of spirantization, entails shortening (c) as well.

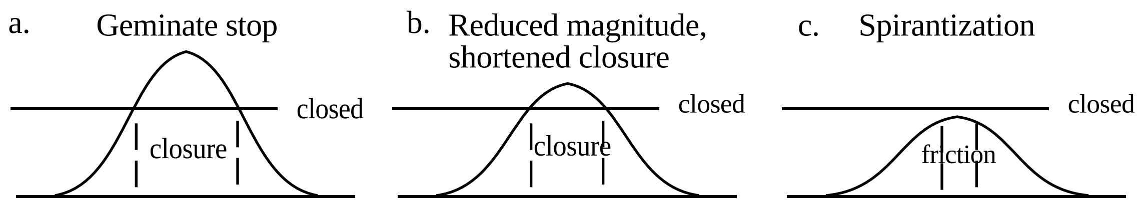


Figure 5-10. Reduction of magnitude and duration.

Such an account is implicit in the Articulatory Phonology framework (see discussion in Chapter 4, section 6.2), where lenition is treated as an operation of gestural reduction, without explicit reference to effort minimization.

However, as argued in Chapter 4, section 6.2, lenition is *not* characterizable as simple gestural reduction: frequently, lenition involves gestural modification beyond mere reduction, and even insertion of gestures, in compensation for gestures which are eliminated or weakened. Therefore, the Articulatory Phonology account offers no satisfactory answer to the question, why can't reduction of the constriction gesture in Figure 5-10a be accompanied by some change in shape of the displacement curve, as in Figure 5-11?

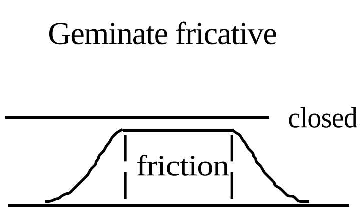


Figure 5-11. Reduction to a prolonged close constriction.

The key observation is that the unattested gestural modification (Figure 5-10a - Figure 5-11) involves a net effort increase, rather than a decrease as in the attested geminate lenition types (Figure 5-10a - b or c). Therefore, reference to effort minimization indeed appears to be crucial to an adequate account of the behavior of geminates under lenition.

Chapter 6:

Effort-Based Contexts

This chapter focuses on two principal generalizations: First, all else being equal, lenition is more likely to occur the more open the segments which flank the target; the commonly observed intervocalic lenition context proves to be a special case of this generalization. Second, all else being equal, lenition is more likely to occur the faster or more casual the speech style. I argue that both phenomena are manifestations of *effort-based contexts*: that is, lenition occurs more readily in these contexts or under these conditions because greater effort is required to achieve a given constriction target in these contexts/conditions. The claim, then, is that the effort-based approach not only permits a unified characterization of lenition processes, and accounts for some of the restrictions on lenition outputs (i.e. the non-stridency and geminate generalizations of Chapters 4 and 5); this approach also naturally accounts for a broad range of contexts and conditions which figure prominently in lenition typology.

Until this point, we have treated LAZY as an indivisible scalar constraint: other constraints were ranked completely above LAZY, blocking certain kinds of lenition, or completely below LAZY, triggering certain kinds of lenition. Under this treatment, languages must make a rather drastic choice: for some feature F, either preserve F in all contexts and under all conditions, no matter how exorbitant the effort expenditure (6-1a), or do not bother to preserve F in any context or under any condition, no matter how minuscule the effort expenditure (6-1b).

(6-1)

a. +F in context K	PRES(F)	LAZY
☞ +F		**********
-F	*!	

b. +F in context K'	LAZY	PRES(F)
+F	*!	
☞ -F		*

In this chapter, we enrich the range of constraint interactions, by decomposing the scalar LAZY constraint into a series of constraints which refer to discrete effort thresholds. With this enrichment, the theory is able to account for the generalizations above in terms of effort-based contexts.

1. APERTURE-CONDITIONED LENITION

1.1. THE GENERALIZATIONS

1.1.1. INTERVOCALIC LENITION. It is fairly well established that intervocalic position is a natural lenition environment. Lass and Anderson (1975), for example, identify /V__V as the "prime" lenition context, citing a number of lenitional sound changes from proto-Romance and proto-Uralic (see also Hyman 1975). Of course, intervocalic lenition is not limited to diachronic cases. In Gujarati (Cardona 1965), for example, voiced aspirated stops optionally spirantize in intervocalic position (6-2a), but not elsewhere (b).

(6-2)	a.	wag^{h}əɳ ~ wa̤ɣəɳ	'tigress'		
		sud^{h}i ~ sṳði	'until'		
		ab^{h}ar ~ a̤βar	'thanks'		
	b.	wag^h	'tiger'	lab^h	'profit'
		b^{h}ar	'burden'	pədd^{h}əti	'method'
		d^{h}wədʒ	'flag'	lʌgb^{h}ʌg	'approximately'
		g^{h}ər	'home'	dʒ^{h}əbb^{h}o	'robe'

Additional synchronic cases of intervocalic lenition appear in the following table:

Table 6-1. Lenition in Intervocalic Position.

Language	Reference	Intervocalic pattern
Bashkir	Poppe 1964	b - β
Basque	Hualde 1993	voiced stops, affricates spirantize
Dahalo	Tosco 1991	b,d - β,ð
English (New Zealand)	Holmes 1994	flapping of /t/
Finnish	Sulkala & Karjalainen 1992	r - ɾ
Gooniyandi	McGregor 1990	p - b
Gosiute Shoshoni	McLaughlin 1989	voiceless stops voice and spirantize
Germanic (Gothic)	Bennett 1980	voiced stops spirantize
Guerzé/Kpelle	Casthelain 1952	ɣ - Ø
Hawaiian	Elbert & Pukui 1979	h - Ø
Kannada	Chisum 1975, Schiffman 1983	k - g, glides - Ø
Kanuri	Lukas 1967	b,g - β,ɣ
Kirghiz	Hebert & Poppe 1961	β - w
Macushi	Abbott 1991	voicing of stops and /s/
Mongolian	Poppe 1970	g,ɣ - Ø
Navajo	Kari 1976	delete palatal, velar fricatives
Nkore-Kiga	Taylor 1985	b - ʋ
Paez	Gerdel 1985	x - h
Pennsylvania German	Kelz 1971	h - Ø
Perigourdin	Marshall 1984	p - b, b- β
Purki	Rangan 1979	d,ɖ - ð,ɾ
Sanuma	Borgman 1986	voicing of stops and affricates, h - Ø
Sekani	Hargus 1988	fricative voicing
Somali	Armstrong 1964	voiced stops spirantize, retroflex flaps
Tojolabal	Furbee-Losee 1976	g,r - ɣ, h,w,j - Ø
Totonac/Misantla	MacKay 1984	optional stop voicing
Turkish	Underhill 1976	v - β or w
Tzeltal	Kaufman 1971	voiced stops spirantize
Urubu-Kaapor	Kakamasu 1986	optional stop voicing
West Greenlandic	Fortescue 1984	q - χ or ʁ
West Tarangan	Nivens 1992	k - ʔ
Yana	Sapir & Swadesh 1960	b - w
Yankunytjatjara	Goddard 1985	optional stop voicing

1.1.2. POST- AND PREVOCALIC POSITION. There are also well-known examples of postvocalic lenition. In Tigrinya (Kenstowicz 1982), for example, postvocalic velar and uvular stops spirantize.

(6-3)

Initial and postconsonantal:		Postvocalic:	
kəlbi	'dog'	ʔa-**x**alɨb	'dogs'
		ʔɨti **x**albi	'the dog'
		ʔɨti **x**albi	'the dog'
		su**χ**ti	'silence'
		mɨbɨrax	'bless'
qətəl-a	'kill-3pl.f. perfect'	tɨ-**χ**ətl-i	'kill-2sg.f. imperfect'
ʕarat-**k**a	'bed-2sg.m.'	kətəma-**x**a	'town-2sg.m.'

Conversely, in Mohawk (Bonvillain 1973), stops undergo voicing lenition in prevocalic position.

(6-4)

Prevocalic:		Final and preconsonantal:	
olii**d**eʔ	'pigeon'	zahse**t**	'hide it! (sg.)'
gaalis	'stocking'	wis**k**	'five'
o**d**ahsa	'tail'	a**p**lam	'Abraham'
de**g**eni	'two'	ohjo**t**sah	'chin'
ojaa**g**ala	'shirt'	d͡ʒi**k**s	'fly'
la**b**ah**b**et	'catfish'	deeze**k**w̥	'pick it up! (sg.)'
s**d**uuha	'a little bit'		
des**d**aʔn̥	'stand up! (sg.)'		
deezekw̥	'pick it up! (sg.)'		

Additional cases of post- and prevocalic lenition from the survey are listed below:

Table 6-2. Lenition in post- and prevocalic position.

Language	Reference	Description
Proto-Greek	Sommerstein 1973	s > h / __V
Blackfoot	Frantz 1971, Proulx 1989	h > Ø / V__, w,j > Ø / __V
Djapu Yolngu	Morphy 1979	laminal stops - j / V__
Efik	Dunstan 1969	voiced stops - fricatives / __V, blocked in initial position
Gallo-Romance	Bourciez & Bourciez 1967	voiceless stops - voiced, voiced stops - fricatives / V__
Gitksan	Hoard 1978	stops - voiced / __V
Nepali	Acharya 1991	k^h - x / V__
Tzeltal	Kaufman 1971	voiced stops - fricatives / V__
Yana	Sapir & Swadesh (1960)	stops - partially voiced / __V

1.1.3. GRADIENT APERTURE CONDITIONING. Less commonly noted, but amply attested, are cases where lenition is conditioned by one or both flanking segments, as in the inter-/post-/prevocalic contexts discussed above, but where the requisite aperture of the flanking segments is either somewhat laxer or stricter than vocalic. (The phonetic meaning of "aperture" is made precise in section 1.2 below; for the moment, it may be equated with the more familiar concept of sonority.)

Laxer-than-vocalic. In the Dravidian language Shina (Rajapurohit 1983), for example, voiced stops spirantize in intervocalic position (6-5a), and in post-flap, prevocalic position (b), but not when preceded by another stop (c).

(6-5)	a. baβo	'father'	səðiː	'monkey'	muɣuɾ	'bowl'
	b. daɾβak	'race'	paɾðaː	'veil'	ɡuɾɣuɾ	'churning rod'
	c. ekbo	'alone'	səkdəɾ	'file (tool)'		

That is, spirantization occurs just in case the aperture of the flanking segments is greater than or equal to that of an [ɾ]. Similarly, in Florentine Italian (Giannelli & Savoia 1979), voiceless stops spirantize, typically to approximants, in the context /V__ {V, liquid, glide} (6-6a); in other environments, spirantization is restricted to casual speech, typically to close fricatives (b):

(6-6)	a.	kaɸ̞o	'head'
		ll ɛɾa ɸ̞jɛna	'(s/he was) full'
		pɾaθ̞o	'meadow'
		pjɛθ̞ɾa	'stone'
		biʃix̞letta	'bicycle'
		i x̞wattɾini	'(the) money'
	b.	e lo spɛɾo (formal), e lo sɸɛɾo (casual)	'I hope'
		la ʃesta (formal), la ʃesθa (casual)	'the basket'
		fresko (formal), fresxo (casual)	'fresh'

That is, spirantization is obligatory, and typically results in greater reduction of constriction degree, when the aperture of the flanking segments is greater than or equal to that of a liquid. Laxer-than-vocalic conditioning is also found in one-sided contexts. For example, in Tiberian Hebrew (Malone 1993), non-emphatic stops spirantized in post-vocalic *and post-glide* position.[77] Additional cases of two- and one-sided laxer-than-vocalic lenition contexts are listed below:

Table 6-3. Lenition in laxer-than-vocalic contexts.

Language	Reference	Description
Malayalam	Mohanan 1986	stops - approximants or flaps / +son ___V (blocked in full and partial geminates)[78]
Djapu Yolngu	Morphy 1979	k,p - w / {V, liquid}__
Lowland Murut	Prentice 1971	b,d,ɡ - β,ɾ,ɣ / {V,j,w,ʔ}__
Mbabaram	Dixon 1991	stops - +voi /+nas__ (oblig.), / {V,l,ɾ}__V (opt.)
Uyghur	Hahn 1991	b - β / V__{l,ɾ,V}

(Note that in none of the above cases are there higher sonority segments occurring in the relevant contexts which fail to condition the relevant lenition process.)

Stricter-than-vocalic. Conversely, in Central dialects of Middle Italian, fortis (voiceless) velar stops became lenis (voiced) in case the preceding vowel was unstressed (6-7a) vs. (b), *or when either of the flanking vowels was low* (c) (Grammont 1939: 163):

[77] Modulo an apparent counterbleeding interaction with a syncope process. Benua (1997: 130-138) has reanalyzed this opaque interaction in terms of an output-output faithfulness effect, along similar lines to the treatment of Tigrinya heteromorphemic geminates in Chapter 5. That is, the velar consonant in [malxee] spirantizes in order to maintain its similarity to other members of its paradigm, including [melex] and [məlaaxiim], in which spirantization is conditioned by the preceding vowel.

[78] According to Mohanan's statement of the rule, spirantization is blocked only after nasals. But the example [sunild̪att ewiɾe] ('where is Sunil Dutt?'), on p. 67, suggests that spirantization is blocked after a homorganic lateral as well.

(6-7)	a.	seˈkuru > siˈguro	'sure'
	b.	aˈmiku > aˈmiko	'friend'
		ˈkaeku > ˈtʃieko	'blind'
	c.	ˈlaku > ˈˈlago	'lake'
		ˈmika > ˈmiga	'crumb'
		lakˈtuka > latˈtuga	'lettuce'

In the case of (6-7c), we observe two-sided conditioning, where the requirement of relatively open flanking segments is stricter than vocalic: the lenition process is sensitive to the *height* of the flanking vowels. Moreover, in Tsou, an aboriginal language of Taiwan, a voiced (implosive) coronal stop lenites to a lateral flap just in case the following vowel is low (Wright 1996a, Wright & Ladefoged 1997):

(6-8)	a.	ɺauja	'maple'	b.	ɗiŋki	'muddy place'
		peːɺa	'be able'		teːɗu	'to arrive on time'
		hoɺatsu	'by the time that'		kaɗi	'knee'
					poɗo	'earthworm'

Moreover, the flapping optionally occurs before mid vowels as well, thus [poɗo] ~ [poɺo]. Additional cases of lenition conditioned by only low, or only non-high, flanking vowels are listed below:

Table 6-4. Lenition in stricter-than-vocalic contexts.

Language	Reference	Description
Chitwan Tharu	Leal 1972	b - β /non-high V__non-high V[79]
Korean	Martin 1992	w - Ø /__non-high V
Mbabaram	Dixon 1991	There is a hierarchy of probability of stop voicing: obligatory in /+nas__, optional in /{V,l,ɾ}__V (see Table 6-3), but with higher probability in /a__V, and lower probability in /{l,ɾ}__V
Sotho	Doke 1957, Grammont 1939	d - l /__ non-high V stops > fricatives /non-high V__V
Yakut	Krueger 1962	k - χ before or after a non-high V

Aperture and degree of lenition. Furthermore, the aperture of adjacent segments may condition not only whether lenition occurs, but also the degree of lenition. It has already been noted that in Florentine Italian, voiceless stops typically spirantize to approximants in the context /V__ {V, liquid, glide}; elsewhere, they typically spirantize to close fricatives, or not at all. In Andalusian Spanish, Romero (1996) observes that lenition of the voiced stops is typically more extreme when one of the flanking vowels is low. Specifically, in an EMMA (electromagnetic mid-sagittal articulometer) study, involving three speakers, Romero found that, though /b,d,g/ all are subject to spirantization non-initially, these "stops" involved a slightly, but consistently, more reduced constriction in the context /a__e than in the context /e__e (pp. 58, 66, 69) (cf. Romero's (1992) similar, albeit less thoroughly tested, findings for Castilian Spanish). This documented tendency in Spanish for greater reduction in the vicinity of more open segments is presumably responsible for the loss of /d/ in the past participle suffix /-ado/, typically realized as [ao] or [aw] in natural speech, whereas the allomorph /-ido/ merely spirantizes to [iðo] or [ið̞o] (Resnick 1975).

[79]The Chitwan Tharu pattern is surmised from a few examples in a rather sketchy grammar, and may therefore be a less solid piece of evidence than the other cases in this table.

Generalization. The particular lenition patterns noted above are subsumed by the following broader generalization:

(6-9) The Aperture Conditioning Generalization

Ceteris paribus, if a consonant C lenites when preceding (or following) X, and X' has aperture greater than or equal to X, then C lenites, to the same extent or to a greater extent, when preceding (or following) X' as well.

The aperture scale which the typology appears to reflect is:

(6-10) low vowels > mid vowels > high vowels > {glides, liquids} > stops . . . > full or partial geminates

(The "..." indicates that geminates are "off the scale," in the sense that reduction of part of a geminate *never* occurs, as discussed in Chapter 5.) It is thus predicted, for example, that if /d/ → [ð] /a__i, /d/ must also spirantize, or reduce further, in the context /a__a; or if /f/ → [h] / ɾ__i, it must also debuccalize, or reduce further, in intervocalic position. It should be noted, however, that the "ceteris paribus" in (6-9) includes a broad array of phonetic factors, which weaken the empirical force of the generalization. For example, in Quechua (Whitley 1979), velar and uvular stops spirantize in coda position, i.e. in the context /V__C but not /V__V, notwithstanding the greater aperture of the latter context. But recall from Chapter 1, section 3.3.2.2, that the coda lenition context is attributable to the impoverished perceptibility of consonants in positions where they lack an audible release. Lenition in coda (or more precisely, unreleased) position thus exemplifies an effect of supervening perceptual considerations, in derogation of the Aperture Conditioning Generalization. Similarly, in Table 6-4 above, it is noted that in Mbabaram

(Dixon 1991), voicing lenition is obligatory in post-nasal position, but optional in intervocalic position, notwithstanding the greater aperture of the latter context. In this case, the supervening consideration is aerodynamic: as discussed in Chapter 5, Hayes & Stivers (1997) demonstrate that voicing is particularly natural in post-nasal obstruents, because the raising of the velum which occurs toward the end of the nasal, and which continues well into the obstruent, serves to expand the oral cavity during the production of the obstruent, thus facilitating the transglottal airflow which is needed to sustain voicing. However, among the non-post-nasal contexts, voicing in Mbabaram is more likely in /a__V position than /i__V position, and more likely in the latter than in /l__V position; hence, modulo post-nasal position, the Mbabaram facts do conform to the Aperture Conditioning Generalization.

This connection between lenition contexts and high aperture flanking segments is by no means a novel observation. Grammont (1939), discussing cases of intervocalic, and stricter-than-vocalic lenition, explicitly invokes the notion of aperture, and anticipates the effort minimization account developed in this Chapter:

> [L]es voyelles, qui ont toujours plus d'aperture que n'importe quelle consonne, tendent à augmenter l'aperture de la consonne; c'est encore de la moindre action (p. 200).
>
> Plus les voyelles sont ouvertes, plus la position qu'elles demandent aux organes est éloignée d'une occlusion et la rend difficile. (p. 163).

Lass and Anderson (1975:162) express roughly the same observation, albeit somewhat less explicitly: “[T]he most likely condition under which lenition will fail is if the segment (otherwise) affected is contiguous to another strong [i.e. highly constricted] segment."

1.2. PHONETIC EXPLANATION

1.2.1. EFFORT-QUA-DISPLACEMENT. Recall from Chapter 2, section 2.4 that, all else being equal, the greater the displacement involved in a gesture, the greater the force required for the gesture, hence the greater the effort cost thereof. In previous chapters, this inference has been applied to comparisons of gestures involving different degrees of constriction: we inferred, for example, that a stop is more effortful than a corresponding (nonstrident) fricative, assuming that both gestures start at and return to the same position, in the same amount of time, because of the greater displacement required for the former:

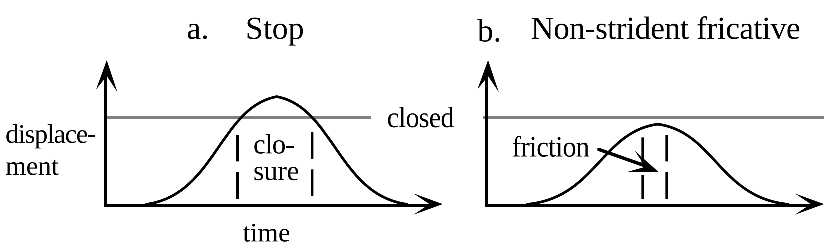


Figure 6-1. Comparison of displacement in consonants of differing constriction degrees.

By the same token, we can compare the relative displacement of gestures involving the same constriction degree, but beginning from (and/or returning to) positions of greater or lesser constriction:

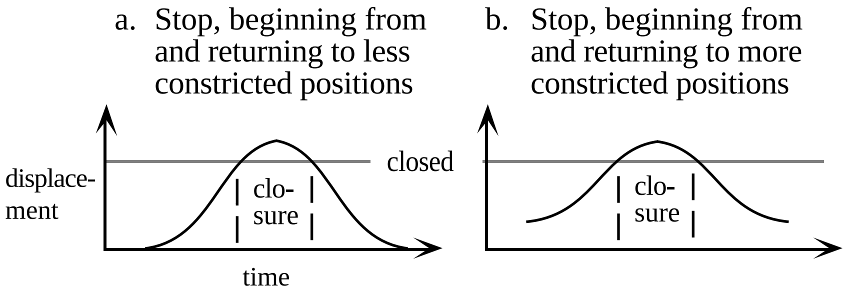


Figure 6-2. Comparison of displacement in consonants of the same constriction degree, but
different starting/ending points.

1.2.2. DISPLACEMENT OF THE SAME ARTICULATOR. The phonetic application of this observation can be seen most directly in cases where the consonantal constriction (i.e. the lenition target) involves the same primary articulator as the flanking segments. Specifically, both vowels and velar consonants involve the tongue dorsum as the primary articulator; thus, the diagrams in Figure 6-2 could be understood as schematics of the vertical displacement of (some point on) the tongue dorsum in, for example, the sequence [aga] (Figure 6-2a) versus [əgə] (b). We can now infer that the [g] incurs a higher effort cost in the context /a__a than in the context /ə__ə, due to the greater tongue dorsum displacement in the former. And by the same reasoning, a dorsal consonant of any given degree of constriction incurs a higher effort cost in the former context than in the latter context.

1.2.3. DISPLACEMENT OF DIFFERENT ARTICULATORS, AND THE ROLE OF THE JAW. But even in cases where the target consonant and the flanking segments involve different primary articulators (labial, coronal, or dorsal), the movement of these supralaryngeal articulators is never completely independent. In the case of the coronal and dorsal articulators, this connection is obvious: one cannot make a closure with the tongue blade without raising the tongue body somewhat, nor vice-versa. More generally, the labial, coronal, and dorsal articulators are all attached to, and ride upon, the jaw. Assume that the jaw is in low position, for an [a], and that the articulator (be it labial, coronal or dorsal) must then move up to form a stop closure, followed by resumption of the [a] position. There are two strategies to consider: either the jaw remains in low position during the stop (Figure 6-3a), in which case the primary articulator must travel further to achieve the closure (relative to the distance it must travel in /ə__ə position, Figure 6-3c); or the jaw must raise during the stop and then return to low position, so that additional displacement of the primary articulator is not required (Figure 6-3b).

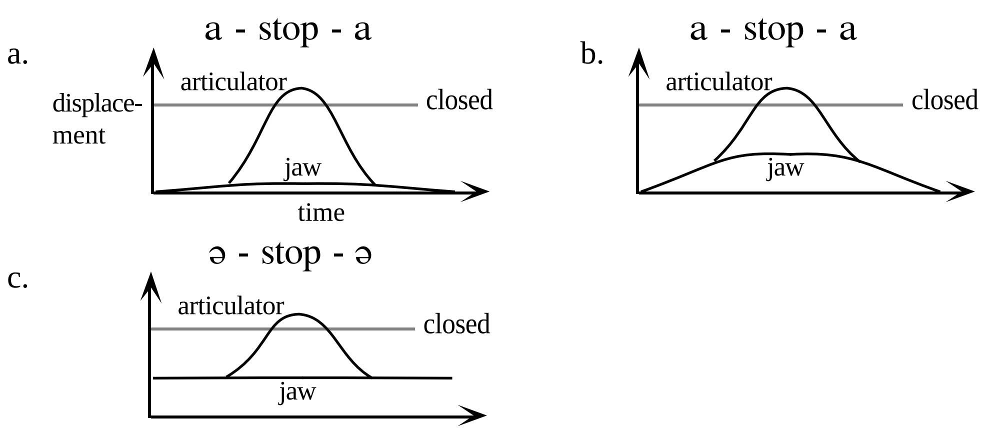


Figure 6-3. Schemata of strategies of jaw/articulator displacement in intervocalic stops.

Or, more typically, there may be some compromise between the pure articulator-raising and pure jaw-raising strategies (cf. the discussion below of Keating et al. 1994). But regardless of how the displacement is apportioned between the jaw and the articulator, the jaw/articulator *ensemble* must travel further to reach the closure and opening targets in an [a-stop-a] sequence than in an [ə-stop-ə] sequence, regardless of the articulator involved. And by extension, we can infer that, all else being equal, the lower the jaw position of the flanking segments, the more displacement, hence effort, is required to achieve any given degree of consonantal constriction.

Thus, in the most general case, where the only link between the articulators of the consonant and those of the flanking segments is the jaw, we may understand the Aperture Conditioning Generalization as referring to *jaw aperture* in particular. And in the case of a more direct link, such as between the tongue blade and body, or, a fortiori, involving the same primary articulator in both the consonant and the flanking segments, the Generalization can be understood as referring to the aperture of whatever articulator is involved in the flanking segments. Consequently, all else being equal, more displacement is involved in, hence there is greater impetus to lenite, any given consonantal constriction, the greater the aperture of the flanking segments. To obtain effort-based lenition contexts from this notion of effort-qua-displacement, it is simply necessary to impose language-specific cut-offs on the amount of effort which may be expended for the preservation of particular auditory properties, as discussed in section 1.3 below.

This essence of this account is anticipated by de Jong et al. (1992), who model intervocalic spirantization and voicing in terms of blending of Articulatory Phonology-style gestures. By increasing the overlap between the vowel and consonant gestures, more drastic blending occurs, resulting in undershoot of the consonantal jaw raising: “Again we observe that the oral closure gesture for the consonant shares at least one articulator [i.e. the jaw] with the tongue-body gesture for the neighboring vowels" (p. 49). This is implicitly an effort-based explanation: with greater overlap, the jaw/articulator unit must move from open to closed position faster; and if this is to be done *without increasing the effort* of the constriction gesture beyond some threshold, displacement must be reduced, hence undershoot/lenition.

1.2.4. ADDITIONAL PHONETIC ISSUES. *Horizontal displacement.* We have simplified the discussion, however, by considering only *vertical* displacement. Of course, the trajectory of a gesture involves displacement in both the vertical and horizontal dimensions; and it is therefore conceivable that the overall displacement, hence effort, might be greater in, for example, an [ege] sequence than in an [aga] sequence, despite the lesser vertical displacement in the former, due to the fronter position of the tongue body in the flanking vowels.[80] However, the trajectories of consonant constriction and opening gestures shown in Romero's (1996) EMMA study indicate that vertical displacement generally predominates over horizontal displacement (that is, the angle of the trajectory is typically well over 45˚), regardless of the flanking vowels. Presumably, this reflects the fact the horizontal displacement is frequently reduced, due to coarticulatory fronting or backing of the consonant's constriction location target, e.g. the typically observed palatalization of velars in front vowel environments ([egʲe] vs. [aga]); and/or coarticulatory fronting/backing of the vowels themselves, e.g. in English [tʰʏ̠t] ('toot') vs. [kʰʉk] ('kook').[81]

Studies of jaw aperture. A further phonetic issue concerns the extent to which jaw aperture actually corresponds to the hierarchy of lenition triggers posited in (6-10). This scale is juxtaposed with the hierarchies observed in the jaw movement study of Lindblom 1983, reflecting measurement of consonant jaw height in /aˈ__ɑː, /ˈɑː__, and /ˈ__aː[82] positions in Swedish, and Keating et al.'s (1994) follow-on study of consonant

[80]A possible instance of such a horizontal displacement effect is discussed in Chapter 7 (variation between lenition to a flap vs. a dentral fricative in Tümpisa Shoshone.

[81]This coarticulatory adjustment of constriction location can itself be regarded as a species of lenition.

[82]The article says ˈa__ː position; but I presume that this is a typographical error, as the text states that measurements were taken "for both final and initial positions of the consonants," and final (/ˈɑː__) position is already accounted for (p. 241).

jaw height in /i__i, /e__e, and /a__a positions, for English and Swedish (both jaw hierarchies are based on averaging of measurements across contexts):

(6-11) Lindblom 1983:

ɑː > r,l > j,v > m,n,ŋ > p,t,k,b,d,g,f > s,ɧ[83] (Swedish)

Keating et al. 1994:

a > e > i > k > b > n > l > f > r > d > t > s (English)

a > e > i > l > k > b > r > n > f > d > t > s (Swedish)

Lenition-trigger hierarchy (repeated from 6-10):

low V > mid V > high V > {glide, liquid} > stop .... > full or partial geminate

Keating et al.'s study, which investigated a variety of flanking vowels (unlike Lindblom's), accords with the lenition-trigger ranking low V > mid V > high V: the jaw aperture for low vowels (averaging across context and language), is 9.15 mm; the aperture for mid vowels is 8.35 mm; and for high vowels, 5.29 mm. Moreover, the lenition-trigger ranking of all vowels below all consonants likewise accords with Keating et al.'s findings: for each consonant, the jaw aperture is less than that of its flanking vowels; and Lindblom's study is consistent with this as well.[84] Furthermore, Lindblom's study accords with the lenition-trigger ranking {liquids, glides} > {nasals, obstruents}.[85] Keating et al.'s study is more equivocal: among the alveolars, liquids indeed have greater jaw aperture than obstruents; though there is a local reversal between [n] and [l] in

---

[83][ɧ] is the IPA symbol for a sibilant fricative with simultaneous postalveolar and velar constriction.

[84]Modulo [h], not shown in (6-10). Lacking supralaryngeal constriction, [h] has roughly the same jaw aperture of that of its flanking vowels.

[85]I conjecture that this /v/ is realized as an approximant, [ʋ̣] rather than a fricative [ʋ̞]. (Note that Swedish does not have a distinct /w/.)

English; and the liquids fall below [k] and [b] (and [f], in English). But Keating et al.'s study only considers intervocalic contexts. Unfortunately, neither study squarely addresses the further question of the influence of consonants on the jaw height of adjacent consonants; but on this point, Lindblom's findings are more apposite than Keating et al.'s, as his hierarchy reflects a broader range of contexts (intervocalic, initial, and final). Presumably this provides a better indication of the basic jaw aperture of the consonant, abstracting away from context.[86] We therefore arrive at the following composite aperture hierarchy (subject to variation, depending on the phonetic details of the sounds in particular languages):

(6-12) Composite aperture ranking, consistent with jaw movement studies and the lenition-trigger hierarchy:

low V > mid V > high V > liquids > glides > nasals > stops > strident fricatives > ... > full or partial geminate

Displacement-based reduction of vowels? This displacement-based account of consonant lenition contexts would appear to predict that vowels likewise should tend to reduce when flanked by low-aperture consonants (cf. Fowler & Saltzman 1993, whose model of consonant and vowel "blending" predicts such effects). I am aware of no such patterns of vowel reduction. However, I do not believe this indicates that the phonetic displacement-based account of lenition contexts is incorrect, but rather that the effects of increased displacement are not symmetric for consonants and vowels. As Keating et al. (p. 419) found,

[86]This is under the assumption that the actual jaw aperture during a consonant cluster will involve some sort of compromise between the target apertures of each member consonant. This is the "canonical" jaw aperture which the consonant wants to impose on its neighbors.

> [T]he effect of Vowel Context on consonant [jaw] height was robust, with consonants higher when between /i/ vowels than between either of the other two vowels. . . . In sum, consonants varied reliably in height as a function of vowel context, whereas vowels showed only a statistical tendency to vary in height as a function of consonant context.

This asymmetry may plausibly be attributed to the higher perceptual salience of vowel height distinctions (as reflected, for example, in the generalization that all natural languages employ vowel height contrasts) relative to consonant constriction distinctions. Alternatively, the asymmetry may reflect the greater importance of prosodic, rather than coarticulatory, factors on the realization of vowel height.

Aperture and voicing lenition. From the discussion of aerodynamic factors in Chapters 2 and 5, it is not immediately apparent why greater aperture of flanking segments should condition voicing lenition. For an obstruent of typical singleton duration to be passively voiced, it is sufficient for it to occur in medial position preceded by a sonorant; and there is no aerodynamic reason why this should be more likely to occur when flanked by low vowels than high vowels. Yet such conditioning of voicing is attested, for example, in Macushi (Abbott 1991), and over a dozen other languages listed in Table 6-1. Recall, however, from Chapter 2 that distinctions in obstruent voicing are typically accompanied by distinctions in constriction duration, i.e. the traditional "fortis/lenis" distinction.

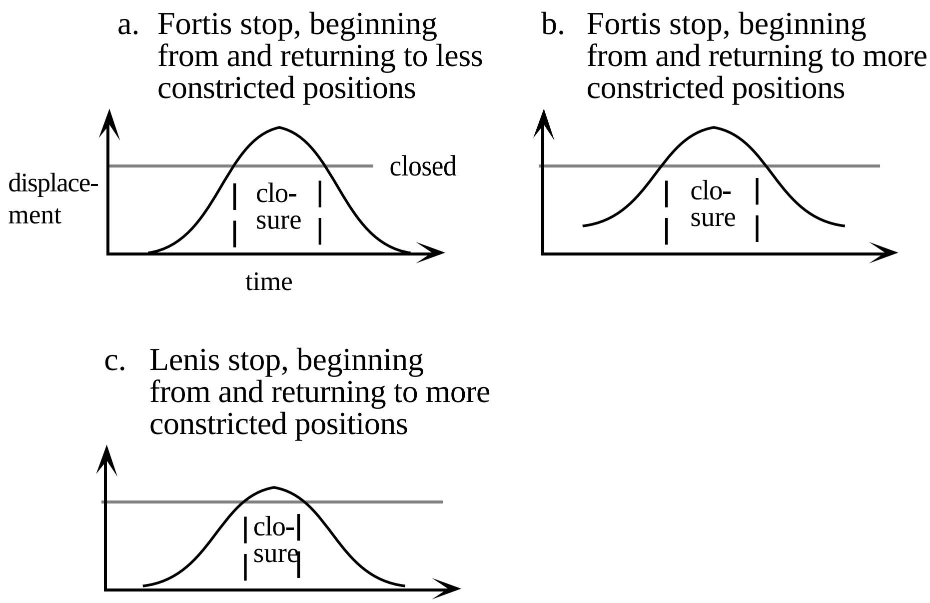


Figure 6-4. Comparison of fortis and lenis stops, with different starting/ending points.

As shown in Figure 6-4, a fortis closure in a high-aperture context (a) involves more displacement than a fortis stop in a lower-aperture context (b), or a lenis stop in a high-aperture context. Thus, there is greater impetus to lenite the fortis stop to a lenis stop in the high-aperture context. I assume, then, that cases of aperture-conditioned "voicing" lenition primarily involve reduction of closure duration, with voicing as a concomitant effect of this temporal reduction.

Laryngeal consonants. Since laryngeal consonants ([h, ʔ]) lack any oral constriction, it is not immediately apparent why greater aperture of flanking segments should condition lenition of such segments. In fact, intervocalic loss of /h/ is amply attested, for example in Hawaiian and a number of other languages listed in Table 6-1. However, such patterns can be explained in terms of aerodynamics rather than displacement. That is, more highly constricted segments tend to facilitate friction, documented by Ohala (1983), and attributed by him to the aerodynamic consequences of higher tongue body position.[87] Specifically, by narrowing the channel of oral airflow, high tongue body position makes it more likely that the airflow through the constriction will become turbulent. The one putative case of intervocalic /ʔ/ deletion in the lenition survey, Gbeya (Samarin 1966), can be reanalyzed as context-free deletion, with blocking in initial position, due a context-sensitive (word-initial) faithfulness constraint.

1.3. FORMAL CAPTURE OF THE DISPLACEMENT EXPLANATION: TWO-SIDED CONTEXTS

1.3.1. CONSTRAINT "BINARIZATION" AND INTERLEAVING. Ostensibly, a central tenet of Optimality Theory, distinguishing it from more mainstream connectionist approaches to constraint satisfaction, is the assumption of strict domination: i.e. a single violation of a higher-ranked constraint weighs more heavily in the evaluation of candidates than the most egregious violation of lower-ranked constraints (see Prince & Smolensky 1993, Chapter 5). Yet, from its inception, OT has countenanced a certain weakening of this claim. A scalar constraint C may be "binarized" into a series of discrete binary constraints (... C_3 » C_2 » C_1), each referring to some degree of violation of C, and other constraints may be interleaved in the series (C_3 » Other Constraint » C_2). With this end-run around strict domination, violation of the Other Constraint weighs more heavily than some minor violation of C (C_2), but not some more serious violation of C (C_3); the grammar thus sets a threshold for violation of C, tolerating minor violations, but prohibiting violations of strength C_3 or greater.

Thus, for example, Prince and Smolensky (1993) binarize the scalar constraint HNUC (syllable nuclei should be maximally sonorous) into a set of discrete "peak affinity" constraints of the form P/x, which in effect penalize segment x (or other

[87]This could be due either to speaker-controlled movement of the tongue body, or due to elevation of the tongue body as a result of a higher jaw position.

segments of equal or lower sonority) in syllable peak position; these constraints have a fixed internal ranking P/x » P/y » P/z, etc., where x is lower in sonority than y, y lower than z, and so on. By interleaving faithfulness constraints, such as PARSE, in the peak affinity scale, Prince and Smolensky obtain language-specific phonotactic cut-offs, such as "nuclei must be vowels or liquids": under the ranking P/n » PARSE » P/l, syllabic liquids (and higher sonority nuclei) are tolerated, in order to satisfy PARSE, while syllabic nasals (and lower sonority nuclei), i.e. more egregious violations of HNUC, are prohibited.

The binarization technique extends straightforwardly to the effort threshold problem:

(6-13) LAZY "binarized":

... LAZY_{n+1} » LAZY_n » LAZY_{n-1} ... (where LAZY_n ="Do not expend effort ≥ *n*")

The ranking within the series is universally fixed: by Pāṇini's Theorem (the OT counterpart to the Elsewhere Condition), a specific constraint, if active, must dominate a more general constraint (Prince & Smolensky 1993, Chapter 5).[88] If *x* refers to a higher effort threshold than *y*, a constraint banning *y* necessarily bans *x* as well (but not vice-versa), thus LAZY_x and LAZY_y stand in the relation of specific-to-general; hence LAZY_x, if active, must dominate LAZY_y.

We can now interleave lenition-blocking constraints within the LAZY series, to obtain language-specific effort/displacement thresholds, corresponding to the sorts of

[88]Modulo certain narrowly defined conditions, identified by Prince (1997), under which a specific constraint may be active on certain inputs, even though outranked by a more general constraint.

aperture-sensitive lenition contexts identified in section 1.1 above. This is illustrated most generally with an abstract hypothetical example:

(6-14)

		LAZY$_{216}$	PRES(F)	LAZY$_{215}$
☞	+F in context K			*
	-F in context K		*!	
	+F in context K'	*!		*
☞	-F in context K'		*	

That is, in context K, let us assume that implementing the feature +F only requires 215 units of effort. Under this ranking, satisfaction of PRES(F) is more important than avoidance of expending this amount of effort, and so the form surfaces with the unlenited specification, +F, in context K. However, in context K', the cost of implementing +F is slightly higher; expenditure of 216 units of effort is *more* costly than violating PRES(F) under this ranking. Consequently, in context K', the lenited specification, -F, is selected. We thus obtain the result that in contexts where some consonant requires more effort, that consonant lenites; whereas it does not lenite in contexts where its articulatory realization is less effortful.

1.3.2. DOMAIN OF EFFORT THRESHOLDS. A remaining question concerns the temporal scope of effort thresholds. Longer words typically involve more articulatory gestures than shorter words, and therefore require more overall effort expenditure. This predicts that, the longer the word, or even utterance, the more extensive the lenition, e.g. we should be able to find lenition patterns such as /ibi/ - [ibi], but /ibibibi/- [iβiβiβi], and /ibibibibibi/ - [iiiiii]. While word length does appear to be one of the factors conditioning gestural reduction (cf. Lehiste's (1977) observation that segments are shorter in long words), the effect is not as strong or invariant as the above approach predicts. Rather, the effect can presumably be explained in terms of the observation that long words are typically realized at a faster speech rate than short words; and as discussed in section 3 below, speech rate conditions lenition.

It is therefore necessary to restrict the scope of the effort threshold to some more local domain, such as the syllable.[89] That is, LAZY$_n$ = "Do not expend effort ≥ n within a syllable." Therefore, assuming the form [ibi] incurs an effort cost of x units, the form [ibibibi] violates exactly the same effort threshold, notwithstanding the greater length of the latter (though it incurs *multiple* violations of this threshold):

(6-15)

		LAZY$_{x+1}$	PRES(cont)	LAZY$_x$	LAZY$_{x-1}$
☞	ibi			*	*
	iβi		*!		*
☞	ibibibi			***	***
	iβibibi		*!	**	***

Therefore if the consonant in /ibi/ does not spirantize, then neither do the consonants in /ibibibi/. This is not say that the *computation* of effort cannot encompass effects which span a syllable boundary. For example, we can still capture the higher effort cost of a geminate, e.g. in [ak.ka], vs. a singleton, e.g. in [a.ka]:

[89]This syllable-based definition of effort threshold domains seems to yield roughly adequate empirical results (modulo the problem of context stabilization, see Chapter 9). In principle, however, one would expect the domain to reflect the time that it takes for the muscle groups involved in each gesture to recover from momentary fatigue. Fatigue, in this sense, refers to depletion of ATP (i.e. muscle "fuel", see Chapter 2) within a motor group, before the bloodstream replenishes the supply. Therefore the domain presumably should be relativized to the proportion of "red" (fast recovering) and "blue" (slow recovering) muscle fibers in the relevant muscle groups (see generally Borden & Harris 1980). A fuller exploration obviously involves careful articulatory experimentation, as well as study of the physiology; and I leave this as a topic for further research.

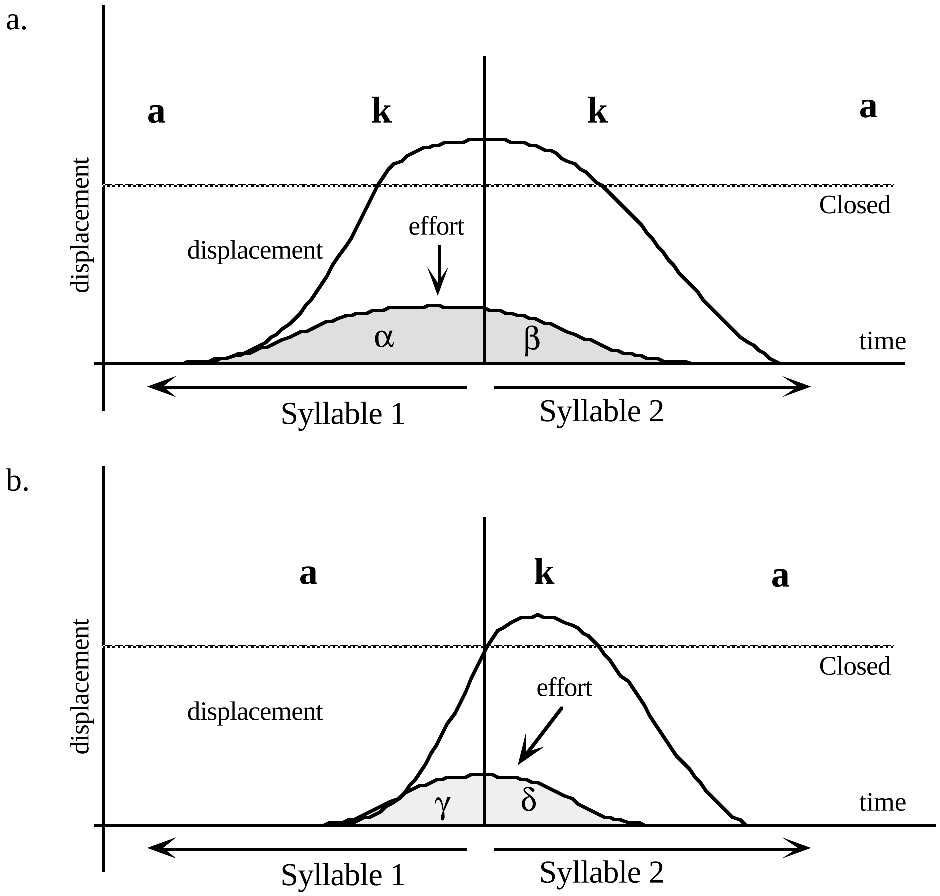


Figure 6-5. Syllabic apportionment of effort cost of consonants in [akka] vs. [aka].

That is, Syllable 1 in Figure 6-5a incurs a substantially higher effort cost than Syllable 1 in Figure 6-5b (α >> γ); Syllable 2 in Figure 6-5a incurs a slightly higher effort cost than Syllable 2 in Figure 6-5b (β > δ). Thus, local (syllable-by-syllable) evaluation of effort thresholds still reflects the greater effort cost of the geminate, though the geminate is split between two syllables. In terms of effort thresholds:

(6-16)

	LAZY$_{\alpha}$	...	LAZY$_{\gamma}$	LAZY$_{\beta}$	LAZY$_{\delta}$
a.ka		...	* (σ_1)	* (σ_1)	* (σ_1) * (σ_2)
ak.ka	* (σ_1)	...	* (σ_1)	* (σ_1) * (σ_2)	* (σ_1) * (σ_2)

Note that, by the logic of effort thresholds, if σ_1 violates LAZY$_{\alpha}$, then σ_1 also violates all lower thresholds, LAZY$_{\gamma,\beta,\delta}$; and the same for σ_2. Henceforth, I refrain from indicating the multiple violations of particular effort thresholds incurred by longer forms, where this is not crucial.

1.3.3. FROM EFFORT THRESHOLDS TO EFFORT-BASED LENITION CONTEXTS. We will now show that intervocalic spirantization and similar effort-based lenition contexts can be obtained by interleaving PRES(cont) (or some other lenition-blocking constraint) at particular points within the LAZY series. Let us label the relevant effort thresholds as follows:

(6-17) a = effort required for b in /i__i position.

b = effort required for b in /r__i position.

c = effort required for β in /i__i position.

d = effort required for β in /r__i position.

It follows from considerations of relative displacement that the following relations hold among these effort levels:[90]

[90]To determine the relation of b to c, we need more information, namely which of the two gestures involves greater displacement; but the analysis is unaffected by this determination.

(6-18) a > b (more displacement required in "a" due to more open preceding segment)

a > c (more displacement required in "a" due to greater constriction degree)

b > d (more displacement required in "b" due to greater constriction degree)

c > d (more displacement required in "a" due to more open preceding segment)

a > d (by transitivity)

Therefore, the ranking must be LAZY$_a$ » {LAZY$_b$, LAZY$_c$} » LAZY$_d$.

Intervocalic. Let PRES(cont) (or some other spirantization-blocking constraint fall between LAZY$_a$ and LAZY$_b$. This ranking yields the pattern of spirantization in intervocalic position, but not elsewhere, as in Bashkir, Basque, Dahalo, Gosiute Shoshoni, Gothic, Gujarati, Nkore-Kiga, Perigourdin, Somali, Tzeltal, and Yana (see Table 6-1).

(6-19)

		LAZY$_a$	PRES(cont)	LAZY$_b$	LAZY$_c$	LAZY$_d$
	...ibi... - ...ibi...	*!		*	*	*
☞	...ibi... - ...iβi...		*		*	*
☞	...rbi... - ...rbi..			*	?	*
	...rbi... - ...rβi..		*!			*

To reiterate, the intuition behind this formalism is that there is greater impetus to lenite more effortful gestures than easier gestures; hence a given consonant may be targeted for lenition in high-effort environments (e.g. intervocalic), to the exclusion of lower-effort environments (e.g. non-intervocalic), but never vice-versa, in accordance with the Aperture Conditioning Generalization.[91]

---

[91]Modulo supervening phonetic considerations, such as perceptual salience and aerodynamic factors, as discussed in section 2.1.

Laxer-than-vocalic. Now let us consider the consequences of ranking the spirantization-blocking constraint at other points along the LAZY series.

(6-20) e = effort required for [b] in /k__i position.

Again, by considerations of relative displacement due to the flanking segments, b > e, and a > e by transitivity (and the ranking of e relative to c and d is undetermined and irrelevant to the analysis). Therefore the ranking is LAZY$_a$ » LAZY$_b$ » LAZY$_e$. If the spirantization-blocking constraint is shifted downward between LAZY$_b$ (the effort required for b in /r__i position) and LAZY$_e$ (the effort required for b in /k__i position), we obtain the pattern of spirantization when flanked by [ɾ], or a higher-aperture segment, attested in Shina (6-4), Florentine Italian (6-5), and Malayalam and Uyghur (Table 6-3).

(6-21)

		LAZY$_a$	LAZY$_b$	PRES(cont)	LAZY$_e$
	...ibi... - ...ibi...	*	*		*
☞	...ibi... - ...iβi...			*	
	...rbi... - ...rbi...		*!		*
☞	...rbi... - ...rβi...			*	
☞	...kbi... - ...kbi...				*
	...kbi... - ...kβi...			*!	

More generally, if the spirantization-blocking constraint is shifted *downwards* in the LAZY series, lenition occurs in a *broader* class of environments. We thus capture the connection between intervocalic position and the laxer-than-vocalic two-sided contexts.

Stricter-than-vocalic. Conversely, if the spirantization-blocking constraint is shifted *upwards* in the LAZY series, lenition occurs in a *narrower* class of environments.

(6-22) f = effort required for b in /e__e position

Because of the greater displacement of the stop gesture when flanked by non-high vowels, f > a; therefore the ranking is LAZY$_f$ » LAZY$_a$. If the spirantization-blocking constraint is ranked between LAZY$_f$ and LAZY$_a$, we obtain the pattern of spirantization when flanked by non-high vowels, attested in Chitwan Tharu (Table 6-4).

(6-23)

		LAZY$_f$	PRES(cont)	LAZY$_a$
	...ebe... - ...ebe...	*!		*
☞	...ebe... - ...eβe...		*	
☞	...ebi... - ...ebi...			*
	...ebi... - ...eβi...		*!	
☞	...rbi... - ...rbi...			
	...rbi... - ...rβi...		*!	

Thus an upwards shift of the spirantization-blocking constraint in the LAZY series results in spirantization which is sensitive to vowel height. In sum, broader or narrower environments for spirantization depend on the ranking of the spirantization-blocking constraint within the LAZY series. As extreme cases, ranking of the spirantization-blocking constraint above the entire LAZY series results in complete absence of spirantization processes in the grammar, while ranking it below the entire LAZY series results in context-free spirantization. But since the ranking of the series itself is fixed, if lenition occurs in an environment involving less displacement, it must also occur in an environment involving more displacement, thus capturing the Aperture Conditioning Generalization.

This mode of analysis straightforwardly extends to all types of lenition in effort-based contexts. As discussed in Chapter 1, section 3, the type of lenition obtaining in a given language depends on the identity of the active lenition-blocking constraints. For example, the Kirghiz pattern of reduction of fricatives to glides (Table 6-1) is obtained by ranking PRES(son) between LAZY$_c$ (effort required for β in /i__i position) and LAZY$_d$ (effort required for β in /r__i position).

(6-24)

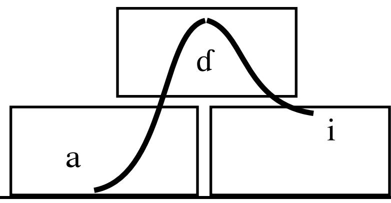
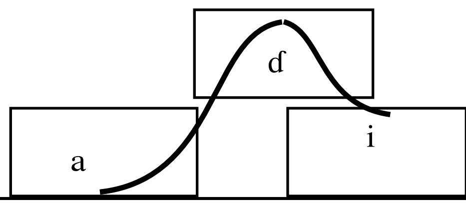
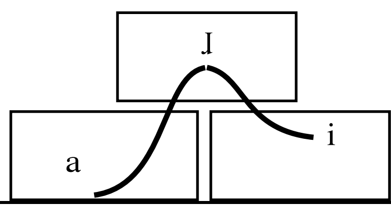
		LAZY$_c$	PRES(son)	LAZY$_d$
	...iβi... - ...iβi...	*!		*
☞	...iβi... - ...iwi...		*	
☞	...rβi... - ...rβi...			*
	...rβi... - ...rẉi.		*!	

That is, the underlying [-son] specification of the input fricative is sacrificed to avoid the relatively high effort of an obstruent in intervocalic position, but not elsewhere. This treatment further extends to patterns whereby lenition occurs in some broad array of contexts, but occurs to a greater extent, or with higher probability, in sub-contexts in which greater effort is required, e.g. Florentine Italian and Castilian Spanish, as discussed in section 1.1.3. Analyses of such patterns are presented in the case studies of Chapters 7 and 8.

1.4. ONE-SIDED CONTEXTS

1.4.1. RIGHT-HAND CONTEXTS. Recall that in Tsou (6-8) the voiced coronal stop lenites before a low vowel, but not after one. That is, the flanking segment to the right, but not to the left, appears to relevant in conditioning the lenition process. This asymmetry poses a problem for the effort-based approach developed above: why does lenition occur in /iɗa/ - [iɹa], but not in its mirror image, [aɗi], since both involve the same overall displacement? Note, however, that reduction in displacement is not the only strategy for avoiding violation of some degree of LAZY. The effort consequences of a closure gesture in a high displacement context can be offset by reducing the velocity of the gesture, without leniting. This is shown in (6-25):

(6-25)

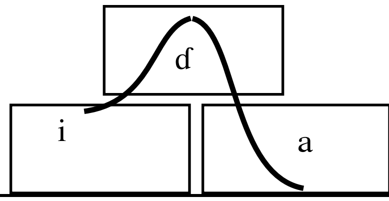
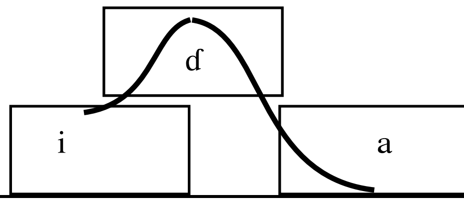
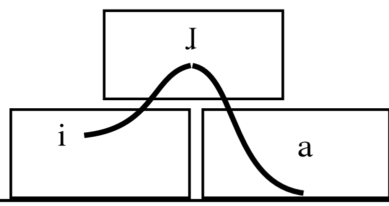
		LAZY$_x$	PRES(son)	LAZY$_y$	LAZY$_z$
a.	High displacement, high velocity: effort = *x*  	*!		*	*
b. ☞	High displacement, lower velocity: effort = *y*  			*	?
c.	Lower displacement, high velocity: effort = *z*  		*!	?	*

Considerations of displacement and velocity dictate that $x > y$ and $x > z$ (and the relation of *y* to *z* is undetermined and irrelevant); therefore the ranking is LAZY$_x$ » {LAZY$_y$, LAZY$_z$}. By reducing the overlap between the tongue-body-lowering gesture of the [a] and the tongue-tip-closure gesture of the [ɗ] (candidate b vs. a), the closure gesture is given more time to reach its target, i.e. the velocity of the gesture is lowered. Consequently, the effort cost of the closure gesture remains below the *x* threshold, and the lenited candidate, c, is avoided.

But what rules out this "slow-down" strategy in /iɗa/ - [iɗa]? If the transition from the stop to the following vowel were attenuated, it would give rise to the percept of an on-glide, i.e. [iɗʲa]. This is presumably a graver violation of faithfulness than the mirror image, [aʲɗi], because of the greater perceptual salience of C-to-V transitions than V-to-C transitions (see Bladon 1986 on the auditory basis of this asymmetry, and Steriade 1993b, 1995b, and Silverman 1995, on its implications for phonological theory). This account can be formalized in terms of a constraint referring to C-to-V transitions, more specific than, therefore universally ranked above, a more general faithfulness constraint:

(6-26) PRES(transition properties/C_V) » PRES(transition properties)

The CV faithfulness constraint is inoperative in tableau (6-25), since the candidates differ with respect to the V-C rather than C-V transitions. With input /iɗa/, however, in tableau (6-27), reduction in overlap of ɗ-a (b) results in a violation of C-V faithfulness.

(6-27)		PRES(transition properties/C_V)	LAZY$_X$	PRES(son)
a.	High displacement, high velocity effort = *x* 		*!	
b.	High displacement, lower velocity: effort = *y* 	*!		
c. ☞	Lower displacement, high velocity: effort = *z* 			*

The only way, then, to avoid a set of gestures that crosses effort threshold *x* (6-27a), is to lenite the stop (candidate c). The difference in Tsou between the realization of a coronal stop before a low vowel (lenited) and after a low vowel (unlenited) is attributable to the C-V transition faithfulness constraint. Since this faithfulness constraint does not apply to the transition from a low vowel into a following stop, the slow-down strategy is available as an alternative to lenition. The aperture of the preceding segment, then, will have an effect on the timing of the closure gesture, but will not condition lenition; whereas the aperture of the following segment does condition lenition. Note that this treatment of right-hand contexts makes strong predictions regarding the duration of formant transitions in lenition contexts. In Tsou, for example, it is predicted that the duration of the closure transitions in a [aɗ] sequence are significantly longer than the closure transitions in [iɗ], and the opening transitions in [ɗi] or [ɺa].

1.4.2. LEFT-HAND CONTEXTS. In light of the presumed asymmetry of CV and VC transitions, the slow-down analysis proposed for right-hand contexts, such as Tsou prevocalic flapping, cannot be extended to left-hand contexts, such as Tigrinya postvocalic spirantization. Rather, such postvocalic contexts can be viewed as the union of two independently attested lenition contexts: intervocalic position and coda position (e.g. Quechua, Whitley 1978, spirantization of coda velars). Recall from Chapter 1 that coda lenition is analyzed in terms of differential ranking of PRES(cont), depending on whether or not it occurs in a context in which it is released (cf. Jun 1995); and as discussed above, intervocalic lenition is attributed to interleaving of faithfulness constraints within the LAZY hierarchy.

(6-28) Tigrinya postvocalic spirantization:

		LAZY_x	PRES(cont w/ release)	$\text{LAZY}_{x\text{-}1}$	PRES(cont)
a.	mɨrak-na - mɨrakna			*!	
☞	mɨrak-na - mɨraxna				*
b. ☞	ʕarat-ka - ʕaratka			*	
	ʕarat-ka - ʕaratxa		*!		*
c.	kətəma-ka - kətəmaka	*!		*	
☞	kətəma-ka - kətəmaxa		*		*

Thus, lenition occurs in coda position (6-28a), and in intervocalic position (c); it is blocked only when the velar is in post-consonantal position (b): in this context, its realization does not violate LAZY$_X$ since the target consonant is not in intervocalic position; and it has a release, so the spirantized candidate is ruled out by PRES(cont w/ release).

2. SPEECH RATE AND REGISTER

2.1. THE GENERALIZATION

It has long been recognized that there is a relation between speech rate, register, and lenition. Specifically, if a consonant lenites at a given rate or register, it also lenites in faster or more casual speech (cf. Zwicky 1972, Donegan and Stampe 1979, Lindblom 1990). More recently, the positive effect of increased speech rate on consonant reduction has been demonstrated for German by Kohler (1991), for American English by Byrd (1994), and for Andalusian Spanish by Romero (1996). In particular, Romero (p. 58) found that, while the spirantization pattern is normally restricted to voiced stops, for one of his three subjects, in fast speech voiceless stops spirantize as well.

As an example, consider Harris' (1969) description of Mexico City Spanish. In slow/careful speech, non-initial voiced stops spirantize (6-29a), except when following a homorganic nasal or lateral (b):

(6-29)

a.	aβa 'bean'	aða 'fairy'	aɣa 'make'
	kalβo 'bald'	--	alɣo 'something'
	aβla 'speak'	aðlateɾes 'lackies'	aɣlomeɾaɾ 'to cluster'
	aɾβol 'tree'	aɾðe 'burn'	aɾɣamasa 'mortar'
	aβɾa 'will have'	paðre 'father'	aɣɾio 'sour'
	xajβo (no gloss)	najðen 'nobody'	kajɣa 'fall'
	aβjeɾto 'open'	aðjestɾaɾ 'to guide'	siɣjendo 'following'
	ewβolja (no gloss)	dewða 'debt'	sewɣma 'zeugma'
	aβwelo 'grandfather'	aðwana 'customhouse'	aɣweɾo 'fortune-teller'
	aðβeɾso 'unfavorable'	aβðomen 'abdomen'	suβɣlotal 'subglottal'
	suβmaɾino 'submarine'	aðmiɾasjon 'admiration'	diaɣnostiko 'diagnostic'
b.	bomba 'bomb'	donde 'where'	gaŋga 'bargain'
		kaldo 'hot'	

In faster or more casual speech (Harris does not distinguish between style and rate effects), spirantization applies to word-initial voiced obstruents as well (6-30a), except when following a homorganic nasal or lateral (as in word-internal spirantization) or utterance initially (b):

(6-30)

a.	beatɾis **b**aβea	(careful)	'Beatriz slobbers'
	beatɾis̬ **β**aβea	(casual)	
b.	un **d**ia	(any style)	'a day'
	el **d**ia	(any style)	'the day'
	beatɾis kanta	(any style)	'Beatriz sings'

That is, the word-internal spirantization pattern is extended to (utterance-medial) word-initial position in faster or more casual speech.

2.2. ANALYSIS

2.2.1. SPEECH RATE. *Analysis of the basic Spanish pattern.* The Spanish word-internal spirantization pattern described above follows from ranking the spirantization-blocking constraints, PRES(cont) and *[+cont,-son,-strid], below the entire LAZY series, making the spirantization context-free. The restriction of spirantization to voiced obstruents requires a further constraint, *[+cont, -son,-strid,-voi] (motivated by the perceptual weakness of voiceless nonstrident fricatives.[92] This constraint (and PRES(voi)) are undominated.

(6-31)

		PRES(voi)	*[+cont,-son, -strid,-voi]	LAZY	PRES(cont)
	aba - aba			**!	
☞	aba - aβa			*	*
☞	apa - apa			***	
	apa - aɸa		*!	**	*
	apa - aβa	*!		*	*

The blocking of spirantization in partial geminates, namely the homorganic nasal-stop and lateral-stop clusters, has been accounted for in Chapter 5.

[92]They have neither the strong friction nor the formant structure (it is obscured by voicelessness) which provides salient place cues in other consonants. The same consideration presumably lies behind the markedness of voiceless sonorants. Alternatively, or perhaps additionally, the dispreference for voiceless non-strident fricatives may follow from the antagonism between the characteristically short interval of close constriction in these consonants (without which they are not articulatorily advantageous, cf. Chapter 4) and the difficulty of actively devoicing a very short consonant (cf. the discussion of voicing and devoicing in Chapters 2 and 5).

Variable initial blocking. To account for the blocking of spirantization utterance-initially, and the variable blocking word-initially, we require two fortition constraints.

(6-32) *#ð̆: *[+cont,-son, -strid] in word-initial position.

*‖ð: *[+cont,-son, -strid] in utterance-initial position.

(Reviewing from Chapter 1, the perceptual rationale for fortition constraints such as these concerns the presumably greater salience of places cues in stops, due to the release burst; the word-initial context presumably reflects the greater importance of initial segments for purposes of lexical access, and the utterance-initial context presumably reflects the need to improve the perceptual robustness of consonants in an otherwise perceptually endangered context, specifically due to the lack of preceding formant transitions.)[93] The utterance-initial fortition constraint (*‖ð), if active, must outrank the word-initial constraint (*#ð) by Pāṇini's Theorem (the former being a subcase of the latter). The utterance-initial constraint, if undominated, will invariantly block utterance-initial spirantization. It is the ranking of the word-initial constraint which accounts for the variability; but first we must consider why this variable blocking of spirantization is sensitive to speech rate.

Phonetic explanation. Fast speech, by definition, involves shortening of articulatory gestures.[94] This shortening can mean one of two things: either the articulator

[93]Alternatively, Steriade (1996) accounts for the generalization of utterance-initial constraints to word-initial cases through the mechanism of paradigm uniformity constraints.

[94]In the Articulatory Phonology framework, it is frequently said that shortening can also be accomplished by increasing the overlap between (otherwise unmodified) gestures. But recall that "gesture" in this sense refers to an abstract task-dynamic unit. If there is any articulatory structure involved in both overlapping gestures (e.g. the jaw in intervocalic consonants), when physically implemented, the articulator must either move faster to achieve its target, or it must undershoot its target. Thus, the result is the same as described above.

reaches the target constriction faster, as in Figure 6-6a, or the constriction itself is shortened (b).

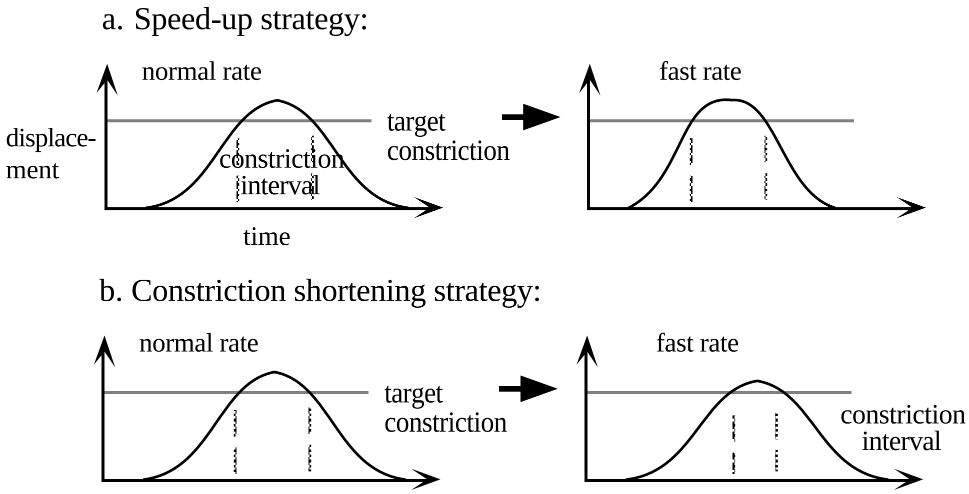


Figure 6-6. Schemata of fast-speech shortening strategies.

First, consider the constriction shortening strategy: by reducing the magnitude of the gesture, temporal reduction of the constriction is achieved. The connection between this strategy of fast-speech shortening and lenition is trivial: the temporal reduction in Figure 6-6b is *itself* a form of lenition. No further discussion is required.

A more interesting problem is to account for the connection between fast speech and lenition under the speed-up strategy (Figure 6-6a). This strategy amounts to an increase in the constriction gesture's velocity. Recall from Chapter 2 section 2.4 that, all else being equal, the greater the velocity of a gesture, the higher the effort cost. Consequently, the effort required to achieve some constriction target at a fast speech rate will be higher than that required for the same constriction at a slower speech rate.

Formal analysis. The solution, once again, relies on the device of effort thresholds. Let x be the effort cost of a [b] in some context at a slow rate of speech. As the velocity increases in faster speech, the effort cost of the [b] in this context increases, say to x+1. Returning to the Spanish data, the rate-sensitive variation now follows from interleaving of the word-initial fortition constraint at this crucial point in the LAZY series.

(6-33)

	Input: beatɾis babea *'fast' speech rate*	LAZY_{x+1}	*#ð	LAZY_{x}
	beatɾis baβea *(fast)*	*!		*
☞	beatɾis βaβea *(fast)*		*	*
	Input: beatɾis babea *'slow' speech rate*			
☞	beatɾis baβea *(slow)*			*
	beatɾis βaβea *(slow)*		*!	

Lest there be any confusion, I stress that I am *not* predicting (contra naturam) that high-effort gestures are more likely to occur in fast speech. Rather, the claim is that an equivalent constriction becomes more costly in fast speech, because greater velocity is required; in particular, it may become too costly relative to a lenition-blocking constraint, as in (6-33), and so the lenited candidate emerges as the winner.

2.2.2. SPEECH REGISTER. Optionality of phonological processes is standardly accounted for in Optimality Theory in terms of partially free constraint ranking (see Anttila 1995, Hayes & MacEachern 1996, and Boersma 1997d, for sophisticated deployments of this idea). If candidate A violates constraint C_1, candidate B violates constraint C_2, C_1 and C_2 are freely ranked with respect to each other, and no other constraints eliminate candidates A or B, then both candidates are equally optimal. However, "optionality" of phonological processes is not the same thing as speech-register conditioning of phonological processes, in that the former is random, whereas the latter presumably can be shown to correlate with sociolinguistic/pragmatic variables. We might attempt to capture this correlation by introducing register-conditioned demotions of lenition-blocking constraints relative to LAZY: for example, *#ð » LAZY$_X$ under normal situations, but in low register, the ranking is reversed.

The problem with this demotion device is that it can only refer to particular lenition-blocking constraints and particular LAZY thresholds, reversing the rankings on an item-by-item basis, so to speak. Lowering of speech register, however, is characterized by an across-the-board shift towards hypoarticulation, which becomes more extreme the lower the register (cf. Lindblom's (1990) notion of a global hypoarticulation parameter). This across-the-board characterization is supported by the case study of Florentine Italian in Chapter 8. What is needed, then, is some means of demoting the entire hierarchy of lenition-blocking constraints relative to the entire LAZY series.

a. High Register:

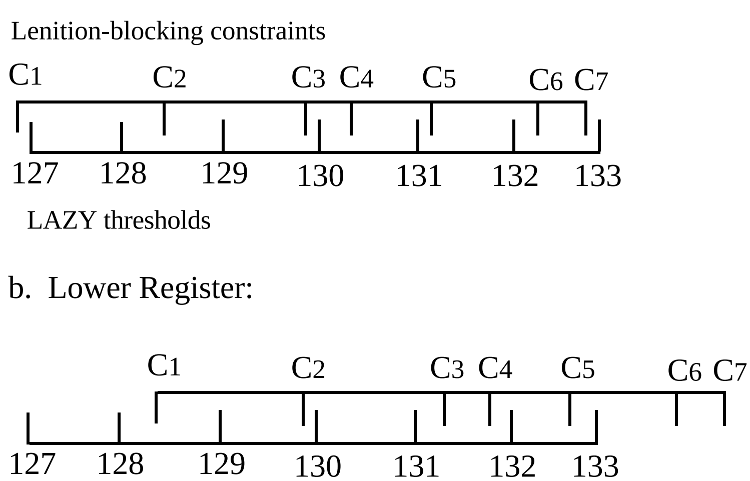


Figure 6-7. Register lowering as demotion of lenition-blocking constraints relative to effort thresholds.

This expresses the intuition that speakers curb hypoarticulation when they are more concerned with accommodating the hearer's perceptual needs, i.e. in high register conditions (e.g. when the hearer is of higher social status, or when the hearer is not a member of the "in-group," hence he requires things to be "spelled out" more explicitly); whereas in lower register, accommodation of the hearer's perceptual needs is globally demoted in importance, resulting in increased hypoarticulation.

Note, however, that demotion of the lenition-blocking constraints relative to the LAZY series is formally identical to promotion of LAZY relative to the lenition-blocking constraints. This result, in turn, is equivalent to augmenting the effort cost of all gestures by some amount, inversely related to register, as follows. Assume that the input to phonological computation contains not only some phonological representation, but also information about the pragmatic context of the speech act, which, for simplicity, we may reduce to a single numerical index *i*: the lower the register, the greater the value of *i*. Let the *register-adjusted effort cost* of a gesture be defined as the effort cost + *i*; and let the definition of the LAZY thresholds be modified as follows:

(6-34) LAZY_n (modified) = Do not expend register-adjusted effort ≥ *n* within a syllable.

Again, to avoid confusion, I stress that I am *not* predicting (contra naturam) that more effortful gestures occur in casual speech. Rather, the augmentation of effort cost by the register index is a technique for globally demoting the lenition-blocking constraints relative to the effort thresholds, thus resulting in lenited gestures in casual speech.

Now, returning to the Spanish variation problem, with the same ranking as in tableau (6-33), we obtain register-sensitive spirantization of the word-initial voiced obstruents.

(6-35)

	Input: beatɾis baβea *i = 1* (lower register)	LAZY$_{x+1}$	*#ð	LAZY$_{x}$
	beatɾis baβea	*!		*
☞	beatɾis βaβea		*	
	Input: beatɾis baβea *i = 0* (higher register)			
☞	beatɾis baβea			*
	beatɾis βaβea		*!	

That is, even if speech rate is held constant, lowering of speech register similarly results in lenition. Thus, this approach captures the generalization that, ceteris paribus, lenition at a given rate and register implies lenition in faster or more casual speech styles.

3. ALTERNATIVE APPROACHES

3.1. AUTOSEGMENTAL/PROSODIC LICENSING.

Since the advent of Autosegmental Phonology (Leben 1973, Goldsmith 1976) and the principle of Prosodic Licensing (Itô 1986), it has standardly been assumed that the conditioning of phonological processes by some trigger external to the target segment reduces either to autosegmental feature spreading, or to Prosodic Licensing, in combination with prosodic well-formedness conditions and some theory of repair strategies (e.g. Itô 1989). Indeed, without this assumption, the standard framework has little in the way of a theory of natural conditioning of phonological processes. Problems with feature-spreading treatments of lenition have already been noted in Chapter 1 section 2.4: specifically, such treatments provide no natural account of two-sided contexts such as intervocalic position; nor do they afford any treatment at all of intervocalic debuccalization or deletion (since these cannot be expressed as the spreading of any feature). These problems are compounded when we take into account the full scope of the Aperture Conditioning Generalization, as documented in section 1.1. If lenition is the spreading of features such as [+voi] or [+cont], why is [a] more likely than [i], or [z], to be a trigger, since all of them bear the feature which is to be spread?

The Prosodic Licensing line of analysis faces similarly grave objections. To wit, "intervocalic position" has no natural prosodic characterization. For example, the /b/'s in /aba/, /agba/, and /aːba/ are all in onset position; and in /aːba/ and in /agba/ they both follow a heavy syllable: why is it then that spirantization of the /b/ frequently occurs in /aba/ and /aːba/, but not in /agba/? Moreover, onset position is generally viewed as prosodically "strong" (cf. the discussion of licensing of features in onset, but not coda, position, in Goldsmith 1990); thus, by the logic of Prosodic Licensing, we ought to observe intervocalic fortition rather than lenition. Furthermore, the device of "ambisyllabic" affiliation of segments appears to have little motivation (pace Kahn 1976), except as a means of recharacterizing intervocalic position and similar contexts in prosodic terms: that is, to restate the observation. Furthermore, the notion of ambisyllabicity does not distinguish between intervocalic singletons, as in /aba/, which do frequently spirantize, and geminates, as in /abba/, which never spirantize, as documented in Chapter 5. But even if we allow ambisyllabicity as a characterization of intervocalic position, this device does not capture the broader generalizations of gradient aperture conditioning: (ceteris paribus) lenition is more likely to occur the more open the flanking segments; and lenition which occurs between open segments is more extreme than lenition which occurs elsewhere. A segment either is ambisyllabic or it is not: there is no notion of /b/ being "more ambisyllabic" in the context /a__a than in the context /i__i. Nor is there any intrinsic connection between ambisyllabicity and aperture.

Finally, the Autosegmental theory of representations encodes no temporal information other than precedence, overlap, and phonemic length distinctions. It is therefore beyond the capacity of Autosegmental Phonology (and, a fortiori, its rule-based precursors) to model speech rate variation at all, let alone the relation between such variation and lenition. The standard device of annotating rules with stylistic conditions (e.g. -son - +cont /V__V *[fast speech]*) merely restates the description. Nor does such an approach capture the across-the-board shift towards hypoarticulation occurring in fast and casual speech: rather, the annotated-rule approach must stipulate this result, process-by-process, and register-by-register. In particular, Autosegmental Phonology cannot capture the observed connection between aperture-based lenition contexts and rate- and register-sensitive conditioning of lenition (clearly exemplified in Florentine Italian, Chapter 8): as speech rate increases or register decreases, the context for lenition expands, from high-aperture contexts to lower-aperture contexts. In contrast, in the effort-based approach, the treatment of aperture and rate/register conditioning is unified: both are instances of effort-based contexts, and the expansion of lenition contexts in fast/casual speech is an automatic consequence of this treatment.

3.2. A NON-EFFORT-BASED OT APPROACH

A number of these problems are faced as well by an Optimality Theoretic approach which does not refer to effort minimization. To capture the Aperture Conditioning Generalization in such an approach, the following set of constraints (at least) appears to be necessary:

(6-36) (5 = low vowel, 4 = non-high vowel, 3 = vowel, 2 = sonorant continuant, 1 = continuant)

Spirantization (with a similar set of *voiceless constraints for voicing lenition)
Two-sided:
*5-stop-5, *5-stop-4, *5-stop-3, *5-stop-2, *5-stop-1
*4-stop-5, *4-stop-4, *4-stop-3, *4-stop-2, *4-stop-1
*3-stop-5, *3-stop-4, *3-stop-3, *3-stop-2, *3-stop-1
*2-stop-5, *2-stop-4, *2-stop-3, *2-stop-2, *2-stop-1
*1-stop-5, *1-stop-4, *1-stop-3, *1-stop-2, *1-stop-1
Right-hand:
*stop-5, *stop-4, *stop-3, *stop-2, *stop-1
Left-hand:
*5-stop, *4-stop, *3-stop, *2-stop, *1-stop
Context-free:
*stop

Reduction to sonorants
Two-sided:
*5-obst-5, *5-obst-4, *5-obst-3, *5-obst-2, *5-obst-1
*4-obst-5, *4-obst-4, *4-obst-3, *4-obst-2, *4-obst-1
*3-obst-5, *3-obst-4, *3-obst-3, *3-obst-2, *3-obst-1
*2-obst-5, *2-obst-4, *2-obst-3, *2-obst-2, *2-obst-1
*1-obst-5, *1-obst-4, *1-obst-3, *1-obst-2, *1-obst-1
Right-hand:
*obst-5, *obst-4, *obst-3, *obst-2, *obst-1
Left-hand:
*5-obst, *4-obst, *3-obst, *2-obst, *1-obst
Context-free:
*obst

Debuccalization
Two-sided:
*5-place-5, *5-place-4, *5-place-3, *5-place-2, *5-place-1
*4-place-5, *4-place-4, *4-place-3, *4-place-2, *4-place-1
*3-place-5, *3-place-4, *3-place-3, *3-place-2, *3-place-1
*2-place-5, *2-place-4, *2-place-3, *2-place-2, *2-place-1
*1-place-5, *1-place-4, *1-place-3, *1-place-2, *1-place-1
Right-hand:
*place-5, *place-4, *place-3, *place-2, *place-1
Left-hand:
*5-place, *4-place, *3-place, *2-place, *1-place
Context-free:
*place

Elision
Two-sided:
*5-C-5, *5-C-4, *5-C-3, *5-C-2, *5-C-1
*4-C-5, *4-C-4, *4-C-3, *4-C-2, *4-C-1
*3-C-5, *3-C-4, *3-C-3, *3-C-2, *3-C-1
*2-C-5, *2-C-4, *2-C-3, *2-C-2, *2-C-1
*1-C-5, *1-C-4, *1-C-3, *1-C-2, *1-C-1
Right-hand:
*C-5, *C-4, *C-3, *C-2, *C-1
Left-hand:
*5-C, *4-C, *3-C, *2-C, *1-C
Context-free:
*C

By virtue of Pāṇini's Theorem, this constraint set cannot be ranked such that, for example, spirantization occurs in inter-obstruent but not inter-vocalic position (*3-stop-3 inherently outranks *1-stop-1). However, a theory such as (6-36) implicitly acknowledges that the conditioning of lenition is driven by a notion of (vertical) articulatory displacement. In failing to reflect this notion directly, it thus gives up on a central goal of the Generative Phonology tradition: namely, the formal capture of phonological explanation. Moreover, this theory says nothing about rate-conditioned lenition; whereas in the effort-based approach, it is the combination of velocity and displacement that makes for effort-based contexts. Finally, the unified notion of effort-based contexts further embraces *horizontal* displacement, thus capturing coarticulatory effects (e.g. in Chapter 7 section 2.3); whereas the non-effort-based OT approach would require yet another slew of constraints to handle such effects.

3.3. ARTICULATORY PHONOLOGY

The idea of jaw-displacement-based lenition contexts fits as naturally into the Articulatory Phonology framework as the phonetically based OT framework. Indeed, as discussed in section 1.2, such an Articulatory Phonology-style account is sketched by de Jong et al. (1992), who model intervocalic spirantization and voicing in terms of blending of Articulatory Phonology-style gestures: overlap between flanking vowels and the consonant results in undershoot of jaw raising, hence undershoot of the constriction target for the consonant. It further captures the relation between increased rate of speech and increased propensity for lenition: as the overlap between closing and opening gestures is increased, constriction gestures of a specified velocity undershoot their original target, hence gestural reduction. The principal problem faced by such an Articulatory Phonology account is the lack of a higher level of description, elegantly and insightfully characterizing such patterns as "Shina voiced stops spirantize in high-displacement contexts, but nasals do not." In Articulatory Phonology, the restriction of undershoot-qua-spirantization to some subset of consonants must be done in terms of lower stiffness parameters[95] for each affected gesture, with different values for each articulator. In Shina, therefore, the stiffness value for each articulator must be set somewhat lower in closure gestures which are not significantly overlapped by velic opening gestures (i.e. nasals). The formalism cannot capture the insight that spirantization is blocked in nasals because the output, a nasal continuant, is *marked*, presumably due to its perceptual indistinctness.[96] In contrast, the effort-based OT approach can capture this insight directly, in terms of blocking by a higher-ranked constraint, *[+nas,+cont]. Furthermore, the Articulatory Phonology approach does not draw any intrinsic connection between speech *register* and lenition.

3.4. "BUT THIS IS ALL MERE PHONETICS"

The final alternative, at least with respect to rate- or register-sensitive lenition processes, is to exclude such processes from the purview of phonological theory, relegating them instead to the "phonetic" component of the grammar (derivationally ordered after the phonology, as in Keating 1984). But regardless of the grammatical component they are assigned to, an explicit treatment of rate- and register-sensitive lenition requires reference to articulatory effort, i.e. displacement and velocity; therefore the theory of natural-language sound patterns must include an effort-based treatment of lenition such as that proposed herein.

---

[95]By lowering the stiffness, the approach to the target is slowed, resulting in undershoot if the articulator fails to achieve closure before the following opening gesture takes over.

[96]That is, the formant excursions are less extreme in a continuant than in a stop, since the articulator is not moving as far: this makes for poorer cues to the consonant's place of articulation. The problem is further compounded by nasalization, which results in "shrinking" of formant frequency distinctions.

Moreover, the division of lenition processes into distinct formal treatments, phonological and phonetic, based on sensitivity to rate and register variation, is particularly unappealing in light of the Spanish facts discussed in the previous section. Stable word-internal and variable utterance-level spirantization appear be a unified phenomenon, applying to the same class of segments, yielding the same outputs, and being subject to the same conditions (e.g. blocking of spirantization in partial geminates). By the same reasoning as that of Halle 1959 (arguing against distinct morphophonemic and allophonic levels), the loss of generality attendant upon the two-level treatment is sufficient grounds to reject this position.

It is acknowledged, however, that the effort-based approach, as developed thus far has difficulty in capturing aperture-based lenition contexts which are *not* sensitive to rate and register variation. We return to this problem in Chapter 9.

Chapter 7:

Tümpisa Shoshone

To exemplify the effort-based approach to lenition in greater depth, this chapter examines the sound pattern of Tümpisa Shoshone, which includes spirantization, voicing, nasal weakening, and elision of laryngeal consonants.[97]

1. DATA

The data for this case study are drawn from Dayley's (1989) grammar of the Tümpisa[98] (also known as Panamint) dialect of Shoshone, a Uto-Aztecan language spoken in the region of the California-Nevada border. As a frame of reference, I present the following chart of consonant "phonemes" for Tümpisa Shoshone:

[97]Other processes described by Dayley include optional devoicing of short vowels between voiceless consonants and in initial unstressed position; nasalization of vowels adjacent to a nasal consonant; palatalization of sibilants and nasals after front vowels; fronting of velars before front vowels; coalescence of /w + a/ to [o] or [u]; lowering of high vowels after [ʔ]; vowel rounding harmony; rounding of velars before round vowels; and a rather complex system of consonant gradation, involving morphemes which Dayley analyzes as ending in abstract /n/ and /"/ (the latter indicating that the morpheme induces gemination of the initial consonant of the following morpheme).

[98]Note that Dayley's /ü/ represents a phonetically central unrounded vowel [ɨ]: thus the word Tümpisa ('Death Valley') is pronounced [tɨmbiʃa].

Table 7-1. Consonant "phonemes" of Tümpisa Shoshone.

labial	coronal	dorsal	labio-dorsal	laryngeal
p pp	t tt	k kk	k^w kk^w	ʔ
	ts tts			
	s			h
m mm	n nn	ŋ	$ŋ^w$	

Note that here and throughout the description of consonant variation below, my use of phonemic terminology and notation (e.g. voiceless stop phonemes realized as voiced fricative allophones) is purely descriptive; it does not imply the assumption that the surface fricatives are uniformly stops in underlying representation. Syllables are maximally CVVC, and the only permissable clusters are full geminates and homorganic nasal + stop clusters. All words end in vowels on the surface, though Dayley posits some word-final /h/'s and /n/'s in UR.

1.1. SPIRANTIZATION

In Tümpisa Shoshone (TS), stops occur in initial position (7-1a), as geminates (b), and following a homorganic nasal (c). Flaps (7-1b) and nonstrident fricatives (a) occur elsewhere.[99]

[99]Dayley characterizes the spirantization/flapping environment as "intervocalic" (or more precisely, /V(h)__V). However, as the language's phonotactics permit no consonant clusters other than geminates and homorganic nasal + stop clusters, and (rarely) [hC] clusters, the spirantization context reduces to the context-free characterization above, subject to blocking in initial position and in full and partial geminates.

(7-1)	a.	puhaɣãntɨ̥	'shaman'
		taβettʃi̥	'sun'
		tuɣʷãnni	'night'
		tsiðoohi̥	'push'
		kɨ̃mmãɣ̃ɨ̃nn̥ḁ	'to come here'
		kʷɨ̃̃jãã	'eagle'
	b.	paɾɨasɨppɨ̥	'ice'
		uttũnn̥ḁ	'to give'
		taβettʃi̥	'sun'
		pũñɨ̃kkḁ	'see, look at'
		ukkʷḁ	'when, if'
	c.	taziũmbi	'star'
		ɨndãw̃ɨ̃ʔi̥	'your little brother'
		tɨppiʃiɸũŋki̥	'stinkbug'

That is, [p,k,kʷ] are in complementary distribution with [β,ɣ,ɣʷ], as [t] is with [ð] (after a front vowel) and with [ɾ] (after a back vowel or /h/, the latter context illustrated by /tɨkkappih tukkʷan/ - [tɨkkappi̥ ɾ̥ukkʷan] 'under the food'). “Initial" position appears to mean utterance-initial, since spirantization applies across word boundaries:

(7-2) pɨe ðuɣʷãnni j̃ããɣ̃ɨ̃nn̥ḁ 'it's already getting dark'

The distribution of sibilants is somewhat more complicated. On the one hand, the affricate /ts/ spirantizes (to [z]), except in initial position and in full or partial geminates, just like the stops, e.g. /motso/ - [mõzõ] ('whiskers'). But unlike the other fricatives, [s] can occur in initial (7-3a) as well as medial (b) position.

(7-3) a. suɾɨ̃mm̥ɪ̥ 'those'

b. paɾɨasɨppɨ̥ 'ice'

Nor is there a contrast between geminate and singleton [s], as there is in the stop and affricate series.

1.2. VOICING AND DEVOICING

The distribution of voicing is also predictable. Stops are voiced following a nasal (7-1c); in initial position, and in geminates, however, they are voiceless (a,b). The fricatives resulting from spirantization are voiced in most contexts. However, the underlying fricative /s/ (i.e. which does not derive by spirantization from /ts/) is realized as voiceless in all contexts. Moreover, utterance-final vowels are optionally devoiced, in which case the preceding consonant (or typically the second half of a geminate nasal, e.g. suɾɨ̃mm̥ɨ̥ ('those')) is devoiced as well (7-4).[100] Furthermore, h + obstruent clusters coalesce to voiceless obstruents (7-4b). In these non-initial singleton obstruents which surface as voiceless, either due to final devoicing, or due to underlying /h/, spirantization is optional; whereas flapping (7-4c) is obligatory, as is spirantization in the non-devoiced case (7-1a).

[100]This utterance-final devoicing optionally takes the form of glottalization of the final vowel, described by Dayley as insertion of [ʔ] plus a voiceless echo vowel. Dayley notes that this devoicing-by-glottalization is most common in uninflected nouns (in final position), and speculates that it may function as an allomorph of the absolutive suffix.

(7-4) a. tahaβi ~ (ɸ/p)i̥ 'snow'

huβiaɾɨɣi ~ ...(x/k)i̥ 'sing'

peðɨ ~ ...(θ/t)ɨ̥ 'daughter'

mõzõ ~ ...(z̥/ts)o̥ 'whiskers'

b. /ohpimpɨ/ - o(ɸ/p)ĩmbɨ 'mesquite tree'

/ɨattɨah ka/ - ɨattɨa(x/k)a[101]

c. /tɨkkappih tukkʷan/ - [tɨkkappi̥ɾ̥ukkʷan] 'under the food'

Note that the devoiced sibilant fricative [z̥], derived from /ts/, does not neutralize with /s/: Dayley describes [z̥] as more "lenis," presumably meaning shorter, than [s].

1.3. Nasal weakening.

The spirantization pattern of obstruents is partially paralleled by nasals: a non-initial singleton labial nasal is realized as [w̃] (7-5a); and a non-initial singleton coronal nasal as [j̃] after a front vowel (b).

(7-5) a. sɨ̃ɨ̃w̃õõɾɨ 'ten'

b. sẽj̃ũ 'therefore'

c. jɨβãni 'autumn'

After a back vowel, Dayley transcribes the coronal nasal as [n], apparently without weaking (7-5c). The velar nasal [ŋ] likewise is not described as weakening (7-6a). Labiovelar [ŋʷ] does not weaken after a front vowel; however, after back vowels [ŋʷ] occurs in free variation with [w̃] (b).

---

[101] In addition, /h + k$^{(w)}$/ sequences can debuccalize to [h$^{(w)}$], but this outcome is restricted to particular suffixes.

(7-6) a. pãŋe 'up'

b. sɔ̃(ŋʷ/w̃)ɔ̃ 'lungs'

Note that there are no geminate velar or labiovelar nasals, unlike the labial and coronal nasals. Also note that vowels are nasalized before and (to a lesser extent) after a nasal consonant.

1.4. ELISION OF LARYNGEAL CONSONANTS

Dayley further describes an optional process of elision of intervocalic /h/ and /ʔ/, e.g. [po(ʔ)ittʃi] ('path'), [ta(h)aβi] ('snow'). However, as /ʔ/ *only* occurs in intervocalic word-medial position to begin with, the /ʔ/ elision can alternatively be viewed as context-free. The distribution of /h/ is somewhat broader. It can occur initially, as in [huβiaɾɨɣi] ('sing'), and before a following glide, as in [tɨkkappi̥h ʝãã] ('on the food'). As noted in section 1.2, h + obstruent clusters coalesce to a devoiced fricative or stop (or a devoiced flap, in the case of coronal stops). Thus, /h/ elision is restricted to intervocalic and pre-obstruent position.

2. ANALYSIS OF SPIRANTIZATION AND FLAPPING

2.1. BASIC SPIRANTIZATION PATTERN

In accordance with the effort-based approach outlined in Chapter 1, the basic TS pattern of context-free spirantization at all places of articulation, subject to blocking in utterance-initial position, follows from the following ranking:[102]

---

[102]Voicing / devoicing of the outputs is addressed in section 2.2 below.

(7-7)	non-initial:	*[+cont,-strid,+cons]/[__...]$_{Utt}$	LAZY	PRES(cont)
	ta**p**ettʃi̥		**!	
☞	ta**β**ettʃi̥		*	*
	tsi**t**oohi̥		**!	
☞	tsi**ð**oohi̥		*	*
	puha**k**ãntɨ̥		**!	
☞	puha**ɣ**ãntɨ̥		*	*
	tu**kʷ**ãnni		**!	
☞	tu**ɣʷ**ãnni		*	*
	initial:			
☞	**p**uhaɣãntɨ̥		**	*
	βuhaɣãntɨ̥	*!	*	
☞	**t**uɣʷãnni		**	*
	ðuɣʷãnni	*!	*	
☞	**k**ĩmmãɣĩnn̥ḁ		**	*
	ɣĩmmãɣĩnn̥ḁ	*!	*	
☞	**kʷ**ĩjãã		**	*
	ɣʷĩjãã	*!	*	

(Here, and in tableaux below where the lenition-blocking constraints need not be interleaved within the LAZY series, I continue the practice from Chapters 1-5 of presenting LAZY as a single scalar constraint (with greater or lesser violations), notwithstanding the binarization of LAZY motivated in Chapter 6.) As discussed in Chapter 3, section 8, since continuancy is allophonic in these obstruents, the constraint system determines the surface value, even if we assume that the underlying specification is contrary to the surface value (as indicated in the above tableaux by assuming PRES(cont) violations even in the winning candidates). The failure of these consonants to lenite further is captured by ranking faithfulness to other features above LAZY, e.g. PRES(cons).

(7-8)

		PRES(cons)	LAZY
☞	p - β		*
	p - Ø	*!	

Moreover, the failure of /t/ to spirantize to [s], and the blocking of spirantization in geminates and homorganic nasal + stop clusters (i.e. partial geminates) instantiate cross-linguistic generalizations, which follow from this effort-based approach, as discussed in Chapters 4 and 5.

2.2. SIBILANTS.

The sibilant fricative /s/ occurs in initial position, hence the utterance-initial fortition constraint above does not prohibit strident continuants; but under the constraint hierarchy above, we incorrectly fail to block spirantization of initial /ts/. I assume that the general property distinguishing /s/ from /ts/ and its allophones [ts,z,z̥,ʒ,ʒ̥] is the shorter duration of strident energy in the latter. This seems plausible, in light of the general observation that voiced fricatives are typically shorter than voiceless (e.g. Nartey 1982), and Dayley's comment that [z̥,ʒ̥] are more "lenis" than [s,ʃ]; moreover, I observe in spectrograms of my own speech that the fricated portion of an sibilant affricate (or t + s cluster) is typically shorter than that of a fricative, presumably due to the more gradual onset of strident energy in the latter. Furthermore, a short strident fricative (with gradual onset of strident friction) is presumably perceptually weaker than an affricate (with abrupt onset of "full-strength" stridency, due to the sudden release of the preceding stop closure), or than a longer strident fricative. I therefore posit a binary feature, [long stridency], which distinguishes the fortis strident fricative [s] ([+long strid]) from the lenis strident fricative [z] or [z̥], and the strident affricate [ts] ([-long strid]), as well as the palatalized variants of all of the above (non-stridents are unspecified for this feature).

The TS sibilant pattern now follows from undominated ranking of PRES(long strid), in combination with another utterance-initial fortition constraint, *[-long strid,+cont] / [__...]$_{\text{Utterance}}$, grounded in the relative perceptual weakness of the shorter non-affricate sibilants.

(7-9)

	initial:	PRES(long strid)	*[-long strid, +cont] /[__...]$_{\text{Utt}}$	LAZY
☞	tsitoohi - **ts**iðoohi̥			***
	tsitoohi - **s**iðoohi̥	*!		**
	tsitoohi - **z̥**iðoohi̥		*!	*
	senu - **ts**ẽj̃ũ	*!		***
☞	senu - **s**ẽj̃ũ			**
	senu - **z̥**ẽj̃ũ	*!		*
	non-initial:			
	motso - mõ**ts**õ			**!*
	motso - mõ**s**õ	*!		**
☞	motso - mõ**z**õ			*
	patɨasɨppɨ - parɨ**ts**ɨppi̥			***!
☞	patɨasɨppɨ - parɨ**s**ɨppi̥			**
	patɨasɨppɨ - parɨ**z**ɨppĩ̥	*!		*

These constraints block initial /ts/ from spirantizing, but do not block spirantization of /ts/ medially, and permit /s/ to surface unchanged both initially and medially.

Finally, note that the absence of a geminate fricative [ss] is reflective of the higher effort cost, hence markedness, of geminate fricatives relative to stops (cf. Chapter 5), and follows from subordination of PRES(cont) to LAZY.

(7-10)

		LAZY	PRES(cont)
	ss - ss	**!	
☞	ss - tts	*	*

Thus, even if an input contains a geminate sibilant fricative, it will neutralize to an affricate in all contexts (degeminated and deaffricated outputs are presumably ruled out by ranking of PRES(length) and PRES(strid) above LAZY).

2.3. VARIATION WITH FLAPPING

A minor elaboration of this analysis further captures the variation between [ð] and [ɾ] as lenited allophones of /t/. Relative to a stop, a flap involves a reduction in magnitude, such that the active articulator makes the briefest of contacts with the passive articulator, while still maintaining non-continuancy (see generally Inouye 1995). Presumably, coronal (specifically, apical) flaps are common, whereas non-coronal flaps are rare or unattested,[103] because of the greater stiffness of the coronal articulator, which allows it to reach its closure target and release the closure relatively quickly, without additional expenditure of energy.

(7-11)

		LAZY	PRES(son)
☞	p, k	**	
	b̆, ğ (extra-short)	***!	*
	t	**!	
☞	ɾ	*	*

I further hypothesize that in TS, the distribution of [ɾ] (after back vowels) vs. [ð] (after front vowels) is due to (LAZY-driven) coarticulation involving the tongue body. That is, in contexts where the non-continuant ([ɾ]) can be achieved without significant tongue body displacement, i.e. following a back vowel (Figure 7-1b), this is done. However, in a front vowel context, the tongue tip is closer to the dental region (Figure 7-

103Margi presents the only known case of labial flaps (see Maddieson & Ladefoged 1996), and dorsal flaps are unattested.

1a); to achieve a flap, therefore, the tongue tip must either be dramatically retroflexed, or the tongue body must be retracted before the flap is made.

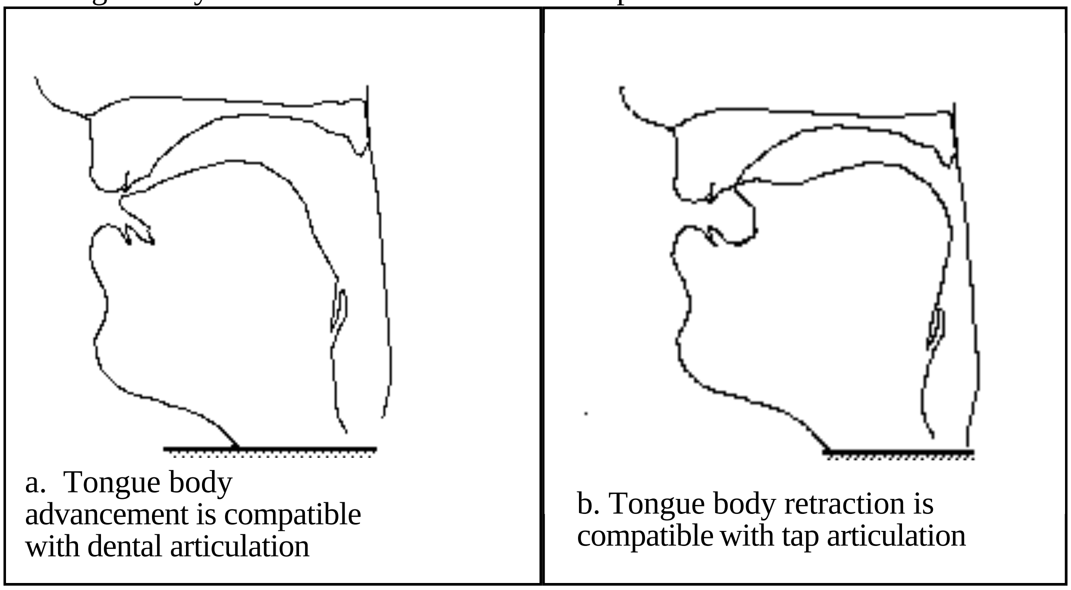


Figure 7-1. Flapping as coarticulatory retraction of the tongue tip.

(A dental flap is presumably not generally feasible, due to typical leakage of airflow through the teeth.) Hence, I assume that a flap is slightly more effortful following a front vowel than following a back vowel. Specifically, let *x* denote the minimum of effort required to achieve a flap following a front vowel, and *y* following a back vowel: for the foregoing phonetic reasons, $x > y$. The TS allophonic flapping pattern now follows from interleaving of a spirantization-blocking fortition constraint, *[+cont,-strid,+cons], between these effort thresholds:

(7-12)

		LAZY$_x$	*[+cont,-strid,+cons]	LAZY$_y$
	at	*!		*
☞	aɾ			*
	að		*!	*
	it	*!		*
	iɾ	*!		*
☞	ið		*	*

Analysis of (devoiced) flapping from coalescence of /h + t/ is deferred until after the general account of voicing and devoicing below.

3. ANALYSIS OF (DE)VOICING

3.1. BASIC PATTERN

The TS context-free devoicing of (full) geminate obstruents is reflective of a cross-linguistic markedness generalization, and follows from the general effort-based aerodynamic account of geminate devoicing/blocking of voicing presented in Chapter 5. TS utterance-initial obstruent devoicing, and voicing in most other contexts, likewise follow from aerodynamic considerations. Recall from Chapter 2 section 4 that obstruents passively devoice in utterance-initial position. Of course, as in the aerodynamic account of geminate devoicing (Chapter 5), this aerodynamic state of affairs may be overcome by intercostal contraction (raising subglottal pressure) or various oral cavity expansion gestures, such as larynx lowering and pharynx expansion (lowering oral pressure); but these additional voicing-enabling gestures carry some additional effort cost. Hence,

(7-13) Utterance-initial voiced obstruent $>_{effort}$ Utterance-initial voiceless obstruent

Moreover, as discussed in Chapter 2, in utterance-medial position, (singleton) obstruents are passively voiced. Thus,

(7-14) Utterance-medial voiceless obstruent $>_{effort}$ Utterance-medial voiced obstruent

TS voicing allophony now follows from the following ranking:

(7-15)

	medial:	LAZY	PRES(voi)
	ta**ɸ**ettʃi̥	**!	
☞	ta**β**ettʃi̥	*	*
	tsi**θ**oohi̥	**!	
☞	tsi**ð**oohi̥	*	*
	puha**x**ãntɨ̥	**!	
☞	puha**ɣ**ãntɨ̥	*	*
	tu**xʷ**ãnni	**!	
☞	tu**ɣʷ**ãnni	*	*
	initial:		
☞	**p**uhaɣãntɨ̥	*	*
	buhaɣãntɨ̥	**!	
☞	**t**uɣʷãnni	*	*
	duɣʷãnni	**!	
☞	**kʲ**immãɣʲinn̥ḁ	*	*
	gʲimmãɣʲinn̥ḁ	**!	
☞	**kʷ**ĩj̃ãã	*	*
	gʷĩj̃ãã	**!	

3.2. h + OBSTRUENT COALESCENCE/DEVOICING

This coalescence/devoicing resulting from /h + obstruent/ clusters can be obtained by disjunctively combining PRES(-voi) and PRES(aspiration) (cf. Kirchner 1996). Elision of the /h/ then follows from ranking plain PRES(asp) below LAZY:

(7-16)

		PRES(-voi) ∨ PRES(asp)	LAZY	PRES(asp)
	...hɸ...		***!	
☞	...ɸ...		**	*
	...β...	*!	*	*

That is, the [ɸ] candidate satisfies the disjunctive constraint, even though the aspiration noise is lost, because the voicelessness of the /h/ is preserved, shifted onto the following obstruent (satisfying PRES(-voi)); whereas the [β] candidate violates both.

3.3. Final devoicing

Finally, TS devoicing of utterance-final syllables may be attributed to abduction of the vocal folds, or increase in inspiratory force (causing subglottal pressure to drop off) (Westbury & Keating 1986), in anticipation of post-utterance breathing. Variable timing of these respiratory gestures relative to the end of the utterance is sufficient to account for the optionality of this process.

3.4. Interaction with spirantization and flapping

Recall that spirantization is optionally blocked in these devoiced obstruents. I attribute this blocking to a fortition constraint, *[+cont,-voi,{-strid∨-long strid}]. This constraint is presumably grounded in the observation that lack of modal voicing tends to obscure the formant transitions (cf. Silverman 1995) associated with these continuants, which are relatively acoustically weak, either because they lack strong friction, as in the nonstridents [ɸ,θ,x], or because the duration of this friction is brief, as in [z̥]. The TS optional blocking of spirantization now follows from free ranking of this constraint with Lazy:

(7-17)

		*[+cont,-voi,{-strid∨-long strid}]	Lazy
☞	tahapi̥		**
☞	tahaɸi̥	*	*
☞	huβiarɨki̥		**
☞	huβiarɨxi̥	*	*
☞	mõtso̥		**
☞	mõz̥o̥	*	*

An additional aspect of the post-/h/ context is that coronals lenite to a voiceless flap, rather than [ð] or [θ], even following a front vowel (cf. section 2.3). Let us assume that the loss of the /h/ in this context results in some phonetic compensatory lengthening of the transition from the preceding vowel into the consonant, preserving something of the duration of the original /hC/ cluster. As a consequence, the tongue tip/tongue body ensemble have a longer time to achieve a non-continuant target.

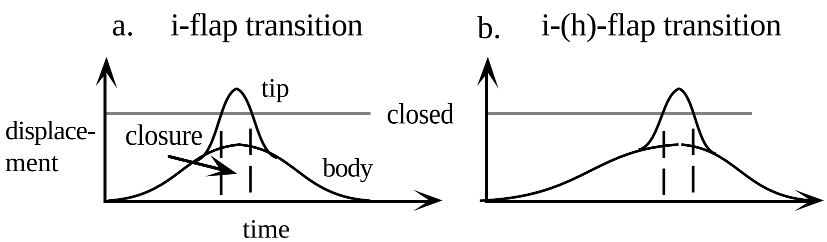


Figure 7-2. Schemata of tongue tip/tongue body ensemble dispacement vs. time, without and with compensatorily attenuated transition.

Because the consonant gesture does not require as great a velocity to reach closure (Figure 7-2b) as it would in the non-attenuated case (a), less effort is required in (b) than in (a) (see Chapter 2 section 2.4). Thus, if the effort required to achieve a flap in Figure 7-2a = *x* (as assumed in example (7-12), the effort required in Figure 7-2b (call it *w*) is somewhat less than *x*.   Under the same ranking as posited in (7-12), with the additional assumption that *[+cont,-strid,+cons] » LAZY$_w$, we obtain the result that the coronal stop lenites to a flap, rather than a continuant, in the post-h context (7-18a), whereas it otherwise lenites to [ð] following a front vowel (b):

(7-18)		PRES(cluster duration)	LAZY$_x$	*[+cont,-strid,+cons]	LAZY$_w$
a.	it - iɾ		*!		*
☞	it - ið			*	
b. ☞	iht - i·ɾ̥				*
	iht - i·θ			*!	
	iht - iθ	*!			

4. Analysis of Nasal Weakening

4.1. Basic Pattern

The foregoing analysis of spirantization can be extended to the nasal weakening facts, with minimal elaboration. Indeed, the general ranking LAZY » PRES(cont), motivated above for TS, results in spirantization of obstruents and nasals alike.[104] TS differs from most other languages (e.g. Spanish, Harris 1969), in which spirantization is restricted to oral noncontinuants, in that TS further subordinates the nasal fortition constraint, *[+nas,+cont], to LAZY.

(7-19)

		LAZY	*[+nas,+cont]	PRES(cont)
☞	kʷĩ**n**ãã	**!		
	kʷĩ**j̃**ãã	*	*	*
	s̃ɨ̃ɨ̃**m**ɔ̃ɔ̃ɾɨ	**!		
☞	s̃ɨ̃ɨ̃**w̃**ɔ̃ɔ̃ɾɨ	*	*	*

As discussed in Chapter 5, the blocking of nasal weakening in full and partial geminates follows from the same considerations as the blocking of spirantization in geminate obstruents. Moreover, utterance-initial nasal weakening is blocked by the same fortition constraint which blocks obstruent spirantization:

(7-20)

		$*[+cont,-strid,+cons]/[\_\_...]_{Utt}$	LAZY
☞	**m**õtso̥		**
	w̃õtso̥	*	*

[104]Loss of closure in a nasal results in a nasalized approximant rather than a fricative (in the absence of dramatically increased subglottal pressure), due to the inhibiting effect of nasal venting on oral pressure, which is necessary to generate fricated airflow.

4.2. VOICING

The failure of the nasals to devoice initially, as the obstruents do, is presumably due to nasal venting of airflow, which prevents significant build-up of oral pressure; hence, initiation of voicing in nasals does not present the same aerodynamic problems as in oral stops. Nasals are therefore passively voiced in all contexts, modulo optional utterance-final devoicing, due to anticipatory glottal abduction, as discussed in section 3.3.

4.3. APPARENT PLACE RESTRICTIONS ON NASAL WEAKENING

Two facts, however, remain to be explained. First, according to Dayley's transcription, the coronal nasal surfaces as [n] after a back vowel, apparently failing to lenite. Given the reduction of the oral coronal stop to a flap in this context, and in light of the generally parallel behavior of obstruent spirantization and nasal weakening in TS, we would expect /n/ to reduce to a nasalized flap, [ɾ̃] in this context. However, without instrumental measurements of duration, it is difficult to distinguish [n] from [ɾ̃], since the other acoustic cues to the stop/flap distinction (e.g. presence of a burst) are absent in nasals.[105] Therefore it seems plausible that these coronal nasals are actually flaps. Assuming this to be the case, the variation between [j̃] and what Dayley transcribes as [n] follows from the same analysis as the variation between coronal fricatives and flaps, in section 2.3.

[105]Particularly since the [n/ɾ̃] distinction is non-phonemic in English, Dayley's native language.

(7-21)

		LAZY$_x$	*[+cont,-strid,+cons]	LAZY$_y$
	an	*!		*
☞	ar̃			*
	aj̃		*!	*
	in	*!		*
	ir̃	*!		*
☞	ij̃		*	*

Second, according to Dayley, the velar nasal /ŋ/ never weakens, though velar stops spirantize. This fact might be attributed to the lowering of the velum during nasalization, decreasing the distance which the tongue body must travel to achieve full closure. Blocking of /ŋ/-weakening would then follow from an interleaved ranking, where *u* denotes the effort required for [ŋ], and *v* denotes the (greater) effort required for a non-velar nasal.[106]

(7-22)

		LAZY$_v$	*[+cont,-strid,+cons]	LAZY$_u$
☞	ŋ			*
	ɰ̃		*!	
	m	*!		*
☞	w̃		*	

However, the presence or absence of complete closure in a velar nasal is a subtle cue. Ohala (1975) (citing House 1957) observes that [ŋ] is acoustically quite close to nasalized vocoids:

> [T]he velar nasal has primarily just a single resonating cavity with a small, perhaps negligible side-cavity, unlike other nasals, and thus negligible anti-resonances with large bandwidths and is more like that of a nasalized vowel than are those of any other nasal.

[106]This ranking is consistent with the previous tableaux, provided that $v > y$ and $x > u$.

It is therefore plausible (again, in spite of Dayley's impressionistic transcription), that these velar nasals are, at least in some cases, a nasalized vocoid (presumably a nasalized dorsal glide, [ɰ̃]) rather than a non-continuant (cf. Trigo 1988 on the "placeless" behavior of many nasals which have been transcribed as [ŋ]).

In fact, the variable weakening of the labiovelar nasal ([ŋʷ] ~ [w̃]) suggests that both scenarios occur in TS. When complete velar closure is achieved, the nasal does not appear to weaken. Hence, /ŋ/ and /ŋʷ/ can surface unlenited. But when velar closure does not occur, due to contextual or pragmatic conditions which raise the effort cost of velar closure in a nasal (cf. Chapter 6), the resulting continuants are heard (by Dayley) as [ŋ] in the case of the plain velar (due to its confusability with [ɰ̃]), and as [w̃] in the case of the labiovelar. Indeed the notion of contextual raising the effort cost of velar closure allows us to understand why the variation in the realization of /ŋʷ/ appears to be limited to the context /V_{+back}__. Presumably, it is easier to achieve closure with the tongue body against the velum when the tongue body is already retracted due to the preceding vowel.

5. ANALYSIS OF LARYNGEAL ELISION

We have already accounted for obligatory elision of /h/ in pre-obstruent position (section 3.4). To account for its optional elision in intervocalic position, we simply need a context-sensitive version of the blocking constraint, PRES(aspiration): specifically, higher ranking for preservation of aspiration noise in contexts where it is followed by a more sonorous segment (see Bladon 1986, Silverman 1995 for the auditory basis for greater salience of quiet-loud vs. loud-quiet transition). Moreover, the variability of /h/ elision in intervocalic position, versus non-elision in pre-glide position, follows from interleaving of the context-sensitive faithfulness constraint within the LAZY series: specifically, between effort thresholds *s* (corresponding to [h] in pre-glide position) and *t* (corresponding to [h] in pre-vocalic position). For reasons discussed in Chapter 6 section 1.2.4, *s* > *t*.

(7-23)

		$LAZY_s$	PRES(asp/__[-cons])	$LAZY_t$	PRES(asp)
	...hɸ...			*!	
☞	...ɸ...				*
☞	...VhV...	*		*	
☞	...VV...		*		*
☞	...Vhw...			*	
	...Vw...		*!		*

Finally, context-free optional elision of /ʔ/ follows from free ranking of PRES(glottalization) relative to LAZY.

(7-24)

		LAZY	PRES(glottalization)
☞	ʔ - ʔ	*	
☞	ʔ - Ø		*

6. SUMMARY

The TS lenition facts can thus be accounted for in terms of the following constraint hierarchy (in Hasse diagram form):

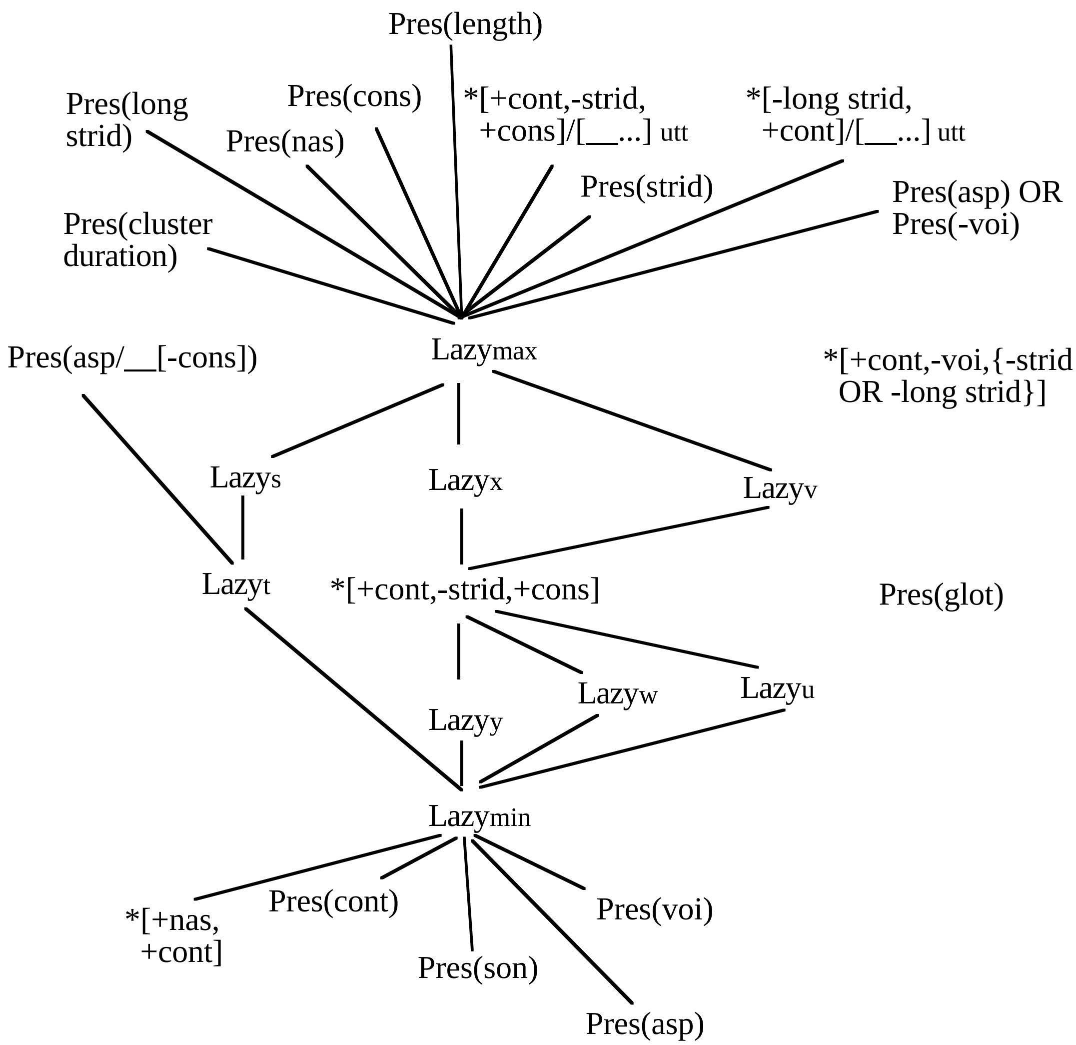


Figure 7-3. Hasse diagram of constraint hierarchy for Tümpisa Shoshone

(LAZY$_{max}$ and LAZY$_{min}$ in Figure 7-3 refer to the highest and lowest effort threshold, respectively, within the LAZY series: this corresponds to rankings in previous tableaux where LAZY (i.e. the whole series) either dominates or is dominated by some lenition-blocking constraint. Constraints which are not connected to the lattice in Figure 7-3 are freely ranked with respect to the other constraints.)

The lenition processes of TS exemplify a number of generalizations discussed and accounted for in previous chapters: specifically, the non-stridency of spirantization outputs (except in the case of underlyingly affricates); blocking of spirantization and flapping in full and partial geminates; blocking of voicing lenition in full, but not partial, geminates. TS also presents a number of plausible cases of effort-based lenition contexts, including the variation between [ð] and [ɾ] (more extreme reduction occurs in contexts where the closer constriction is more effortful, in this case due to coarticulation with the preceding vowel); and (perhaps optional) blocking of weakening in velar and labiovelar nasals (due to the reduced displacement required for closure when the velum is lowered). Moreover, this case study demonstrates that the effort-based approach presented in this dissertation is not only capable of accounting for isolated typological generalizations, but can also offer a coherent and comprehensive analysis of the detailed lenition patterns of a particular language.

Moreover, it is worthwhile in particular to contrast the foregoing analysis of TS with conceivable rule-based alternatives, which permit no unified expression of the spirantization operation (-cont - +cont, i.e. constriction reduction) and the flapping operation (-son - +son, i.e. temporal reduction).[107] The effort-based analysis, on the other hand, is unified, in the sense that both the flapping process and the spirantization process are driven by the same constraint, LAZY, and the choice between the spirantized and flapped outputs follows from a single, consistent constraint hierarchy.

[107] See generally Inouye 1995 for a review of the phonetic and phonological arguments against treatment of flaps as continuants (though Inouye does propose that flaps are [+cont] at their edges, i.e. tripartite contour segments, as noted in Chapter 1.

Chapter 8:
Florentine Italian

Florentine Italian, as described in Giannelli and Savoia (1979), displays extensive lenition of all its consonants. Most notably, the stops, under various phonetic and stylistic conditions, can be spirantized to fricatives or approximants, voiced, debuccalized, or dropped entirely. On the other hand, certain aspects of Florentine lenition appear to be stable across speech rates and registers, such as lenition of intervocalic affricates [tʃ, dʒ] to fricatives [ʃ,ʒ]. Yet the stable and variable patterns are intimately bound up with one another, requiring a unified analysis. I will show that both can be analyzed in terms of conflicts between avoidance of certain levels of effort expenditure vs. lenition-blocking constraints. In particular, Florentine illustrates the following general points, anticipated in previous chapters:

- Effort thresholds correspond to the observed contexts in which lenition occurs, as well as constraining the possible outcomes of lenition.
- Effort thresholds correspond to the speech rate and register conditions under which lenition is more likely to occur. The higher the rate or the lower the register, the broader the contexts under which lenition occurs, and the more extreme the lenition which occurs in a given context.
- The notion of effort minimization is essential to a unified treatment of the various types of lenition.

The facts are drawn from Giannelli and Savoia's (1979) detailed sociolinguistic study of Florentine and its relation to the dialects of the surrounding region of Tuscany (henceforth "G&S"). In brief, spirantization of voiceless stops (e.g. [kasa] / [la xasa], 'house / the house') distinguishes Florentine from the rest of Tuscan, which instead frequently voices these stops ([la gasa]).[108] Within Florence itself, a characterization of the dialect is complicated by a certain amount of diglossia, particularly among the educationally and economically privileged classes, between Florentine and Standard Italian (the latter being nearly equivalent to the Tuscan dialect), with concomitant inhibition of spirantization. I will therefore confine the description to the more distinctively local sociolect, termed *uso conservativo* by G&S, spoken by working-class and older middle-class speakers (born before 1940-45) in Florence and immediately adjoining rural areas, which displays the most extreme consonant lenition processes.

1. DATA.

As a frame of reference, I present the following chart of consonant "phonemes" for Florentine Italian:

Table 8-1. Consonant "phonemes" of Florentine Italian.

Short:

lab	lab-dent	dent	post alv	pal-alv	pal	vel
p		t				k
b		d				g
				tʃ		
				dʒ		
	f		s			
	v					
m		n				
		l				
			ɾ			
w					j	

Long:

lab	lab-dent	dent	post alv	pal-alv	pal	vel
pp		tt				kk
bb		dd				gg
			tts	ttʃ		
				ddʒ		
	ff		ss			
mm		nn				
		ll			ʎʎ	
			rr			

[108] It is therefore ironic that this feature of Florentine proper is traditionally referred to as “Gorgia Toscana.” On the other hand, Giannelli and Savoia report that Florentine spirantization has been spreading to surrounding dialects in recent decades.

Note that here and throughout the description of consonant variation below, my use of phonemic terminology and notation (e.g. stop phonemes realized as approximant allophones) is purely descriptive; it does not imply the assumption that the surface approximants are uniformly stops in underlying representation.

1.1. OBLIGATORY SPIRANTIZATIONS

1.1.1. VOICELESS STOPS. In intervocalic position, the voiceless stops /p,t,k/ are obligatorily spirantized, typically to approximants [ɸ̞,θ̞,x̞]. The environment is, more precisely, intervocalic with an optional intervening liquid or glide, word-internally or across a word boundary: i.e. /V__(#)({+son,-nas})V (henceforth “weak position"; and its complement, "strong position").

(8-1)

	Florentine	Orthography	Gloss
a.	kaɸ̞o	*capo*	'head'
	plastiha / la ɸ̞lastiha	*(la) plastica*	'(the) plastic'
	pentola / la ɸ̞entola	*(la) pentola*	'(the) pot'
	pɾeso / el anno ɸ̞ɾeso	*(l'hanno) preso*	'(they have) taken'
	pjɛna / ll ɛɾa ɸ̞jɛna	*(era) piena*	'(s/he was) full'
b.	pɾaθ̞o	*prato*	'meadow'
	pjɛθ̞ɾa	*pietra*	'stone'
	tjɛne / e lo θ̞jɛne	*(lo) tiene*	'(s/he) has (it)'
	tavola / la θ̞avola	*(la) tavola*	'(the) table'
	tɾave / la θ̞ɾave	*(la) trave*	'beam'
c.	amix̞o	*amico*	'friend'
	pɔx̞o	*poco*	'little'
	biʃix̞letta	*bicicletta*	'bicycle'
	kasa / la x̞asa	*(la) casa*	'(the) house'
	koltɛlli / i x̞oltɛlli	*(i) coltelli*	'(the) knives'
	kwattɾini / i x̞wattɾini	*i quattrini*	'(the) money'

Indeed, G&S report that many Florentine speakers have difficulty producing voiceless stops, particularly [k], in weak position, e.g. when attempting to imitate Standard Italian.

Some further details:

• The [θ̞] allophone in (2b) is typically post-dental, though a slightly noisier interdental realization occurs as well, particularly among women and rural speakers. Among younger speakers of *uso conservativo*, /k/ in weak position (2c) typically reduces to [h] rather than [x̞]: [amiho], [la hasa], etc.

• Though weak position spans word boundaries, Nespor and Vogel (1982) have noted that it does not span the boundary of an intonational phrase, i.e. “the domain over which an intonation contour is spread."

(8-2) [$_I$ le hase haɾine] [$_I$ kostano molto haɾe in ameɾiha]
'Cute houses are very expensive in America'
[$_I$ ki ai visto hosti]
'Who did you see there?

• Geminates are immune to this obligatory spirantization. These geminates derive from underlying geminates (8-3a), or by *radoppiamento sintattico* (b), a mutation whereby certain function words, such as the determiner *il*, induce gemination of a following word-initial consonant (for a fuller description of *radoppiamento sintattico* , see Nespor and Vogel 1982).[109]

(8-3)	a.	tappalo	*tappalo*	'plug it up'
		gɾatta	*gratta*	'scratch'
		sekko	*secco*	'dry'
	b.	i ppane	*il pane*	'the bread'
		i kkorridojo	*il corridoio*	'the corridor'

Spirantization of /g/. Voiced velar stops in weak position obligatorily spirantize as well:

[109]In addition, the alveolar affricate [tts] and palatal lateral [ʎʎ] are predictably long.

(8-4)	Florentine	Orthography	Gloss
	gamba / la ɰamba	*(la) gamba*	'(the) leg'
	gjanda / la ɰjanda	*(la) ghianda*	'(the) acorn'
	gɾatːa / e si ɰɾatta	*(si) gratta*	'(it) scratches'
	e seɰa	*sega*	's/he mows'
	i ssuɰo	*il sugo*	'the juice'

Spirantization of Affricates. Finally, the affricates /tʃ/ and /dʒ/ obligatorily spirantize in weak position:[110]

(8-5)				
	a.	peʃe	*pece*	'pitch'
		baʃo	*bacio*	'kiss'
		paʃe	*pace*	'peace'
		tʃena / la ʃena	*(la) cena*	'dinner'
	b.	e ɾiʒetta	*rigetta*	'reject'
		faʒano	*fagiano*	'pheasant'
		dʒoɾni / i ʒoɾni	*(i) giorni*	'(the) days'
		dʒɔx̞a / e ʒɔx̞a	*gioca*	's/he plays'

Note that this process neutralizes /tʃ/ with /ʃ/: for example, *la cena* ('the dinner') and *la scena* ('the scene') are both realized as [la ʃena].

[110]Giannelli and Savoia state the context as /V__V, rather than the "weak position" described above; but the phonotactics are such that affricate + liquid or affricate + glide clusters do not arise.

The post-alveolar affricate /tts/ (orthographic *z* or *zz*) does not participate in this spirantization, because it is predictably long: e.g. [pittsa] ('pizza'), [pattsjente] (*paziente*, 'patient').

1.2. Variable Lenition: Weak Position

1.2.1. Voiceless Stops. Though lenition of voiceless stops is obligatory in weak position, the *degree* of lenition varies along a scale, from a close fricative all the way to Ø, depending on certain phonetic and pragmatic conditions:

Close fricatives. In careful (though not necessarily slow) speech (*stile accurato*), particularly in the context of new discourse information, the stops tend to lenite to close fricatives [ϕ,θ̝,x̝] rather than approximants. These close fricatives are "easily perceptible as stops," though a spectrographic analysis shows them to be continuants.

Close/open approximants. In slow but natural speech (*stile lento ma naturale*), the stops tend to lenite to close (but unfricated) approximants; whereas at a faster speech rate or lower register, the approximant allophones tend to be more open. The open approximants are transcribed as [ϕ̞,θ̞,x̞].[111] The preference for close or open approximants also varies somewhat by idiolect.

Debuccalization.

• Velar. It has already been noted that younger speakers of *uso conservativo* tend, at least in moderate, natural speech styles, to debuccalize the velar stop to [h], whereas

[111] G&S use the older, but equivalent, IPA "open" diacritic [ϕ̩,θ̩,x̩]), and transcribe the closer approximants as [ϕ,θ,x] (since the fricatives are already distinguished from the approximants by the "close" diacritic [ϕ̝,θ̝,x̝]).

older speakers tend to spirantize to [x̞] under equivalent circumstances. However, even older speakers debuccalize in careless speech (*stile trascurato*). A further factor favoring a debuccalization is occurrence in what G&S call the "body of the phrase" (*corpo di frase*), defined as “a syllable followed by at least one stressed syllable within the phrase," i.e. not in the phrase-final foot:

(8-6)	ˈlː ɛɾa ˈʃɛx̞a	*but frequently*	ˈlː ɛɾa ʃɛha ðа un ˈɔcːo
	era cieca		*era cieca da un occhio*
	'she was blind'		'she was blind in one eye'

I view this tendency for phrase-final blocking of debuccalization as an effect of the common phonetic process of phrase-final lengthening. That is, a given speech style may be just fast and casual enough to trigger debuccalization in the "body" of the phrase; but in the final foot, the speech rate slows down to the point that debuccalization is blocked. Thus, the *corpo di frase* context is merely a special case of the inhibiting effect of slow speech rate on lenition.

• Bilabials and Dentals. Debuccalization of /p/ and /t/ is rarer than with /k/; moreover, /p/ debuccalizes somewhat more readily than /t/. Otherwise, the factors favoring debuccalization of /k/, i.e. low register and fast rate (including the *corpo di frase* context), likewise favor debuccalization of /p/ and /t/.

Finally, G&S note that the voiceless stops can also sporadically debuccalize to an allophone which they transcribe as [h̞]; however, it is unclear precisely how this [h̞] differs from [h] (and [ɦ]), therefore I disregard it below.

Elision. In the lowest registers and fastest rates, the voiceless stops tend to elide completely.[112] G&S note that elision is also common in "emotive" exclamations, e.g. [ʃɛx̞o] ('blind') but [o kke sse ʃɛo!] (*O che sei cieco!*, 'How blind you are!'). Such emotive exclamations can, of course, be viewed as a special case of low register.

In addition, G&S claim that the velar stop tends to elide when flanked by vowels which are identical with respect to height, frontness, and tenseness[113] or when followed by a [+back] vowel; however, the examples they give do not all conform to this generalization:

(8-7)		Frequent	Rarer	Orthography	Gloss
	(a)	baa	bax̞a	*baca*	'worm'
		ɸɔo	ɸɔx̞o	*poco*	'little'
		ʃɛe	ʃɛhe	*cieche*	'blind-f.pl.'
		bua	bux̞a	*buca*	'hole'
		le ɔrna	le hɔrna	*le corna*	'the horns'
		la assetta	la x̞assetta	*la cassetta*	'the box'
		i oltɛlli	i x̞oltɛlli	*i coltelli*	'the knives'
	(b)	i bahini	i βaini	*i bachini*	'the little worms'
		e sɔno x̞inaθ̞i	e sɔno inaθ̞i	*sono chinati*	'they are leaning'
		ʃɛha	ʃɛa	*cieca*	'blind-f.sg.'

---

[112]So G&S report, based on auditory impression, apparently supplemented with some spectrographic analysis. It is, of course, possible that an articulatory study would reveal vestigial gestures, which are merely accoustically "hidden" (cf. Browman and Goldstein 1990), paralleling Nolan's (1992) and Barry's (1985) findings for consonant reduction in English, and Jun's (1995) findings for Korean.

[113]G&S use the homologous feature [peripheral] (*periferico*).

Thus, [ʃεe] preferrably undergoes elision in spite of non-matching [tense] specifications, and [ʃεx̞a] preferrably does not undergo elision, although the following vowel is [+back].

Rather, the data are consistent with the simpler generalization that elision of /k/ tends to be blocked before high vowels (8-7b). This context can therefore be understood in terms of the tendency of high vowels to facilitate devoicing and frication, as discussed in Chapter 6, section 2.2.[114]

Voicing. Finally, under the same rate and register conditions in which the voiceless stops elide, they may instead undergo voicing.

(8-8)	e (ɰ̞)orre sεmpɾe	*corre sempre*	's/he is always running'
	e lo (ð̞)εŋgo io	*lo tengo io*	'I have it'
	e ʃe sta(ð̞)o lui	*c'é stato lui*	'it was him
	e lo (β̞)ɾεnde lui	*lo prende lui*	'he takes it'

According to G&S, the approximant derived by the voicing process optionally remains "[+tense]," thus distinct from underlying voiced ("lax") stops, which can also lenite to voiced approximants (see section 1.2.2 below). However, "tenseness" cannot refer to greater constriction degree, for all these approximant allophones, both tense and lax, are described as having extremely open constrictions. Presumably, then, G&S's "tenseness" distinction refers not to a distinction in the consonant itself, but to some external cue to the voicing contrast, most likely the duration of the preceding vowel (cf. Steriade 1995 on the role of such external cues in voicing contrasts). Thus, since this "tenseness"

---

[114] I offer no explanation for the failure of a *preceding* high vowel to inhibit elision, e.g. in [bu(x̞)a].

distinction does not concern the degree of consonantal lenition per se, I henceforth disregard it.

1.2.2. VOICED STOPS. Obligatory spirantization of /g/ has already been noted in section 1.1. In slow, careful speech, /b/ and /d/ tend not to spirantize; however, in faster, more casual speech, they can spirantize as well.

(8-9)	a.	dɔɾme / e {d/ð̞}ɔɾme	*dorme*	'(s/he) sleeps'
		korridojo ~ korrið̞ojo	*corridoio*	'corridor'
	b.	beve / e beve ~ e β̞eve	*beve*	'(s/he) drinks
		debole ~ deβ̞ole	*debole*	'weak'

Like the voiceless stops discussed above, the actual degree of lenition varies, depending on speech rate and register. Among the voiced stops, only the velar tends to debuccalize in careless speech, to [ɦ].

(8-10)	la {ɰ/ɦ}amba	*la gamba*	'the leg'
	la {ɰ/ɦ}janda	*la ghianda*	'the acorn'
	e si {ɰ/ɦ}ɾatta	*si gratta*	's/he scratches'
	e se{ɰ/ɦ}a	*sega*	's/he mows'
	i ssu{ɰ/ɦ}o	*il sugo*	'the juice'

All voiced stops, however, can reduce to Ø in lower stylistic levels, typically not in the phrase-final foot, paralleling the behavior of the voiceless stops.

(8-11)	e ʃ a lle (ɰ/ɦ)ɔθ̞e ˈɾosse	*ha le gote rosse*	's/he has red cheeks'
	la mi(ð̞)olla (ð̞)i ˈppane	*la midolla di pane*	'the crumb of bread'
	e (β̞)ɾontola ˈsɛmpɾe	*brontola sempre*	's/he is always grumbling'

1.2.3. THE BEHAVIOR OF GEMINATES. In extremely fast, careless speech, when flanked by vowels, even geminate stops can undergo spirantization. G&S describe these as "very constricted realizations, easily perceived as stops" (*realizzazioni molto chiuse e facilmente percepibili come occlusive*).

(8-12)	ll e bbɾutto / ll e bbɾuθ̞θ̞o	*è brutto*	'he's ugly'
	ll e ssekko / ll e ssex̞x̞o	*è secco*	'it's dry'
	tappalo / taɸɸalo	*tappalo*	'plug it up'
	tɾe ddiθ̞i / tɾe ð̞ð̞iθ̞i	*tre diti*	'three fingers'
	tɾe ggalletti / tɾe ɣ̞ɣ̞alletti	*tre galletti*	'three biscuits'

Geminate affricates can spirantize as well in careless speech:.[115]

(8-13)	i bɾattʃi ~ i bɾaʃʃi	*le braccia*	'the arms'
	ɛɾa ɸɛddʒo di lui ~ ɛɾa ɸɛʒʒo di lui	*era peggio di lui*	's/he was worse than him'
	tu ttʃeni oɾa ~ tu ʃʃeni oɾa?	*ceni ora?*	'are you dining now?'

[115]Giannelli and Savoia do not discuss the behavior of the (predictably long) alveolar affricate [tts], (e.g. [pattsjente] (*paziente*, 'patient'), though this omission may be inadvertent. I assume that this geminate affricate patterns with the aveopalatal geminate affricates with regard to lenition.

However, as discussed in Chapter 5, as this is a very fast speech phenomenon, I assume that these “geminate" spirants are not phonetically long, though other cues to the consonant length contrast are apparently preserved. Henceforth, I therefore transcribe these shortened, spirantized “geminates” as singletons.

1.2.4. LENITION OF OTHER CONSONANTS IN WEAK POSITION.

Strident Fricatives. In weak position, strident fricatives /f,v,s,ʃ/, as well as the (obligatorily spirantized) affricates /tʃ,dʒ/, can lenite to corresponding approximants [f̞,ʋ,s̞,ʃ̞,ʒ̞], particularly in non-phrase-final feet:

(8-14)	e lo {f/f̞}anno loɾo	*lo fanno loro*	'they do it'
	{f/f̞}aθ̞elo voi	*fatelo voi*	'you-pl. do it'
	e l a s{f/f̞}ilaθ̞o	*l'ha sfilato*	's/he has paraded it'
	e ʃi ʋeð̞o	*ci vedo*	'I see there/it'
	e ʃi staʋo	*ci stavo*	'I was there'
	e me lo ðaʋa	*me lo dava*	's/he was giving it to me'
	ũ llo {s/s̞}aɸ̞eva	*non lo sapeva*	's/he didn't know it'
	e lo ɸ̞i{ʒ/ʒ̞}ava {ʒ/ʒ̞}u	*lo pigiava giù*	's/he was stepping on it'
	e uŋ kwɔʃ̞e	*non cuoce*	'it doesn't cook'

G&S note that these approximants are still consonantal; thus the approximant [ʒ̞] (derived from /dʒ/) is still somewhat closer than the glide [j].

Moreover, /v/ can elide entirely in careless speech:

(8-15)	e ʃi eð̞o	*ci vedo*	'I see there/it'
	e ʃi stao	*ci stavo*	'I was there'
	e me lo ð̞aa	*me lo dava*	's/he was giving it to me'
	e uɲ tʃi ɔ	*non ci vo*	'I'm not going there'

Furthermore, G&S report that the voiceless stridents may undergo voicing in careless speech.

Nasals. The nasals /m,n/ reduce to nasalized approximants [β̞̃,ɹ̃] in weak position[116] in natural speech styles, particularly in a non-phrase-final foot.

(8-16)	e veŋgo ðoβ̞aɹ̃i	*vengo domani*	'I'm coming tomorrow'
	e lo {m/β̞}aɲdʒa	*lo mangia*	's/he eats it'
	i ppa{n/ɹ̃}e	*il pane*	'the bread'

Moreover, in more casual speech registers, nasals in weak position may elide altogether.

(8-17)	e veŋgo ðõãi seɾa	*vengo domani sera*	'I'm coming tomorrow evening'
	e lõ ãɲdʒa lui	*lo mangia lui*	'he eats it'
	i ppãẽ sekko	*i pane secco*	'the dry bread'

Note that the persistence of nasalization on the adjacent vowels in (8-17) is not tied to the loss of the nasal consonant; rather, there is a general process whereby nasalization spreads bidirectionally from [+nasal] segments to adjacent vowels and glides, iteratively,

[116] "Intervocalic" position according to Giannelli and Savoia, but nasal + liquid and nasal + glide sequences do not arise, cf. fn. 2.

which is consistently reflected in G&S's transcriptions, though I have in most cases omitted this from the transcriptions, as it is not directly relevant to the analysis of lenition.

Liquids. The short lateral /l/, and the short and long rhotics /ɾ,rr/ can lenite to approximants [l̞,ɹ,ɹɹ] in weak position, in natural speech styles, preferrably in a non-phrase-final foot.

(8-18)	la ʃe{ɾ/ɹ}a	*la cera*	'the wax'
	e ɸ̞aɾte ~ e ɸ̞aɹθ̞e	*parte*	'part'
	i kko{rr/ɹɹ}iðojo	*il corridoio*	'the corridor'
	{ɾ/ɹ}imaniʃi	*rimanici*	'stay here'
	{l/l̞}evaθ̞i̞	*levati*	'get up'
	ʎʎ e ddoltʃe ~ ʎʎ e ddol̞ʃe	*è dolce*	'it's sweet'

Note that the durational distinction between the rhotics is maintained under this reduction. The geminate laterals /ll, ʎʎ/ appear not to undergo this reduction.

1.2.4. SUMMARY: WEAK POSITION. The weak-position consonant variation described above can be summarized in terms of the following chart, where levels A through K represent a conflation of speech rate and register factors: level A corresponds to the slowest, most careful speech style; B is somewhat faster or more casual (or a modicum of both); and so on, up to K, the fastest, most careless level.[117]

[117] These rate and register factors should in principle be expressed in terms of continuous numerical values. I am forced to use a sequence of discrete levels here, since G&S's data, which I am attempting to model, are qualitative rather than quantitative. Further note that Table 8-2 is not taken directly from G&S, but is rather my summary of this complex set of rate- and register-sensitiveprocesses.

Table 8-2. Lenition variation according to rate/register.

	A	B	C	D	E	F	G	H	I	J	K
/k/	x̝	x	x̞			h				ɰ	Ø
/t/	θ̝	θ	θ̞						h	ð̞	Ø
/p/	ɸ̝	ɸ	ɸ̞					h		β̞	Ø
/g/	ɣ̝	ɣ	ɰ						ɦ		Ø
/b,d/	b,d		β̝,ð̝	β̞,ð̞							Ø
/tʃ,dʒ/	ʃ,ʒ				ʃ̞,ʒ̞					ʒ̞	
/v/	v				ʋ						Ø
/ʃ,s,f/	ʃ,s,f				ʃ̞,s̞,f̞					ʒ̞,z̞,ʋ	
/m,n/	m,n		β̃,ɹ̃								Ø
/ɾ,rr,l/	ɾ,rr,l				ɹ,ɹɹ,l̞						
/kː,tː,pː/	kː,tː,pː										ɸ̝,θ̝,x̝
/bː,dː,gː/	bː,dː,gː										β̝,ð̝,ɣ̝
/dːʒ/	dːʒ										ʒ
/tːʃ/	tːʃ										ʃ

The lower the register, or the higher the rate, the greater the reduction. Moreover, the probability of occurrence of particular allophones is reflected by the number of levels that the allophone occupies: for example, /p/ debuccalizes in levels I and J, whereas /t/ debuccalizes only in J, thus reflecting the lower probability of /t/ - [h] in fast/careless speech.

Note that Table 8-2 pertains to the older sociolect, which lenites /k/ to [x̞] in intermediate levels. The younger sociolect has no [x̞] allophone, therefore /k/ must debuccalize at level C (and, I presume, /g/ debuccalizes at a somewhat earlier level as well):

Table 8-3. Lenition variation according to rate/register, younger speakers.

	A	B	C	D	E	F	G	H	I	J	K
/k/	x̝	x	h							ɰ	Ø
/g/	ɣ̝	ɣ	ɰ			ɦ					Ø

1.3. LENITION IN STRONG POSITION

In faster, more casual speech styles, the environment for lenition expands beyond weak position.

Both voiceless and voiced stops can also spirantize in strong positions, particularly if the previous consonant is a continuant.

(8-19) Voiceless:

a.	[ɪ {k/x̞}uʃilo	*cucilo*	'sew it'
	lo ʃeɾ{k/x}a	*lo cerca*	's/he is looking for it'
b.	[ɪ tjɛni /[ɪ θ̞çɛni	*tieni*	'you have it'
	la θ̞oɾ{t/θ}a	*la torta*	'the cake'
	la ʃes{t/θ}a	*la cesta*	'the basket'
c.	[ɪ {p/ɸ}ɔɾtalo	*portalo*	'carry it'
	e lo s{p/ɸ̞}ɛɾo	*lo spero*	'I hope'

(8-20) Voiced

a.	[ɪ {g/ɣ}ɾattalo	*grattalo*	'scrape it'
	e si s{g/ɣ}ottʃola	*si sgocciola*	'it drips'
b.	[ɪ {d/ð}aʎʎelo	*daglielo*	'give it to him'
	lo z{d/ð}entaθ̞o	*sdentato*	'toothless'
	e si skɔɾda di θ̞utto ~ e si skɔɹða ði θ̞utto	*si scorda di tutto*	's/he remembers everything'
c.	[ɪ βevilo θ̞e	*bevilo te*	'drink it up'
	lo z{b/β}udella	*lo sbudella*	's/he guts it'

Moreover, after [s], G&S note that close fricatives are commonly found: [lo sɸ̞ɛɾo], [la sθ̞essa] (*la stessa*, 'the same-f.sg.').

In addition, affricates frequently spirantize in strong position:

(8-21)	[$_{I}$ (t)ʃeɾkalo	*cercalo*	'look for it'
	[$_{I}$ (d)ʒiɾalo	*giralo*	'turn it'
	una hweɾtʃa ~ una hweɹʃa	*una quercia*	'a live oak'
	e si spɔɾ(d)ʒe sɛmpɾe	*si sporge sempre*	's/he is always peaking'
	e lo maɲ(d)ʒa lui	*lo mangia lui*	'he eats it'

This strong position affricate spirantization is likewise tied to register and rate, again with some inhibition in the phrase-final foot.

Finally, G&S report that lenition of nasals occurs in strong position in careless speech registers.

(8-22)	[$_{I}$ βaɲdʒalo	*mangialo*	'eat it'
	[$_{I}$ {n/ɹ̃}askondilo	*nascondilo*	'hide it'
	e lo z{m/β}onta	*lo smonta*	's/he dismounts it'
	e lo z{n/ɹ̃}ɔttʃola	*lo snocciola*	's/he pits it'

In sum, the lenitions which are obligatory in weak position apply only in fast/casual speech in strong position; and the lenitions which apply in fast/casual speech in weak position do not apply in even faster/more casual speech, or not at all, in weak position. It is as if the set of reductions were "bumped up," as it were, several levels to the right, relative to their behavior in weak position, Table 8-2.

However, it should be clear that the binary distinction between strong and weak position in G&S's characterization of the data, is merely a rough approximation to reality. It has already been noted that elision of /k/ is sensitive to the height of the following vowel; and that, within strong position, the degree of reduction of stops is sensitive to the constriction degree of the preceding consonant. I assume that these cases are not isolated, but rather suggest that these reduction processes are generally gradiently sensitive to the aperture of adjacent segments, as well as to speech rate and register. "Strong" vs. "weak" position, then, should be understood as two points along a continuum involving openness of flanking segments, as discussed in Chapter 6.

1.4. SUMMARY OF DATA

Several of the lenition processes described above are stable across speech rates and registers: namely, spirantization (or further reduction) of the voiceless stops in weak position, spirantization (or further reduction) of the voiced velar stop in weak position, and spirantization of affricates in weak position. However, these stable lenitions are clearly part and parcel of more general system of variable consonant reduction, which applies in a broader range of contexts, to a larger class of consonants, and results in more drastic reductions,

A. the faster the speech rate,

B. the lower the register, and

C. the more open the flanking segments.

The other conditions noted by G&S, e.g. the *corpo di frase* context, and emotive exclamations, plausibly reduce to rate and register conditions as well.

We shall see that this sea of variation, as well as the islands of stability, admit of a unified analysis in terms of interleaved effort thresholds and faithfulness or fortition constraints.

2. Effort Values

Before proceeding to the analysis, I must lay out my assumptions concerning the relative effort involved in these consonants, in the relevant contexts and stylistic conditions, based on considerations (displacement, velocity, precision, etc.) identified in previous chapters. Since the number of effort comparisons is large, the use of letter variables for effort levels, as in previous chapters, becomes unwieldy. I therefore must use numerical values; however, it should be borne in mind that these values are in abstract units, deduced from general considerations of what sorts of gestures are more or less effortful in particular contexts, as in previous chapters; the values do not reflect actual measurements. Furthermore, whereas with letter variables it is possible to have partial orderings of effort levels ($x > z$ and $y > z$, but the relation of x to y is unknown), the use of numbers commits us to a total ordering.

Here, then, are the set of effort values:

Table 8-4. Hypothetical effort values (in abstract units) for consonant allophones in weak position.

	A	B	C	D	E	F	G	H	I	J	K
β̝β̝,ð̝ð̝,ɣ̝ɣ̝,zz,ʒʒ,vv	106	111	117	123	129	135	142	149	157	164	173
ɸ̝ɸ̝,θ̝θ̝,x̝x̝,ss,ʃʃ,ff	105	110	116	122	128	134	141	148	155	163	171
ddʒ	103	108	114	119	125	131	138	145	152	160	168
ttʃ	100	105	110	116	122	128	134	141	148	155	163
bb,dd,gg	93	98	103	108	113	119	125	131	137	144	151
kk,pp,tt	90	95	99	104	109	115	121	127	133	140	147
tʃ	96	101	106	111	117	123	129	135	142	149	156
dʒ	95	100	105	110	115	121	127	134	140	147	155
ʃ,s,f	91	96	100	105	111	116	122	128	134	141	148
ʒ,z,v,rr	90	95	99	104	109	115	121	127	133	140	147
k,p,t	85	89	94	98	103	108	114	120	126	132	138
b,d,g,m,n	75	79	83	87	91	96	101	106	111	116	122
ɸ̝,θ̝,x̝,l	74	78	82	86	90	94	99	104	109	115	121
β̝,ð̝,ɣ̝,ɾ	73	77	80	85	89	93	98	103	108	113	119
ɸ,θ,x	70	74	77	81	85	89	94	98	103	109	114
ɸ̞,θ̞,x̞,f̞,s̞,ʃ̞	65	68	72	75	79	83	87	91	96	101	106
h	60	63	66	69	73	77	80	84	89	93	98
β,ð,ɣ,ɹɹ	60	63	66	69	73	77	80	84	89	93	98
β̞,ð̞,ɣ̞,ʋ,z̞,ʒ̞,β̞̃,ɹ̃,ɹ,l̞	55	58	61	64	67	70	74	77	81	85	91
ɦ	50	53	55	58	61	64	67	70	74	78	81
Ø	0	0	0	0	0	0	0	0	0	0	0

Table 8-5. Hypothetical effort values (in abstract units) for consonant allophones in strong position.

	A	B	C	D	E	F	G	H	I	J	K
β̝β̝,ð̝ð̝,ɣ̝ɣ̝,zz,ʒʒ,vv	81	85	89	94	98	103	109	114	120	126	132
ɸ̝ɸ̝,θ̝θ̝,x̝x̝,ss,ʃʃ,ff	80	84	88	93	97	102	107	113	118	124	130
ddʒ	78	82	86	90	95	100	105	110	115	121	127
ttʃ	75	79	83	87	91	96	101	106	111	116	122
bb,dd,gg	68	71	75	79	83	87	91	96	100	105	111
kk,pp,tt	65	68	72	75	79	83	87	91	96	101	106
tʃ	71	75	78	82	86	91	95	100	105	110	116
dʒ	70	74	77	81	85	89	94	98	103	109	114
ʃ,s,f	66	69	73	76	80	84	88	93	98	102	108
ʒ,z,v,rr	65	68	72	75	79	83	87	91	96	101	106
k,p,t	60	63	66	69	73	77	80	84	89	93	98
b,d,g,m,n	50	53	55	58	61	64	67	70	74	78	82
ɸ̝,θ̝,x̝,l	49	51	54	57	60	63	66	69	72	76	80
β̝,ð̝,ɣ̝,ɾ	48	50	53	56	58	61	64	68	71	74	78
ɸ,θ,x	45	47	50	52	55	57	60	63	66	70	73
ɸ̞,θ̞,x̞,f̞,s̞,ʃ̞	40	42	44	46	49	51	54	56	59	62	65
h	35	37	39	41	43	45	47	49	52	54	57
β,ð,ɣ,ɹɹ	35	37	39	41	43	45	47	49	52	54	57
β̞,ð̞,ɣ̞,ʋ,z̞,ʒ̞,β̞,ĩ,ɹ,l̞	30	32	33	35	36	38	40	42	44	47	49
ɦ	25	26	28	29	30	32	34	35	37	39	41
Ø	0	0	0	0	0	0	0	0	0	0	0

Letters A-K correspond to the rate/register levels in Table 8-2. The hypothetical values assigned reflect the following assumptions:

(a) The faster or more casual the speech style, the greater the effort cost (cf. Chapter 2, 6). *Each level above A is augmented by five percent of the value in the preceding column.*

(b) The closer, more fortis, or longer the constriction of the consonant, i.e. the greater the displacement of the constriction gesture, the greater the effort cost (cf. Chapters 2, 4, 5, 6). *Fricatives are assigned a value 3 points higher than close* *approximants. Otherwise, a consonant of constriction degree x is 5 points higher than a corresponding consonant of constriction degree x-1.*

(c) The more open the flanking segments, the greater the effort cost (cf. Chapter 6). I further assume that there is substantially less overlap between gestures belonging to distinct intonational phrases than phrase-internally (Byrd & Saltzman, forthcoming); therefore in phrase-initial position a given constriction can be achieved with a lower velocity, hence less effortful, gesture. Simplifying this contextual continuum into G&S's binary weak vs. strong distinction, *a consonant in weak position has an effort cost 25 points higher than the same consonant in strong position.*

(d) Strident fricatives and affricates are more effortful (*5 points higher*) than corresponding stops (cf. Chapter 4).

(e) Voiced geminate stops (and affricates) are more effortful (*3 points higher*) than their voiceless counterparts (cf. Chapter 5).

(f) Voiceless fricatives are slightly more effortful (*1 point higher*) than their voiced counterparts.[118]

(g) Modulo (e) and (f), voiceless consonants are more effortful (*10 points higher*) than voiced consonants.

[118]The slight difference in effort for the fricatives, compared to the more substantial difference for other consonants (f), reflects Ohala's (1983) observation that voiced fricatives involve strict aerodynamic requirements: if oral pressure is too high, voicing ceases; but if oral pressure falls too low, friction ceases (cf. Chapter 5 section 3.1.4.2). I assume therefore that maintenance of voicing in fricatives involves more precise control of airflow than in stops or approximants.

(h) Nasals are equivalent in effort to voiced stops (assuming that the effort of velic opening or closing is negligible).

(i) The trill [r] is equivalent in effort to a [z], as both involve a precise partial constriction, with voicing.

(j) The lateral [l] is slightly less effortful than a stop, since the constriction need not entirely block oral airflow: I therefore equate its effort with that of a voiceless fricative.

3. ANALYSIS

3.1. THE STABLE SPIRANTIZATIONS

3.1.1. VOICELESS STOPS. Spirantization of stops in weak position at level A is captured in terms of the following ranking:

(8-23) Weak position, level A (p,t,k=85, b,d,g=75, ɸ,θ̠,x̠ =74, β̠,ð̠,ɣ̠=73)

		LAZY75	*-strid, +cont,+cons	PRES(cont)
	p,t,k - p,t,k	*!		
☞	p,t,k - ɸ,θ̠,x̠		*	*
	g - g	*!		
☞	g - ɣ̠		*	*

(The subordination of PRES(cont) to *[-strid,+cont,+cons] ensures that the spirantization is allophonic, i.e. that when considerations of effort do not demand a continuant, namely in the earlier levels in strong position, no nonstrident continuant can surface; rather it must strengthen to a stop, even if underlyingly [+cont] (and, a fortiori, if already [-cont] underlyingly).

(8-24) Strong position, level A (p,t,k=60): fortition to stops

		LAZY75	*-strid, +cont,+cons	PRES(cont)
☞	ɸ̞,θ̞,x̞ - p,t,k			*
	ɸ̞,θ̞,x̝ - ɸ̞,θ̞,x̞		*!	

Thus, despite the use of phonemic terminology in the description above, the assumption of phonemic representation, i.e. that these consonants are uniformly stops (or uniformly continuants) in UR proves to be superfluous to an adequate formal analysis. The reduction is to a fricative, rather than an approximant, at this level, due to another fortition constraint banning sonorant continuants, ranked above the effort cost of the fricatives at this level.

(8-25) Weak position, level A (ɸ̝,θ̝,x̝=74, ɰ=73)

		*[+son,+cons, +cont]	LAZY74	LAZY73
☞	ɸ̞,θ̞,x̞		*	*
	ɸ̝,θ̝,x̝	*!		
☞	ɣ̞			*
	ɰ	*!		

3.1.2. VOICED STOPS. The ranking thus far predicts, however, that all the voiced stops should spirantize as well, whereas only /g/ does so. Velar stops are characterized by a noisy release (seen on a spectrogram as multiple bursts), whereas labials and coronal stops typically have a crisp release. Thus, velar stops are somewhat less acoustically distinct from continuants than coronal or labial stops are. I posit the feature [crisp release]: labial and coronal stops are [+crisp rel]; continuants and velar stops are [-crisp rel]. Spirantization of the labials and coronals, then, is blocked by ranking PRES(crisp rel) above LAZY75.

(8-26) Weak position, level A (b,d,g=75, β,ð,ɣ=73)

		PRES(crisp rel)	LAZY75	PRES(cont)
	g - g		*!	
☞	g - ɣ			*
☞	b,d - b,d		*	
	b,d - β,ð	*!		*

3.1.3. AFFRICATES. The low ranking of PRES(cont) further entails spirantization of the affricates at this level.

(8-27) Weak position, level A (tʃ=96, dʒ=95, ʃ=91, ʒ=90)

		LAZY95	LAZY91	...	LAZY75	PRES(cont)
	tʃ,dʒ - tʃ,dʒ	*!	*	*	*	
☞	tʃ,dʒ - ʃ,ʒ		(* for ʃ)	*	*	*

Blocking of further reduction is attributable to PRES(strid):

(8-28) Weak position, level A (ʃ=91, ʒ=90, ʃ=65, ʒ=55)

		PRES(strid)	LAZY91
☞	tʃ,dʒ - ʃ,ʒ		(* for ʃ)
	tʃ,dʒ - ʃ,ʒ	*!	

3.1.4. STABILITY OF THESE SPIRANTIZATIONS. Since the effect of lower registers (or faster rates) is to increase the effort cost of the gestures comprising each consonant, the foregoing stops and affricates, which spirantize in weak position at level A, must, a fortiori, spirantize in weak position at all faster rates and lower registers of speech. Therefore, the spirantization (or further reduction) of these stops and affricates is obligatory at all levels in weak position.

3.2. THE VARIABLE REDUCTIONS, WEAK POSITION

We will now consider the reductions in weak position, level by level:

Table 8-6. Weak position chart (repeated from Table 8-2)

	A	B	C	D	E	F	G	H	I	J	K
/k/	x̞	x	x̝			h				ɰ	Ø
/t/	θ̞	θ	θ̝						h	ð̝	Ø
/p/	ϕ̞	ϕ	ϕ̝					h		β̝	Ø
/g/	ɣ̞	ɣ	ɰ						ɦ		Ø
/b,d/	b,d		β̞,ð̞	β̝,ð̝							Ø
/tʃ,dʒ/	ʃ,ʒ				ʃ̝,ʒ̝					ʒ̝	
/v/	v				ʋ						Ø
/ʃ,s,f/	ʃ,s,f				ʃ̝,s̝,f̝					ʒ̝,z̝,ʋ	
/m,n/	m,n		β̃,ɹ̃								Ø
/ɾ,rr,l/	ɾ,rr,l				ɹ,ɹɹ,l̝						
/kː,tː,pː/	kː,tː,pː										ϕ̞,θ̞,x̞
/bː,dː,gː/	bː,dː,gː										β̞,ð̞,ɣ̞
/dːʒ/	dːʒ										ʒ
/tːʃ/	tːʃ										ʃ

3.2.1. LEVEL B. At this level, the stops reduce to approximants rather than fricatives. This result is obtained by ranking *[+son,+cons,+cont] below the effort cost of the fricatives at this level:

(8-29) Weak position, level B (ϕ̞,θ̞,x̞=78, ɣ̞=77, ϕ,θ,x=74)

	LAZY77	*[+son,+cons,+cont]	LAZY75
ϕ̝,θ̝,x̝	*!		*
☞ ϕ,θ,x		*	
ɣ̝	*		*
☞ ɣ		*	

Assume a feature [close], which distinguishes the close approximants [ϕ,θ,x,β,ð,ɣ] from the open ones [ϕ̞,θ̞,x̞,β̞,ð̞,ɰ]. The voiced velar reduces all the way to an open approximant; however, this is blocked in the voiceless approximants at this level, by a fortition constraint requiring the voiceless approximants to be [+close]:

(8-30) Weak position, level B (ϕ,θ,x=74, ϕ̞,θ̞,x̞=65)

	*[-voi,-close]	LAZY74
☞ ϕ,θ,x		*
ϕ̞,θ̞,x̞	*!	
ɣ		*!
☞ ɰ		

3.2.2. LEVEL C. At this level, the voiceless approximants likewise reduce to open variants, /b/ and /d/ spirantize, and the nasals /m,n/ spirantize to nasalized approximants (and /k/ debuccalizes in the younger sociolect).

Reduction to [-close] follows from the ranking:

(8-31) Weak position, level C (ϕ,θ,x=77, ϕ̞,θ̞,x̞=72

	LAZY77	*[-voi,-close]	LAZY74
ϕ,θ,x	*!		*
☞ ϕ̞,θ̞,x̞		*	

(For the idiolects in which the close approximants are somewhat more prevalent (let us say they occur at level C as well as B, and reduce to open approximants at level D), the ranking would be LAZY81 » *[-voi, -close] » LAZY77.)

Spirantization of /b,d/ at this level follows from the ranking:

(8-32) Weak position, level C (b,d,g=83,β̞,ð̞,ɰ=80)

		LAZY83	PRES(crisp rel)	LAZY79
	b,d - b,d	*!		*
☞	b,d - β̞,ð̞		*	*

The outcome is a fricative rather than an approximant, due to a distantial faithfulness constraint, PRES(crisp rel)∨PRES(son):

(8-33) Weak position, level C (β̞,ð̞,ɰ=80, β̞,ð̞,ɰ=61)

		PRES(crisp rel)∨PRES(son)	LAZY80
☞	b,d - β̞,ð̞		*
	b,d - β̞,ð̞	*	

Spirantization of the nasals follows from the ranking:

(8-34) Weak position, level C (m,n=83, β̞̃,ɹ̃=61)

		LAZY83	*[+nas,+cont]	LAZY79
	m,n - m,n	*!		*
☞	m,n - β̞̃,ɹ̃		*	

The ranking of *[+nas,+cont] above LAZY79 ensures that the spirantization of the nasals is blocked at earlier levels.

Debuccalization of /k/ in younger idiolects follows from the ranking:

(8-35) Weak position, level C (x̞=72, h=66)

		PRES(cor,lab)	LAZY72	PRES(dors)
	k - x̞		*!	
☞	k - h			*
☞	b,d - β̞,ð̞		*	
	b,d - h	*!		

3.2.3. LEVEL D. At this level, the voiced stops spirantize to open approximants rather than fricatives.

Reduction to open approximants follows from the ranking:

(8-36) Weak position, level D (β̞,ð̞,ɣ̞=85, β̞,ð̞,ɰ=64)

		LAZY85	PRES(crisp rel)∨PRES(son)	LAZY82
	b,d - β̞,ð̞	*!		*
☞	b,d - β̞,ð̞		*	

3.2.4. LEVEL E. At this level, the strident fricatives reduce to nonstrident approximants, and the liquids reduce to approximants.

Reduction to nonstridents follows from the ranking:

(8-37) Weak position, level E (s,ʃ,f=111, ʒ,z,v=109, s̞,ʃ̞,f̞=79, ʒ̞,z̞,ʋ=67)

		LAZY109	PRES(strid)	LAZY105
	tʃ,dʒ,s,ʃ,f,v - ʒ,s,ʃ,f,v	*!		*
☞	tʃ,dʒ,s,ʃ,f,v - ʒ̞,s̞,ʃ̞,f̞,ʋ		*	

The ranking of PRES(strid) above LAZY105 ensures that stridency is maintained at earlier levels.

Reduction of the liquids follows from the rankings:

(8-38) Weak position, level E (l=90, l̞=67)

		LAZY90	*[+lat,+cont]	LAZY86
	l - l	*!		*
☞	l - l̞		*	

(8-39) Weak position, level E (r=109,ɹɹ=73, ɹ=67)

		PRES(long)	LAZY109	*[+rho,+long, -trill]	LAZY104
	rr - rr		*!		*
☞	rr - ɹɹ			*	
	r - ɹ	*!		*	

(8-40) Weak position, level E

		LAZY89	*[+rho,-long, -tap]	LAZY85
	ɾ - ɾ	*!		*
☞	ɾ - ɹ		*	

3.2.5. LEVEL F. At this level, /k/ debuccalizes in the older sociolect. This follows from the ranking:

(8-41) Weak position, level F (x̝=83, h=77)

		LAZY83	PRES(dors)	LAZY79
	k - x̝	*!		*
☞	k - h		*	

3.2.6. LEVEL G. No further reductions at this level.

3.2.7. LEVEL H

At this level, the voiceless labial debuccalizes:

(8-42) Weak position, level H (ɸ̝=91, h=84)

		LAZY91	PRES(lab)	LAZY87
	p - ɸ̝	*!		*
☞	p - h		*	

3.2.8. LEVEL I. At this level, /g/ debuccalizes to [ɦ] (in the older sociolect). This is entailed by the ranking LAZY81 » PRES(dors) » LAZY79:[119]

(8-43) Weak position, level I (ɰ̝=81, ɦ=74)

		LAZY81	PRES(dors)	LAZY79
	g - ɰ̝	*!		*
☞	g - ɦ		*	

Moreover, /t/ debuccalizes:

[119]The ranking LAZY83 » PRES(dors) » LAZY79 has already been established in the analysis of /k/ debuccalization.

(8-44) Weak position, level I (θ̪=96, h=89)

		LAZY96	PRES(cor)	LAZY91
	t - θ̪	*!		*
☞	t - h		*	

On the other hand, we must prevent debuccalization of the voiceless approximants derived from stridents [ʃ̪,s̪,f̪], despite the violability, at this level, of all the PRES(place) constraints. This requires another disjunctively combined faithfulness constraint, PRES(strid)∨PRES(place):

(8-45) Weak position, level I (ʃ̪,s̪,f̪=96, h=89)

		PRES(strid)∨ PRES(place)	LAZY96	PRES(cor)	PRES(lab)
☞	tʃ,ʃ,s - ʃ̪,s̪		*	*	
	tʃ,ʃ,s - h	*!			
☞	f - f̪		*		*
	f - h	*!			

3.2.9. LEVEL J. At this level, all voiceless approximants become voiced:

(8-46) Weak position, level J (h=93, ʃ̪,s̪,f̪=101, β̪,ð̪,ɰ,ʒ,z̪,ʋ=85)

		LAZY93	PRES(voi)	LAZY89
	p,t,k - h	*!		*
☞	p,t,k - β̪,ð̪,ɰ		*	
	tʃ,ʃ,s,f - ʃ̪,s̪,f̪	*!		*
☞	tʃ,ʃ,s,f - ʒ,z̪,ʋ		*	

Further reduction to [ɦ] (debuccalization + voicing) is blocked by a distantial faithfulness constraint, PRES(-voi)∨PRES(place).

(8-47) Weak position, level J (β̞,ð̞,ɰ,ʒ,z̞,ʋ=85, ɦ=78)

		PRES(-voi)∨ PRES(place)	LAZY85
☞	p,t,k - β̞,ð̞,ɰ		*
	p,t,k - ɦ	*!	
☞	tʃ,ʃ,s,f - ʒ,z̞,ʋ		*
	tʃ,ʃ,s,f - ɦ	*!	

3.2.10. LEVEL K. At this final level, the singleton stops and nasals delete.

(8-48) Weak position, level K (β̞,ð̞,ɰ,β̞̃,ɹ̃=91, ɦ=81, Ø=0)

		LAZY91	PRES(-voi)∨ PRES(place)	PRES(place)	LAZY81	PRES(seg)
	p,t,k - β̞,ð̞,ɰ	*!			*	
	p,t,k - ɦ		*		*!	
☞	p,t,k - Ø		*			*
	b,d,g - β̞,ð̞,ɰ	*!			*	
	b,d,g - ɦ			*	*!	
☞	b,d,g - Ø			*		*
	m,n - β̞̃,ɹ̃	*!			*	
	m,n - ɦ			*	*!	
☞	m,n - Ø			*		*

For the remaining consonants, elision must be blocked. For the liquids, this result can be obtained by having PRES(lateral) and PRES(rhotic) in undominated position:

(8-49) Weak position, level K (l̞=91, ɹɹ=98, ɹ=91)

		PRES(lat)	PRES(rho)	LAZY98	LAZY91
☞	l - l̞				*
	l - Ø	*!			
☞	rr - ɹɹ			*	*
	rr - Ø		*!		
☞	ɾ - ɹ				*
	ɾ - Ø		*!		

For the stridents, the results are more complicated: /v/ does delete, whereas the other stridents do not. Acoustic stridency, i.e. noise intensity, however, is not binary: we can recognize degrees of noisiness (cf. Flemming 1995):

<---------------------------- +strid | -strid ----------------->
{s,ʃ} > {z,ʒ,f} > v > {ɸ,θ,x} > {β,ð,ɣ}
<--------- +more strid | -more strid ----------------------->

Figure 8-1. Stridency scale.

Thus, in addition to the standard feature [strident], which draws the line between [v] and [ɸ,θ,x], I posit a feature [more strident], which distinguishes the sibilants and [f] from [v] and the non-strident fricatives. Blocking of deletion of the [+more strident] fricatives can now be attributed to a distantial faithfulness constraint, Pres(+more strident)∨ Pres(place).

(8-50) Weak position, level K (ʒ,ẓ,ʋ=91)

		Pres(+more strid)∨Pres(place)	Lazy91
☞	tʃ,ʃ,ʒ,s,f - ʒ,ẓ,ʋ		*
	tʃ,ʃ,ʒ,s,f - Ø	*!	
	v - ʋ		*!
☞	v - Ø		

Finally, the geminates stops and affricates, which resist lenition at all earlier levels, succumb to spirantization (and, I assume, degemination, cf. sec. 1.2.3) at this level. Resistance to spirantization at earlier levels is attributable to high ranking of Pres(long):

(8-51) Weak position, level J (and earlier) (bb,dd,gg=144, β̞β̞,ð̞ð̞,ɰɰ=173, β̞,ð̞,ɰ=85)

		LAZY173	PRES(long)	LAZY144
☞	bb,dd,gg - bb,dd,gg			*
	bb,dd,gg - β̞β̞,ð̞ð̞,ɰɰ	*!		*
	bb,dd,gg - β̞,ð̞,ɰ		*!	

Spirantization at level K follows from subordination of PRES(long) to LAZY147:

(8-52) Weak position, level K (bb,dd,gg=151, pp,tt,kk=147, β̞,ð̞,ɰ=121, ɸ̞,θ̞,x̞=106)

		LAZY147	PRES(long)
	bb,dd,gg - bb,dd,gg	*!	
☞	bb,dd,gg - β̞,ð̞,ɰ		*
	pp,tt,kk - pp,tt,kk	*!	
☞	pp,tt,kk - ɸ̞,θ̞,x̞		*

Further reduction of the geminates is blocked by an (undominated) disjunctively combined faithfulness constraint:

(8-53) Weak position, level K (β̞,ð̞,ɰ=121, β̞,ð̞,ɰ=91)

		PRES(long) ∨PRES(son)	LAZY121
☞	bb,dd,gg - β̞,ð̞,ɰ		*
	bb,dd,gg - β̞,ð̞,ɰ	*!	
	bb,dd,gg - Ø	*!	

Note that the ranking, LAZY147 » PRES(long) » LAZY144, established above entails that the geminate affricates must spirantize slightly earlier than the stops, namely at level I, since their effort cost is higher. While G&S's description does not specifically note that the geminate affricates spirantize more readily than the stops, this result is not inconsistent with their description, and I will assume that it is correct.

3.2.11. SUMMARY: WEAK POSITION. In the foregoing analysis, the following rankings of effort minimization, faithfulness, and fortition constraints yields the pattern above of rate- and register-sensitive reduction in weak position:

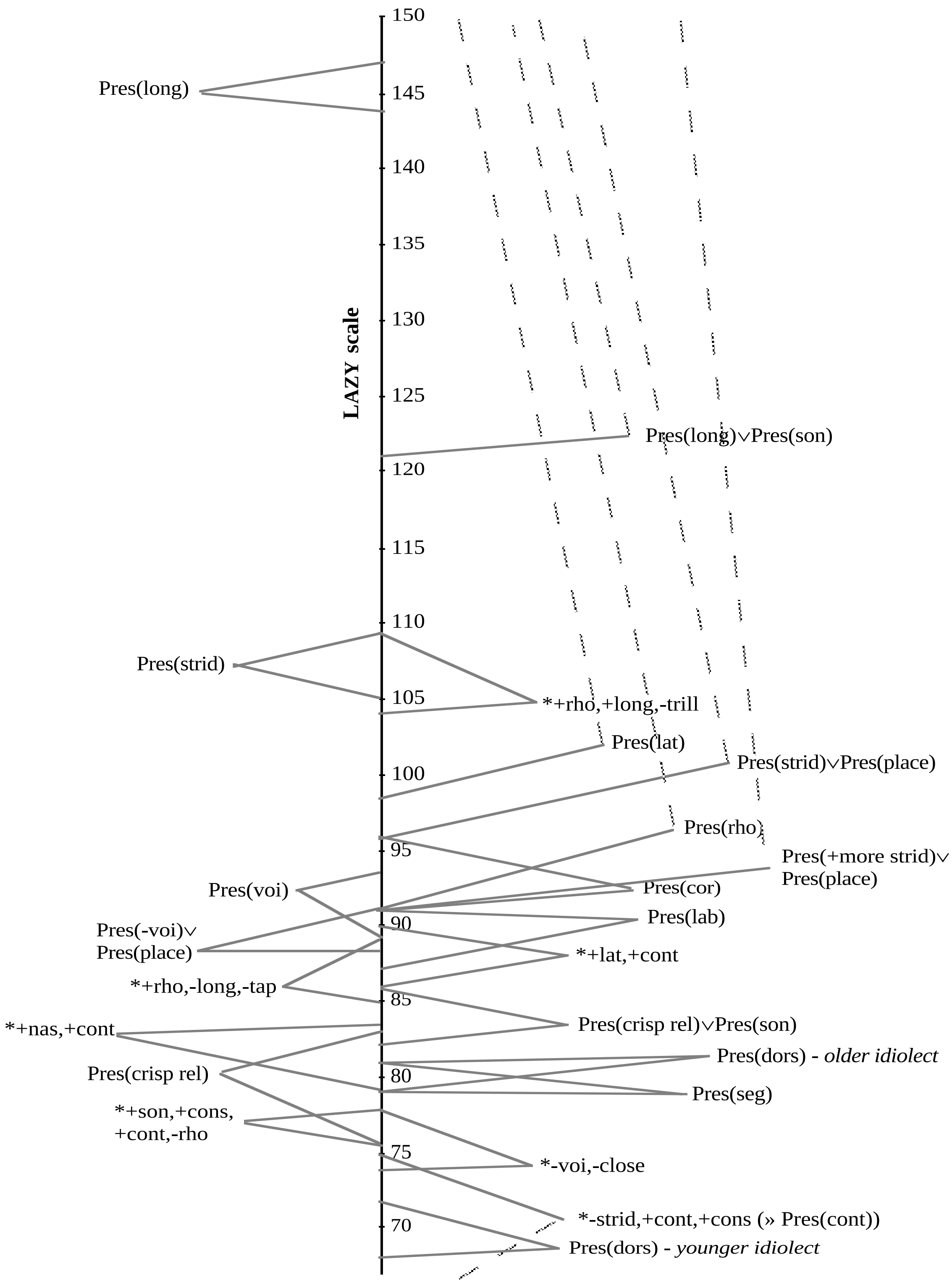


Figure 8-2. Crucial ranking ranges of lenition-blocking constraints, relative to LAZY, for Florentine lenition variation.

In this ranking diagram, the diagonal lines connecting the faithfulness and fortition constraints to the LAZY scale indicate the range within the LAZY hierarchy that a particular constraint occupies. For example, *[-voi,-close] is crucially ranked above LAZY_{74} and below LAZY_{78}; PRES(long)∨PRES(son) is ranked above LAZY_{121}, but is not crucially dominated by any other constraint.

3.3. LENITION OUTSIDE OF WEAK POSITION

The behavior of consonants in strong position follows from the same set of constraint rankings. Since the effort cost of consonants in strong position is lower than in weak position, lenition processes which apply at level A in weak position apply only at later levels in strong position; and lenition processes which apply later levels in weak position do not apply at all in strong position.

Let us assume that in strong position, voiceless stops spirantize at level F (the mid-point in our rate/register scale): that is, spirantization is blocked at level E and earlier. This means that *[-strid,+cont,+cons] dominates LAZY_{73}, the effort cost of the voiceless stops in strong position at level E (we have already established the ranking LAZY_{75} » *[-strid,+cont,+cons]).

(8-54)

		LAZY_{75}	*-strid, +cont,+cons	LAZY_{73}
☞	p,t,k *(level E)*			*
	ɸ,θ̠,x̠ *(level E)*		*!	
	p,t,k *(level F)*	*!		*
☞	ɸ,θ̠,x̠ *(level F)*		*	

This ranking entails that the affricates, which are more effortful than the stops, must spirantize at a somewhat earlier level, namely level B for /dʒ/ and level C for /tʃ/. Moreover, the voiced velar stop, which is less effortful than the voiceless stops, cannot spirantize until level J.

By similar reasoning, applying the ranking in Figure 8-2 to the chart of effort values for strong position, the following picture emerges:

Table 8-7. Lenition variation in strong position:

	A	B	C	D	E	F	G	H	I	J	K
/k/	k					x̞					x
/t/	t					θ̞					θ
/p/	p					ɸ̞					ɸ
/g/	g									ɰ	
/b,d/	b,d										β̞,ð̞
/dʒ/	dʒ		ʒ								
/tʃ/	tʃ	ʃ									
/v/	v										
/ʃ,s,f/	ʃ,s,f										
/m,n/	m,n										β̞̃,ɹ̃
/ɾ,rr,l/	ɾ,rr,l										
/kk,tt,pp/	kk,tt,pp										
/bb,dd,gg/	bb,dd,g g										
/ddʒ/	ddʒ										
/ttʃ/	ttʃ										

That is, rate/register-sensitive spirantization of the stops, affricates and nasals in strong position, as described in section 1.3.

I reiterate that G&S's binary distinction between strong and weak position is probably a simplification: a more careful analysis would presumably reveal gradient sensitivity to the openness of the flanking segments, as the effort-based approach predicts. Thus, in place of a set of effort values for strong position, and a set of higher values for weak position, the effort values of a given consonant should increase gradiently as the openness of the flanking segments increases. Consequently, the more open the flanking segments, the greater the propensity of the consonant to lenite.

4. CONCLUSION

We have seen that both stable lenition processes and rate/register-sensitive consonant reduction processes can be accounted for in a unified manner, in terms of ranking of LAZY thresholds versus faithfulness and fortition constraints. In contrast, within standard phonological frameworks, the Florentine lenition pattern cannot be given a unified characterization: for this set of processes does not correspond to the spreading, addition, or deletion of a particular feature or feature-geometric node; nor can the across-the-board shift towards hypoarticulation at faster rates and lower registers be captured within standard frameworks (see Chapter 6).

Chapter 9:

Conclusion: Stabilization of Lenition Patterns

In the previous chapters, I have presented a system of constraints which captures certain generalizations concerning lenition typology, and relates these to plausible assumptions regarding the relative articulatory cost of particular gestures, or their cost in particular contexts and conditions. To review:

- Lenition patterns receive a unified treatment in terms of conflict between effort minimization on the one hand, and auditorily based faithfulness and fortition constraints on the other. This is formalized as a series of effort thresholds (LAZY_{x} » $\text{LAZY}_{x\text{-}1}$ » etc.), interleaved with constraints on faithfulness to auditory features, and fortition constraints (presumably perceptually based), in an Optimality Theoretic constraint hierarchy.
- The non-stridency of synchronically spirantized stops falls out from consideration of the isometric tension required to produce the precise partial constriction needed for strident friction.
- The resistance of geminate stops to lenition likewise falls out from the isometric tension required to make a partial constriction for extended duration, and from the greater effort required to sustain voicing in geminate obstruents.
- The primacy of lenition in intervocalic and similar environments follows from the greater displacement, hence effort, required to achieve some constriction target when flanked by high sonority segments.
- The greater propensity for lenition in fast speech follows from the greater acceleration required to achieve a constriction target in a shorter amount of time.

- The greater propensity for lenition in casual speech is captured by augmenting the base effort cost of gestures in lower registers of speech. By globally lowering the ranking of faithfulness and fortition constraints relative to effort thresholds (by boosting effort cost), gestures cross higher effort thresholds, hence more lenition occurs.
- Finally, this approach was exemplified in more comprehensive analyses of lenition patterns in Tümpisa Shoshone and Florentine Italian.

1. The Stabilization Problem: Tigrinya Spirantization

While the constraint system proposed in previous chapters naturally handles the sort of variable lenition seen in Florentine, there remains the problem of capturing context-sensitive patterns which are insensitive to rate and register variation. The problem is illustrated by Tigrinya spirantization (Kenstowicz 1982):

(9-1)	a.	kətəma-xa	'town-2sg.m.'
		ʕarat-ka	'bed-2sg.m.'
		qətəl-ki	'kill-2sg.f. perfect'
		mɨrax-na	'calf-3sg.f.'
	b.	kəlbi	'dog'
		ʔa-xalɨb	'dogs'
		ʔɨti xalbi	'the dog'
	c.	qətəl-a	'kill-3pl.f. perfect'
		tɨ-χətl-i	'kill-2sg.f. imperfect'
	d.	fəkkəra	'boast'
		qətəl-na-kka	'we have killed you (masc.)'

As shown in (9-1) velar and uvular stops, except for tautomorphemic geminates (9-1d), spirantize in post-vocalic position, morpheme-internally, word-internally and across word boundaries.

The post-vocalic environment, as discussed in Chapter 6, is analyzed as the union of coda and intervocalic environments. First consider the coda (i.e. unreleased) context:

(9-2)

		PRES(cont w/ release)	LAZY	PRES(cont, unreleased)
	mɨrak-na - mɨrakna		**!	
☞	mɨrak-na - mɨraxna		*	*
☞	ʕarat-ka - ʕaratka		**	
	ʕarat-ka - ʕaratxa	*!	*	

There is no need to identify a particular threshold of effort as triggering this lenition: PRES(cont) may be ranked below the whole LAZY series, and PRES(cont/released) above it; therefore there is no problem in accounting with this approach for stable spirantization in coda position.

However, stable spirantization in intervocalic position, e.g. in [ʔa-xalɨb], poses a problem. The ranking in (9-2) cannot be correct: some degree of LAZY must dominate PRES(cont w/ release), otherwise spirantization in [ʔa-xalɨb] would be blocked.

(9-3)

		LAZY$_x$	PRES(cont/ released)	LAZY$_{x-1}$	PRES(cont)
	mɨrak-na - mɨrakna			*!	
☞	mɨrak-na - mɨraxna				*
☞	ʕarat-ka - ʕaratka			*	
	ʕarat-ka - ʕaratxa		*!		
	kətəma-ka - kətəmaka	*!		*	
☞	kətəma-ka - kətəmaxa		*		

Now that the lenition-blocking constraint, PRES(cont/released), falls between two effort thresholds, LAZY$_x$ and LAZY$_{x-1}$, it is predicted that, in slower or more formal speech, spirantization in [kətəmaxa] would be blocked.

(9-4)

		LAZY$_x$	PRES(cont w/ release)	LAZY$_{x-1}$
☞	kətəma-ka - kətəmaka *(slow/careful)*			*
	kətəma-ka - kətəmaxa *(slow/careful)*		*!	

But Michael Kenstowicz (p.c.) reports that his consultant exhibited no such variation in the spirantization pattern.

More generally, by tying lenition contexts such as intervocalic position to effort thresholds, we have achieved a natural explanation for why consonants are more likely to lenite in these environments than elsewhere, and why such lenition is often rate- and register-sensitive. But effort thresholds are a two-edged sword: for the actual effort required to achieve some constriction target is highly variable from token to token; therefore, if some effort threshold is truly the factor conditioning spirantization, the spirantization pattern should be variable, with higher probabilities when flanked by low vowels or in fast speech, and lower probabilities when adjacent to a lower-sonority segment. The problem, then, is how lenition conditioned by effort thresholds relates to a stable pattern of intervocalic lenition (or, in the case of Tigrinya, post-vocalic, by addition of coda position to the conditioning contexts).

One response to this problem is simply to question whether such stable patterns actually exist. While it is possible that many reports of ostensibly stable context-sensitive lenition merely reflect idealization of the data in the descriptive grammar, Kenstowicz is clear that this is not the case in Tigrinya. Thus this skepticism does not suffice as a general response to the problem.

2. Selection of Phonetic Variants

An idea put forward by Lindblom et al. (1995) (elaborating an earlier proposal of Ohala 1981) is that stabilization results from selection of a standard form of a lexical item from among the variant phonetic realizations of that item (where the variation is due to some system of phonetic constraints on the speech production system, similar to the general approach of this dissertation).[120] Let us illustrate this idea in terms of the Tigrinya root [mɨrax] ('calf'). In pre-pausal and coda (i.e. pre-consonantal) position (e.g.

[120]This notion of selection of variants can readily be cast in terms of the OT notion of Lexicon Optimization (Prince and Smolensky 1993, ch. 9): that is, the UR of a lexical item is selected so as to minimize violations (specfically of faithfulness constraints) in its mapping to PR; this principle ensures that, modulo alternations, the UR is identical to the PR. The further assumption of the Lindblomian idea is that, in the face of phonetic variation, Lexicon Optimization results in selection of a UR which is identical to the most common phonetic realization of the lexical item.

[mɨraxǁ] and [mɨrax-na] respectively) the velar consonant stably spirantizes, because it is unreleased (see (9-2)). Moreover, in intervocalic position (e.g. [mɨrax-V]), the stop has a high probability of spirantizing (see (9-3)). Thus, in any morphosyntactic context, 'calf' would, more likely than not, be realized with a velar fricative rather than a stop. Let us further assume that, due to this preponderance of realizations as [mɨrax] rather than [mɨrak], the hearer/learner selects [mɨrax] as the standard form for 'calf': that is, the form which is stored in the lexicon, and which serves as the input to future productions of this lexical item. As seen in (9-5), the realization of the velar consonant in words containing the 'calf' root is now stably [+cont], regardless of rate or register.

(9-5)

		LAZY$_x$	PRES(cont w/ release)	LAZY$_{x\text{-}1}$
	mɨrax-a - mɨraka *(fast/casual)*	*!	*	*
☞	mɨrax-a - mɨraxa *(fast/casual)*			
	mɨrax-a - mɨraka *(slow/careful)*		*!	*
☞	mɨrax-a - mɨraxa *(slow/careful)*			

Once this sound change occurs, the spirantization is "locked in" for this lexical item. The input, having changed to /mɨrax/, through the mechanism of selection of a standard form, no longer contains a [-cont] specification to be faithful to, and so [x] no longer reverts to a stop in careful or slower speech.

Patterns such as Tigrinya post-vocalic spirantization appear, however, to be too *systematic* to be handled through this mechanism of selection of phonetic variants. Consider the second person singular suffix [-ka]. The [-xa] allomorph only occurs when this suffix attaches to vowel-final stems. Given Tigrinya's abundance of consonant-final roots (including presumably common words such as [ʕarat] 'bed' and [qətəl] 'kill'), it is implausible to suppose that the [-xa] realization has a much broader distribution than [-ka]. Thus, spirantization in this suffix, in forms such as [kətəma-xa], is incorrectly predicted to be variable, as we have already seen in (9-4). One might plausibly abandon the assumption of a purely morphemic lexicon, assuming instead that inflected forms such as [kətəmaxa] are themselves listed. Now we are no longer considering the variants of all forms containing the [-ka]/[-xa] suffix, but only the variant realizations of [kətəmaxa] itself. The result, of course, is predominant presence of [x], thus yielding a UR /kətəmaxa/ by this selection mechanism, hence stable realization of the spirant.

However, this move does not solve the systematicity problem in general. For spirantization applies stably across word boundaries as well, e.g. in [ʔɨti xalbi] ('the dog'). One might assume that common phrases such as this determiner + noun collocation are stored as well, but we then predict a distinction in spirantization behavior depending on the token frequency of the phrase: in high frequency phrases with a post-vocalic velar consonant, spirantization should be stable, but variable in lower frequency phrases. Again, Kenstowicz observes no such variation. More generally, the stabilization effect of this selection-of-variants operates item by item, and so any phonetic or pragmatic factors tending to inhibit or promote spirantization in particular lexical items should result in deviations from the pattern of stable post-vocalic spirantization. For example, a word used predominantly in formal discourse settings, containing a velar consonant flanked by high vowels, would presumably have a preponderance of surface tokens in which spirantization is blocked (due to the inhibiting effect of the high vowels combined with high register), eventually resulting in a UR with /k/. This /k/ is then predicted to spirantize only in fast/careless speech, due to the direct effect of the phonetic constraint system (specifically LAZY). Conversely, a word used predominantly in informal settings, containing a velar in the context /r__a, should have a preponderance of surface tokens in which spirantization applies (due to the promoting effects of the following low vowel combined with low speech register), yielding a UR with /x/. Once Lexicon Optimization works this change, /x/ is predicted to spirantize stably, contrary to the observed pattern of stable spirantization only in postvocalic position. In sum, this mechanism gives us stabilization of varying phonetic patterns on an item-by-item basis. It seems likely that patterns of phonetic variation do in some cases harden into just such an item-by-item sound change. For example, American English flapping is optional in low frequency words (compare ['fɔɹɾi] ('forty') vs. ['dɔəɹɾi] ~ ['dɔəɹti] ('Daugherty,' an Irish-American surname); and flapping strikes me as quite unlikely in the case of completely nonce forms, e.g. ephlorty [ə'flɔɹti]. See also Fidelholtz 1975 on the sporadic instantiation of vowel reduction patterns in the English lexicon. Nevertheless, stable, systematic patterns such as Tigrinya post-vocalic spirantization also are attested; therefore, some additional mechanism is needed to account for them.

3. Effort under Canonical Conditions

To account for such stable, systematic lenition patterns, I instead propose a further series of effort threshold constraints, as follows:

(9-6) LAZY(STABLE)$_n$: Do not implement an articulatory gesture if the effort cost, *under canonical conditions,* of the syllable containing that gesture ≥ *n.*

That is, the LAZY(STABLE) constraints refer to the effort required, not in the actual articulatory realization of a lexical item, but that item's realization under canonical conditions, which we may understand to include moderate speech rate, normal loudness, upright orientation of the head, absence of nasal congestion, absence of chewing gum, and more generally, the conditions normally observed in speech, for all performance conditions which might affect the effort cost. Moreover, since LAZY(STABLE) refers to base effort cost rather than register-adjusted effort cost (see Chapter 6), it is not sensitive to register variation.

The variable or stable status of a lenition pattern conditioned by some effort-based context thus turns upon the ranking of the (rate- and register-sensitive) LAZY constraints versus the LAZY(STABLE) constraints. With an active LAZY constraint, we obtain a variable pattern, with a high probability of lenition in intervocalic position in a normal speech style, but blocking of lenition in slow and careful speech, as shown in (9-7a); whereas with an active LAZY(STABLE) constraint, the lenition pattern is insulated from rate, and register variation (as well as variation in other sorts of pragmatic conditions) (9-7b).

(9-7)
a.

	LAZY$_x$	PRES(cont w/ release)	LAZY (STABLE)$_x$
☞ kətəma-ka - kətəmaka *(slow/careful)*			*
kətəma-ka - kətəmaxa *(slow/careful)*		*!	
kətəma-ka - kətəmaka *(moderate/natural)*	*!		*
☞ kətəma-ka - kətəmaxa *(moderate/natural)*		*	

b.

	LAZY (STABLE)$_x$	PRES(cont w/ release)	LAZY$_x$
kətəma-ka - kətəmaka *(slow/careful)*	*!		
☞ kətəma-ka - kətəmaxa *(slow/careful)*		*	
kətəma-ka - kətəmaka *(moderate/natural)*	*!		*
☞ kətəma-ka - kətəmaxa *(moderate/natural)*		*	

4. REMAINING QUESTIONS

Although the LAZY(STABLE) solution adopted in the previous section affords a way of accounting for stable lenition patterns, it is problematic in certain respects. First, it is conceptually unsatisfying to propose that the speech production system is concerned with minimizing effort required for production under canonical speech conditions. One might reasonably propose that the system computes effort solely with respect to canonical speech conditions, as a means of simplifying the effort computation; but why should it do so, when it must compute the effort required for production in actual tokens in any event?

Second, the LAZY(STABLE) proposal offers no account of how a language might diachronically progress from a variable lenition pattern to a stable pattern: the "phonologization" result is obtained by brute-force ranking of LAZY(STABLE) constraints above the LAZY constraints. It is equally possible under this proposal to progress from a stable to a variable pattern, by promotion of the LAZY constraints relative to the LAZY(STABLE) constraints, though "phoneticization" of a phonological process is, as far as I am aware, a completely unattested diachronic development.

A potentially more explanatory, though as yet incompletely developed, approach to the stabilization problem involves the idea of pattern generalization through analogy to existing lexical items. That is, the system of phonetic constraints results in a pattern of variable lenition in speech production, which comes to be reflected in particular lexical items through the selection-of-variants mechanism, as in section 2, above; but once the lexicon comes to contain a substantial number of items reflecting some fairly consistent lenition pattern (e.g. post-vocalic spirantization of singleton dorsal stops, as in Tigrinya), the pattern is then systematically extended to subsequent outputs of the system, by analogy to the lexical items which instantiate the pattern. This notion of analogical extension thus affords a way of handling *systematic* patterns (unlike the bare selection-of-variants proposal); and since the analogical extension mechanism is independent of the effort considerations which originally induced the pattern, the pattern is thus stable across rate and register. The remaining program, which I leave to future research, is to devise, and explore the implications of, an explicit constraint (interacting with the sort of phonetic constraint system proposed in previous chapters of this dissertation) which expresses the notion that outputs of the phonology are well-formed to the extent that they conform to phonological patterns strongly instantiated by similar lexical items.

APPENDIX:

Survey of Lenition Patterns

Language[121]	Reference	Summary of phonological pattern(s)
Afar	Bliese 1981	word-final degemination
Amele	Roberts 1987	p ~ f non-initially
American English	Kahn 1976	Flapping of /t,d/ in "ambisyllabic" (i.e. intervocalic post-stress) position)
Ancient Greek	Bubenik 1983	voiceless aspirated stops - voiceless fricatives (context-free)
Ancient Greek	Hayes 1986	"deaspiration" of the first half of a geminate stop
Ancient Greek	Sommerstein 1973	s > h /__V
Ancient Greek	Sommerstein 1973	t -s /V__i
Ancient Greek	Sommerstein 1973	w,j,h -Ø /V__V
Andalusian Spanish	Romero 1996	non-initial spirantization of voiced stops except after a homorganic nasal or lateral; but greater reduction following a low vowel, and in faster speech
Ao	Gurubasave-Gowda 1967	stops voice /V__V or after a voiced C
Apalai	Koehn & Koehn 1986	k -g /__#
Apatani	Abraham 1985	b -β /V__V+back
Arbore	Harris 1990	debuccalization of coda ejectives
Assamese	Goswami 1966	labial fricatives restricted to wd final position; in some dialects, voiced and voiceless aspirates spirantize to voiced and voiceless fricatives, respectively, except for d^h and g^h
Ayt Ndhir Tamazight Berber	Saib 1972	geminate stops occlusivize (reanalysis of diachronic singleton spirantization), non-geminate pharyngealized stops become voiced
Babine	Story 1984	s,x - z,ɣ stem-finally
Badimaya	Dunn 1988	stop voicing, except finally; in certain words, initial stops have a greater tendency to be voiceless, greater tendency of velar stop to devoice. d,d^j -ð,ʒ /V__V
Balti	Rangan 1975	ð only occurs medially; ɣ occurs initially and medially; f nonfinally

[121] Languages from the lenition survey of Lavoie 1996 appear in italics.

Language	Reference	Summary of phonological pattern(s)
Bashkir	Poppe 1964	b -β /V__V.
Basque	Hualde 1993	k -ɣ word-finally, spirantize voiced fricatives /V__V
Blackfoot	Frantz 1971, Proulx 1989	w,j -Ø syl-initially or after syl-initial C, delete h /V__V, pre-C, tʃ - s before obstruents, x > ss, tk > ssk, fk > ssk (context unclear)
Bontoc	Reid 1971	spirantize and devoice voiced stops syllable-initially
British English	Milroy, Milroy & Hartley 1994	t -ʔ /V__V.and sometimes /V__l
British English	Milroy, Milroy &Hartley 1994	stops delete in final position
Bulgarian	Scatton 1983	VN -nasalized V(N) /__+cont, -son voicing assim. in medial clusters, devoicing at end of phon. word optional degem. in loanwords across morph. boundary c,x - ʒ,ɣ/__ -son,+voi
Burmese	Maran 1971	VN -nasalized V, p,t,k - ʔ/__# stop voicing (contexts unclear)
Canadian English	de Wolf & Hasebe-Ludt 1987	t -ɾ /V__V in post stress syllable (opt.)
Canela-kraho	Popjes & Popjes 1986	l - ɾ medially, stops - voiced in medial position, j, x -h initially
Car Nicobarese	Das 1977	r -d / n,l __
Cardiff English	Collins & Mees 1990	tendency for voiced stops to spirantize medially, to devoice word finally (voicing contrast shifted to length of preceding V
Carrier	Story 1984	ɣ,g,ɢ,g^w > j,j,ɣ,w word-finally
Castilian Spanish	Romero 1992	spirantized segment is most likely to be an approximant; after a fricative, it often resembles a fricative; after a stop, it can appear as a stop.
Catalan	Hualde 1992	vd stops - approximants in same lenition environments as in Spanish
Chitwan Tharu	Leal 1972	elision of intervocalic h; p^h in free variation w/ spirant b spirantizes when flanked by non-high vowels ɖ -ɾ intervocalically, wd-finally, same for d^h; dz -z intervocalically

Language	Reference	Summary of phonological pattern(s)
Chuckchee	Kenstowicz 1989	k -ɰ in coda, stop -nasal / __ nasal, regressive place assimilation in nasal clusters, opaque interactions of these processes
Cockney English	Adresen 1968	t -ʔ /V__V and after n,m,l
Cuna	Keating et al. 1983	Voicing in medial position, blocked in geminates
Dahalo	Tosco 1991	b -β, d - ð̆V__V
Danish	Bauer et al. 1980	p,t,k -b,d,g; b,d,g - β̞,ð̞,ɰ medially
Djabugay	Patz 1991	d -ɾ / stressed V__V (tendency stronger w/ flanking low vowels) r -ɾ in free variation
Djapu Yolngu	Morphy 1979	laminal stops (dental and palatoalveolar) -ð̞,j; k,p -w /{V,liquid}__ (w/ lexical exceptions)
Düzce Shapsug West Circassian	Smeets 1984	in clusters a voiceless final member has relatively weak aspiration, wheras a nonfinal member is pronounced with less articulatory force than in onset position. The nonfinal member usually lacks aspiration or glottal release of its own. Voiced initial consonants have partially devoiced variants
Efik	Dunstan 1969	b -β, d - ɾ, g- ɣ in noninitial pre-V position
Egyptian Arabic	Harrel 1957	all onset consonants are fortis, coda lenis
English (London, Leeds and Fife dialects)	Harris 1990	t -ʔ word-finally
Estonian	Harms 1962	non-labiodental obstruents - lenis /V__V, /V:__#, /R__, /__RV (R = sonorant)
Faroese	Lockwood 1977	stops -voiced / V__V or / __liquid stops -breathy voiced / -voi__liquid (in some dialects)
Farsi	Samareh 1977	-voi -aspirated /#__ and in stressed syllable. +voi -partially voiced in unstressed initial and final position. q -ɣ in free variation. All obstruents devoice /__#
Finnish	Sulkala & Karjalainen 1992	r - ɾ /V__V̧ delete h, k word-finally

Language	Reference	Summary of phonological pattern(s)
Florentine Italian	Giannelli & Savoia 1979	Spirantization of singleton -voi stops, g, and affricatives, in /V__({liquid, glide})V position (obligatory). Increasing reduction of all consonants, from fricatives to Ø, and lenition outside of /V__({liquid, glide})V position, in faster/more casual speech. Geminates spirantize only in very fast speech.
Gallo-Romance	Bourciez & Bourciez 1967	voiceless stops - voiced, voiced stops- fricatives / V__
Gbeya	Samarin 1966	ʔ -Ø /V__V.
Georgian	Aronson 1989	q' ~ q$^{x'}$ ~ χ v -w /__V and in coda
Gitksan	Hoard 1978	p -b /__V
Gojri	Sharma 1979	geminate stops - fortis singletons except in onset of stressed syllable; fortis stops may also be allophones of lenis (voiced) stops in initial or final position.
Gondi	Tyler 1975	k,r,c - h/V__V
Gooniyandi	McGregor 1990	p -b /V__V. (unclear on behavior of t)
Gosiute Shoshoni	McLaughlin 1989	c - z or ʒ after front V, voiceless stops spirantize to voiced fricatives /V__V
Gothic (Germanic)	Bennett 1980	spirantization of voiced stops /V__V
Guayabero	Keels 1985	d -θ word-finally
Guerzé	Casthelain 1952	ɣ -Ø /V__V.
Gujarati	Cardona 1965	b^{h},d^{h},ɖ^{h},g^{h} - β,ð,ɾ,ɣ/V__V ɖ -ɾ /__# (both processes optional)
Haitian Creole	Tinelli 1981	delete final glides (ij -),i final ʒ -j
Halabi	Singh 1977	retroflex stops -ɾ except initially and following homorganic nasal
Hausa	Klingenheben 1928	b,d,g - w,r,w in coda
Hawaiian	Elbert & Pukui 1979	delete h /V__V.

Language	Reference	Summary of phonological pattern(s)
Huallaga Quechua	Weber 1983	k sometimes has strong vocalic release wd.-finally; q -g ~ ɣ /V__V x /__ -voi obs Ø /__ # g~ɣ~x~: / in certain morphemes when followed by some other suffix. b -β as in Spanish no /d/ except in spanish loans, some borrowed as d, others as ɾ. cluster simplifications: kk -k kq -kg llq - ɬg nq -ŋg qq -g ~ ɣ (w/ compensatory lengthening of preceding V) qm -:m~g^{e}m~x^{e}m qn -g^{e}n~x^{e}n qr -g^{e}r qw - gw tq - t^{e}g
Icelandic	Thráinsson 1979	voiceless geminate stops reduce to h + stop
Kabardian	Colarusso 1988	spirantiz ejective affricates (context unclear)
Kabylie Berber	Chaker 1983	all non-geminates spirantize except after homorganic nasal or lateral
Kagate	Hoehlig 1978	b optionally spirantizes g -'lenis' /V__V (unclear what is meant by 'lenis' here) r -ɾ /V__V
Kaliai-Kove (Kandoka-Lusi)	Counts 1969	b,g -β,ɣ in free variation
Kanakuru	Newman 1974	p,t,k -w/p’,ɾ,ɣ intervocalically (w/ lexical exceptions)
Kannada	Chisum 1975	k -g /V__V, delete glide /V__V, word-initial p > h
Kanuri	Lukas 1967	b -v, g - ɣ/V__V
Karao	Brainerd 1994	p,t,d,tʃ,dʒ,g^{w},q -ɸ,θ,l,r,j,w,χ /V__V
Kashmiri	Kachru 1969	ʈ,ɖ -ɾ /V__V, /__#
Kirghiz	Hebert & Poppe 1963	β -w /V__V

Language	Reference	Summary of phonological pattern(s)
Korean	Martin 1992	h -Ø /+voi__+voi w - Ø /__V-high lenis stops -+voi /+voi__+voi
Kuna, Paya	Pike, Forster & Forster 1986	lenis stops -voiced fricatives /V__V
Kupia	Christmas & Christmas 1975	p -ɸ (opt.), ʈ -ɾ (opt.), ɖ -ɾ (oblig.) V __ V -stress Also flapping of d occurs /__C {-son or -cor}
Ladakhi	Koshal 1976	b,d,g ~ β,ð,ɣ /__V except in initial position; no medial or final retr. d; g -ɣ obligatory
Lama	Ourso & Ulrich 1990	p -w word-finally
Lamani	Trail 1970	ɖ -ɾ except initially or after a homorganic nasal or lateral
Laotian	Morev et al. 1979	stops are unreleased ('implosive') in final position glides - V in coda; aspirates and voiced stops prohibited except in onset.
Latin American Spanish	Lipski 1984	s - h/V__Vor word-finally in polysyllables
Latin American Spanish	Resnick 1975	past participle ado - aw
Lezgian	Haspelmath 1993	pV - b/__#
Limbu	van Driem 1987	unaspirated stop voicing /+nas__, /ʔ__, and /V__V. p^h,t^h,k^h -b^h,d^h,g^h /V__V (optional) t -ʔl /__# or /__{ʔ,h} b -w /V__V or after (non-homorganic) nasal (opt.)
Liverpool English	Wells 1982	t,d -θ̠,ð̠ word-finally and /V__V
Lomongo	Hulstaert 1961	b -Ø / V+___V(dialectal variation, sporadic lexical items showing b -)w
Lotha	Acharya 1983	t -d /n__ k -g /V__V (apparently with exceptions)
Lowland Murut	Halle & Clements 1983	b,d,g - β,ɾ,ɣ / -cons (#) ___
Lumasaaba	Brown 1972	p,t,k -β,ɾ,ɣ (context unclear)
Macushi	Abbott 1991	voice stops & /s/ post-nasally & /V__V.
Maidu	Shipley 1963	deglottalize ejective stops finally

Language	Reference	Summary of phonological pattern(s)
Malayalam	Mohanan 1986	stop -+voi /+son ___ V or +nas ___; ʈ -ɾ, other stops -voiced approximants /+son,-nas __ V (voicing, flapping and spirantization blocked in geminates; sensitive to rate or register of speech; word-internal exceptions to these processes)
Manobo	Reid 1971	voiced stops spirantize (environment unclear)
Maori	Bauer 1993	k -x (context-free)
Mataco-Noctenes	Claesson 1994	ʔ - h word-finally
Maxakalí	Gudschinsky et al. 1970	geminate stop - glide + singleton stop
Mbabaram	Dixon 1991	stop voicing / +nas__, optionally /{V,liquid}__V, with strongest tendency after [a], weakest tendency after liquid, blocked in initial and final position
Mexico City Spanish	Harris 1969	b,d,g -β,ð,ɣ except word or utterance initially and after a homorganic nasal or later r -ɾ except word-initially
Miami-Illinois	Costa 1991	voiceless fricatives - h /__-voi stop
Middle Chinese	Pulleyblank 1984	x -h (context unclear)
Middle Egyptian	Callender 1975	t -ʔ word-finally
Middle Italian (Central dialects)	Grammont 1939	voiceless velar stops - voiced in case the preceding vowel was unstressed or when either of the flanking vowels was low .
Middle Korean	Ramsey 1991	b -β /V__V.
Modern Irish	Ó Siadhail 1972	Dialectal variation regarding debuccalization or other sonorization of "palatalized" consonants; but the description seems to be driven more by spelling than phonetics; not clear what "slender" and "broad" mean phonetically, author refrains from explaining (except to say it's complicated).
Modern Welsh	Halle & Clements 1983, Oftedal 1985	p,t,k -b,d,g; b,d,g - v,ð,Ø; m,r̥,ɬ - v,r,l (morphological consonant mutations). p,t,k -f,θ,x ("spirantizing" mutation). p,t,k -m̥,n̥,ŋ̊; b,d,g - m,n,ŋ ("nasalizing" mutation)
Moghamo	Stallcup 1978	P,T,K - β,ɾ,ɣ (context unclear)
Mohawk	Halle & Clements 1983	stop -voi /__V
Mongolian	Poppe 1970	b -v, g - ɣ, q- x /V__V, ɣ -Ø /V__V
Navaho	Kari 1976	x -h non-initially, ɣ,j - Ø /V__V

Language	Reference	Summary of phonological pattern(s)
Nepali	Bandhu 1971, Acharya 1991	p^{h} -ɸ /V__ b -'lax' b /V__V b^{h} -β /__# $ɖ^{h}$ -$ɾ^{h}$ / except /#__ c^{h} -h /V__V k^{h} -x V__ h -Ø /V__C b^{h},d^{h},g^{h} -b,d,g except word-intially
New Zealand English	Holmes 1994	t -ɾ /V__V.
Newari	Nanda 1971	h -Ø /V__V(inducing breathiness on the following vowel)
Nez Perce	Aoki 1970	c - s before n or w, delete h post-C, k -x, q - before k, q, n, l, and in final position
Nkore-Kiga	Taylor 1985	b -ʋ /V__V.
Northern Corsican	Oftedal 1985	p,t,k -b,d,ɟ/g; b,d,ɟ/g -w,Ø,j,w; ts,tʃ -dz,dʒ (contexts unclear)
Numic	Ramer 1993	k -x, h or ɣ /V__V.
Nyawaygi	Dixon 1979	d -ɾ except after a homorganic nasal
Old English	Kabell & Laridsen 1984	voice interdental fric. after weakly stressed syllable
Old Turkic	Hitch 1989	p -b, k - g medially, g -ɣ /V__V, medial d -ð -j
Oscan, Umbrian	Buck 1904	t > d > Ø /__#, d > voiced approx. /V__V kt > ht, pt >ht
Páez	Gerdel 1985	x -h /V__V
Panyjima	Dench	ʈ -ɾ /V_V r -ɾ or ɹ /V_V no r /__# (neutralizes to ʈ?)
Pattani	Sarma 1982	p^{h}-p opt. medially and finally, particularly in unstressed syllables. Same for b-p. Same for other places of articulation. Flapping of retr. d in coda and intervocalically. h is weak intervocallically; pre-pausal stops, when they occur at all, areunreleased. Voiced stops in coda tend to be glottalized.
Pawnee	Parks 1976	delete h word-initially, delete r word-finally
Pengo	Burrow & Bhattacharya 1970	ʈ,ɽ > z /V__V

Language	Reference	Summary of phonological pattern(s)
Pennsylvania German	Kelz 1971	pf -b /V__V & /__#, pf -p before V, b -β, d - ð, g-ɣ /V__V, delete h /V__V
Périgourdin	Marshall 1984	p -b /V__V, b -β /V__V.
Pipil	Campbell 1985	w - h word-finally or /__C
Proto-Ainu	Vavin 1993	g > h(probably /V__V)delete h (/V__V ?)
Proto-Bantu	Halle & Clements 1983	b,d,g - β,l,ɣ except following a homorganic nasal
Proto-Germanic	Meillet 1970, Prokosch 1933	Stage 1: b^h,d^h,g^h -β,ð,ɣ Later: p^h,t^h,k^h -f,θ,x
Ptolemaic Greek	Teodorsson 1977	g > ɣ (context-free)
Punjabi	Gill 1969	All stops are lenis in medial position after centralized vowels. After peripheral vowels, geminates degeminate. p^h -ɸ in casual speech, particularly /V__V and /__#
Purki	Rangan 1979	dental d -ð /V_V alv. d - r/V_V no allophones of -ant sounds, even velars, but note no /G/
Quechua	Whitley 1979	k,q -x,χ in coda
Russian	Wade 1992	regressive voicing assim., optional degem. word-internally
Sa'idi Egyptian Arabic	Khalafallah 1969	All consonants have lenis alllophones
Saek	Gedney 1993	g ~ ɣ
Sanuma	Borgman 1986	p -b, t -ɾ, k - g ts dƶ /V__V, delete h /V__V.
Sawai	Whistler 1992	d -ɾ /V__V & /__C (i.e. post-vocalic)
Sekani	Hargus 1988	s,ɬ,j,x,ʍ -z,l,j,ɣ,w when prefixed (prob. = /V__V)
Senoufo	Mills 1984	voice stops in unstressed medial position, b -β, d - ɾ, g- ɣ in unstressed medial syllables
Serbo-Croatian	Partridge 1991	regressive voicing assimilation in obstruent clusters
Sestu Campidanese (Sardinian) Italian	Smith et al. 1990	voiceless stops - +cont +voi (context unclear) voiced stops - +cont (morphological mutation)

Language	Reference	Summary of phonological pattern(s)
Shina	Rajapurohit 1983	bdg sprantize V__V or r__V, retroflex d flaps
Somali	Armstrong 1964	vd stops - approximants/V__V., esp. after stress
Sotho	Doke 1957, Grammont 1939	d -l /__ [-high] V stops - fricatives /[-high] V__V
Southern Italian	Oftedal 1985	partial voicing of voiceless stops, non-initially
Southern Tati (Chali dialect)	Yar-Shater 1969	q -ɣ (optionally) obstruents devoiced before a /__-voi, initially, and when long; h -Ø except initially and in onset of stressed syllable, results in compensatory lengthening of preceding syllable
Southern Tati (Eshtehardi and Xiarah dialects)	Yar-Shater 1969	v -w after o, œ, and a
Southern Tati (Takestani dialect)	Yar-Shater 1969	g -ɣ /o__o (= ɒ__ɒ?)
Tahltan	Nater 1989	t -d, s - z, ts dz, tʃ -dʒ, ɬ - l,k -g, x - ɣ /V__V, word final
Taiwanese	Hsu 1995	P,T,K - β,ð,ɣ except word initially
Tamazight Berber	Abdel-Massih 1971	b -β (optionally) k,g -x,ɣ (Ayt Ayache dialect) processes are context free, blocked in geminates
Tamil	Annamalai 1975	voiceless stops spirantize to voiced fricatives between sonorant & V-initial word
Tamil	Keating et al. 1983	Voicing and spirantization in medial position, blocked in geminates
Tatar	Poppe 1963	b -β, d - ɾ, g- ɣ (context-free)
Tatar	Poppe 1963	g -ɣ / mostly V__V
Tauya	MacDonald 1990	k,k^{w} -ʔ,ʔ^{w} / non-initially; t -ɾ / V__
Thai	Noss 1964	sonorants shorten and weaken after a long V
Tiberian Hebrew	Malone 1993	nonemphatic stops - fricatives post-V or G, word-final degemination
Tiberian Hebrew	Malone1993	spirantization of non-geminate non-emphatics /-cons__ (counterbled by syncope)
Tigrinya	Schein 1981, Kenstowicz 1982	k,q -x,χ /V__, blocked in geminates
Toba Batak	Hayes 1986	p,t,k -ʔ in coda
Tojolabal	Furbee-Losee 1976	r,g -ɣ /V__V (opt.), w,j,h -Ø /__#

Language	Reference	Summary of phonological pattern(s)
Totonac, Misantla	MacKay 1984	q -χ word-finally, stops optionally voice /V__V
Tsou	Wright & Ladefoged 1994, Wright 1996	voiced (implosive) coronal stop lenites to a lateral flap just in case the preceding vowel is low
Tümpisa Shoshone	Dayley 1989	P,T,K - β/ɾ,ð,ɣ m,n,ŋ^{W} -w̃,ỹ/ɾ̃,[w̃] except utterance initially, and in full or partial geminates; ɾ/ð variation in coronals conditioned by backness of preceding V. ʔ,h -Ø (morphologized context, phonological factors unclear)
Turkana	Dimmendaal 1983	t -s before non-back vowel
Turkish	Bayraktaroglu 1992	v -w /noninitially
Turkish	Underhill 1976, Swift 1963	delete or "glide" velar fricative /V__V, v -β /V__V
Tzeltal	Kaufman 1971	spirantization of voiced stops /V__V, and word/morpheme-finally, after V
Uradhi	Dixon 1979	p,t,k -β,ð,ɣ (context unclear)
Urdu	Beg 1988	retroflex stops -ɽ except initially and following homorganic nasal
Urubu-Kaapor	Kakamasu 1986	stops optionally voice /V__V
Uyghur	Hahn 1991	b -β /V__ {V, liquid} k,g -x /V__C g -ɣ / in coda of wd-initial syllable and /V__V q -χ in coda ɢ -ʁ / #__ and V__V, ɢ -χ /__ -voi C; all processes blocked in full or partial geminates
Uzbek	Sjoberg 1963	q -x non-initially, p ~ ɸ especially medially b ~ w medially
Vietnamese	Emeneau 1951	g -ɣ except when preceding word ends in ŋ
Warndarang	Heath 1980	T^{j} - j K,P -w /V__V
West Greenlandic	Fortescue 1984	q -χ or ʁ /V__V.
West Tarangan	Nivens 1992	g -w, dʒ -j medially in an unstressed syllable, k -ʔ /V__V
Wiyot	Teeter 1964	deaspirate stops wd-finally

Language	Reference	Summary of phonological pattern(s)
Yakut	Krueger 1962	p -b /V__V (/p/ is rare) t -d /V__#V k -χ before and after non-high vowels; no voiced geminate obstruents
Yana	Sapir & Swadesh 1960	partly voice neutral stops before V, b -w /V__V
Yankunytjatjara	Goddard 1985	stops optionally voice /V__V
Yindjibarndi	Wordick 1982	p -w /V__V k -w /u __ k -Ø /V__V t̪ -ð̪ /V__V tʲ - j /V__V glides - Ø /V__V t,ʈ - ɾ r -j /a__i r - Ø V__V
Yonkalla	Berman 1990	c >s (context unclear)
Yuman	Wares 1968	v > approximant in a pre-stress syllable, c >s (context unclear)

REFERENCES

Abbott, M. (1991) Macushi, in D. Derbyshire and G. Pullum (eds.), Handbook of Amazonian Languages, v. 3, Mouton, Berlin.

Abdel-Massih, E. (1971) A Reference Grammar of Tamazight (Middle Atlas Berber), Center for Near Eastern and North African Studies, University of Michigan, Ann Arbor.

Abraham, P. (1985) Apatani Grammar, Mysore: Central Institute of Indian Language.

Acharya, J. (1991) A Descriptive Grammar of Nepali and an Analyzed Corpus, Washington, D.C.: Georgetown University Press.

Acharya, K. (1983) Lotha Grammar, Mysore, Central Institute of Indian Languages.

Adrésen, B. (1968) Pre-Glottalization in English Standard Pronunciation, Oslo: Norwegian Universities Press.

Al-Mozainy, H. (1981) Vowel alternations in a Bedouin Hijazi Arabic dialect: abstractness and stress. Doctoral Dissertation. University of Texas, Austin.

Anderson, S. (1969) West Scandinavian Vowel Systems and the Ordering of Phonological Rules, Doctoral Dissertation, MIT.

Anderson, S. (1981) Why Phonology Isn't Natural, LI 12:4, 493-539.

Annamalai, E. (1975) Phonology of Tamil Nouns, in H.F. Schiffman and C.M. Eastman (eds.), Dravidian Phonological Systems, Seattle: University of Washington and Institue for Comparative and Area Studies.

Aoki, H. (1970) Nez Perce Grammar, Berkeley: University of Califomia Press.

Archangeli, D. (1984) Underspecification in Yawelmani Phonology and Morphology, Doctoral Dissertation, MIT.

Archangeli, D. and D. Pulleyblank (1986) The Content and Structure of Phonological Representations, Ms., University of Arizona and University of British Columbia.

Archangeli, D. and D. Pulleyblank (1994) Grounded Phonology, Cambridge, MA: MIT Press.

Armstrong, L. (1964) The Phonetic Structure of Somali, Gregg Press, London.

Armstrong, L. (1967) The Phonetic and Tonal Structure of Kikuyu, London: Dawsons of Pall Mall.

Aronson, H. (1989) Georgian: a Reading Grammar, Columbus, OH, Slavica.

Bakovic, E. (1995) Strong Onsets and Spanish Fortition, MIT Working Papers in Linguistics, Proceedings of SCIL 6.

Bandhu, C. (1971) Nepali Segmental Phonology, Kirtipur: Summer Institute of Linguistics and Tribhuvan University.

Bansal, R. (1990) The Pronunciation of English in India, in S. Ramsaran (ed.), Studies in the Pronunciation of English, New York: Routledge.

Barry, M. (1985) A Palatographic study of Connected Speech Processes, Cambridge Papers in Phonetic and Experimental Linguistics, 4, 1-16.

Barry, M. (1991) Temporal Modelling of Gestures in Articulatory Assimilation, Proceedings of the XIIth International Congress of Phonetic Schiences, 4, 14-17.

Bat-El, Outi (1986) Extraction in Modern Hebrew Morphology, M.A. Thesis, UCLA.

Bauer, L. (1988) What is Lenition? Joumal of Linguistics 24, 381-392.

Bauer, L., J. Dienhart, H. Hartvigson, L. Jakobsen (1980) American English Pronunciation: Supplement, Comparison with Danish. Gyldendalsk Boghandel, Copenhagen.

Bauer, W. (1993) Maori, London: Routledge.

Bayraktaroglu, A. (1992) Colloquial Turkish, London, Routledge.

Beckman, J. (1997) Positional Faithfulness, Doctoral Dissertation, UMass-Amherst.

Beg, M. (1988) Urdu Grammar: History and Structure, New Delhi: Bahri Publications.

Bennett, W. (1980) An Introduction to the Gothic Language, New York: The Modern Language Association.

Benua, L. (1995) Identity Effects in Morphological Truncation, in Papers in Optimality Theory, J. Beckman, S. Urbanczyk & L. Walsh (eds.), 77-136, UMass-Amerst GLSA.

Benua, L. (1997) Transderivational Identity: Phonological Relations between Words, Doctoral Dissertation, U.Mass-Amherst.

Berman, H. (1990) An Outline of Kalapuya Historical Phonology, International Journal of American Linguistics 56, 27-559.

Bhatia, T. (1993) Punjabi, London: Routledge.

Bhattacharya, P. (1977) A Descriptive Analysis of the Boro Language, Gauhati, Gauhati University.

Bladon, A. (1986) Phonetics for Hearers, in G. McGregor (ed.) Language for Hearers, Oxford: Pergamon.

Bliese, L. (1981) A Generative Gramar of Afar, Norman, Oklahoma: SIL.

Bloomfield, L. (1956) Eastem Ojibwa: Grammatical Sketch, Texts and Word List, Ann Arbor: University of Michigan Press.

Boersma, P. (1997a) Elements of Functional Phonology, ms., Instituut voor Fonetische Wetenschappen.

Boersma, P. (1997b) Sound Change in Functional Phonology, ms., Instituut voor Fonetische Wetenschappen.

Boersma, P. (1997c) Inventories in Functional Phonology, ms., Instituut voor Fonetische Wetenschappen.

Boersma, P. (1997d) How We Learn Variation, Optionality and Probability, ms., Instituut voor Fonetische Wetenschappen.

Bonvillain, N. (1973) A Grammar of Akwesasne Mohawk, Ottawa: National Museum of Man.

Borden, G. & K. Harris (1980) Speech Science Primer: Physiology, Acoustics and Perception of Speech, Baltimore: Williams and Wilkins.

Borgman, D. (1986) Sanuma, in D. Derbyshire and G. Pullum (eds.), Handbook of Amazonian Languages, v. 3, Mouton, Berlin.

Bourciez, E. & J. Bourciez (1967) Phonétique Française: Etude Historique, Paris: Editions Klincksieck.

Brainerd, S. (1994) The phonology of Karao, Canberra: Pacific Linguistics, Australian National University.

Branstine, Z. (1991) Stop/Spirant Alternations in Spanish: On the Representation of Contrast, Studies in the Linguistic Sciences 21:1, 1-22.

Bright, W. (1957) The Karok Language, Berkeley: University of California Press.

Bright, W. (1968) A Luiseño Dictionary, Berkeley: University of California Press.

Browman, G. & L. Goldstein (1989) Articulatory Gestures as Phonological Units, Phonology 6: 201-252.

Browman, G. & L. Goldstein (1990) Tiers in Articulatory Phonology, with some Implications for Casual Speech, in J. Kingston and M.. Beckman (eds.) Papers in Laboratory Phonology I, Cambridge University Press.

Browman, G. & L. Goldstein (1992) Articulatory Phonology: an Overview, Phonetica, 49, 155-180.

Brown, G. (1972) Phonological Rules and Dialect Variation: A Study of the Phonology of the Lumasaaba. Cambridge: Cambridge University Press.

Bubeník, V. (1983) The Phonological Interpretation of Ancient Greek: A Pandialectal Analysis, Toronto: U. Toronto Press.

Buck, C. (1904) A Grammar of Oscan & Umbrian. Boston: Ginn & Co./The Atheneum Press.

Burrow, T. & S. Bhattacharya (1970) The Pengo Language, Oxford: Clarendon Press.

Burzio, L. (1997) Strength in Numbers, paper presented at Hopkins OT Workshop/University of Maryland Mayfest 1997, Baltimore.

Bwantsa-Kafungu, S. P. (1970) Esquisse Grammaticale de Lingala, Kinshasa: Publications de l'université Lovanium.

Bybee, J. (1995) Regular Morphology and the Lexicon, Language and Cognitive Science 10, 425-55.

Byrd, D. (1994) Articulatory Timing in English Consonant Sequences, Doctoral Dissertation, UCLA.

Byrd, D. & E. Salzman (forthcoming) Intragestural Dynamics of Multiple Phrasal Boundaries, Journal of Phonetics.

Callaghan, C. A. (1965) Lake Miwok Dictionary. Berkeley: University of California Press.

Callaghan, C. A. (1984) Plains Miwok Dictionary. Berkeley: University of California Press.

Callender, J . (1975) Middle Egyptian, Malibu: Undena.

Campbell, L. (1985) The Pipil Language of El Salvador. Berlin: Mouton.

Cardona, G. (1965) A Gujarati Reference Grammar, Philadelphia: University of Pennsylvania Press.

Carr, P. (1991) Weakening on Tyneside: Postlexical Neutralisation, in I. van Ostade & J. Frankis (eds.), Language Usage and Description. Amsterdam: Rodopi.

Casali, R. (1995) Labial Opacity and Roundness Harmony in Nawuri, Natural Language and Linguistic Theory 13, 949-663

Casali, R. (1996) Resolving Hiatus, Doctoral Dissertation, UCLA.

Casthelain, R. (1952) La Langue Guerzé. Dakar: Institut Français d'Afrique Noire.

Chaker, S. (1983) Un parler berbere d'Algerie (Kabylie). Aix en Provence: Publications Université de Provence.

Charney, J.0. (1993) A Grammar of Comanche. Lincoln, NE: University of Nebraska Press.

Chisum, G. (1975) Systematic Phonology of the Kannada Noun, in H. Schiffman and C. Eastman, eds., Dravidian Phonological Systems, University of Washington Press.

Cho, Y. (1990) Parameters of Consonantal Assimilation, Doctoral Dissertation, Stanford University.

Choi, J. (1992) Phonetic Underspecification and Target Interpolation: an Acoustic Study of Marshallese Vowel Allophony, Doctoral Dissertation, UCLA.

Chomsky, N. & M. Halle(1968) The Sound Pattern of English, New York, Harper and Row.

Christdas, P (1988) The phonology and morphology of Tamil, Doctoral Dissertation, Cornell University.

Christensen, J. & Christensen, S. (1992) Kisar Phonology. in D. Burquest & W. Laidig (eds.), Phonological Studies in Four Languages of Maluku. Arlington, TX: The University of Texas and the Summer Institute of Linguistics.

Christmas, R. & J. Christmas (1975) Kupia phonemic summary, Kathmandu: Summer Institue of Linguistics and Institute of Nepal and Asian Studies, Tribhuvan University.

Churma, D. On 'On Geminates,' ms., SUNY-Buffalo.

Churma, D.G. (1988) Consonant Gradation in Fula Suffixes: The Ugly Truth. Studies in African Linguistics 19, 35-74.

Claesson, K. (1994) A Phonological Outline of Mataco-Noctenes, International Journal of American Linguistics 60, 1-38.

Clark, J. & C. Yallop (1990) An Introduction to Phonetics & Phonology, Oxford: Basil Blackwell.

Clements, G. & S. Keyser (1983) CV Phonology: A Generative Theory of the Syllable, Cambridge, Mass.: MIT Press.

Clements, G. (1985) The Geometry of Phonological Features, Phonology Yearbook 2, 223-52.

Clements, G. (1986) Compensatory Lengthening and Consonant Gemination in Luganda, Phonology Year, in L. Wetzels and E. Sezer (eds.), Studies in Compensatory Lengthening, 37-77, Dordrecht: Foris.

Clements, G. (1988) Toward a Substantive Theory of Feature Specification, Proceedings of NELS 18, 79-93.

Clements, G. (1990) The Role of the Sonority Cycle in Core Syllabification, in J. Kingston and M. Beckman (eds.) Papers in Laboratory Phonology I, 283-333, Cambridge University Press.

Cohn, A. (1990) Phonetic and Phonological Rules of Nasalization, Doctoral Dissertation, UCLA.

Colarusso, J. (1988) The Northwest Caucasian Languages: A Phonological Survey, New York & London: Garland.

Cole, J. & C. Kisseberth (1994) An Optimal Domains Theory of Harmony, ms., University of Illinois, Urbana-Champaign.

Collins, B. & I. Mees (1990) The Phonetics of Cardiff English, in English in Wales: Diversity, Conflict and Change, Nikolas Coupland (ed.), 87-103, Clevedon: Multilingual Matters.

Connelly, M.J. (1991) The Phonological Interaction of Lenition and Secondary Articulation in the Cols Fhairrge Dialect of Irish. The Georgetown Journal of Languages & Lingusitics 2, 11-48.

Conteh, P., Cowper, E. & Rice, K. (1986) The Environment for Consonant Mutation in Mende. in G.J. Dimmendaal (ed.), Current Approaches to African Linguistics, vol. 3. Dordrecht: Foris Publications.

Costa, D. J. (1991) The Historical Phonology of Miami-Illinois Consonants. International Joumal of American Linguistics 57, 365-93.

Counts, D. (1969) A Grammar of Kaliai-Kove, Honolulu, U. Hawaii Press.

Crazzolara, J.P. (1960) A Study of the Logbara (Ma'di) Language. London: Intemational African Institute.

Cressey, W. W. (1980) Sobre la abstracción en la fonología generativa y ciertos fenómenos del castellano. in J.M. Guitart & J. Roy (eds.) La estructura fónica de la lengua castellana. Barcelona: editorial Anagrama.

Crosswhite, K. (1997) Avoidance of Homophony in Trigrad Bulgarian Vowel Reduction, Paper presented at Southwestern Workshop on Optimality Theory, University of California, Los Angeles.

Crumrine, L. S. (1961) The Phonology of Arizona Yaqui. Tucson: The University of Arizona Press.

Das, A. (1977) A Study of the Nicobarese Language, Calcutta, Anthropological Survey of India, Govt. of India.

Dayley, J. (1989) Tümpisa (Panamint) Shoshone Grammar, University of California Publications in Linguistics, vol. 115., UC Press.

Dell, F. & M. Elmedlaoui (1985) Syllabic Consonants and Syllabification in Imdlawn Tashlhiyt Berber, J. African Languages and Ling. 7, 105-130.

Dench, A. (1991) Panyjima, in R. Dixon and B. Blake (eds.) Handbook of Australian Languages, v. 4, 125-244, Oxford, Oxford University Press.

Derbyshire, D. (1985) Hixkaryana & Linguistic Typology. Arlington, TX: The Summer Institute of Linguistics.

Diagne, P. (1971) Grammaire de Wolof Moderne, Paris:: Présence Africaine.

Dimmendaal, G. (1983) The Turkana Language, Dordrecht: Foris.

Dimmendaal, G. J. (1983) The Turkana Language. Dordrecht: Foris.

Dixon, R. & B. Blake (1979) Handbook of Australian Languages, v. 3, Amsterdam: J. Benjamins.

Dixon, R. & B. Blake (1991) Handbook of Australian Languages, v. 4, Oxford, Oxford University Press.

Dixon, R. (1977) A Grammar of Yidiɲ, Cambridge: Cambridge University Press.

Dixon, R. (1991) Mbabaram, in Dixon & Blake (1991), 349-402.

Dixon, R.M.W. (1972) The Dyirbal Language of North Queensland, Cambridge: Cambridge University Press.

Dixon, R.M.W. (1977) A Grammar of Yidiny. Cambridge: Cambridge University Press.

Docherty, G. & A. Ladd (1992) Papers in Laboratory Phonology II: Gesture, Segment, Prosody, Cambridge University Press.

Doke, C. (1957) A Textbook of Southern Sotho Grammar, Longmans, Green and Co., London.

Donegan, P. & D. Stampe (1979) The Study of Natural Phonology, in D. Dinnsen (ed.) Current Approaches to Phonological Theory, 126-73, IU Press, Bloomington.

Driem, G. van (1987) A Grammar of Limbu, New York: Mouton.

Dunn, L. (1988) Badimaya, a Western Australian Language, in Papers in Australian Linguistics, Canberra: Pacific Linguistics, Australian National University.

Dunstan, E. (1969) Twelve Nigerian Languages, New York: Africana Publishing Corp.

Elbert, S. & M. Pukui (1979) Hawaiian Grammar, Honolulu: U. Hawaii Press.

Elmedlaoui, M. (1993) Gemination and Spirantization in Hebrew, Berber, and Tigrinya: a Fortis-Lenis Module Analysis, Linguistica Communicatio V, 1: 121-176.

Emeneau, M.B. (1951) Studies in Vietnamese (Annamese) Grammar, Berkeley: University of California Press.

Evans, N. (1995) Current Issues in the Phonology of Australian Languages, in Goldsmith (1995).

Flemming, E. (1995) Auditory Features in Phonology, Doctoral Dissertation, UCLA.

Flemming, E. (1997) Phonetic Optimization, Compromise in Speech Production, paper presented at Hopkins OT Workshop/University of Maryland Mayfest 1997, Baltimore.

Foley, J. (1977) Foundations of Theoretical Phonology, Cambridge University Press.

Foley, L.M. (1980) Phonological Variation in Westem Cherokee. New York & London: Garland.

Fortescue, M. (1984) West Greenlandic, Croom Helm, London.

Fowler, C. & E. Saltzman (1993) Coordination and Coarticulation in Speech Production, Language and Speech 36, 171-195.

Frantz, D. (1971) Toward a Generative Grammar of Blackfoot, Norman, OK: U. Oklahoma & SIL.

Frisch, S. (1996) Similarity and Frequency in Phonology, Doctoral Dissertation, Northwestern University.

Furbee-Losee, L. (1976) The Correct Language: Tojolabal. New York: Garland.

Gafos, A. (1997) On the Proper Characterization of 'Nonconcatenative' Languages, ms., Johns Hopkins University.

Gedney, W. (1993) The Saek Language, Ann Arbor, U. Mich. Press.

Gerdel, F. (1985), Páez: Pitch and Stress in the Phonological Word and Phrase, in R. Brend, ed., From Phonology to Discourse, SIL.

Giannelli, L. & L. Savoia (1979) Indebolimento Consonantico in Toscana, Revista Italiana di Diallettologia 2: 23-58.

Gill, H. (1969) A reference grammar of Punjabi , Patiala: Punjabi University.

Goddard, C. (1985) A Grammar of Yankunytjatjara. Alice Springs, Australia: Institute for Aboriginal Development.

Goldinger, S. (1997) Perception and Production in an Episodic Lexicon, in Johnson & Mullenix (1997).

Goldsmith, J. (1976) Autosegmental Phonology, Doctoral Dissertation, MIT.

Goldsmith, J. (1990) Autosegmental and Metrical Phonology, Oxford, Basil Blackwell.

Goldsmith, J. (1995) The Handbook of Phonological Theory, Oxford, Basil Blackwell.

Goodman, M. (1964) A Comparative Study of French Creole Dialects, the Hague: Mouton.

Gorbet, L.P. (1976) A Grammar of Diegueño Nominals. New York: Garland.

Gordon, M. & I. Maddieson (1996) The Phonetics of Ndumbea, UCLA WPP.

Gordon, M. (in progress) Syllable Weight: Phonetics, Phonology and Typology, Doctoral Dissertation, UCLA.

Goswami, G. (1966) An Introduction to Assamese Phonology, Poona: Deccan College.

Grammont, M. (1933) Traité de Phonetique, Delgrave, Paris.

Greenberg, J. (1948) The Tonal System of Proto-Bantu, Word 4:3, 196-208.

Gregores, E. & Suarez, J. (1967) A Description of Colloquial Guarani. The Hague: Mouton.

Grijzenhout, J. (1995) Irish Consonant Mutation and Phonological Theory, Doctoral Dissertation, Universiteit Utrecht, OTS Dissertation Series.

Gudschinsky, S., H. Popovich & F. Popovich (1970) Native Reaction and Phonetic Similarity in Maxakalí Phonology, Language 46:1, 77-88.

Guerssel, M. (1977) Constraints on Phonological Rules, Linguistic Analysis 3, 267-305.

Gurubasave-Gowda, K. (1975) Ao Grammar, Mysore: Central Institute of Indian Languages.

Hahn, R. (1991) Spoken Uyghur, Seattle: University of Washington Press.

Halle, M. & G. Clements (1983) Problem Book in Phonology , MIT Press

Halle, M. (1959) The Sound Pattern of Russian, the Hague: Mouton.

Halle, M. (1971) The Sound Pattern of Russian. The Hague: Mouton (first printing 1959)

Hargus, S. (1988) The Lexical Phonology of Sekani, Garland, New York.

Harms, R. (1962) Estonian Grammar, Bloomington, Indiana U. Uralic and Altaic series, v. 12.

Harrel, R. (1957) The phonology of Colloquial Egyptian Arabic, NY, Council of Learned Societies.

Harris, James (1969) Spanish Phonology, Cambridge Mass., MIT Press.

Harris, James (1983) Syllable Structure and Stress in Spanish: a Nonlinear Analysis, Cambridge Mass., MIT Press.

Harris, James (1984) La espirantización en castellano y su representación fonológica autosegmental, in Bartra et al. (eds.) Estudis Grammaticals, Universitat Autònoma de Barcelona, Bellaterra.

Harris, John (1990) Segmental Complexity and Phonological Government, Phonology 7:2, 255-300.

Haspelmath, M. (1993) A Grammar of Lezgian. Berlin: Mouton de Gruyter.

Hayes, B. & M. MacEachern (1996) Folk Verse Form in English, ms., UCLA.

Hayes, B. & T. Stivers (1997) A Phonetic Account of Post-Nasal Voicing, ms., UCLA.

Hayes, B. (1986) Inalterability in CV Phonology, Language 62:2, 321-352.

Hayes, B. (1989) Compensatory Lengthening in Moraic Phonology, LI 20, 253-306.

Hayes, B. (1997) Phonetically-Driven Phonology: The Role of Optimality Theory and Inductive Grounding, ms., UCLA.

Heath, J. (1980) Basic Materials in Warndarang: Grammar, Texts and Dictionary, Pacific Linguistics Series B, No. 72, Australian National University.

Hebert, R. & N. Poppe (1963) Kirghiz Manual, Bloomington, Indiana U. Press.

Hitch, D. (1989) Old Turkic b' and p', in K. Sagaster & H. Einer (eds.) Religious and Lay Symbolism in the Altaic World and Other Papers, Proceedings of the 27th Meeting of the Permanent International Altaist Conference.

Hoard, J. (1978) Obstruent Voicing in Gitksan: Some Implications for Distinctive Feature Theory, in E. Cook & J. Kaye (eds.), Linguistic Studies of Native Canada, Vancouver: UBC.

Hock, H. (1991) Principles of Historical Linguistics (2d ed.), Mouton de Gruyter, Berlin.

Hoehlig, M. (1978) Kagate Phonemic Summary, Kathmandu: Summer Institute of Linguistics and Institute of Nepal and Asian Studies, Tribhuvan University.

Hoff, B.J. (1968) The Carib Language. The Hague: Martinus Nijhoff.

Holmes, J. (1994) New Zealand Flappers: An Analysis of T Voicing in New Zealand English. English World Wide 15, 195-224.

Hooper, J. (1976) An Introduction to Natural Generative Phonology, New York: Academic Press.

House, A. (1957) Analog Studies of Nasal Consonants, J. Speech & Hearing Disorders 22, 190-204.

Hsu, C. (1995) A Phonetically-Based Optimality Theoretic Account of Consonant Reduction in Taiwanese, M.A. Thesis, UCLA.

Hualde, J. (1991) Basque Phonology, London: Routledge.

Hualde, J. (1992) Catalan, London: Rouledge.

Hualde, J. (1993) Topics in Souletin Phonology, in Generative Studies in Basque Linguistics, John Benjamins, Amsterdam.

Huffman, M. (1990) Implementation of Nasal: Timing and Articulatory Landmarks, Doctoral Dissertation, UCLA.

Hulstaert, G. (1961) Grammaire du Lomongo, Tervuren, Belg., Musee Royale de l'Afrique Centrale.

Hyman, L. (1975) Phonology: Theory & analysis. New York: Holt, Rinehart & Winston.

Hyman, L. (1985) A Theory of Phonological Weight, Dordrecht: Foris.

Inkelas, S. & Y. Cho (1991) Inalterability as Prespecification, Language 69:3, 529-574.

Inkelas, S. (1994) The Consequences of Optimization for Underspecification, Ms., Berkeley.

Inouye, S. (1995) Trills, Taps and Stops in Contrast and Variation, Doctoral Dissertation, UCLA.

Itô, J. (1986) Syllable Theory in Prosodic Phonology, Doctoral Dissertation, University of Massachusetts, Amherst.

Itô, J. (1989) A Prosodic Theory of Epenthesis, NLLT 7, 217-259.

Itô, J., A. Mester, and J. Padgett (1995), Licensing and Redundancy: Underspecification in Optimality Theory, LI.

Jacobs, H. & L. Wetzels (1988) Early French Lenition: a Formal Account of an Integrated Sound Change, in H. van der Hulst and N. Smith (eds.) Features, Segmental Structures and Harmony Processes, 105-129, Dordrecht, Foris.

Jacobs, H. (1994) Lenition in Optimality Theory, ms. Nijmegen University/Free University, Amsterdam.

Jaeger, J. (1978) Speech Aerodynamics and Phonological Universals, BLS 4, 311-329.

Janda, R. (1987) On the Motivation for an Evolutionary Typology of Sound Structural Rules, Doctoral Dissertation, UCLA.

Johnson, K. & J. Mullenix (1997) Talker Variability in Speech Processing, San Diego: Academic Press.

Johnson, K. (1997) Speech Perception without Speaker Normalization, in Johnson & Mullenix (1997).

Jong, K. de, M. Beckman & J. Edwards (1993) The interplay between prosodic structure and coarticulation, Language and Speech 35, 45-58.

Joseph, B. D. & I. Philippaki-Warburton (1987) Modern Greek. London: Croom Helm.

Jun, J. (1995) Perceptual and Articulatory Factors in Place Assimilation: an Optimality Theoretic Approach, Doctoral Dissertation, UCLA.

Ka, O. (1994) Wolof Phonology and Morphology, Lanham, Maryland: University Press of America.

Kabell, I. & H. Laridsen (1984) On the Evidence of the Earliest Danish Works on English Pronunciation. in I. Kabell, H. Lauridsen & A. Zettersten (eds.) Studies in Early Modem English Pronunciation. Copenhagen: University of Copenhagen.

Kachru, B. (1969) A Reference Grammar of Kashmiri, Urbana, U. Ill.

Kahn, D. (1976) Syllable-Based Generalizations in English Phonology, Doctoral Dissertation, MIT.

Kakamasu, J. (1986) Urubu-Kaapor, in D. Derbyshire and G. Pullum (eds.), Handbook of Amazonian Languages, v. 3, Mouton, Berlin.

Kaplan, L.D. (1982) Consonant Alternations in Inupiaq Eskimo. International Joumal of American Linguistics 48, 385-393.

Kari, J. (1976) Navajo Verb Prefix Phonology, Garland, New York.

Kaufman, T. (1971) Tzeltal Phonology and Morphology. Berkeley: University of California Press.

Kaun, A. (1994) The Typology of Rounding Harmony, Doctoral Dissertation, UCLA.

Kawasaki, H. (1982) An Acoustical Basis for Universal Constraints on Sound Sequences, Doctoral Dissertation, UC-Berkeley.

Kaye, J., Lowenstamm, J., and J.-R. Vergnaud (1985) The Internal Structure of Phonological Elements: a Theory of Charm and Government, Phonology Yearbook 2, 305-28.

Keating, P. (1984) Phonetic and Phonological Representation of Stop Consonant Voicing, Language 60:2, 286-319.

Keating, P. (1990) The Window Model of Coarticulation, in J. Kingston, ed., Papers in Laboratory Phonology I, Cambridge University Press.

Keating, P., B. Lindblom, J. Lubker & J. Kreiman (1994) Variability in Jaw Height for Segments in English VCVs, J. Phonetics 22, 407-422.

Keating, P., W. Linker & M. Huffman (1983) Patterns in Allophone Distribution for Voiced and Voiceless Stops, J. Phonetics 11, 277-290.

Keels, J. (1985) Guayabero: Phonology and Morphophonemics. in R.M. Brend (ed.), From Phonology to Discourse: Studies in Six Colombian Languages. Summer Institute of Linguistics.

Kelz, H. (1971) Phonologische Analyse des Pennsylvaniadeutschen. Hamburg: Helmut Buske Verlag.

Kenstowicz, M. (1982) Gemination and Spirantization in Tigrinya. Studies in the Linguistic Sciences 12:1, 103-122.

Kenstowicz, M. (1995) Base Identity and Uniform Exponence: Alternatives to Cyclicity, ms., MIT.

Kenstowicz, M. and C. Kisseberth (1979) Generative Phonology, New York: Academic Press.

Khalafallah, A. (1969) A descriptive grammar of Sa?idi, Egyptian Colloquial Arabic, Te Hague, Mouton.

Kimball, G.D. (1991) Koasati Grammar. Lincoln, NE & London: University of Nebraska Press.

Kiparsky, P. (1973) How Abstract is Phonology, in O. Fujimura (ed.) Three Dimensions of Linguistic Theory, TEC Corp., Tokyo.

Kiparsky, P. (1981) Vowel Harmony, ms., MIT.

Kiparsky, P. (1982) Lexical Phonology and Morphology, in I. Yang (ed.) Linguistics in the Morning Calm, 2: 3-91, Seoul: Hanshin.

Kiparsky, P. (1991) Blocking in Underived Environments, ms., Stanford University.

Kirchner, R. (1993) Turkish Vowel Harmony and Disharmony: an Optimality Theoretic Account, paper presented at Rutgers Optimality Workshop I.

Kirchner, R. (1995) Contrastiveness Is an Epiphenomenon of Constraint Ranking, Proceedings of the Berkeley Linguistics Society 21.

Kirchner, R. (1996) Synchronic Chain Shifts in Optimality Theory, Linguistic Inquiry 27:2, 341-350 (1996).

Kirchner, R. (1997) Contrastiveness and Faithfulness, Phonology 14:1, 83-111.

Klatt, D. (1975) Vowel Lengthening Is Syntactically Determined in a Connected Discourse, Journal of Phonetics 3(3), 129-140.

Klingenheben, A. (1928) Die Silbenauslautsgesetze des Hausa, Zeitschrift für Eingeborenen Sprachen 18, 272-297.

Koehn, E. & S. Koehn (1986) Apalai. in D. Derbyshire & G. Pullum (eds.), Handbook of Amazonian Languages, Vol. 1. Berlin & Amsterdam: Mouton de Gruyter.

Kohler, K. (1991) The Phonetics/Phonology Issue in the Study of Articulatory Reduction, Phonetica 48, 180-192.

Koshal, S. (1976) Ladakhi phonetic reader, Mysore: Central Institute of Indian Languages.

Kraft, C. & M. Kraft (1973) Introductory Hausa, Berkeley: UC Press.

Krueger, J. (1962) Yakut Manual, Bloomington, IU Press.

Kucera, H. (1961) 7he Phonology of Czech. The Hague: Mouton & Co.

Kuepper, K. J. (1992) Place Variation in the Consonant System: The Rhenish Velarization Revisited. Canadian Joumal of Linguistics 37, 17-40.

Ladefoged, P. & I. Maddieson (1996) The Sounds of the World's Languages, Blackwell, Oxford.

Lakoff, G. (1993) Cognitive Phonology, in J. Goldsmith (ed.), The Last Phonological Rule, University of Chicago Press.

Lass, R. & J. Anderson (1975) Old English Phonology, Cambridge University Press, Cambridge.

Lass, R. (1980) On Explaining Language Change, Cambridge University Press.

Lavoie, L. (1996) Consonant Strength: Results of a Data Base Development Project, Working Papers of the Cornell Phonetics Laboratory 11, 269-316.

Leal, D. (1972) Chitwan Tharu Phonemic Summary, Kirtipur: Summer Institute of Linguistics and Institute of Nepal Studies, Tribhuvan University.

Leben, W. (1973) Suprasegmental Phonology, Doctoral Dissertation, MIT.

Lehiste, I. (1977) Isochrony Reconsidered, J. Phonetics 5, 253-263.

Lepschy, A. & G. Lepschy (1977) The Italian Language Today. London: Hutchison.

Li, F. K. (1977) A Handbook of Comparative Tai. Honolulu: University of Hawaii Press.

Liljencrants, J. & B. Lindblom (1990) Numerical Simulation of Vowel Quality Systems, Language, 48, 839-862.

Lindblom, B. (1983) Economy of Speech Gestures, in MacNeilage, P. (ed.), Speech Production, New York, Springer Verlag.

Lindblom, B. (1990) Explaining Phonetic Variation: a Sketch of the H&H Theory, in W. Hardcastle and A. Marchal (eds.) Speech Production and Speech Modelling, 403-439.

Lindblom, B. (1963) Spectrographic Study of Vowel Reduction, Journal of the Acoustical Society of America 35, 1773-81.

Lindblom, B., S. Guion, S. Hura, S. Moon, & R. Willerman (1995) Is Sound Change Adaptive? Rivista di Linguistica, 7:1, 5-37.

Lipski, J. (1984) On the Weakening of /s/ in Latin American Spanish. Zeitschrift für Dialektologie und Linguistik 5 1: 1.

Lockwood, W. (1977) An Introduction to Moden Faroese, Føroya Skúlabókagrunnur, Tórshavn.

Lombardi, L. (1991) Laryngeal Features and Laryngeal Neutralization, Doctoral Dissertation, University of Massachusetts, Amherst.

Loos, E.E. (1969) The Phonology of Capanahua and Its Grammatical Basis. Norman, OK: University of Oklahoma & Summer Institute of Linguistics.

Lorimer, D.R.L. (1935) The Burushaski Language. Oslo: H. Aschchoug & Co.

Lozano, M. (1979) Stop and Spirant Altemations: Fortition and Spirantization Processes in Spanish Phonology. Ohio State Dissertation. Distributed by Indiana University Linguistics Club.

Lukas, J. (1967) A Study of the Kanuri Language: Grammar and Vocabulary, Dawsons of Pall Mall, London (reprint of 1937 original).

Macaulay, M. & Salmons, J.C. (1995) The Phonology of Glottalization in Mixtec. Intemational Journal of American Linguistics 61, 38-61.

MacDonald, L. (1990) A grammar of Tauya, New York: Mouton de Gruyter.

MacEachern, M. (1996) Laryngeal Similarity Effects in Quechua and Aymara, WCCFL 15.

MacKay, C. (1984) A Sketch of Misantla Totonac Phonology, International J. of Amer. Ling. 60, 369-419.

MacWhinney, B. & J. Leinbach, Implementations are not Conceptualizations: Revising the Verb Learning Model, *Cognition* 40:121-157.

Maddieson, I. & V. Anderson (1994) The Phonetic Structure of Iaai, UCLAWPP .

Maddieson, I. (1984) Patterns of Sounds. Cambridge: Cambridge University Press.

Maghway, J. (1995) Annotated Iraqw Lexicon, Tokyo: Institute for Study of Languages and Cultures of Asia and Africa.

Mallinson, G. (1986) Rumanian. Dover, NH: Croom Helm.

Malone, J. (1993) Tiberian Hebrew Phonology, Winona Lake, Indiana: Eisenbrauns.

Manley, T. (1972) Outline of Sre Structure, Honolulu, U. Hawaii Press.

Maran, L. (1971) Burmese and Jingpo: a study of tonal linguistic processes, Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics.

Marshall, M. (1984) The Dialect of Notre Dame de Sanilhac: a Natural Generative Phonology. Stanford French and Italian Studies 31.

Martin, S. (1992) A Reference Grammar of Korean, Tokyo, Charles E. Tuttle Co,, Inc.

Mascaró, J. (1983) Continuant Spreading in Basque, Catalan, and Spanish, in M. Aronoff and R. Oehrle (eds.), Language Sound Structure, Cambridge, MIT Press.

Mascaró, J. (1987) A Reduction and Spreading Theory of Voicing and Other Sound Effects, ms. Universitat Autònoma de Barcelona.

McCarthy, J. & A. Prince (1994) The Emergence of the Unmarked: Optmality in Prosodic Morphology, in Mercè Gonzàlez (ed.) Proceedings of NELS (U-Mass GLSA) 24: 333-379.

McCarthy, J. (1979) Formal Problems in Semitic Phonology and Morphology, Doctoral Dissertation, MIT.

McCarthy, J. (1981) The Representation of Consonant Length in Hebrew, LI 12, 322-327.

McCarthy, J. (1995) Remarks on Phonological Opacity in Optimality Theory, ms., U.Mass. at Amherst.

McCarthy, J. and A. Prince (1995) Faithfulness and Reduplicative Identity, Ms., University of Massachusetts, Amherst, and Rutgers University, New Brunswick, New Jersey.

McGregor, W. (1990) A Functional Grammar of Gooniyandi, John Benjamins, Amsterdam.

McLaughlin, J. E. (1989) A Note on the Change of Strident to Non-Strident in Gosiute Shoshoni. International Journal of American Linguistics 55, 240-247.

Meillet, A. (1970) General Characteristics of the Germanic Languages (W. Dismukes, tr.) Coral Gables, Florida: University of Miami Press.

Miller, W. (1965) Acoma Grammar and Texts. Berkeley and Los Angeles: UC Press.

Mills, E. (1984) Senoufo Phonology: Discourse to Syllable. Summer Institute of Linguistics.

Milroy, J., Milroy, L., & Hartley, S. (1994) Glottalisation in British English. English World-Wide 15.

Mithun-Williams, M. (1976) A Grammar of Tuscarora. New York: Garland.

Mohanan, K. (1986) The Theory of Lexical Phonology, D. Reidel (Kluwer), Dordrecht

Moon, S. & B. Lindblom (1994) Interaction between duration, context, and speaking style in English stressed vowels, JASA 96, 40-45.

Morev, N., A. Moskalyov, & Y. Plam (1979) The Lao Language, Moskow, USSR Academy of Sciences.

Morphy, F. (1979) The Djapu dialect of Yolngu, in Dixon & Blake (1979).

Mosel, U. & Hovdhaugen, E. (1992) Samoan Reference Grammar. Oslo: Scandinavian University Press.

Mous, M. (1992) A Grammar of Iraqw, Doctoral Dissertation, Leiden University.

Mrayati, M., R. Carré & B. Guérin (1988) Distinctive Regions and Modes: A New Theory of Speech Production, Speech Communication 7, 257-286.

Munro, P. & D. Gaye (1991) Ay Baati Wolof: a Wolof Dictionary, UCLA Occasional Papers in Linguistics.

Myers, J. (1996) Canadian Raising and the Representation of Gradient Timing Relations, paper presented at MCWOP, University of Illinois, Urbana-Champaign.

Nagaraja, K. (1985) Khasi, a Descriptive Analysis, Pune, Deccan College Post-Graduate & Research Institute.

Nanda, I. (1971) Conversational Newari, Katmandu, SIL and Institute for Nepal Studies, Tribhuvan U.

Nartey, J. (1982) On Fricative Phones and Phonemes: Measuring the Phonetic Differences Within and Between Languages, UCLA Working Papers in Phonetics 55.

Nater, H. F. (1989) Some Comments on the Phonology of Tahltan. International Journal of American Linguistics 55, 25-42.

Nelson, W. (1980) Performance Bounds in Speech Motor Control of Jaw Movements, Ms., Bell Laboratories.

Nespor, M. & I. Vogel (1982) Prosodic Domains of External Sandhi Rules, in H. van der Hulst & N. Smith (eds.) The Structure of Phonological Representations (Part I), 225-255, Dordrecht: Foris.

Newman, P. (1974) The Kanakuru Language, Leeds, Yorkshire: Institute of Modern English Language Studies and the West African Linguistic Society.

Ní Chiosáin, M. (1991) Topics in the Phonology of Irish, Doctoral Dissertation, UMass. Amherst.

Nivens, R. (1992) A Lexical Phonology of West Tarangan. in D. A. Burquest & W. D. Laidig (eds.), Phonological Studies in Four Languages of Maluku. Arlington, TX: Summer Institute of Linguistics and University of Texas at Arlington Publications in Linguistics, 127-227.

Nolan, F. (1992) The Descriptive Role of Segments: Evidence from Assimilation, in Docherty and Ladd (1992).

Nordbustad, F. (1985) Iraqw Grammar, ms.-affiliation unknown.

Noss, R. (1964) Thai Reference Grammar, Washington D.C., U.S. Foreign Service InstituteBeckman, M. et al. (1992) The Interaction of Coarticulation and Prosody in Sound Change, Language and Speech 35, 45-58.

Ó Siadhail, M. (1972), Modern Irish, Cambridge U. Press

Oftedal, M. (1985) Lenition in Celtic and in Insular Spanish: the Secondary Voicing of Stops in Gran Canaria, Oslo: Universitetsforlaget.

Ohala, J. (1975) Phonetic Explanations for Nasal Sound Patterns, in C. Ferguson, L. Hyman, J. Ohala (eds.) Nasalfest: Papers from a symposium on nasals and nasalization, Stanford: Language Universals Project, 289-316.

Ohala, J. (1981) The Listener as a Source of Sound Change, Papers from the Parasession on Language and Behavior, CLS, 178-203.

Ohala, J. (1983) The Origin of Sound Patterns in Vocal Tract Constraints, in P. MacNeilage (ed.)The Production of Speech, 189-216, New York, Springer-Verlag.

Ohala, J. (1992) Alternatives to the Sonority Hierarchy for Explaining Segmental Sequential Constraints, Papers from the Parasession on the Syllable, Chicago Linguistic Society, 319-338.

Oller, D. (1973) The Duration of Speech Segments: the Effect of Position in Utterance and Word Length, Journal of the Acoustical Society of America 54, 1235-73.

Olmsted, D.L. (1964) A History of Palaihnihan Phonology. University of California Press, Berkeley & Los Angeles.

Orgun, O. (1995) Correspondence and Identity Constraints in Two-Level Optimality Theory, Ms., U.C.-Berkeley.

Osborne, C. R. (1974) The Tiwi Language. Canberra: Australian Institute of Aboriginal Studies.

Osthoff, H. & K. Brugmann (1878) Morphologische Untersuchungen auf dem Gebiete der indogermanischen Sprachen.

Ourso, M. & C. Ulrich (1990) Sonorant Strengthening in Lama, Studies in the Linguistic Sciences 20: 1, 136-147.

Padgett, J. (1991) Stricture in Feature Geometry, Doctoral Dissertation, UMass-Amherst.

Painter, C. (1970) Gonja: A Phonological and Grammatical Study. Bloomington: Indiana University Press.

Paradis, C. (1988) On Constraints and Repair Strategies, Linguistic Review 6:1, 71-97.

Parker, G.J. (1969) Ayacucho Quecha Grammar and Dictionary. The Hague: Mouton.

Parker, S. (1994) Coda Epenthesis in Huariapano. International Journal of American Linguistics 60, 95-119.

Parks, D. (1976) A Grammar of Pawnee, New York: Garland.

Partridge, M. (1991) Serbo-Croat Practical Grammar nd Reader, Beograd, Prosveta.

Pater, J. (1996) *NC. Proceedings of the North East Linguistic Society 26, Graduate Linguistic Student Association, University of Massachusetts, Amherst, 227-239.

Patz, E. (1991) Djabugay, in Dixon & Blake (1991), 245-348.

Paul, H. (1889) Prinzipien der Sprachgeschichte.

Payne, D.L. (1981) The Phonology and Morphology of Axininca Campa. Arlington, TX: University of Texas and Summer Institute of Linguistics.

Pierrehumbert, J. (1980) The Phonology and Phonetics of English Intonation, Doctoral Dissertation, MIT.

Pierrehumbert, J. (1994) Knowledge of Variation, in CLS 30.

Pierrehumbert, J. and D. Talkin (1992) Lenition of /h/ and Glottal Stop, in Docherty and Ladd (1992), 90-117.

Piggott, G.L. (1980) Aspects of Odawa Morphophonemics. New York: Garland.

Pike, E., D. Forster & W. Forster (1986) Fortis vs. Lenis Consonants in the Paya Dialect of Kuna, in B. Elson (ed.), Language n Global Perspective: Papers in Honor of the 50th Anniversary of the Summer Institute of Linguistics, Dallas: SIL.

Popjes, J. & Popjes, J. (1986) Canela-kraho. in D. Derbyshire & G. Pullum (eds.) Handbook of Amazonian Languages, vol. 1. Berlin/Amsterdam: Mouton de Gruyter.

Poppe, N. (1963) Tatar Manual: a Descriptive Grammar and Texts with a Tatar-English Glossary, Bloomington: Indiana University, Uralic and Altaic series v. 25.

Poppe, N. (1964) Bashkir Manual. Indiana University Publications.

Poppe, N. (1970) Mongolian Language Handbook, Center for Applied Ling., Washington, D.C.

Port, R. and M. O'Dell (1985) Neutralization of Syllable-Final Voicing in German, Journal of Phonetics 13, 455-472.

Prentice, D. (1971) The Murut Languages of Sabah, Canberra: Pacific Linguistics, Australian National University.

Prince, A. & P. Smolensky (1993) Optimality Theory: Constraint Interaction in Generative Grammar, ms. Rutgers University, University of Colorado.

Prince, A. (1975) The Phonology and Morphology of Tiberian Hebrew, Doctoral Dissertation, MIT.

Prince, A. (1997) Stringency and Anti-Paninian Hierarchies, handout from lecture at LSA Institute, Ithaca, NY, June 26, 1997.

Prokosch, E. (1933) An Outline of German Historical Grammar, New York, Oxford University Press.

Proulx, P. (1989) A Sketch of Blackfoot Historical Phonology. International Journal of American Linguistics 55, 43-82.

Pulleyblank, E.G. (1984) Middle Chinese: A Study in Historical Phonology. Vancouver: University of British Columbia Press.

Rajapurohit, B. (1983) Shina Phonetic Reader, Mysore: Central Institute of Indian Langauges.

Ramer, A. M. (1993) On Lenition in Some Northern Uto-Aztecan Languages. International Joumal of American Linguistics 59, 334-341.

Ramsey, S. (1991) Proto Korean and the Origins of Korean Accent, in W. Goltz & M. Shapiro (eds.), Studies in tthe Historical Phonology of Asian Languages, Amsterdam: Benjamins.

Rangan, K. (1975) Balti phonetic reader, Mysore: Central Institute of Indian Langauges.

Rangan, K. (1979) Purki phonetic reader, Mysore: Central Institute of Indian Langauge.

Read, C. (1975) Children's Categorization of Speech Sounds in English, National Council of Teachers of English Research Report no. 17, Urbana, IL: ERIC Clearinghouse on Language and Linguistics.

Rehg, K. & D. Sohl (1981), Ponapean Reference Grammar, Honolulu: U. Hawaii Press.

Reid, L. (1971) Philippine Minor Languages: Word Lists and Phonologies, U. Hawaii Press.

Resnick, M. (1975) Phonological Variants and Dialect Identification in Latin American Spanish, Mouton, the Hague.

Rigsby, B. & Ingram, J. (1990) Obstruent Voicing and Glottalic Obstruents in Gitksan. International Journal of American Linguistics 56, 251-263.

Rischel, J. (1974) Topics in West Greenlandic Phonology, Akademisk Forlag, Copenhagen.

Ristinen, E.K. (1960) An East Cheremis Phonology. in American Studies in Uralic Linguistics, vol. 1. Indiana University.

Roberts, J. (1987) Amele, London: Croom Helm.

Robins, R.H. (1958) The Yurok Language. Berkeley & Los Angeles: University of California Press.

Romero, J. (1992) An Experimental Analysis of Spirantization in Spanish, paper presented at ASA, New Orleans, LA.

Romero, J. (1996) Gestural Organization in Spanish: an Experimental Study of Spirantization and Aspiration, Doctoral Dissertation, University of Connecticut.

Rood, D. S. (1976) Wichita Grammar. New York: Garland.

Rothenberg, M. (1969) The Breath-Stream Dynamics of Simple-Released-Plosive Production, Bibliotheca Phonetica 6, Basel: S. Karger.

Rumelhart, D. & J. McClelland (1986) On learning the past tenses of English verbs, in J. McClelland and D. Rummelhart, eds., *Parallel Distributed Processing,* v. 2, Cambridge, MA: MIT Press

Sagey, E. (1986) The Representation of Features and Relations in Non-Linear Phonology, Doctoral Dissertation, MIT.

Saib, J. (1977) A phonological study of Tamazight Berber: dialect of the Ayt Ndhir, Doctoral Dissertation, UCLA.

Salmond, A. (1974) A Generative Syntax of Luangiua. The Hague: Mouton.

Saltzman, E. & J. Kelso (1987) Skilled Actions: A Task Dynamical Approach, Psych. Rev. 94, 84-106.

Saltzman, E. & K. Munhall (1989) A Dynamic Approach to Gestural Patterning in Speech Production, Ecological Psych. 14, 333-382.

Samarin, W. (1966) The Gbeya Language, Berkeley: UC Press.

Samarin, W. J. (1967) A Grammar of Sango. The Hague: Mouton.

Sánchez, J.R.M. (1981) La Lengua Kamen.ntzá: Fonética, Fonología, Textos. Bogotá: Publicaciones del Instituto Caro y Cuervo.

Sapir, E. & Swadesh, M. (1960) Yana Dictionary. Mary Haas (ed.) Berkeley: University of California Press.

Sapir, J.D. (1965) A Grammar of Diola-Fogny. Cambridge: Cambridge University Press.

Sarma, D. (1982) Tibeto-Himalayan Linguistics: a Descriptive Analysis of Pattani, a Dialect of Lahaul, Hoshiarpur: Vishveshvaranand Vishva Bandhu Institute of Sanskrit and Indological Studies, Panjab University.

Saul, J. (1980) Nung Grammar, Dallas, SIL and U. Tex. at Arlington.

Sauvageot, S. (1965) Description synchronique d'un Dialecte wolof: le Parler du Dyolof (Mémoires de l'Institut Fondamental d'Afrique Noire 73), Dakar: IFAN.

Scatton, E. (1983) Reference Grammar of Modern Bulgarian (1983), Columbus, Ohio, Slavica

Schein, B. & D. Steriade (1986) On Geminates, LI 17:4, 691-744.

Schein, B. (1981) Spirantization in Tigrinya, in H. Borer & J. Aoun (eds.), MIT WPL 3.

Schiffman, H. (1983) A Reference Grammar of Spoken Kannada, University of Washington Press.

Selkirk, E. (1980) Prosodic Domains in Phonology: Sanskrit Revisited, in M. Aronoff and R. Oehrle (eds.), Language Sound Structure, Cambridge, MIT Press.

Sharma, J. (1979) Gojri Phonetic Reader, Mysore: Central Institute of Indian Languages.

Sherzer, J. (1975) A Problem in Cuna Phonology, J. Ling. Assoc. of the Southwest 1, 45-53.

Shipley, W. (1963) Maidu Texts and Dictionary, Berkeley: University of Californian Press.

Shryock, A. (1995) Investigating Laryngeal Contrasts: An Acoustic Study of the Consonants of Musey, Doctoral Dissertation, UCLA.

Siebert, F.T. (1989) A Note on Quapaw. Intemational Journal of American Linguistics 55, 471-476.

Silverman, D. (1995) Phasing and Recoverability, Doctoral Dissertation, UCLA.

Singh, R. (1977) Survey of Halabi in Madhya Pradesh, Delhi: Office of the Registrar General, Census of India 1971 monograph series 1, no. 5.

Sjoberg, A. (1963) Uzbek Structural Grammar, Indiana University, Bloomington.

Skousen, R. (1989) Analogical Modeling of Language, Dordrecht: Kluwer.

Smeets, H. (1984) Studies in West Circassian Phonology and Morphology, Leiden, Hakuchi Press.

Smith, C. (1992) The Timing of Vowel and Consonant Gestures, Doctoral Dissertation, Yale University.

Smith, N., M. Biers, R. Rod, R, Bolognesi, H. Humbert, F. van der Leeuw (1990) Lenition in a Sardinian Dialect, Certamen Phonologicum II.

Smolensky, P. (1993) Harmony, Markedness, and Phonological Activity, paper presented at Rutgers Optimality Workshop 1.

Sohn, H. (1994) Korean. London: Routledge.

Sommerstein, A. (1973) The Sound Pattern of Ancient Greek, Oxford: Basil Blackwell.

Stallcup, K. (1978) A Comparative Perspective on the Phonology and Noun Classification of Three Cameroon Grassfields Bantu Languages: Moghamo, Ngie, and Oshie, Doctoral Dissertation, Stanford University.

Stampe, D. (1972) How I Spent My Summer Vacation, Doctoral Dissertation, U. Chicago.

Steriade, D. & B. Schein (1998) On Geminates, LI 17:4, 691-744.

Steriade, D. (1982) Greek Prosodies and the Nature of Syllabification, Doctoral Dissertation, MIT.

Steriade, D. (1987) Redundant Values, in A. Bosch et al. (eds.) 23rd CLS , Part II: Parasession on Autosegmental and Metrical Phonology .

Steriade, D. (1993a) Closure, Release, and Nasal Contours Phonetics and Phonology 5, 401-470.

Steriade, D. (1993b) Positional Neutralization, ms., UCLA.

Steriade, D. (1995a) Underspecification and Markedness, in Goldsmith (1995).

Steriade, D. (1995b) Neutralization and the Expression of Contrast, ms., UCLA.

Steriade, D. (1996) Paradigm Uniformity and the Phonetics-Phonology Boundary, paper presented at Lab. Phon.

Stevens, K. (1972) The Quantal Nature of Speech: Evidence from Articulatory-Acoustic Data, in E. David & P. Denes (eds.) Human Communication: a Unified View, New York: McGraw-Hill, 51-66.

Stevens, K., S. Keyser & H. Kawasaki (1986) Toward a Phonetic and Phonological Theory of Redundant Features, in J. Perkell & D. Klatt (eds.) Invariance and Variability in Speech Processes, Hillsdale: Lawrence Erlbaum.

Stone, M. (1996) Laboratory Techniques for Investigating Speech Articulation, in W. Hardcastle and J. Laver (eds.), The Handbook of the Phonetic Sciences, 11-32, Oxford: Blackwell.

Story, G. (1984) Babine and Carrier Phonology -- A Historically Oriented Study, Norman, Oklahoma: SIL.

Street, J. (1962) Khalkha Structure, Uralic and Altaic Series 24, Indiana University Publications in Linguistics, Bloomington.

Sulkala, H. & M. Karjalainen (1992) Finnish, Routledge, London.

Svantesson, J. (1983) Kammu Phonology and Morphology, Lund, CWK Gleerup.

Swadesh, M. (1969) Elementos del Tarasco Antigua. Universidad Nacional Autónoma de

Swift, L.B. (1963) A Reference Grammar of Modern Turkish. Bloomington: Indiana University.

Taylor, C. (1985) Nkore-Kiga. London: Croom Helm.

Taylor, D. (1977) Languages of the West Indies. Baltimore: The Johns Hopkins University Press.

Teeter, K. (1964) The Wiyot Language, Berkeley: UC Press.

Teodorsson, S.-T. (1977) The Phonology of Ptolemaic Koine. Göteburg: Acya Universitatis Gothoburgensis.

Thráinsson, H. (1978) On the Phonology of Icelandic Pre-Aspiration, Nordic Journal of Linguistics 1, 3-54.

Tinelli, H. (1981) Creole Phonology, the Hague: Mouton.Tosco, M. (1991) A Grammatical Sketch of Dahalo, Helmut Buske Verlag, Hamburg.

Tosco, M. (1991) A Grammatical Sketch of Dahalo. Hamburg: Helmut Buske Verlag.

Trail, R. (1970) The Grammar of Lamani, Norman, Oklahoma: Summer Institute of Linguistics and University of Oklahoma.

Trigo, L. (1988) On the Phonological Derivation and Behavior of Nasal Glides, Doctoral Dissertation, MIT.

Trubetzkoy, N. (1939) Grundzüge der Phonologie. Prag. [Bd. 7 der Travaux du Cercle Linguistique de Prague.] English transl. by C. Baltaxe (1969) Principles of Phonology, Berkeley, U.C. Press.

Tyler, S.A. (1975) Gondi /h/ with Some Notes on Number and Gender in Proto-Central Dravidian. in H.F. Schiffman & C.M. Eastman (eds.), Dravidian Phonological Systems. Seattle: University of Washington & Institue for Comparative and Area Studies.

Ulving, T. (1953) Consonant Gradation in Eskimo. Intemational Journal of American Linguistics 19, 45-52.

Underhill, R. (1976) Turkish Grammar. Cambridge: MIT Press.

Valdman, A. (I 970) Basic Course in Haitian Creole. Bloomington: Indiana University.

Vavin, A. (1993) A Reconstruction of Proto-Ainu. Leiden: Brill.

Vogt, H. (1940) The Kalispel Language. Oslo: I Kommisjon Hos Jacob Dybwad.

Wade, T. (1992) A Comprehensive Russian Grammar (1992), Oxford, B. Blackwell.

Wares, A. (1968) A Comparative Study of Yuman Consonantism, the Hague:: Mouton.

Weber, D. (1983) A Grammar of Huallaga (Huanuco) Quechua, Berkeley, UC press.

Weldert, A. (1975) Componential Analysis of Lushai Phonology. Amsterdam: John Benjamins.

Wells, J. (1982) Accents of English, v. 2, Cambridge (Cambridgeshire): Cambridge University Press.

Westbury, J. & P. Keating (1986) On the Naturalness of Stop Consonant Voicing, Journal of Linguistics, 22, 145-166.

Whistler, R. (1992) Phonology of Sawai, in D. Burquest, W. Laidig (eds.), Phonological Studies in Four Languages of Maluku, Arlington, Texas: SL and U-Texasat Arlington Publications in Linguistics, 7-32.

Whitely, W. (1958) A Short Description of Item Categories in Iraqw, Kampala: East African Institute of Social Research.

Whitley, M. (1978) Generative Phonology Workbook, Madison: U. Wisconsin Press.

Wolf, G. de & Hasebe-Ludt, E. (1987) The Relationship of Speakers' Language Attitudes to Intervocalic Voicing in Fonual and Informal Speech. in Proceedings of the 14th International Congress of Linguists II, 1562-1567. Berlin: Akademie-Verlag.

Wordick, F. (1982) The Yindjibarndi language, Canberra: Pacific Linguistics, Australian National University.

Wright, R. & P. Ladefoged (1997) A Phonetic Study of Tsou, Bulletin of the Institute of History and Philology, Academia Sinica 68, 987-1028.

Wright, R. (1996) Consonant Clusters and Cue Preservation in Tsou, Doctoral Dissertation, UCLA.

Yar-Shater, E. (1969) A Grammar of Southern Tati Dialects, Columbia University Publications in Near and Middle Eastern Studies, series B 3, the Hague: Mouton.

Zec, D. (1988) Sonority Constraints on Prosodic Structure, Doctoral Dissertation, Stanford University.

Zipf, G. (1949) Human Behavior and the Principle of Least Effort, Cambridge, Mass., Addison-Wesley.

Zonneveld, W. (1978) A Formal Theory of Exceptions in Generative Phonology. Lisse: The Peter de Ridder Press.

Zvelebil, K. (1970) Comparative Dravidian Phonology. The Hague: Mouton.

Zwicky, A. (1972) On Casual Speech, CLS 8.