UNIVERSITY OF CALIFORNIA
SANTA CRUZ

AN EXPERIMENTAL APPROACH TO DEBUCCALIZATION AND SUPPLEMENTARY GESTURES

A dissertation submitted in partial satisfaction of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

in

LINGUISTICS

by

Jeremy O'Brien

June 2012

The Dissertation of Jeremy O'Brien
is approved:

Professor Grant McGuire, Chair

Professor Jaye Padgett

Professor Armin Mester

Dean Tyrus Miller
Vice Provost and Dean of Graduate Studies

Table of Contents

List of Figures

Abstract

An experimental approach to debuccalization and supplementary gestures

by

Jeremy O'Brien

Debuccalization is a weakening phenomenon whereby various consonants reduce to laryngeals. Examples include Spanish s-aspiration (s → h word-finally) and English t-glottalization (t → ʔ syllable-finally). Previous analyses of debuccalization view it as a lenition process that deletes or manipulates formal phonological features. This dissertation frames debuccalization based on the articulatory gestures involved rather than features. In the default case, debuccalization processes delete oral gestures and leave behind laryngeal gestures. This is captured in an Optimality Theoretic grammar by the low ranking of faithfulness to oral gestures, and the high ranking of faithfulness to laryngeal gestures. The motivation for changing the consonant is a markedness constraint which favors more articulatorily efficient consonants. When this constraint outranks oral faithfulness, consonants debuccalize.

This constraint system allows us to account for many cases of debuccalization well, but it highlights other cases where the laryngeal gestures of the debuccalized consonant and the fortis version are not identical—cases which show supplementary gestures. These supplementary gestures require an explanation, and three competing analyses are considered: perceptual faithfulness, dissimilation, and neutralization avoidance. The perceptual faithfulness analysis claims that the laryngeal gesture is changed to make the resulting sound more similar to the underlying sound. The dissimilation analysis states that the laryngeal gesture is changed to make it less like its neighboring sounds. Finally, the neutralization avoidance analysis claims that laryngeal gestures must change to avoid neutralization with other phonemes.

I first demonstrate that all three analyses are able to account for the case study of Indonesian k-debuccalization (k $\rightarrow$ ʔ in coda position). Moreover, all three are consistent with the principles of Optimality Theory and have plausible motivations. As such, experimental data is used to provide evidence for or against the competing analyses. A perceptual experiment was designed to verify the perceptual faithfulness and dissimilation analyses, evaluating the predictions each analysis makes regarding perceptual similarity. Some evidence is found for the perceptual faithfulness account of Indonesian. An artificial grammar learning experiment was also performed, which was used to test the possibility of a learning bias of avoiding neutralization in phoneme inventories, and the evidence from this experiment provides support for the neutralization avoidance analysis.

Acknowledgments

I am grateful to a number of people who have helped make this dissertation become a reality. My advisor, Grant McGuire, met with me countless times to let me bounce around ideas and to guide me through the entire dissertation-writing process. Together we discussed the hundreds of choices involved in designing and implementing experiments, and he always helped me see the pros and cons of each choice. Jaye Padgett has been extremely supportive in both the research and the personal aspects of being a graduate student. He pushed me to critically analyze each part of any chain of reasoning. He has also been a source of strength and kindness in difficult times. Armin Mester has taught me to analyze tough phonological problems with optimism, and it has been my pleasure to work with him as a teaching assistant in *four* phonology courses.

The Department of Linguistics at UCSC is a great source of pride for me, and I am so glad to be a part of this community. The students, faculty, and staff have supported me and cheered me on from the beginning. My compatriots in the Phonetics Lab—Ryan Bennett, Paul Willis, and Allan Schwade—were always willing to talk shop with me, providing references and interesting ideas when I would hit a wall. Also, I appreciate and thank all the participants who spent their time taking my experiments.

Finally, I need to express my gratitude to my friends and family. My friends, both linguists and non-linguists, have helped keep me sane throughout the whole dissertation process. My friends and family have given me so much love, and I wouldn't be here today without them—in particular, my Aunt Barb. When my mother died, she took me in and gave me all of her heart. She is the most caring and giving person I know, and her love and faith in me have kept me going.

I am honored to have had such a wonderful collection of people supporting me.

Chapter 1

Introduction

Debuccalization is a weakening phenomenon whereby various consonants with oral constriction reduce to laryngeal consonants. There is an extensive literature on the phonetics and phonology of lenition in general (Lavoie 1996, Kirchner 2001, Gurevich 2004, Bauer 2008, Gess 2009, among others), and within this literature debuccalization is explored as a subtype of lenition. For the most part, the focus in the literature is on the similarity between types of lenition processes. This dissertation continues in that tradition, by analyzing debuccalization with many of the same tools that other lenition processes are analyzed with, but it also looks for those attributes that may set debuccalization apart.

Previous analyses (Sections 1.1 and 1.4) often cast debuccalization as a lenition process that deletes or manipulates formal phonological features. The present analysis focuses more on the articulatory gestures involved. We define debuccalization as a weakening that, in the default, deletes oral gestures and leaves laryngeal gestures alone. This accounts for many cases of debuccalization well, but it highlights other cases where the laryngeal gestures of the debuccalized consonant and the fortis version are not identical. Such cases, which we term **supplementary gestures**, need to be explained, and three competing analyses are considered—perceptual faithfulness, dissimilation, and neutralization avoidance. Experimental evidence is used to decide between the alternative analyses.

1.1 What is debuccalization?

1.1.1 Defining aspects of debuccalization

Debuccalization is often defined as the loss of oral place of articulation. The term itself incorporates the Latin root *bucca*, meaning mouth or cheek, so it is akin to de-oralization. It is a type of sound change or alternation where a consonant no longer has any obstruction in the oral tract, and the sound that results is a laryngeal consonant ([h], [ɦ], or [ʔ]). This overall point of view is shared by most authors that work on lenition. Well accepted examples of debuccalization include Spanish s-aspiration (s → h word-finally) and English t-glottalization (t → ʔ syllable-finally).[1]

There are several edge-cases that arguably could be considered debuccalization. For instance, if a process leaves behind the secondary articulation of voicelessness, breathy voice, or creaky voice onto an adjacent vowel (as opposed to leaving behind a fully-fledged laryngeal consonant), this might be a type of debuccalization. Some authors (for instance, de Lacy 2002) discuss the debuccalization of nasal consonants to [N], the so-called phonologically placeless nasal, often realized phonetically as [ŋ], uvular [ɴ], or a nasalized off-glide. Gildea 1995 includes the velar fricative [x] as a possible outcome of debuccalization. This is generally not considered debuccalization elsewhere in the literature, unless an argument can be made for [x] being 'placeless' within the phonology of the language. In my opinion, nasal absorption (VN → Ṽ) may also fall under the purview of debuccalization, if not in name then at least in spirit. It is also possible that processes that change consonants to pharyngeals or epiglottals would be considered a type of debuccalization—the lenition literature does not mention such processes either way.

For the present study, we will assume a more restrictive definition of debuccalization, given in (1). This is so we can include most of the cases that everyone

[1] The environments given for these rules are simplified—they vary by dialect and register.

agrees are debuccalization while excluding those cases that only a few researchers would categorize as such.

(1) **Debuccalization** is any sound change or synchronic alternation that turns an oral consonant into a laryngeal consonant ([h], [ɦ], or [ʔ]).

There is also the question of what role weakening plays. The concept of debuccalization is almost intrinsically linked to the concept of lenition. Debuccalization is usually seen as a sub-type of lenition, alongside intervocalic voicing, degemination, spirantization, gliding, and the ultimate form of lenition, deletion. Hock 1991 argues that these lenition processes have more in common than just a traditional designation. Lenition processes result in forms that are easier to articulate. The lenition processes cross-linguistically pattern alike. They target similar classes and have similar environments, namely in inter-sonorant position and syllable/word-final position. Hock also argues that processes in these positions typically do not go up the hierarchy of weakening, only down (e.g. h $\rightarrow$ θ would not be found word-finally, but θ $\rightarrow$ h would).

Thus, our definition of debuccalization could include Hock's criteria for lenition. Because laryngeal consonants typically have a subset of the gestures that oral consonants do, it is straightforward to argue that they are articulatorily easier. But should we limit our investigation to only those processes that take place between sonorants or at the ends of prosodic constituents?

The choice made here is to not limit the scope of debuccalization in this way. Instead, a process will be categorized as debuccalization based on its resulting form, not its environment, but we will still assume that these processes are motivated by articulatory ease. The implications of this choice are discussed in Sections 1.4 and 5.2.

1.1.2 Feature geometry and gestures

The approach to debuccalization in this dissertation will mostly be in terms of gestural loss. Oral gestures delete, leaving behind laryngeal gestures in the gestural score. Of course, the removal of oral gestures does not mean that the oral tract is doing nothing. For instance, when there is an intervocalic [h], the tongue is still performing the gestures of the previous and following vowels. From the point of view of Articulatory Phonology (Browman & Goldstein 1986), debuccalization is the loss of oral gestural *targets*, not the loss of all oral movement. The consonant [h] does not provide an oral target, but there is still oral movement, which is the interpolation of oral targets for preceding and following sounds (c.f. Beckman & Pierrehumbert 1986 for tonal patterns in Japanese).

However, most previous work on the phonology of debuccalization makes use of feature systems. From the perspective of feature geometry, debuccalization involves delinking some node (usually Place), while retaining the features associated with laryngeal specification (see McCarthy 1988, who cites Goldsmith 1981 and Clements 1985). For example, Iverson 1989 argues that debuccalization involves delinking certain nodes, but leaving behind the Laryngeal node and features related to continuancy. Under this view, fricatives debuccalize to [h], because what is left is the Laryngeal node and [+continuant]. The Laryngeal node and [+continuant] is sufficient to characterize [h], and so no further features need to be added. Stops, on the other hand, debuccalize to [ʔ], because what remains is Laryngeal and [−continuant]. This is schematized in Figure 1.1.

The strong version of this claim—that all fricatives only debuccalize to [h], and all stops only debuccalize to [ʔ]—is falsified by the typological survey in Section 1.2. However, I agree with Iverson 1989 that the *default*[2] debuccalization pattern for most fricatives is to [h]. Most fricative targets of debuccalization are

[2] *Default* here means that there are no mitigating constraints that would change the debuccalized form. One of the main goals of this dissertation is to investigate what happens when debuccalization results in a laryngeal sound that is unexpected if we simply deleted oral gestures (i.e. non-default debuccalization).

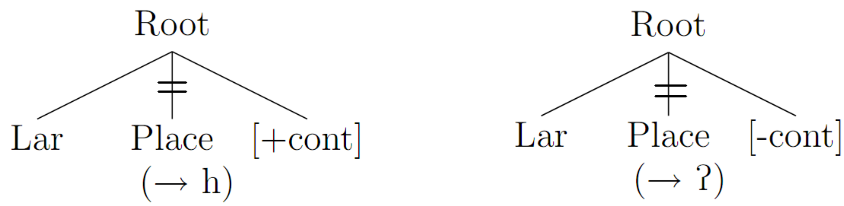

Figure 1.1: The feature geometric view of debuccalization described in Iverson 1989 (figure adapted from Padgett 1995)

voiceless. If you take a voiceless consonant and remove the oral gestures (leaving the laryngeal gestures alone), then this would result in a period of voicelessness with a spread glottis, something much like [h].[3] I disagree with Iverson's claim that plain stops debuccalize in the default case to [ʔ], though. Instead I argue that the default is usually [h], for the same reason given above for fricatives.

This leads us to an important distinction between the deletion of oral gestures and the deletion of features. In the arguments that follow in this paper, we will be viewing default debuccalization as loss of oral gestures.[4] From the perspective of deleting features, there appears to be disagreement over exactly what should be deleted in default debuccalization. It could be the Supralaryngeal node (Clements 1985), the traditional Place node (McCarthy 1988, Iverson 1989), or the Place node with stricture features (Padgett 1995). Debuccalization can even involve more complicated feature manipulation—for Lass 1976, debuccalization involves copying the [continuant] feature from the oral submatrix of features onto the laryngeal submatrix of features, followed by the deletion of the oral submatrix. These points of view mostly differ with respect to whether or not [continuant] has a say in the default rule of debuccalization. For us, the preservation of continuancy from non-lenited form to lenited form is not the real issue. The question of whether or not default debuccalization should result in a continuant should only

[3]See section 2.2 and Keating 1988 for further discussion on this issue.

[4]More precisely, the unfaithful mapping of oral gestures in the gestural score of the input, where the input oral gestures correspond to no gestures in the output form.

take into account the laryngeal gestural score of the non-lenited form. Part of the motivation for this is because speakers can freely combine constriction in the oral cavity and constriction in the larynx. The feature [continuant] often collapses this distinction or is ambiguous with respect to how laryngeal constriction should be treated. By framing the issue using gestures, we are able to emphasize the constriction of the vocal folds over oral constriction, with the intention of settling the continuancy question in a principled manner.

On the other hand, the use of gestures in the present analysis has many similarities to the use of phonological features. Like features, gestures here are meant to be formal units that the phonology can manipulate, and they can abstract over many pronunciations of the same word, not necessarily being tied to a particular utterance. And because, for the most part, features have a small number of articulatory gestures associated with them, the two systems will make similar predictions. In particular, the present conception of debuccalization is very similar to that of Padgett 1995. Under Padgett's view, stricture features belong to the Place node, so when Place is deleted, features like [continuant] should also delete. Thus, both approaches rely on laryngeal specifications, not general stricture features like [continuant]. They also make similar predictions when it comes to the default debuccalization of voiceless stops and fricatives. The primary difference between the two perspectives is that of stricture at multiple points in the vocal tract. The use of articulatory gestures is almost equivalent to saying there are stricture-like features on the Place node and on the Laryngeal node, while Padgett's system puts stricture features only on the Place node. Overall though, given the complexities of the feature [continuant] discussed above, the use of gestures seems warranted.

1.2 Typological survey of debuccalization processes

The following tables and related examples provide an overview of the typology of debuccalization processes. Most of the data is from Lavoie 1996 and de Lacy 2002, who both provide tables of debuccalization processes. For all tables, * indicates that the source is via Lavoie 1996, ♯ indicates that the source is via Gurevich 2004 (often correcting entries from Lavoie 1996), and † indicates the source is via de Lacy 2002.

The sound processes in Table 1.1 (and continued in Table 1.2) are ones that have [h] as the resulting sound. Some of the processes are historical sound changes, and others are synchronic alternations. Table 1.3 shows sound processes that result in [ɦ], and Table 1.4 shows those processes that result in [ʔ].

Language	Reference	Debuccalization pattern
Ainu	Poser 2001	p t k tʃ ɾ → h in coda position
Ainu	Vovin 1993*	*g > h
Awa	Loving & Loving 1966 via Smith 2007	obstruents → h in coda position (static generalization?)
Babine	Story 1984*	x > h stem-finally
Canela-Krahô	Popjes & Popjes 1986*	j x > h initially
English (Scots)	Lass 1976	θ → h intervocalically (optionally)
Florentine Italian	Giannelli & Savoia 1979 via Kirchner 2001	k → h (younger speakers, moderate speech rate, between vowel and liquid/vowel)
Gondi	Tyler 1975*	consonants > h intervocalically
Irish	Padgett 1995	s t → h (morphologically governed)
Japanese	Shibatani 1990	s p → h initially and intervocalically
Kannada	Schiffman 1983*	p > h word-initially
Kashaya	Buckley 1997	q qw → h in coda position
Kirundi	Goldsmith 1990	voiceless stops → h after nasals
Liverpool English	Watson 2001	t → h after short, unstressed vowels (monosyllabic words must be function words)
Miami (Illinois)	Costa 1991*	s x θ ʃ tʃ ç > h before voiceless stops
Middle Chinese	Pulleyblank 1984*♯	χ > h (Southern dialects)
Navaho	Kari 1976*♯	x → h medially
Nepali	Bandhu & Dahal 1971♯	ts^h → h intervocalically
Oscan and Umbrian	Buck 1904*	k p > h before t
Pipil	Campbell 1985*	w → h word finally and before C
Proto-Greek	Sommerstein 1973*	s > h before V
Páez	Gerdel 1985*♯	x → h everywhere except between /k/ and /i i̯/, optional word-initially

Table 1.1: Examples of debuccalization processes to [h]

Language	Reference	Debuccalization pattern
Sanskrit	Whitney 1889	s r > h finally (possibly syllable-finally)
Slave	Rice 1989 via Smith 2007	consonants → h in coda position
Spanish (Latin American)	Lipski 1984*	s → h intervocalically and word-finally in polysyllabic words
Spanish (Penin. dialects)	Morris 2000	s → h in coda position before [-voi] or [+son] sounds (other dialects: s → preaspirated geminate, s → geminate)
Tiriyó	Meira 2001	obstruents > h in coda position (all obstruents are voiceless)
Yoruba	Akinlabi 1992	w j → h before nasalized homorganic vowels
Yucatec Maya	Lombardi 1990	stops → h before homorganic stops and affricates

Table 1.2: Examples of debuccalization processes to [h] (continued)

Language	Reference	Debuccalization pattern
Florentine Italian	Giannelli & Savoia 1979 via Kirchner 2001	g → ɦ in fast speech (between vowel and liquid/vowel)
Ukrainian	Czaplicki 2006	ɣ → ɦ in onset position

Table 1.3: Examples of debuccalization processes to [ɦ]

Language	Reference	Debuccalization pattern
Arbore	Harris 1990♯	ejective and implosive stops → ʔ before non-identical consonant (optionally)
Arekuna Carib	Edwards 1978†	k → ʔ in coda position
Burmese	Lass 1976	p t k tʃ > ʔ word-finally
English (British)	Milroy et al. 1994*	t → ʔ intervocalically and sometimes pre-laterally
English (Cockney)	Andrésen 1968*♯	voiceless stops → ʔ intervocalically and before n m l
English (London, Leeds & Fife)	Harris 1990♯	t → ʔ word finally
Ethiopian Semitic languages	McCarthy 1988	p' t' k' → ʔ (environment unclear)
Indonesian	Lapoliwa 1981	k → ʔ in coda position
Kagoshima Japanese	Kaneko & Kawahara 2002	stops and affricates → ʔ (and nasals → N) in coda position
Kashaya	Buckley 1994†	plain stops → ʔ in coda position (but this might be wrong, see Buckley 1997)
Makassarese	Aronoff et al. 1987†	k → ʔ in coda position
Muher Gurage	Rose 2000	ejective k' → ʔ post-vocalically
Tahitian	Coppenrath & Prévost 1975†	k > ʔ
Takelma	Sapir et al. 1922 via Linguist List message (Paul Fallon)	k' $k^{w'}$ → ʔ before x
Tauya	MacDonald 1990♯	k k^{w} → ʔ $ʔ^{w}$ non-initially
Toba Batak	Hayes 1986♯	p t k → ʔ before consonants
Ulu Muar Malay	Hendon 1966†	stops → ʔ in reduplicant codas
West Tarangan	Nivens 1992*♯	k → ʔ intervocalically (word-internally where both vowels are non-high, fast speech)
Yamphu	Rutgers 1998†	t → ʔ in coda position (but assim. to following obstruents)

Table 1.4: Examples of debuccalization processes to [ʔ]

1.2.1 Additional examples

The following examples are attested debuccalization patterns, but for various reasons they are difficult to add to the tables above. The examples in (2) lack environments, they all come from Austronesian languages, and the examples come from an unpublished source. The examples in (3) do have an environment, but it is a difficult one to describe—Gildea 1995 uses historical reconstruction to account for the seemingly strange environment. Because the analysis of one Cariban language depends on the others, I have left them together in one place. Likewise, the examples in (4) are very closely related to each other, so separating them into different tables would obfuscate a generalization.

(2) Gess 2009 cites Nivins (p.c.), who cites Blust 1990, in giving many languages that have k-glottalization: Dobel, Lola, Yalahatan, Fordata, Luang, Kisar, Hawaiian.

(3) Cariban languages, when a suffix begins with -CV (e.g. -CV, -CVC, -CVCV), and the suffix is added to a final consonant[5] (from Gildea 1995):

a. Hixkaryana: some instances of k and ʃ → h
b. Makushi: at least some obstruents → h, ʔ (depending on source)
c. Panare: obstruents (p t k tʃ s) → h (ʔ before nasals)
d. Apalaí: at least some obstruents → ʔ
e. Carib: obstruents → h (ʔ before nasals)
nasals → ʔ before nasals

(4) Klamath (Barker 1964 via Clements 1985)

a. l̥ → h after n and l
b. l' → ʔ after n and l

[5]The environment is not this simple, because there are complications that result from syncope and from the historical basis of this environment.

Language	Debuccalization pattern
Proto-Indo-European	*/p/ > Armenian h initially
Proto-Indo-European	*/ǵʱ/ > Latin h
Proto-Indo-European	*/s/ > Iranian h
Proto-Indo-European	*/s/ > Brythonic Celtic h (most frequently initially)
Proto-Indo-European	*/k/ > Germanic h
Proto-Indo-European	*/ǵʱ/ > Sanskrit ɦ
Proto-Uralic	*/k/ > S. Ostyak h, Hungarian h, Yurak h initially
Proto-Uralic	*/tʃ/ > Finnish h initially
Proto-Uralic	*/ʃ/ > Finnish h initially
Proto-Uralic	*/k/ > Yurak, Yenisei Samoyed h intervocalically
Proto-Dravidian	*/p/ > ɦ in Brahmin dialects of Kannada initially
Proto-Dravidian	*/c/ > h in Pengo, Kuvi
Proto-Dravidian	*/k/ > h in Manda, Kui initially

Table 1.5: Examples of historical debuccalization processes from Lass 1976

Finally, there are the debuccalization sound changes in Table 1.5. These sound changes are represented in Lass 1976 as a long list with example words after each change. As such, I thought it best to present it in a similar way here, so as to show the proto-language and the attested derived language, and also to keep the proto-languages together. Furthermore, some of the proto-forms (namely Proto-Indo-European */ǵʱ/ and Proto-Dravidian */c/) are provided with symbols that may or may not be phonetically similar to the corresponding IPA symbols. This is fine, especially given that reconstructed forms often don't have a single precise phonetic description, due to the nature of the comparative method and the lack of direct evidence. Therefore, the data is given in a single place, with the warning that it is repeated more or less directly from Lass 1976.

1.3 Processes that might be viewed as debuccalization, but are not

Many varieties of English have a phenomenon called glottal reinforcement[6], where voiceless stops become reinforced by a glottal closure gesture. Because glottal reinforcement involves the *addition* of laryngeal gestures without the concomitant removal of oral ones, it does not appear to be a case of debuccalization (nor a case of lenition in general, depending on your point of view). Note, however, that a glottally reinforced stop can serve as the input to debuccalization ([ˀt] → [ʔ]), although it doesn't have to. In particular, Milroy et al. 1994 provide evidence that the two processes are distinct in some varieties of English. They pattern differently in terms of social behavior—gender and class distinctions will favor one process over the other. This leads the authors to believe that they represent two distinct processes, with glottal reinforcement being a local, older phenomenon, and glottal replacement a newer import from the south of England. If the two processes are indeed distinct in English, then glottal replacement cannot completely depend on glottal reinforcement as a source of explanation. Even if it were dependent on it, though, the glottal reinforcement itself is not a form of debuccalization.

In the framework of Articulatory Phonology, there is some literature on consonantal weakenings of nasals and liquids that resemble debuccalization (Browman & Goldstein 1995, Sproat & Fujimura 1993). To date, these papers have not been incorporated into the general discussion of debuccalization and lenition. In Browman & Goldstein 1995 and Sproat & Fujimura 1993, the timing of various articulatory gestures depends on syllabic position. For all the segments under consideration, there is a so-called consonantal gesture (the gesture which is asso-

[6]This phenomenon has many other names, including the ambiguous name 'glottalization'. To maintain clarity of discussion, we will use the terms 'glottal replacement' ([t] → [ʔ]) and 'glottal reinforcement' ([t] → [ˀt]). The terminology is relatively inconsistent, but within the literature this is the most clear terminology that I have come across.

ciated with a closure in the oral cavity, like the tongue-tip gesture of /l/) and a vocalic gesture (tongue-body gesture or velum-lowering gesture). These gestures are closely aligned in onset position (in terms of simultaneity), but in coda position the vocalic gesture tends to occur first, and the consonantal gesture tends to weaken in addition to being timed later. The misalignment and weakening of vocalic gestures in coda position looks like debuccalization, because it could eventually result in the removal of those gestures. In utterance-final position, gestures that are timed later and later will plausibly become inaudible. Perhaps this is how some debuccalization processes come about—oral gestures being timed later and later until they finally disappear.

This explanation based on misalignment may provide insight in some cases of coda debuccalization, but it certainly is not the whole story. For one thing, in the examples of /l/ weakening in coda position, the vocalic gesture still remains, and this gesture can definitely be considered an oral gesture. A weakened coda /l/ may become velarized, or it may eventually result in some type of vowel (l-vocalization), but it does not weaken to a laryngeal consonant. Furthermore, many cases of debuccalization processes occur intervocalically (as in Cockney English, Nepali, West Tarangan, etc.). According to their analysis, intervocalic position should be an excellent place for the timing of multiple gestures. Thus, the Sproat & Fujimura 1993 system of gesture misalignment is similar to debuccalization, but it has enough differences that we will pursue other approaches.

1.4 More background on debuccalization

1.4.1 Is debuccalization a unified phenomenon?

Other than the definition given in Section 1.1, is there anything else uniting debuccalization processes; any overarching generalizations or implicational universals? The environment for debuccalization does not appear to be the key. That is because debuccalization processes occur in many different types of environments—

word-initially (Kannada), word-finally (Pipil), intervocalically (Páez), in coda position (Indonesian), and before and after various other sounds. Even if we restrict the environments to the positions Hock 1991 argues for, there is still nothing particular to the environments of debuccalization as compared to other lenition processes.

Lavoie 1996 gives a few generalizations that we may comment on:

> "Except for glides and one instance of [g], all of the debuccalized stops or fricatives were voiceless. Some glides, such as [j], may be debuccalized. Fricatives usually become [h]. The voiceless velar fricative very frequently debuccalizes. All of the segments that debuccalized to glottal stops were stops to begin with." (p. 290)

The tables above are expanded versions of Lavoie's tables, providing us with more evidence to support or refute Lavoie's generalizations. We have another example of [g] debuccalizing, but this time to [ɦ] (Florentine Italian). There are other examples of voiced sounds debuccalizing: Ainu *[g] > [h] and /ɾ/ → [h], Sanskrit [r] > [h], and Ukrainian /ɣ/ → [ɦ]. Even so, the new evidence supports the claim that voiceless sounds debuccalize much more frequently than voiced sounds. The other generalizations still hold—stops and fricatives become [h], but stops are the only sounds that debuccalize to [ʔ].[7] Lavoie points out that [x] debuccalizes frequently, but the fact of the matter is [k] also frequently debuccalizes, as do other velars ([g], [ɣ]).

One way to investigate the properties of debuccalization is to follow these generalizations. If there can be a unified explanation for why debuccalization processes across the world's languages pattern this way, that serves as possible evidence for debuccalization as a unified phenomenon. If there fails to be a unified explanation, then perhaps debuccalization is simply an externally defined category, a grab-bag of sound changes and alternations that appear on the surface to be the same, but are in fact mostly unrelated.

[7] In the related languages of Panare and Carib, we find fricatives debuccalizing to [ʔ]. However, they only do so before nasal stops, which supports the weaker claim that fricatives do not debuccalize to [ʔ] unless assimilating to nearby (oral or nasal) stops.

Smith 2007 proposes that lenition can be divided into two basic types. The first type involves a segment becoming less marked, and generally affects coda consonants ("neutralization-to-the-unmarked" lenition). The second type of lenition involves segments becoming more sonorous (and possibly more marked), and this type generally applies intervocalically ("sonority-increasing" lenition). A segment like [h] can be interpreted as a high sonority consonant, in line with the view of [h] as a glottal glide or a voiceless vowel. It could also be interpreted as being of lower sonority, because it has frication noise and sometimes patterns with fricatives. There are other cases of segments having ambiguous featural specifications. For example, Mielke 2005 discusses the ambiguity of the feature [continuant] with respect to /l/ and /n/, where the feature values vary depending on the language/process in question. To give Smith's approach the benefit of the doubt, we will treat glottal segments as highly sonorous for the time being. Debuccalization then appears to make consonants less marked *and* more sonorous, compatible with either lenition type. If the debuccalized segment is more sonorous than the oral segment it replaces, then either explanation could hold. From this perspective, the type of a debuccalization process depends strictly on the environment and the proposed motivation for the process. Assumedly, coda debuccalization is neutralization-to-the-unmarked, while intervocalic debuccalization is sonority-increasing.[8]

This proposal can be falsified in two ways: (a) if there is debuccalization in a strong position that is not also intervocalic or intersonorant, or (b) if there is debuccalization in intervocalic position that *decreases* sonority. Both of these are attested in the tables above, although there is only one example of type (b), and it is not a very clear-cut example at that, as the Yoruba case could simply be dissimilation. Examples are shown in Table 1.6.

[8] If we fail to give Smith 2007 the benefit of the doubt when it comes to the sonority of [h], then we end up with a situation where [h] is decreasing in sonority in intervocalic position. Thus, those varieties of debuccalization would fail to be categorized in *either* of Smith's lenition types.

(a) Debuccalization in a strong position (non-intervocalic)	
Kannada	p > h word-initially
Canela-Krahô	j x > h initially
Irish	s t → h (morphologically governed)
Ukrainian	ɣ → ɦ in onset position

(b) Debuccalization in intervocalic position that decreases sonority	
Yoruba	w j → h before nasalized homorganic vowels (not nec. intervocalically)

Table 1.6: Examples of debuccalization that do not conform to the lenition typology in Smith 2007

Based on the attested examples of debuccalization patterns above, it appears that Smith 2007's division of lenition processes may not apply to every example of debuccalization. Some types of debuccalization occur in strong positions. It is possible, however, that the historical changes are the result of telescoping (see Subsection 1.4.3). The Irish example is also a result of telescoping—vowel-final particles used to precede the words, so the environment for lenition was intervocalic, but later the particles were dropped. The change also occured word-internally, but there are no alternations in that position. (Padgett, p.c.) Thus, it is not clear if these word-initial debuccalization cases are good counter-examples to Smith's approach.

If we adopt the distinction between neutralization-to-the-unmarked and sonority increasing lenition, it may not provide a lot of explanatory coverage. That is to say, there does not appear to be much of a difference between debuccalization processes in weak positions and those in intervocalic position, as shown in the typology from Section 1.2. Both positions include debuccalization to [h] and [ʔ], and both positions target stops, fricatives, and the occasional sonorant consonant.

Smith's approach is reasonable and is consistent with almost all of the data. Thus, I do not dismiss it out of hand, but it still does not provide a large amount of insight into our typology.[9]

1.4.2 Perceptual data on debuccalization to [ʔ]

In a perceptual study I conducted (O'Brien 2010), stop consonants at three places of articulation ([p], [t], and [k]) were compared to glottal stop in terms of confusability. The results showed that the pair [t ʔ] was more confusable than [p ʔ]. The pair [k ʔ] trended between the other two pairs in confusability. This led to a prediction that debuccalization processes where [p] becomes [ʔ] should be rare, while processes turning [t] or [k] to [ʔ] should be more common. Table 1.4 confirms this prediction—some glottal replacement processes target [t] or [k], some processes target all stops and affricates, but no such process targets [p] to the exclusion of other stops. Note that it is not surprising that some processes affect [p] in addition to other stops. The prediction from the perceptual study is due to *relative* confusability, not absolute confusability. The sound [p] is less like [ʔ] than other stops, but it might still be more like [ʔ] than non-stop consonants are.

These results provide support for the idea that debuccalization processes are influenced by perception. In part due to this previous line of inquiry, the three competing analyses of supplementary gestures in Chapter 2 are oriented towards perceptual differences.

1.4.3 Initial motivation vs. synchronic formalism

Kirchner 2001 and Bauer 2008 provide discussion about the distinction between initial motivation for a sound change or alternation, and the later development of that process over generations of speakers. Bauer explains this with an

[9]Chapter 2 provides an example analysis of debuccalization in coda position, and it assumes that this is a type of effort-reduction in weak position. Intervocalic debuccalization could be viewed in the same way, or it could be captured with a combination of effort-reduction and the drive to increase sonority. It is unclear how to capture the few purported cases of word-initial debuccalization, or if we even want to capture them in a similar way.

example: In a sound change of the type [d] > [ð], the first generation of speakers fail to achieve the stop closure target for [d], hence the initial motivation makes for a lenition process under his definition. Subsequent generations, argues Bauer, analyze the [ð] as the actual target. This means that the change to [ð], which might require more precise motor control and therefore could be seen as a more difficult segment, is *motivated* by articulatory ease but does not always *result* in articulatory ease. Bauer emphasizes the process and the motivation for that process, and avoids categorizing a process as lenition based on the final outcome.

An approach that highlights initial motivation becomes more difficult when attempting to account for synchronic alternations. Kirchner, for instance, uses requirements of articulatory ease as a synchronic motivation for alternations within an OT grammar. Explanations like Bauer's are unavailable to such an analysis—if [ð] is more difficult to *reliably* produce than [d], and the grammar controls such a reliable alternation, then articulatory ease alone cannot account for the alternation. This is not necessarily a problem for Kirchner, however. In his mass-spring model of articulation, fricatives like [ð] are more articulatorily efficient than stops like [d].[10] If we had a clear metric for articulatory ease, then the difference between Kirchner 2001 and Bauer 2008, with respect to spirantization, could be empirically evaluated. As it stands, Kirchner's mass-spring model is the closest thing we have to verification.[11]

It might turn out that all debuccalization processes (and all lenition processes in general) are rooted in articulatory underachievement. But it is definitely not the case that all debuccalization processes, as defined in section 1.1, are in and of themselves articulatory underachievement. Sometimes supplementary articulation is added in non-default debuccalization, either in the initial development of

[10] However, stridents like [s] are *less* efficient than [ð]—Kirchner assumes "that strident fricatives (Figure 4-2) require a relatively precise, sustained close constriction, in order to generate highly turbulent airflow" (p. 111, 1998 ms.). This enables the analysis to partially capture the fact that stops never lenite to stridents without first affricating. For exact values of articulatory efficiency in Kirchner's analysis of Florentine Italian, see pp. 271−272, 1998 ms.

[11] Moreover, because Spanish spirantized obstruents are often transcribed as open fricatives or approximates, the question of the precision of the articulatory target is also up for debate.

the process, or in its eventual evolution. Following this line of reasoning, there might be interesting implications for synchronic analyses of debuccalization phenomena. “New” debuccalization processes, those that have not strayed very far from simple articulatory underachievement, should be more easily captured using models that incorporate articulatory difficulty (Kirchner’s model, Liljencrants & Lindblom 1972, Lindblom 1990, etc.). The older the debuccalization process, the more “cruft” it has the opportunity to accumulate—generations of speakers reanalyze the input to come upon different articulatory targets and different generalizations, and this telescoping is not what simple articulatory underachievement predicts.

When analyzing synchronic alternations, we will assume that the drive to minimize articulatory effort is the impetus for debuccalization. Such an assumption privileges “new” debuccalization processes over “old” ones, but it is necessary to get the analysis off the ground.

1.5 Overview of the dissertation

The analysis of debuccalization and supplementary gestures advocated for in this dissertation is given in Chapter 2. The general idea is that debuccalization involves deletion of oral gestures. This is captured in an Optimality Theoretic grammar by the low ranking of faithfulness to oral gestures, and the higher ranking of a markedness constraint that favors more articulatorily efficient consonants.

In the default case, the laryngeal gestures of the consonant remain in the debuccalized form. This is reflected by high ranking of faithfulness to laryngeal gestures. In some situations, though, the laryngeal gestures of the debuccalized form are different from the strong form of the consonant. These supplementary gestures must be introduced to satisfy some other constraint. Chapter 2 explores this line of reasoning in detail. We develop three competing analyses—perceptual faithfulness, dissimilation, and neutralization. As we will show, all three analyses are plausible, in that they successfully account for the example debuccalization process and they are consistent with general principles of OT. Experiments are used to provide evidence for or against the competing analyses.

Chapters 3 and 4 detail two such experiments. Chapter 3 reports on a perceptual experiment designed to verify the plausibility of the perceptual faithfulness and dissimilation analyses of supplementary gestures. The perceptual similarity of various consonants is tested, including the laryngeal consonants, by way of a speeded *same-different* discrimination task. The reaction time results from this task are used to test predictions made by the competing analyses. Additionally, multidimensional scaling and hierarchical cluster analyses are used to visualize the similarity of the sounds in question.

In Chapter 4, an artificial grammar learning experiment is performed in order to evaluate the neutralization avoidance analysis of supplementary gestures. Two nearly-identical artificial languages are used, one with a neutralizing rule, and the other with a non-neutralizing rule. By comparing how well learners acquire and use these rules, the experiment is able to investigate whether there is a learning bias towards avoiding neutralization in phoneme inventories.

Chapter 5 concludes.

Chapter 2

Analyzing Debuccalization and Supplementary Gestures

2.1 Introduction

We have defined debuccalization as the removal of oral gestures, without much being said about the addition of other gestures. Kirchner 2001, however, provides evidence that some debuccalization processes involve supplementary gestures, in addition to the simple removal of oral gestures. Viewing debuccalization as a phenomenon from this perspective allows us to explain some of the generalizations from the earlier typological survey of debuccalization processes (Section 1.2). This approach also brings up the question as to what precisely motivates the inclusion of these supplementary gestures. As such, the explanation given in Kirchner 2001, based on perceptual faithfulness, is compared to other possible explanations, which are based on dissimilation from neighboring sounds and the need to avoid neutralization.

2.2 Default debuccalization and supplementary gestures

Some examples of debuccalization can be analyzed as the simple removal of oral gestures. For example, in some varieties of American English there is variation between a pre-glottalized voiceless alveolar stop [ˀt] and a glottal stop with no alveolar closure [ʔ]. The difference between these two sounds can be seen as removing an oral gesture: [ˀt] debuccalizes to [ʔ] by losing all oral gestural targets.

However, for other alternations the picture is not so simple. For an alternation with glottal stop and a *non*-pre-glottalized stop (for instance intervocalic English [t] or West Tarangan [k]), it is not obvious what the sound would be like if there were no oral gestural targets. The vocal folds during [t] are usually spread apart (Keating 1988), thus it is likely that [t] without oral gestures is similar to [h].[1] If we take the sequence [ata] and remove all oral gestures from the gestural score of [t], then the result would be vocal fold vibration (voicing), followed by some amount of voicelessness, followed by more vocal fold vibration. For a voiced sound like [d], it appears that the sound that would result from a lack of oral gestures would simply be voicing (in other words, intervocalic [d] would not be a segment at all if its oral gestures were eliminated). In the environment of being adjacent to voiced segments, voiced consonants should "debuccalize" to zero, and voiceless sounds should debuccalize to [h] by default. However, if a voiced sound is voiced with breathy voicing (as Gess 2009 argues for [ɣ]), then [ɦ] would be the result of losing oral gestures. Any other type of debuccalization must be the result of something different from (or in addition to) the simple removal of oral gestural targets.

[1] It is possible that [h] requires a bit more laryngeal constriction than a voiceless oral obstruent does. Here we are claiming that [h] and the voicelessness of obstruents are alike in their laryngeal gestures, enough so that [h] satisfies laryngeal gestural faithfulness constraints when compared to an input voiceless obstruent.

(5) Working hypothesis of default debuccalization:
Assuming that effort minimization is the cause of debuccalization, then, all things being equal, the result of debuccalization should be the same as removing oral gestures (usually [h]). If this is not the case, and there is a laryngeal gesture that is not in the original consonant, then this supplementary gesture must be there for some other grammatical reason.

One way to formalize this idea in Optimality Theory is with a constraint that prohibits oral gestures in surface forms, *OralGesture. While relatively straightforward, such an approach would mean that debuccalization processes and other lenition processes would be captured in different ways. If the motivation for debuccalization were encoded as a constraint on articulatory difficulty, then this would formally unite many lenition processes (a trait that Kirchner 2001 finds desirable). It would also make the connection between effort minimization in debuccalization and the formal implementation of it stronger. For these reasons, the overall analysis of debuccalization in this dissertation makes use of a general articulatorily motivated constraint—Kirchner's Lazy constraint against articulatorily difficult segments.[2]

Thus, Lazy is the markedness constraint that provides the motivation for leniting a consonant. In the default case, debuccalization processes delete oral gestures and leave behind the laryngeal gestures. In such a situation, there is a low ranking of faithfulness to oral gestures, and a high ranking of faithfulness to laryngeal gestures. When Lazy is ranked between these two types of faithfulness, then a consonant will undergo debuccalization. This is schematized in the tableau in (6), illustrating the general approach to Spanish s-aspiration with the word *mas* 'more'.

[2] The constraints *OralGesture and Lazy make similar predictions with respect to the analysis here, so the reliance on Lazy is not crucial.

(6)

/mas/	Faith-Laryngeal	Lazy	Faith-Oral
a. mas		*!	
b. ☞ mah			*

This OT formulation treats debuccalization just like any other lenition process. Effort minimization constraints require the sound to be easy to articulate, and other faithfulness and markedness constraints make further demands on the output consonant. As the typology in Section 1.2 and the discussion in Section 1.4 suggest, the class of debuccalization processes is heterogeneous and difficult to systematize. Therefore, the OT treatment of debuccalization (as "nothing special") is compatible with our understanding of debuccalization as a diverse grouping of processes. Debuccalization processes are united in the relative ranking of Lazy over faithfulness to oral gestures, but they are not special in any other way from other lenition processes.

As stated above, Kirchner 2001 argues that certain examples show that supplementary gestures must be added (pp. 120–123, 1998 ms.).[3] In Florentine Italian /g/ can be realized as [ɦ] (acquiring a breathy-voicing gesture, according to Kirchner), and in many English varieties intervocalic /t/ is realized as [ʔ] (acquiring a glottal closure gesture). Kirchner believes that the supplementary gestures are added for perceptual reasons, so that the lenited form will sound more like the non-lenited one. This is not the only logically possible explanation, however, and below we explore how to evaluate this claim and other claims that could explain supplementary gestures.

2.3 Three analyses

Let's use Kirchner 2001's Florentine Italian g-debuccalization case. If we remove the velar constriction gesture of [g], the result is a vowel that is the interpo-

[3]Kirchner 2001 may not have been the first to argue that debuccalization processes are sometimes accompanied by further modifications. Padgett 1995 cites an unpublished manuscript, Bessell 1993, which argues that some cases of debuccalization cannot be attributed to the laryngeal specification of the original consonant.

lation of adjacent vowels—continuous voicing with no gestural target for the oral articulators.[4] But the debuccalization process in Florentine Italian adds a supplementary gesture of slight glottal abduction to get breathy voicing. Kirchner believes this is to make the resulting sound more like [g], in terms of "satisfying perceptually based faithfulness constraints" (p. 121, 1998 ms).[5] I refer to this as the **perceptual faithfulness analysis**. According to Kirchner, lenition in general and debuccalization in particular is "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" (p. 123, 1998 ms). Because fortition presumably is not involved in the debuccalized form, this leaves the grammar's faithfulness constraints as the source of supplementary gestures.

(7) Perceptual faithfulness analysis:
All supplementary gestures in debuccalization processes are due to perceptual faithfulness constraints. These constraints demand the presence of supplementary gestures to make the debuccalized form more perceptually similar to the non-debuccalized form.

The general idea of faithfulness to the fortis form is taken by Kirchner and others working within OT. However, it is not the only logically possible motivation; the supplementary gesture could be added for other reasons.

Perhaps there is a constraint at play that forces the sound to be different from the vowels around it (independent of the resulting sound being more like [g] or less like [g]). Then, [ɦ] would be one possible laryngeal segment that is in some respects different from the surrounding vowels. Under this view, the supplementary glottal gesture wasn't added to make the resulting sound more

[4]Ignoring for the time being Gess 2009's criticisms of this view.

[5]In Kirchner's tableaux, this particular example is not fleshed out because [ɦ] is the most articulatorily efficient form in any position, except for zero. Thus, the [ɦ] could be the result of perceptual similarity to [g], or it could be the most efficient articulation that still allows for there to be a phonological consonant present.

like [g], but rather to make it less like the vowels around it. Moreover, this drive to make the lenited [g] less like its neighboring vowels could be driven by a need to maintain prosodic structure, allowing for weakening but keeping the CV structure of the non-lenited form intact. This line of reasoning will be referred to as the **dissimilation analysis**.

(8) Dissimilation analysis:
All supplementary gestures in debuccalization processes are caused by the need for the lenited sound to be sufficiently different from its surrounding sounds.

The dissimilation analysis above is syntagmatic in nature, with the basis of comparison being the segments in the immediate environment. An alternative proposal would be more paradigmatic in nature. The drive to add supplementary gestures comes from a need to maintain phonemic contrast and avoid neutralization. This is the **neutralization avoidance analysis**.

(9) Neutralization avoidance analysis:
All supplementary gestures in debuccalization processes are there to prevent the debuccalized consonant from neutralizing with another phoneme in the language.

All three analyses are, in principle, possible accounts for supplementary gestures. In the following subsections, we will evaluate the analyses by formalizing them in Optimality Theory and comparing the predictions they make for languages.

2.4 Comparing the analyses with an example: Indonesian coda k-debuccalization

To make the comparison of these three analyses more concrete, we will use the example of Indonesian coda k-debuccalization. This example demonstrates how each analysis uses different motivations for the supplementary glottal stop gesture, and it further demonstrates the different predictions the analyses make.

2.4.1 Shared properties of the three competing analyses

All three analyses, at least in their incarnations here, have several commonalities. They are all formalized in Optimality Theory, and the motivation for reducing coda /k/ is the same—a markedness constraint assumed to be rooted in effort minimization. This is not completely transparent, however, because only the voiceless velar stop is affected. The other consonants fail to participate in the debuccalization process. (Table 2.1 shows the consonant inventory of Indonesian.)

	Labial	Alveolar	Palatal	Velar	Glottal
Plosive	p b	t d		k ɡ	ʔ
Fricative	(f)	s (z)	(ɕ)	(x)	h
Affricate			tɕ dʑ		
Nasal	m	n	ɲ	ŋ	
Liquid		l r			
Approximant	w		j		

Table 2.1: Indonesian consonant inventory, adapted from Lapoliwa 1981. Segments in parentheses indicate loan-phonemes.

There are many ways to allow debuccalization of /k/ but prevent it for other segments, but the direction pursued here is to assume faithfulness to other places of articulation, faithfulness to voiced segments, and faithfulness to loan-phonemes. To prevent debuccalization in onset position, we also assume a positional faithfulness constraint (Beckman 1998). With these faithfulness constraints in place, a general-purpose effort-minimization constraint like Kirchner 2001's Lazy constraint can successfully target just coda /k/. The examples below demonstrate this approach using OT tableaux.

Faithfulness to other places of articulation

(10)

/selamat/	Preserve(cor)	Lazy
a. ☞ selamat		*
b. selamaʔ	*!	

(11)

/tidak/	Preserve(cor)	Lazy
a. tidak		*!
b. ☞ tidaʔ		

The other shared assumption of these analyses involves the articulatory difficulty of the different debuccalization options. Kirchner 2001 provides effort values for various segments in weak positions. As such, we can assume that [k] is the most difficult, [x] is less difficult, and [h] is the easiest to articulate. Kirchner provides no effort values for [ʔ], so we will assume it is more effortful than [h] but still relatively easy.[6]

(12) Partial hierarchy of articulatory difficulty:
k > x > ʔ > h > ∅[7]

In the tableaux below, the constraint Lazy will only evaluate the final consonant (e.g. it will not penalize an initial [t]). Lazy assigns five violations to [k], four to [x], and so on down the hierarchy. Kirchner 2001 uses a similar reckoning system for Lazy in the first few chapters of the dissertation. In later chapters, he

[6]The assumption that [ʔ] is more effortful than [h] is compatible with the explanation of supplementary gestures given above. That is to say, if it were the other way around, then Lazy would push all segments to lenite to [ʔ]. Instead, we have Lazy pushing the segment toward [h], and one of the other constraints (perceptual faithfulness, dissimilation, or neutralization avoidance) disallowing [h] in favor of [ʔ]. Therefore, the ordering of [ʔ] > [h] is crucial to this analysis, because otherwise Lazy would do the same work as the other constraints. Alternatively, *OralGesture can take the place of Lazy without the use of the proposed hierarchy.

[7]This symbol is used to indicate the absence of a consonant.

uses a family of LAZY constraints that allows for other constraints to be ranked between LAZY constraints—like Kirchner, I will use the simpler system unless the family of constraints is called for.

(13) LAZY: Assign violation marks to articulatorily difficult segments. The number of violation marks received corresponds to the difficulty of the segment as proposed in the hierarchy in (12).

2.4.2 Perceptual faithfulness analysis

As explained in Section 2.2, in the default case of debuccalization laryngeal gestures are faithfully mapped to the output, but oral gestures fail to do so. Supplementary gestures are those laryngeal gestures that cannot be explained by this default mapping. For the perceptual faithfulness analysis, supplementary gestures are caused by faithfulness to *perceptual* aspects of the underlying consonant, not faithfulness to the *gestures* of the underlying consonant. Thus, the ranking of perceptual faithfulness over gestural faithfulness sometimes results in supplementary gestures. This can be seen in the tableau below.

(14)

/tidak/		PerceptDist < X	Lazy	Pres(LarGesture)
a.	tidak		***!*	
b.	tidax	*!	***	
c. ☞	tidaʔ		**	*
d.	tidah	*!	*	
e.	tida	*!		*

The highest ranked constraint is the perceptual faithfulness constraint, PERCEPTUALDISTANCE < X. This constraint demands that an input consonant be perceptually similar to its corresponding output consonant. It is violated when the two corresponding consonants are at a perceptual distance of X or greater, where X is a particular distance in the perceptual space. It is also violated when the consonant is deleted altogether (as long as the perceptual distance between the form with the consonant and the form without is sufficiently large, like comparing

[ak] with [a]).

(15) PERCEPTUALDISTANCE < X: Assign one violation when corresponding consonants are more perceptually distinct than X, where X is an experimentally determined distance in perceptual space.

The perceptual experiment in Chapter 3 will provide the evidence for a perceptual map where [k] and [ʔ] are relatively close together, and where [x] [h] and ∅ are relatively far from [k]. The perceptual faithfulness analysis relies on the perceptual map looking something like the map in Figure 2.1.

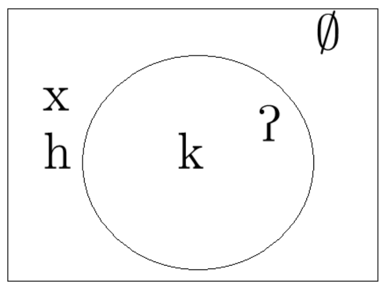

Figure 2.1: Proposed perceptual map for the perceptual faithfulness analysis of Indonesian. The circle shows similar sounds to [k], where the radius of the circle is X. The sounds [k] and [ʔ] are sufficiently close together to avoid a violation of PERCEPTUALDISTANCE < X, but substitution of any other sound will result in a violation of that constraint.

The next constraint in the tableau is LAZY, defined above in (13). With PERCEPTDIST < X leaving only candidates a and c, LAZY is able to decide in favor of the less effortful glottal stop. The generalization that comes from this constraint interaction is that, if you are going to lenite, lenite to the most articulatorily efficient consonant possible that is still somewhat perceptually faithful to the input consonant.

The lowest-ranked constraint is the gestural faithfulness constraint, called PRESERVE(LARYNGEALGESTURE). This constraint is violated when the laryngeal gestures of the input /k/ are not faithfully realized in the output. The winning candidate c violates this constraint due to the supplementary glottal stopping gesture. The deletion candidate e also violates the constraint, but this is unimportant because candidate e is out of the running due to a fatal violation of PerceptDist < X.

(16) Preserve(LaryngealGesture): Assign one violation when corresponding consonants do not share the same laryngeal gesture.

There is reason to believe that laryngeal gestures and oral gestures need to be treated differently at times, and that their respective faithfulness constraints must be freely rankable. In default debuccalization, laryngeal faithfulness is ranked higher than oral faithfulness. For example, Spanish /mas/ → [mah] violates the oral version of the faithfulness constraint, Pres(OralGesture), while still satisfying Pres(LarGesture). The opposite ranking is attested in other examples. When vowels undergo devoicing, this results in a violation of Pres(LarGesture), but it does not result in a violation of Pres (OralGesture) (e.g. Japanese /ʃita/ → [ʃi̥ta]). Processes that spread breathy voice or creaky voice also have the same general ranking as devoicing.

At this point, it might be useful to show a factorial typology of the constraints under discussion here. Table 2.2 demonstrates the various outcomes of constraint rankings. The table outlines the difference between default debuccalization (where Pres(LarGesture) is ranked high) and supplementary gestures (where it is ranked lower, and some other constraint influences the laryngeal gesture of the output form). Also included is the case of devoicing discussed above, and the general cases of deletion and surfacing faithfully.

Returning to the constraint interaction in Indonesian: The ranking of effort reduction over gestural faithfulness ensures that lenition of /k/ happens. The even higher ranking of perceptual faithfulness ensures that the lenited sound has the supplementary glottal stopping feature.

Default debuccalization	Pres(Lar) > Lazy > Pres(Oral)
Supplementary gestures in debuccalization	Percept Faith, Dissim, or Neu Avoidance > Lazy > Pres(Lar), Pres(Oral)
Devoicing	Pres(Oral) > Markedness > Pres(Lar)
Deletion	Lazy > Pres(Lar), Pres(Oral)
Fully faithful	Pres(Lar), Pres(Oral) > Lazy

Table 2.2: The factorial typology of relevant constraints

The perceptual faithfulness analysis is similar in many respects to Steriade 2001's P-map approach. The P-map hypothesis states that perceptual differences project contextual feature faithfulness constraints. Thus, if a segment is going to undergo an alternation in response to a markedness constraint, then it must do so in ways that minimize the differences in perception, as related to the environment that it occurs in. For example, in response to a constraint banning voiced obstruents word-finally, there is an option to devoice a final /b/, or to make it into a nasal. If [b p] is more similar than [b m] in that position, then this would project the constraint ranking Ident[nasal]/V_] >> Ident[voice]/V_], and would therefore prevent nasalization as a solution to the markedness constraint (example from Steriade 2001, p. 4). In many ways this is like the perceptual faithfulness analysis outlined above. The debuccalized form must satisfy markedness constraints (here, Lazy) while still remaining perceptually close to the underlying form. In terms of perception, the underlying form and output form may deviate only minimally from each other. However, instead of using perceptual data to project feature faithfulness constraints, PerceptDist < X holds directly over the segments themselves. The comparisons are still relativized to a context, but features are not used to mediate between the laboratory data and the formal analysis. This is advantageous for the present approach, because it is not entirely clear which acoustic or articulatory dimension is important here. The use of PERCEPTDIST < X allows us to be agnostic about which features are important, while still making use of perceptual similarity in the constraint system.

The general schema of debuccalization—LAZY over PRES(LARGESTURE)—along with constraints like PERCEPTDIST < X provides us with the ability to answer many of the typological generalizations given in Section 1.4. Glottalized consonants will debuccalize to [ʔ] in the default. On the other hand, when a non-glottalized sound debuccalizes to [ʔ], then there must be a reason for the supplementary stop gesture because the default debuccalization would be to [h] or [ɦ]. If the original segment is a stop, this might be sufficient reason for a supplementary gesture of glottal closure, due to the possible perceptual similarity of oral stops and [ʔ]. Assuming that they are more perceptually distant from [ʔ], fricatives and glides do not have as much of a reason to add a glottal closure. Thus, stops will sometimes debuccalize to [ʔ], but fricatives will not. If a voiced sound loses oral gestures, then the voiced sound would debuccalize to zero unless there is a reason for it not to. This helps explain why debuccalization frequently targets voiceless sounds—in the default case, voiceless sounds debuccalize, but voiced sounds delete ("debuccalize" to zero) unless there is a reason not to. Debuccalization's strong tendency to target voiceless sounds in turn explains why it targets obstruents, due to the dearth of voiceless sonorants. These predictions/explanations of the typology of debuccalization are outlined in Table 2.3.

2.4.3 Dissimilation analysis

In the dissimilation analysis, the constraint in charge of making sure the final /k/ and the preceding vowel are dissimilar enough is ADJACENTPERCEPTUALDISTANCE > Y, which is violated when two adjacent sounds are not perceptually dissimilar enough. Like the previously discussed PERCEPTDIST < X, this constraint uses perceptual distance values to assign violations.

Voiceless stop	Default debuccalization to [h] Non-default to [ʔ] to be more perceptually similar
Vless fricative	Default debuccalization to [h] Less reason to become [ʔ], if fricatives are less similar to [ʔ]
Glottalized C	Default debuccalization to [ʔ]
Voiced sound	Default debuccalization to ∅ Non-default to a laryngeal that is perceptually similar to the voiced sound

Table 2.3: The typology of debuccalization predicted by the perceptual faithfulness analysis

(17) AdjacentPerceptualDistance > Y: Assign one violation when adjacent consonants are more perceptually similar than Y, where Y is an experimentally determined distance in perceptual space.

There are significant differences with these two constraints, however. AdjPerceptDist > Y is a markedness constraint, comparing two adjacent output segments, while PerceptDist < X is a faithfulness constraint, comparing corresponding input and output segments. They also differ in the minimality / maximality requirement. AdjPerceptDist > Y requires two sounds to be more *distinct*, while the earlier constraint required two sounds to be more *similar*.

Like the perceptual faithfulness analysis, the dissimilation analysis of supplementary gestures makes certain predictions about perceptual distance. In order for the explanation to work, the relevant perceptual map must look something like Figure 2.2. Such a map would be consistent with a perceptual system that treats sounds with periods of silence in the speech stream as very distinct from uninterrupted sounds.

With AdjPerceptDist > Y ensuring that adjacent segments are dissimilar, final [ax] and [ah] are disallowed in this example. High-ranking Max also prevents deletion of the final /k/. Thus, the grammar chooses the most articulatorily easy sound, provided that it is not too similar to a vowel. In the example below, this

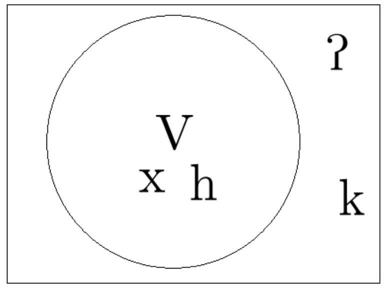

Figure 2.2: Proposed perceptual map for the dissimilation analysis of Indonesian. The circle shows similar sounds to a single vowel, where the radius of the circle is Y. The sounds [k] and [ʔ] are sufficiently far from a vowel to satisfy the constraint, but [x] and [h] are too perceptually similar.

sound is [ʔ].

(18)

/tidak/			AdjPerceptDist > Y	Max	Lazy
a.		tidak			***!*
b.		tidax	*!		***
c.	☞	tidaʔ			**
d.		tidah	*!		*
e.		tida		*!	

The dissimilation analysis does not provide the same type of typological explanation that the perceptual similarity analysis does. This is because the constraint AdjPerceptDist > Y does not directly care about the nature of the input segment—it will treat input stops, fricatives, and sonorants the same way as long as the output sounds are sufficiently distinct. While the constraint may be consistent with the facts, and while it is a plausible force at work, the dissimilation analysis is in some ways less interesting and less useful because of this lack of connection to the typology.

2.4.4 Neutralization avoidance analysis

In the previous analyses, some markedness or faithfulness constraint came into play to prevent /k/ from being realized as [h] or from deleting. In the neutralization avoidance analysis, the force that prevents this is a constraint against neutralization, *Merge.

(19) *Merge: No word in the output has multiple correspondents in the input. (Padgett 2003a, p. 57)

The neutralization avoidance analysis uses *Merge to prevent the merger of /k/ with that of /h/ or ∅. *Merge is a constraint from the version of dispersion theory found in Padgett 2003a. The tableau in (20), which evaluates multiple forms at the same time, demonstrates how this constraint works. The words /tak/, /tah/, and /ta/ are not necessarily lexical items of the language, but rather are abstract representations of possible words in Indonesian. The subscripts indicate correspondence between underlying forms of lexical items and their output forms. Candidate a realizes all forms faithfully, resulting in many violations of Lazy. Candidate b reduces /k/, but not as much as candidate c does. Candidates d and e result in the neutralization of phonemes, and therefore are ruled out by high-ranking *Merge. Candidate c is the winner because it reduces /k/ as much as it can without causing neutralization.

(20)

/tak$_1$ tah$_2$ ta$_3$/			*Merge	Lazy
a.		tak$_1$ tah$_2$ ta$_3$		****! *
b.		tax$_1$ tah$_2$ ta$_3$		*** *!
c.	☞	taʔ$_1$ tah$_2$ ta$_3$		** *
d.		tah$_{12}$ ta$_3$	*!	*
e.		ta$_{13}$ tah$_2$	*!	*

For this analysis to successfully account for the debuccalization phenomenon, *Merge must not care about /k/ merging with the loan-phonemes /x/ and /ʔ/; it must care only about /k/ merging with the native phoneme /h/ or deleting (merging with the absence of a consonant). For this reason, the tableau does not evaluate underlying forms like /tax/ or /taʔ/. The evaluation of such forms is the role of some other stratum of the grammar.

2.5 Falsifying the analyses

If these analyses are sufficiently useful to us as explanations, then some type of experimental or typological information should be able to verify or falsify them. Below are some ways in which the analyses could be falsified.

(21) To falsify the perceptual faithfulness analysis:
find an example of a debuccalization process where experimental results of perceptual distance are inconsistent with choosing a candidate with supplementary gestures.

(22) To falsify the dissimilation analysis:
find an example of a supplementary gesture that fails to make the resulting sound dissimilar to the surrounding sounds, by some consistent metric of similarity.

(23) To falsify the neutralization avoidance analysis:
find an example where the supplementary gesture in the debuccalized consonant fails to prevent a merger with another phoneme.

(24) To falsify any of the three analyses:
find an example where the OT constraints and ranking necessary to choose the supplementary gesture candidate are inconsistent with the rest of the grammar.

Although we might not have all the information available, we can evaluate some of the above falsifications using the typology of debuccalization given in the tables in Section 1.2.

First, a list of supplementary gestures from our typology:

(25) Sounds where a supplementary h-like gesture is added

a. Ainu ɡ > h
b. Ainu ɾ → h (if the sound is indeed a tap)

c. Canela-Krahô j > h

d. Pipil w → h

e. Sanskrit r > h

f. Yoruba j, w → h

(26) Sounds where a supplementary breathy voicing gesture is added

a. Florentine Italian g → ɦ

(27) Sounds where a supplementary glottal stop gesture is added:

a. all of the examples in Table 2.3, except those involving glottalized stops

For many of these examples, there are lines of inquiry that appear to be counter-examples to one of the competing analyses, but on closer inspection are compatible with all of them. For instance, when [j] > [h] in Canela-Krahô, it is not entirely obvious that they are perceptually similar. It seems like it is not perceptual similarity in terms of specific perceptual cues, but a drive to reduce [j] while still keeping it a consonant. Perhaps under the perceptual faithfulness analysis, [h] is perceptually similar to [j] only with respect to the consonant-vowel contrast, but is similar nonetheless. It could also be consistent with the dissimilation analysis (the leniting consonant avoiding becoming a vowel near other vowels), or with a drive to avoid neutralization with zero. The same general explanation can be told through the lenses of three different analysis.

A counter-example to the dissimilation analysis *might* exist with the stops → [ʔ] processes. This would hinge on the idea that [ʔ] is more like a vowel than [h]. The experiment in Chapter 3 investigates this and many other claims related to perceptual similarity, providing a way to evaluate the perceptual faithfulness and dissimilation analyses. The results from the experiment were inconclusive with respect to this falsification of the dissimilation analysis.

The neutralization avoidance analysis might be falsified by the stops → [ʔ] processes, if any of those languages also fail to have [h] as an allophone in the same position as the debuccalized stops. In terms of verification, the experiment in Chapter 4 provides evidence for the plausibility of neutralization avoidance as a bias in language.

The three competing analyses are formulated to be possible explanations for **all** debuccalization phenomena. It might be that some cases of debuccalization are best handled with perceptual faithfulness, and other cases are best handled with neutralization or dissimilation. While logically possible, it is in our interest to look at the strongest version of each analysis, and to choose a weaker version only in the face of evidence.

2.6 Other aspects

2.6.1 Positional Faithfulness

In Section 1.4 we discussed the approach to lenition advocated by Smith 2007. The conclusion was that some lenition processes occur in strong positions that are not also intersonorant, going against the predictions of neutralization-to-the-unmarked vs. sonority-increasing lenition. (An example of this is Kannada word-initial [p] > [h].) However, it should be noted that the present approach also fails to capture these odd cases. If lenition is due to effort minimization, and if positional faithfulness gives preference to preserving gestures in strong positions over weak ones, then we do not have a complete explanation for Kannada aspiration or the other examples. If effort minimization dictates that initial [p] should lose its bilabial gesture, becoming [h], then positional faithfulness should also allow debuccalization in weaker positions.

2.6.2 (In)compatibility with Harmonic Serialism

The overall analysis of debuccalization, and the three competing hypotheses regarding supplementary gestures, are all couched in single-stratum Optimality

Theory. Is this a crucial requirement of the analysis, or would it work just as well in Harmonic Serialism?

The Preserve(LarGesture), Preserve(OralGesture), and Lazy constraints are all conceivably compatible with Harmonic Serialism, as are the dissimilation and neutralization avoidance constraints. The caveat to this is the relationship between Lazy and incremental changes. Depending on the criteria for what defines a single incremental change, there is the possibility that Lazy would push the segment in question into a local minimum of articulatory ease. That is to say, there might be an easier form available, but it would require going through a more difficult form on the way there. If we consider the deletion of oral gestures, or the replacement of one laryngeal gesture with another, as a single operation, then this is unlikely to be a problem for analyzing debuccalization processes with these constraints, but it might need to be verified on a case-by-case basis.

The perceptual faithfulness constraint PerceptualDistance $<$ X, on the other hand, is completely incompatible with Harmonic Serialism. This constraint does not just militate against changes of a particular type from one point to another, rather it compares the total perceptual distance between the absolute input and the absolute output. In a derivational system with gradual changes, the constraint would make very different predictions, because it would allow quite distinct changes as long as each intermediate step is perceptually similar to the last. Such a system might be a valuable tool for other analyses, but it does not properly cash out the intuition that the perceptual faithfulness analysis rests on.

The conclusions related to Harmonic Serialism also hold true, I believe, for Serial Optimality Theory. Once the initial break with the underlying form is made, it does not matter how many derivational stages there are, because PerceptDist $<$ X no longer has access to information regarding that underlying form. The other constraints should be as compatible with Serial OT as they are with Harmonic Serialism.

2.6.3 Counter-arguments by Gess 2009

Running counter to the entire approach of this chapter is Gess 2009, who puts forth the claim that all of the cases of supplementary gestures in lenition from Kirchner 2001 are invalid. The intervocalic English case is argued to involve glottally reinforced [t], and so the glottal stop is not supplementary. Gess argues that West Tarangan /k/ is actually a uvular that results in a raising of the larynx, and this automatic raising is what causes the glottal stop, not a grammatically added gesture. The debuccalization of glides in Canela-Krahô and Pipil, are, according to Gess, probably morpho-phonemic or even morpho-syntactic category marker variants.

Gess 2009's argument for Florentine Italian is the following:

> "It is likely, therefore, that the actual production target for the voiced velar is one with a critical gesture rather than one with full closure, and that full closure is a grammatically controlled process associated with strong position (cf. claims regarding the Spanish alternation between stops and spirants . . .)." (p. 233)

Gess is using a two-level phonological system, with the output of the Lexical Phonology serving as the input to the Post-Lexical Phonology. In his system, the production target (the output of the Lexical Phonology) for the phoneme under question is /ɣ/. He argues that /ɣ/ includes some amount of glottal spreading in its gestural score, in order to keep continuous airflow and maintain frication. As such, the glottal spreading is already there, so it is not a supplementary gesture added in the lenition process.

While I do not dismiss Gess's claims out of hand, it appears that they are not compatible with other evidence. Gess may be on the wrong track when it comes to English, because the two processes, glottal reinforcement and glottal replacement, are not necessarily in a dependency relationship. As discussed in Section 1.3, there is evidence that the two processes are distinct (Milroy et al. 1994), and thus Gess's approach is less likely to be a comprehensive explanation. I believe many of his arguments are dependent on the two-level phonological system with production targets, which is not always compatible with a one-level phonology like the one used here. While a debuccalization process may not involve supplementary gestures in the Post-Lexical Phonology, the end result is a synchronic alternation that involves supplementary gestures overall. There is an alternation between an oral consonant and a laryngeal consonant, and they have different laryngeal gestures—using derivational levels does not prevent the two sounds from being related by alternation. In the end, we should keep the counter-arguments in mind, but there is still productive research to be done on the analysis of supplementary gestures.

Chapter 3

Perceptual experiment

3.1 Introduction

The experiment reported on in this chapter is a speeded *same-different* discrimination task, designed to investigate the perceptual similarity of various consonants in different prosodic environments. The reaction time results are used as a measure of perceptual similarity, which can then in turn be used to create perceptual maps through the use of multidimensional scaling.

Put another way, an experiment of this type allows for verification of the proposed perceptual maps in Figures 2.1 and 2.2, thereby supporting or contradicting the perceptual faithfulness and dissimilation analyses of supplementary gestures. The neutralization avoidance analysis cannot be evaluated in this manner, so a different experiment (Chapter 4) is necessary.

The comparisons in Table 3.1 and Table 3.2 summarize what is needed to evaluate the various claims about perceptual similarity.

The first prediction comes almost directly from Kirchner 2001. His analysis of Florentine Italian, as it is explained in the prose, depends on [g] being more like [ɦ] than it is like a vowel. Perceptual faithfulness also makes claims about other languages in the survey. The case study of Indonesian k-debuccalization, discussed in detail in Chapter 2, relies on the idea that [k] in coda position is more

Claim	Predicted Distance		
Florentine Italian g debuccalizes to ɦ rather than ∅ because g and ɦ are similar.	ga	ɦa	
	ga		a
Indonesian k debuccalizes to ʔ because ʔ is more similar to k than [h].	ak	aʔ	
	ak		ah
Cross-linguistically, fricatives fail to debuccalize to ʔ while stops are allowed to. This is because fricatives are less like ʔ than stops are.	ʔa	ka	
	ʔa		xa
	aʔ	ak	
	aʔ		ax
	aʔe	ake	
	aʔe		axe

Table 3.1: Perceptual distances predicted by the perceptual faithfulness analysis

Claim	Predicted Distance		
Indonesian k debuccalizes to ʔ because ʔ is more dissimilar to vowels than h is.	a	ah	
	a		aʔ

Table 3.2: Perceptual distances predicted by the dissimilation analysis

like glottal stop than it is like [h]. An analysis of English glottal replacement using perceptual faithfulness would, like Indonesian, require oral stops to be more like [ʔ] than they are like [h]. If this were not the case, then the perceptual faithfulness analysis would be called into question.

In the discussion of typological claims, it was proposed that fricatives do not debuccalize to [ʔ] because there is no reason for the supplementary glottal stop gesture. However, this proposal needs empirical support. Specifically, it was claimed that oral stops are more like [ʔ] than fricatives are like [ʔ], and the present experiment is capable of evaluating that claim.

The dissimilation analysis also makes claims about perception. In particular, this analysis is only compatible with supplementary glottal stop gestures if [ʔ] is less like surrounding vowels and consonants than [h] is. The basic idea of the dissimilation analysis is broad enough to include articulatory or feature-based metrics of similarity, but if we look at a version of the dissimilation analysis rooted in perception, then perceptual distance is useful in verifying it.

While there have been previous experiments in perceptual confusability (Singh et al. 1972, Shepard 1980, among many others), the sounds and environments used in these previous studies are not tailored enough to the present investigation. Thus, the perceptual similarity experiment described in this chapter provides the opportunity to evaluate the predictions of the particular analyses under issue, but the results can still be used by other researchers when they are interested in a subset of these consonants.

3.2 Methods

3.2.1 Stimuli production

To verify the predictions made above, a discrimination task was performed. The sounds that were compared were [k g x ɣ ʔ h ɦ ∅]. These sounds were chosen because they are representative of the manners of articulation found in the debuccalization alternations we are interested in. Place of articulation is not an integral part of the claims to be evaluated, so for simplicity, the consonants under investigation were all velars and laryngeals. Velars were chosen because they are a common target for debuccalization phenomena (see Section 1.4), and labials and coronals were avoided to prevent place of articulation from being a confounding factor.[1]

k	g	x	ɣ
ʔ	∅	h	ɦ

Table 3.3: Consonants that were compared

In the experiment, participants were presented with two sounds over head-

[1]Place of articulation was the focus of a previous study, mentioned in Subsection 1.4.2.

phones. They were then asked to respond as to whether they were the same sound or different sounds. Response (*same* or *different*) and response time were recorded. The sounds were embedded in nonsense carrier words, in order to provide different environments for comparison (CV, VC, and VCV). For example, in comparing [k] and [h] intervocalically, the participants heard [ake ahe], and they were required to say if they were the same sounds. Likewise, [k] and [ʔ] were compared as [ake aʔe]. If, in aggregate, participants are able to distinguish the first pair more easily than the second pair, then this provides evidence that the supplementary glottal stop gesture could be motivated by perceptual concerns.

The various conditions for the experiment are given in Table 3.4. There were 36 consonant comparisons to be made.[2] Of these comparisons, 28 were *different* comparisons, and 8 were *same* comparisons. In order to match the expectation that there will be similar numbers of *same* and *different* trials, the comparison types were weighted. *Same* comparisons were given three trials for every *different* comparison's one trial. Each sound comparison was given in both orders (e.g. [ake ahe] and [ahe ake]). Three prosodic positions were tested—word-initial onset, intervocalic, and coda position. This resulted in 312 total trials.

Num.	Condition
52	Comparisons of two sounds (28 *different*, 8 x 3 = 24 *same*)
2	Orders of comparison
3	Blocks by prosodic position: CV, VC, VCV
312	Total comparisons (21 minutes if averaging 15 responses/min)

Table 3.4: Conditions for the perceptual distance experiment

In terms of the words themselves (as opposed to the comparison trials), there were 23 different nonsense words that served as the stimuli for the experiment.

[2]This is given by the formula $n(n+1)/2$, because the experiment sometimes asks the participants to compare a sound to itself.

These nonsense words are shown in Table 3.5. In the experiment, comparisons were only made between sounds in the same row of Table 3.5 (e.g. [ake] and [age] were compared in a single trial, but [ake] and [ga] were never compared). The word [a] is a member of both the CV and VC blocks.

ˈka	ˈga	ˈxa	ˈɣa	ˈʔa	ˈha	ˈɦa	ˈa
ˈak	ˈag	ˈax	ˈaɣ	ˈaʔ	ˈah	ˈaɦ	(ˈa)
ˈake	ˈage	ˈaxe	ˈaɣe	ˈaʔe	ˈahe	ˈaɦe	ˈae

Table 3.5: Stimuli for the perceptual experiment

For the purposes of eliciting natural stimuli, the ideal situation would be to use a language that had all seven of the consonants to be compared, in onset, intervocalic, and coda position. Furthermore, this language would also need to allow onsetless syllables, both word-initially and in vowel hiatus configurations. Unfortunately, I have been unable to find such a language. The closest sound inventories I could find are Egyptian Arabic and Ukrainian, shown below. (Inventories come from Khalafallah 1969 and Pugh & Press 1999, respectively.) Note that there could be phonotactic restrictions or allophonic processes that would further complicate the use of these languages as a source of stimuli.

Egyptian Arabic			
k	g	x	ɣ
ʔ	∅	h	

Ukrainian			
k	g	x	
	∅		ɦ

Table 3.6: Language inventory subsets that are closest to the required sounds

Because none of the inventories in Table 3.6 matches the required sounds, a linguist with phonetic training and with a background in Arabic[3] was used for

[3]The linguist had experience with speaking both Modern Standard Arabic and Egyptian

the production of stimuli. Bisyllabic words were given initial stress, and stops in final position were released. The linguist who supplied the stimuli produced each word several times, avoiding list intonation and keeping duration approximately equivalent for all words within a row. The five best tokens were chosen as stimuli for the experiment. The first and last token in each group of elicitations were automatically excluded. Tokens were judged impressionistically based on similarity to other tokens of the same type, on how clear the pronunciation was (without being hyper-articulated), and on the clarity of the recording.

These tokens were then normalized for duration, pitch, and amplitude. Duration normalization was performed with the Audacity sound editing program, using the Change Tempo function from the SoundTouch sound processing library. Stimuli of the forms CV and VC were 0.490 seconds in duration, while the VCV stimuli were 0.700 seconds in duration. Next, the pitch values of the stimuli were edited using Praat (Boersma & Weenink 2010). A set of pitch contours was created (one for the CV and VC stimuli, another for the VCV stimuli). For both contours, the pitch was falling, starting at approximately 150 Hz then lowering to approximately 125 Hz. The stimuli were then resynthesized with the new pitch tracks, using the overlap-add resynthesis function. Finally, the stimuli were amplitude normalized with a Praat script, which used the scale function to normalize the sound files to the same peak amplitude.

3.2.2 Experiment

The experiment performed was a speeded *same-different* (AX) discrimination task. Participants heard two nonsense words through headphones, and were asked to judge if the words were the same or different. The experiment was created and run using the E-Prime computerized experiment software suite (Psychology Software Tools, Pittsburgh, PA).

In some of the *same* trials, two non-identical tokens of a single word were used.

Arabic.

For instance, token 3 of the word [ake] might be compared with token 5 of that word, and the desired response would be *same*. Because of the controlled nature of the stimulus production, and the subsequent duration, pitch, and amplitude normalization, various tokens of the same word were nearly identical. Still, the inclusion of this in the training block provided possible evidence to the participant as to what level of similarity constituted the *same* and *different* categories.

Participants were instructed to listen carefully to the two nonsense words. If they thought that the two words were the same, they were instructed to press the 1 button on the button box. If they were different, they were instructed to press the 5 button. The inter-stimulus interval (ISI) was 250 ms. The participants were instructed to respond as quickly as possible, preferably with a response time of less than one second (reckoning from the beginning of the second word). If the participant failed to respond within 2.5 seconds, then the trial automatically stopped with a suggestion to respond more quickly.

The reason for implementing this experiment as a speeded AX discrimination task was to attempt to access lower-level auditory processing. The goal was to investigate the raw perceptual differences between the various consonants. In order for the results to have cross-linguistic significance, and not just significance for English, the task should ideally key in to the part of the perceptual system that is language-independent. Babel & Johnson 2010 provide evidence that speeded AX tasks, like the present one, are able to avoid language-particular effects. Their finding is consistent with previous experimental evidence (Pisoni & Tash 1974) that argues for multiple levels of perception, with different information available at each level (acoustic attributes vs. phonetic features).

In terms of organization, the experiment consisted of four blocks. The first block was a training block, and the remaining three were testing blocks—one for each of the prosodic positions under investigation (CV, VC, VCV). The order of the testing blocks was randomized by participant.

In the training block, participants were given feedback as to whether they

Phase	Order	Block
Training	First	Training
Testing	Random by participant	Testing CV
		Testing VC
		Testing VCV

Table 3.7: Blocks in the perceptual distance experiment

answered correctly or not. They were also shown the response time for each trial and the overall average accuracy. There were eight comparisons in this block: two CV, three VC, and three VCV. Four of the comparisons were *same* and four were *different*. The *same* trials all compared non-identical tokens of the same word.

After the training block came the testing phase, which consisted of three testing blocks. Within each block, there were 104 comparisons, all given in random order. That is, each trial could differ with the next in terms of consonants to be compared and the order in which they were presented. In the testing trials, participants were not given accuracy feedback, but they were given response time feedback.

The experiment lasted approximately 25 minutes. Afterward, participants were asked to complete a form on language background information, as well as an exit survey related to the study.

3.2.3 Participants

There were 28 participants, all undergraduate students at the University of California, Santa Cruz. They received course credit for participation in the experiment.

The participants' ages ranged from 19 to 22 years, with a mean age of 20.4 years. Of the participants, 20 were women and 8 were men.

Most of the participants were native speakers of English. Some participants learned other languages from birth (Chinese, Japanese, Russian, Vietnamese, and Spanish), but they all reported the same proficiency or better proficiency in English. The other languages that participants reported having experience with were French, Hindi, Italian, Portuguese, and Swedish.

Participants reported no current speech or hearing disorders. One participant reported previous difficulty in the acquisition of [ɹ] until the age of six, but this was not an ongoing speech disorder.

3.3 Results

All of the statistical analyses were performed using the R statistics software package (R Development Core Team, 2011), and all of the visualizations based on the experimental results were also created using this software.

For the visualizations and statistical tests that follow, the important data points are response time values for correct *different* responses. There were 8,718 total responses, and of these, 4,694 were for *different* trials. From these responses, 4,408 were correct, giving an accuracy rate of 93.9% for *different* trials. Summary statistics are given below.

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
242	629	745	801	901	2468

8736	Total # of trials
4694	# of *different* trials
4408	# of correct *different* trials

Table 3.8: Summary statistics for experimental results (in ms, unless otherwise specified)

In Table 3.9, the accuracy of each comparison is given, in % correct. Each number reflects the accuracy of the comparison pooled across the factors participant, block, and order of presentation. Overall, performance was extremely high. The only comparisons that were below 90% were voiceless vs. voiced glottal fricative (58%), glottal stop vs. no consonant (74%), and [h] vs. no consonant

(87%).

∅/ʔ	∅/g	∅/ɣ	∅/h	∅/ɦ	∅/k	∅/x
74	98	95	87	94	98	99
ʔ/g	ʔ/ɣ	ʔ/h	ʔ/ɦ	ʔ/k	ʔ/x	g/ɣ
96	96	95	99	96	95	95
g/h	g/ɦ	g/k	g/x	ɣ/h	ɣ/ɦ	ɣ/k
96	97	97	97	97	98	97
ɣ/x	h/ɦ	h/k	h/x	ɦ/k	ɦ/x	k/x
92	58	98	95	98	94	98

Table 3.9: Accuracy (in % correct) by comparison type

In terms of accuracy between blocks, participants were correct on 92.7% of the trials in the CV block, 95.5% of the trials in the VC block, and 93.5% of the trials in the VCV block. Lenition processes tend to occur in weak positions and intervocalic positions (see section 1.4.1), but here there is no evidence that the participants fared better in the CV block than the others.

3.3.1 Multidimensional scaling

In the experiment, each of the segments in Table 3.3 was compared to each other segment. This allows for perceptual distance to be calculated for all eight segments. The reaction time data is interpreted as an indirect measure of perceptual similarity—the longer the reaction time for correct responses to a comparison, the more similar the two sounds in question are. The inverse of this value, then, is a measure of perceptual distance. This perceptual distance information was used as the basis for two scientific visualization techniques—multidimensional scaling (MDS) and hierarchical cluster analysis.

In order to obtain values for perceptual distance, the reaction time data was first divided by experimental block (the prosodic position that the sound comparison was made: CV, VC, or VCV). Within each block, the mean reaction time for each sound pair comparison was calculated, disregarding order of comparison or inter-speaker variability. The inverse of this mean was then used as the perceptual distance metric.

The MDS analyses were performed using R's *cmdscale* function. Two-dimensional MDS perceptual distance maps are given in Figures 3.1 – 3.3.

In the MDS map of the CV block, Figure 3.1, we see a clustering of the velars in the top half of the map (with a relatively high value of coordinate 2). The glottals cluster together as well (with low values for coordinate 2), being separated from [ɣ] by their lower coordinate 1 values. The absence of an onset consonant (∅) patterns in a similar way to the glottal onset consonants. The fricatives generally have a value of zero or greater for coordinate 1.

In the next block (VC), [x] patterns with the glottal consonants, with similar coordinates as [ɦ]. The difference between fricatives and stops seems to be captured by both coordinates, such that fricatives are in the lower-right half of the space, and the stops are in the upper-left (and with no coda consonant patterning with the stops). Finally, coordinate 2 might correspond to voicing, with lower values for voiced consonants.

The MDS map for the VCV block shows an extreme distancing of [k] from the rest of the sounds. The velar fricatives appear in almost the same place in the map, indicating that the voicing contrast was difficult to perceive in this environment. The fricatives appear in the upper-right half of the space, while the stops are in the lower-left half.[4]

Three-dimensional maps are shown in Figures 3.5 – 3.7. The conclusions reached regarding the two-dimensional maps are largely the same for the three-dimensional ones, except for the discussion of particular coordinate values and the interpretation of axes.

[4]Although this might appear to be different from the VC block, note that the direction of mapping coordinates is arbitrary. The three MDS analyses were performed independently of each other, so there is no reason to believe coordinate 1 in the VC block should not roughly correspond to coordinate 2 in the VCV block, for instance.

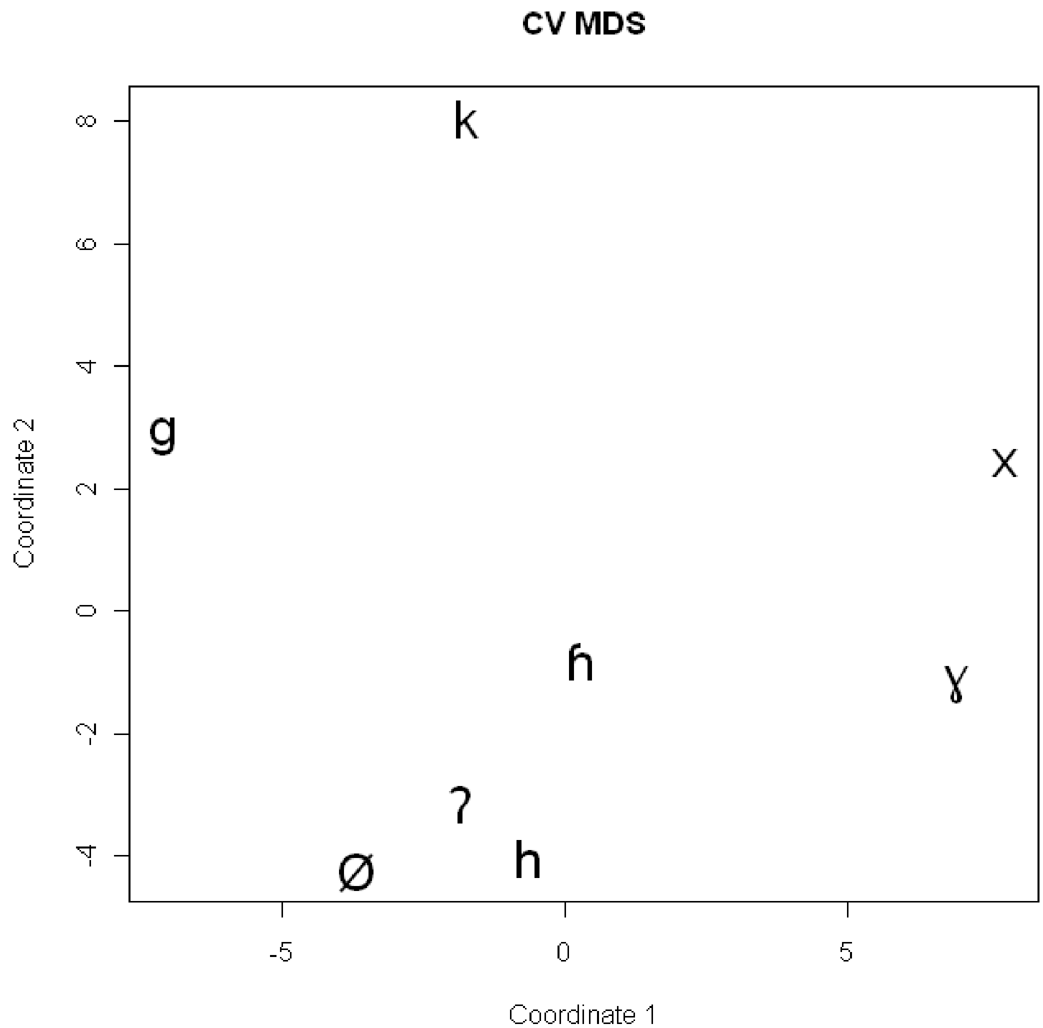

Figure 3.1: Multidimensional scaling plot, CV block

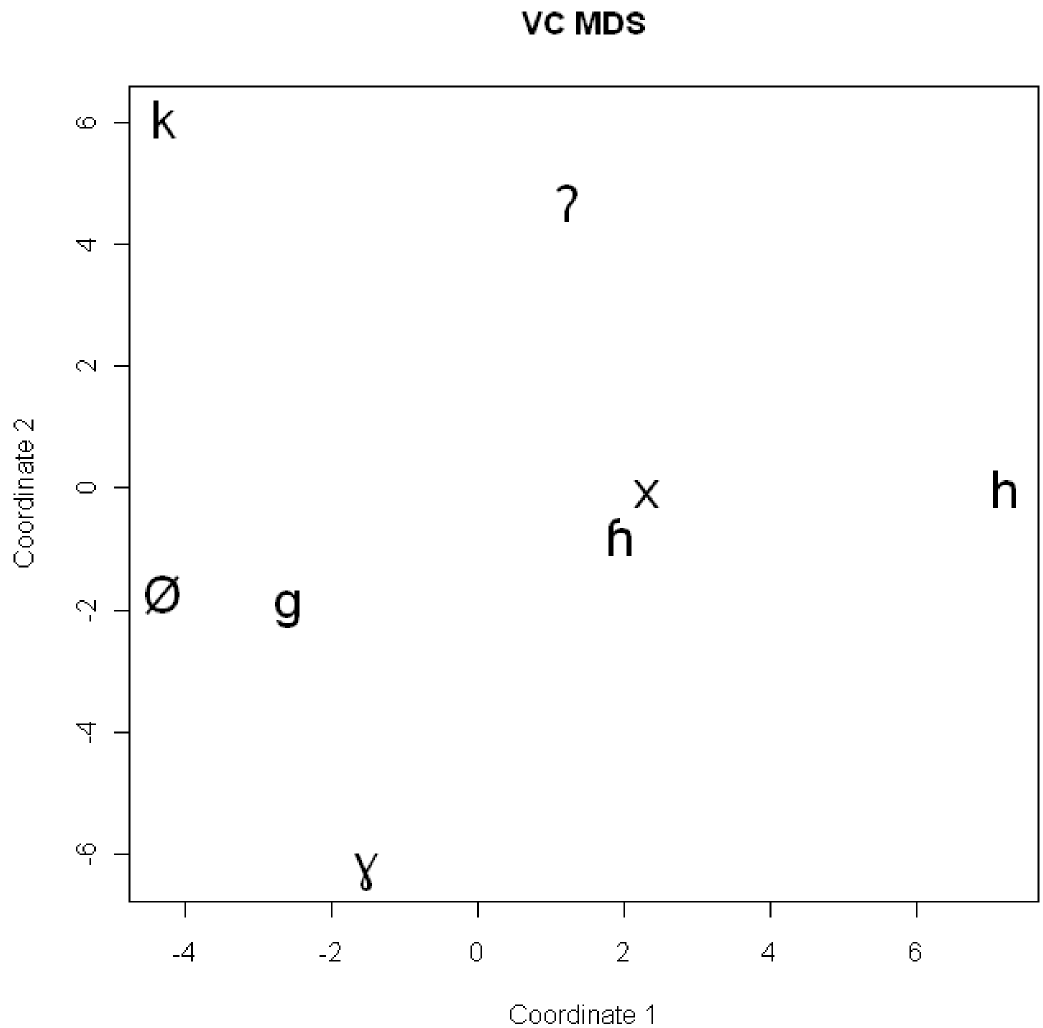

Figure 3.2: Multidimensional scaling plot, VC block

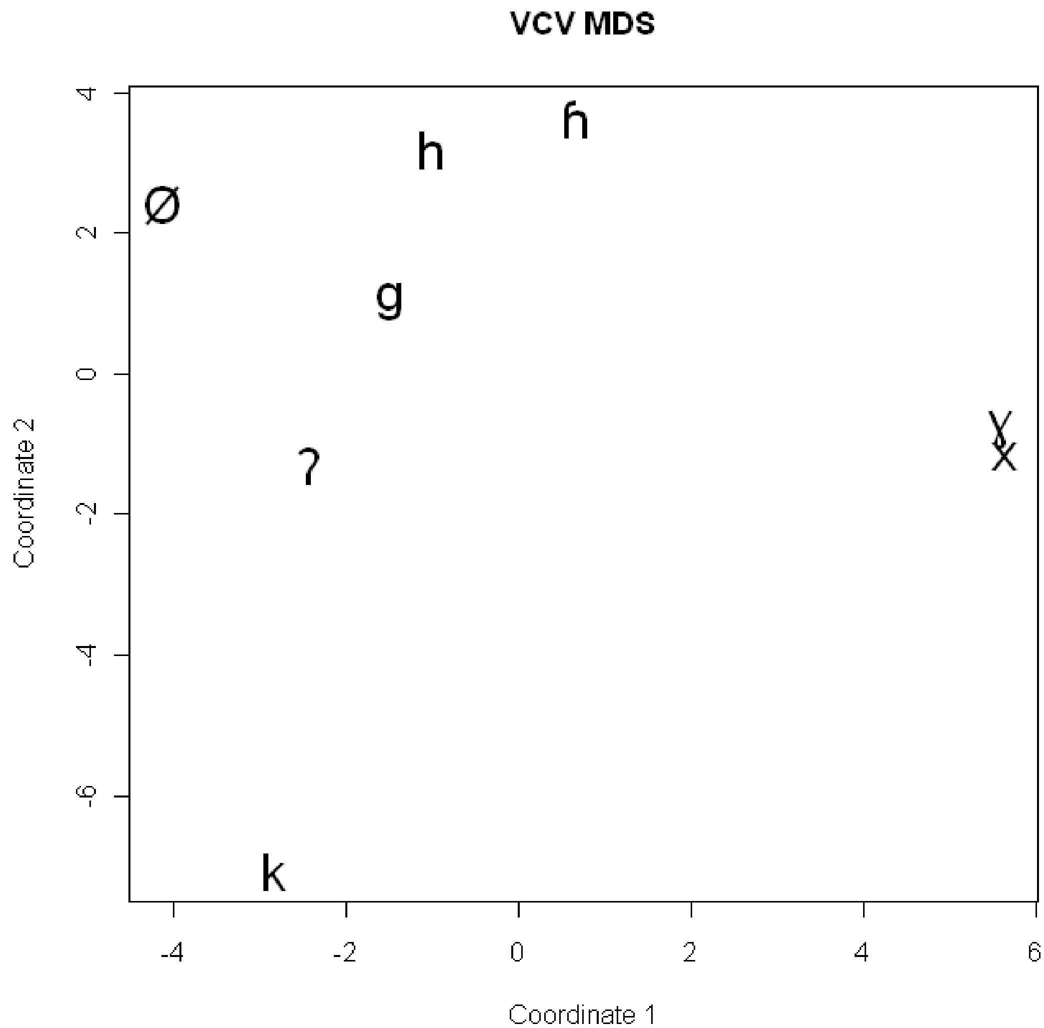

Figure 3.3: Multidimensional scaling plot, VCV block

The proposed perceptual maps from Chapter 2 are reproduced in Figure 3.4, so that comparisons can be made with the two-dimensional VC MDS map.

3.3.2 Hierarchical cluster analysis

The hierarchical cluster analyses were also performed in a similar way to the MDS analyses, with one analysis for each experimental block, using the perceptual distance data described above. The analyses were made with the *hclust* function in R, making use of the complete linkage clustering method.

The cluster for the CV block is remarkably similar to a traditional feature-based analysis of consonants. The velar stops, velar fricatives, and glottal fricatives all pair their voiced and voiceless counterparts at the lowest level. The higher-level hierarchical structure is less obvious, but the glottals (and no onset consonant) all form a category. The primary deviation from features is the grouping of the velar stops with the glottals to the exclusion of the velar fricatives.

The VC block did not give such an interpretable hierarchical clustering. None of the clusters appear to be organized in terms of place, manner, or voicing. Still, [k] and [ʔ] form a cluster by themselves, providing support for the Perceptual Faithfulness analysis of Indonesian k-debuccalization.

More like the first block, the VCV block provides a relatively recognizable cluster analysis. The velar fricatives cluster together, the glottal fricatives cluster together, and [ʔ] and no intervocalic consonant form a clustered pair. The placement of [k] and [g] in the hierarchy is less in line with our understanding of traditional features. Still, the hierarchical cluster analyses are quite similar to our intuitions about these sounds, especially considering they were created using only response times to the *same-different* discrimination task.

3.3.3 Differences in reaction time

The visualizations shown above provide some idea of how perceptually similar the various sounds are to each other, but they do not provide a statistically

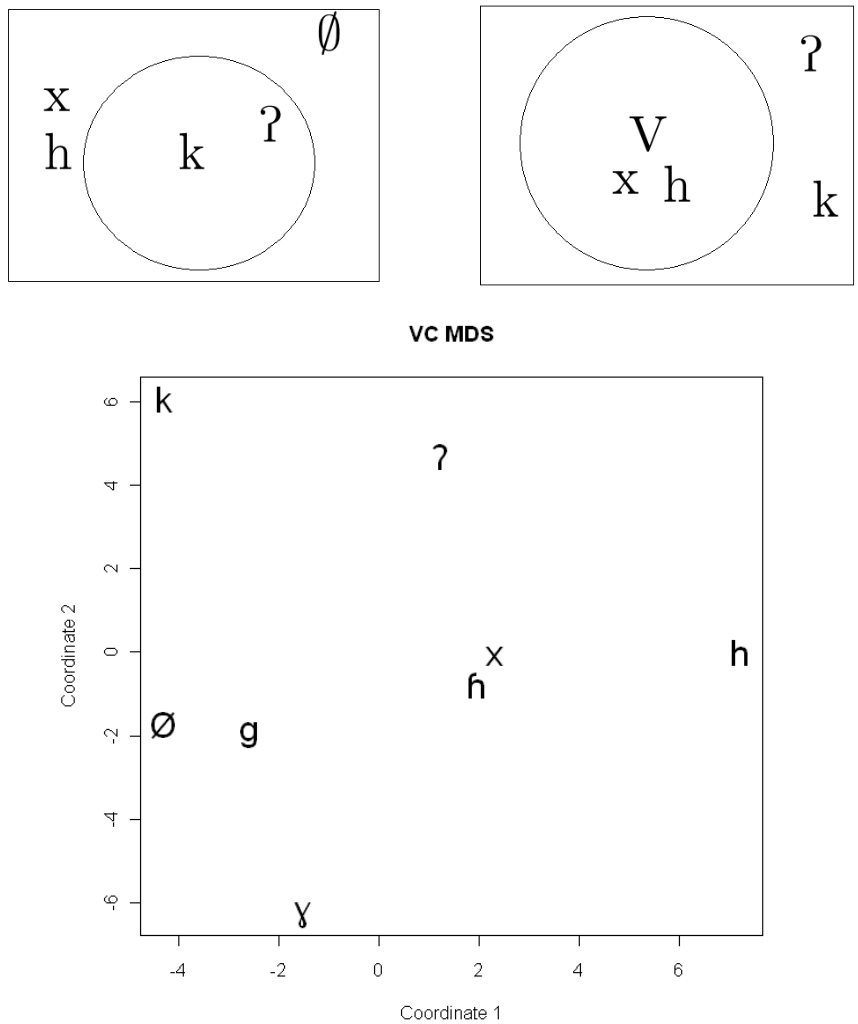

Figure 3.4: Proposed perceptual maps for the perceptual faithfulness (left) and dissimilation (right) analyses of Indonesian, reproduced from Chapter 2. The V in the right map is equivalent to ∅. Also included is the VC MDS map for comparison.

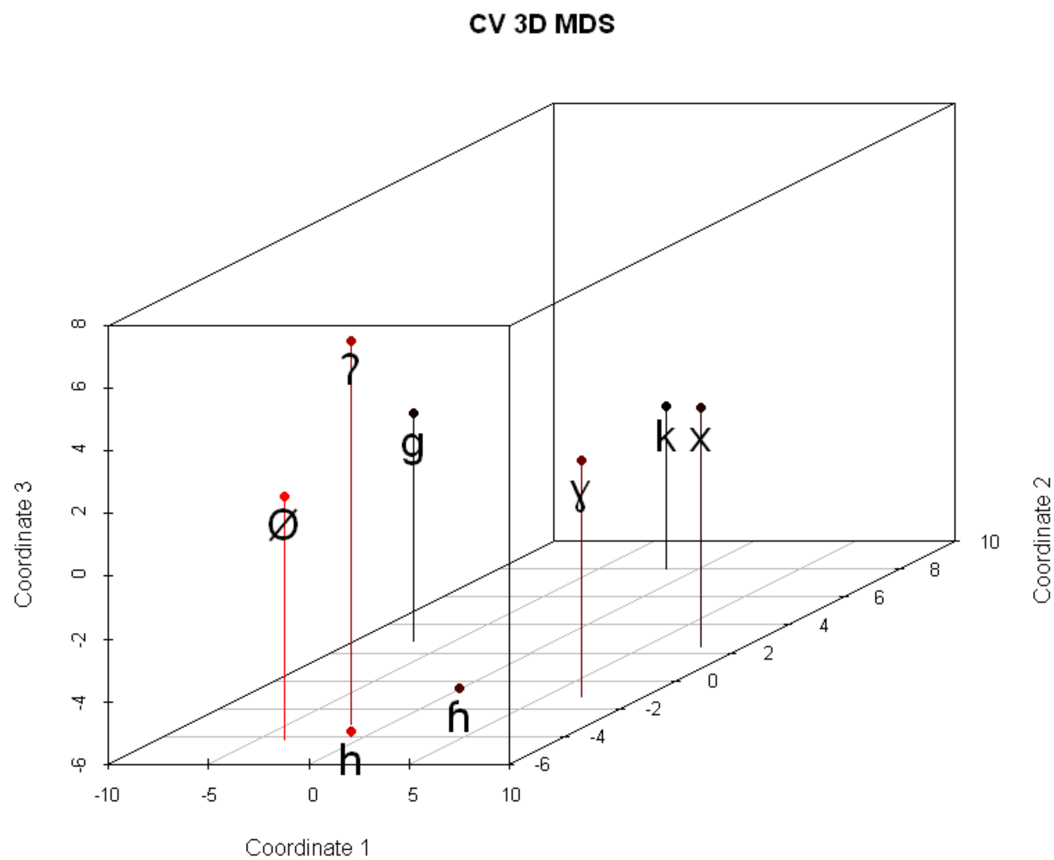

Figure 3.5: 3D multidimensional scaling plot, CV block

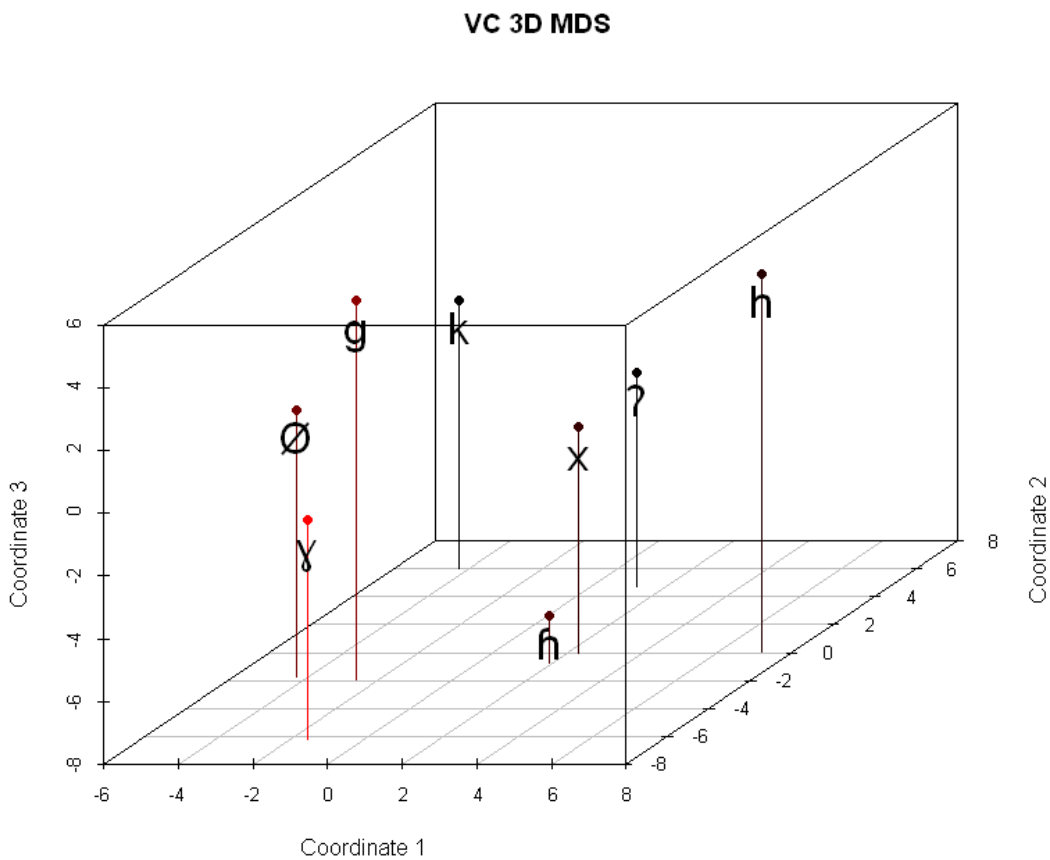

Figure 3.6: 3D multidimensional scaling plot, VC block

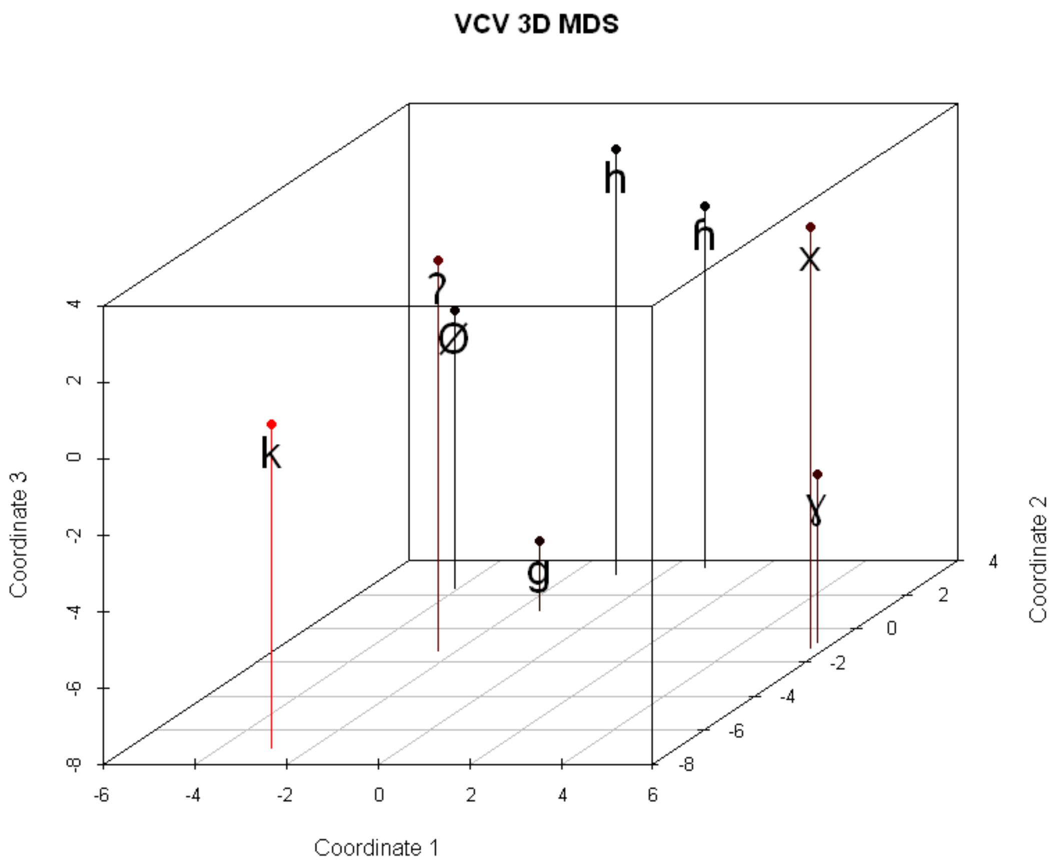

Figure 3.7: 3D multidimensional scaling plot, VCV block

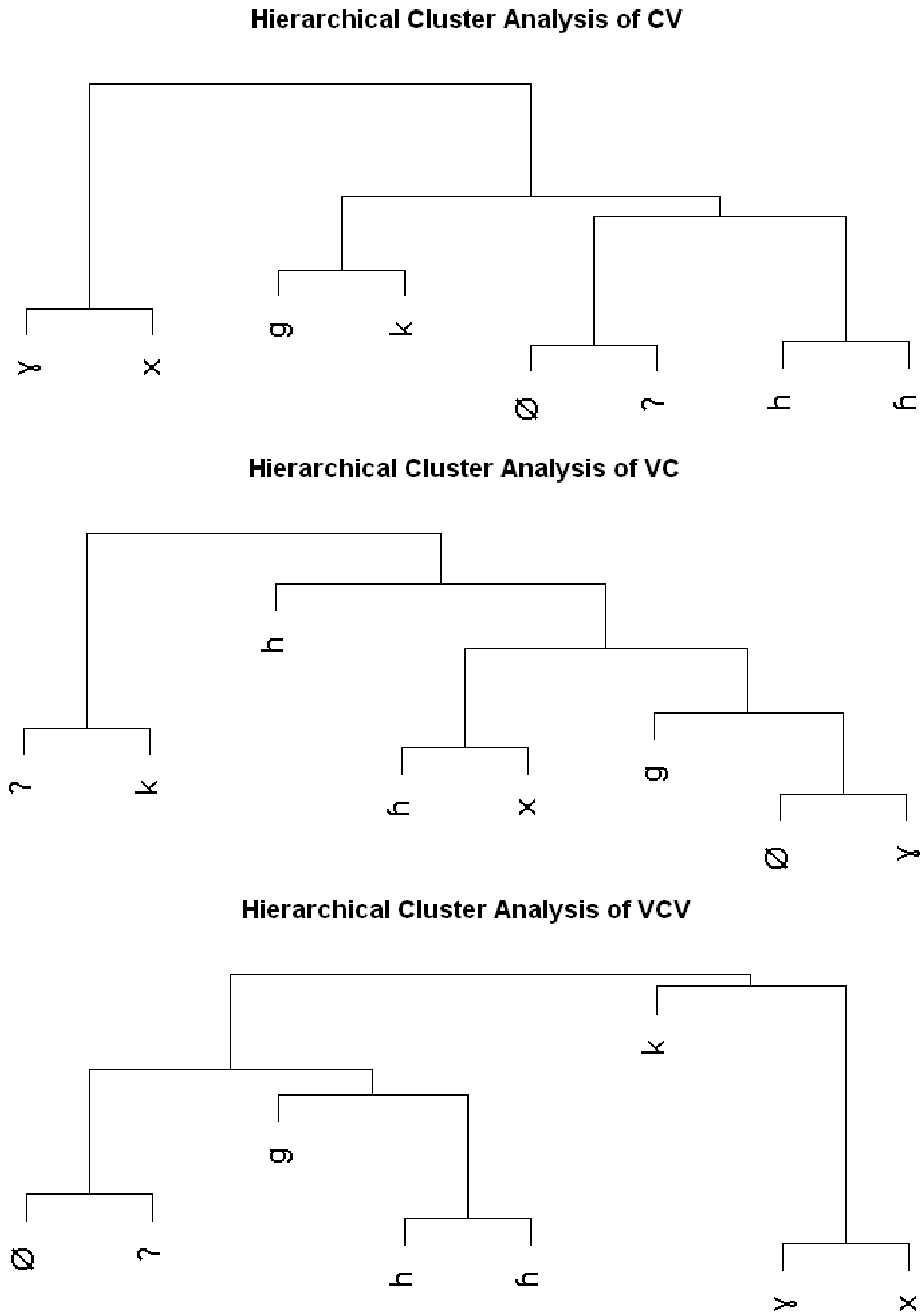

Figure 3.8: Hierarchical Cluster Analyses for the three experimental blocks

verifiable method of comparing perceptual distance. Such a comparison can be made using the response time data directly, as opposed to the derived metrics used earlier.

Recall the predicted differences in perceptual distance from Tables 3.1 and 3.2, repeated in Table 3.10 in a more summarized version.

Claim (with Experimental Block)	Predicted Distance		
Florentine Italian g debuccalizes to ɦ rather than ∅ because g and ɦ are similar. (CV)	ga	ɦa	
	ga		a
Indonesian k debuccalizes to ʔ because ʔ is more similar to k than [h]. (VC)	ak	aʔ	
	ak		ah
Cross-linguistically, fricatives fail to debuccalize to ʔ while stops are allowed to. This is because fricatives are less like ʔ than stops are. (All three blocks)	ʔa	ka	
	ʔa		xa
	aʔ	ak	
	aʔ		ax
	aʔe	ake	
	aʔe		axe
Indonesian k debuccalizes to ʔ because ʔ is more dissimilar to vowels than h is. (VC)	a	ah	
	a		aʔ

Table 3.10: Perceptual distances predicted by the perceptual faithfulness and dissimilation analyses

In order to test these predictions, t-tests were performed on the mean response times. Like before, response times for correct *different* responses were used. Before running the statistical tests, the response times were preprocessed using a natural log transform. The resulting distribution was normal, as seen in the histogram in Figure 3.9. The response times were separated into three data sets (one for each experimental block: CV, VC, and VCV). Then the mean log response time for each pair of comparisons for each participant was found, disregarding order of presentation. The test performed for each comparison was a two-tailed paired t-test. The input to the t-test consisted of two vectors, with each vector corresponding to a sound-pair comparison. The elements in each vector corresponded to the mean response time for a participant. Thus, the fourth element of

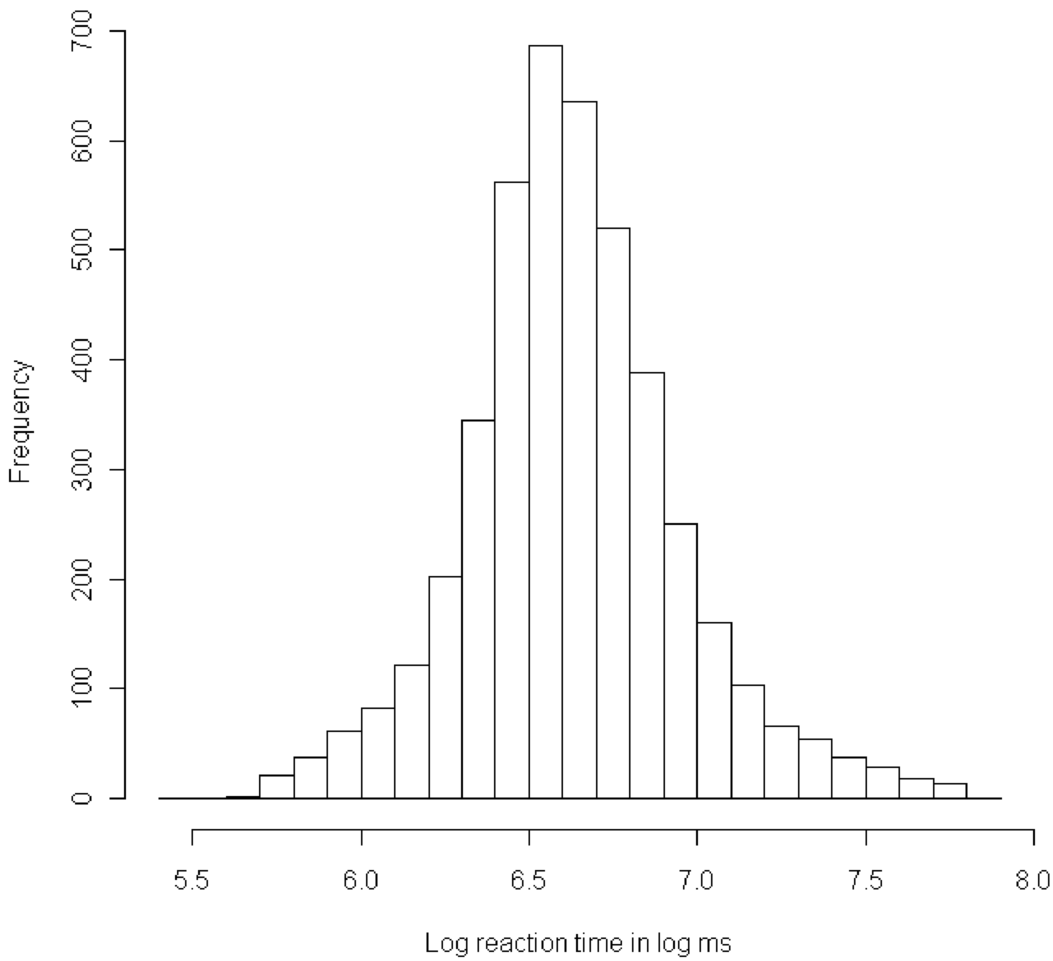

Figure 3.9: Histogram of log reaction times for correct *different* responses

the CV-[g ɦ] vector corresponded to the fourth participant's mean comparison of [ga ɦa] and [ɦa ga]. In this way, the two vectors that were compared were paired by participant.

Resulting p-values are given in Table 3.11, and possibly significant p-values are Bonferroni corrected for the six tests performed. The mean log response times are also shown graphically in Figure 3.10. A significant difference in response time was found for [ak aʔ] vs. [ak ah]. This difference is still significant after Bonferroni correction (corrected p = 0.0040), and the difference in mean response time is in the direction predicted ([ak aʔ] has longer mean response times than [ak ah]). The other comparisons did not reach significance.

Predicted Distance			Block	Result
ga	ɦa		CV	$p = 0.24$
ga		a		
ak	aʔ		VC	$p = 0.00066$
ak		ah		p x 6 = 0.0040*
ʔa	ka		CV	$p = 0.8$
ʔa		xa		
aʔ	ak		VC	$p = 0.74$
aʔ		ax		
aʔe	ake		VCV	$p = 0.82$
aʔe		axe		
a	ah		VC	$p = 0.26$
a		aʔ		

Table 3.11: Results of paired t-tests based on predicted differences

Therefore, the perceptual experiment verified one of the claims made by the perceptual faithfulness analysis of supplementary gestures. This was the claim related to Indonesian k-debuccalization—/k/ debuccalizes to [ʔ] instead of [h] because the pair [k ʔ] is more perceptually similar in coda position than the pair [k h]. We now have experimental evidence for the plausibility of this explanation. Abstracting away from the difference between coronal and velar place of articu-

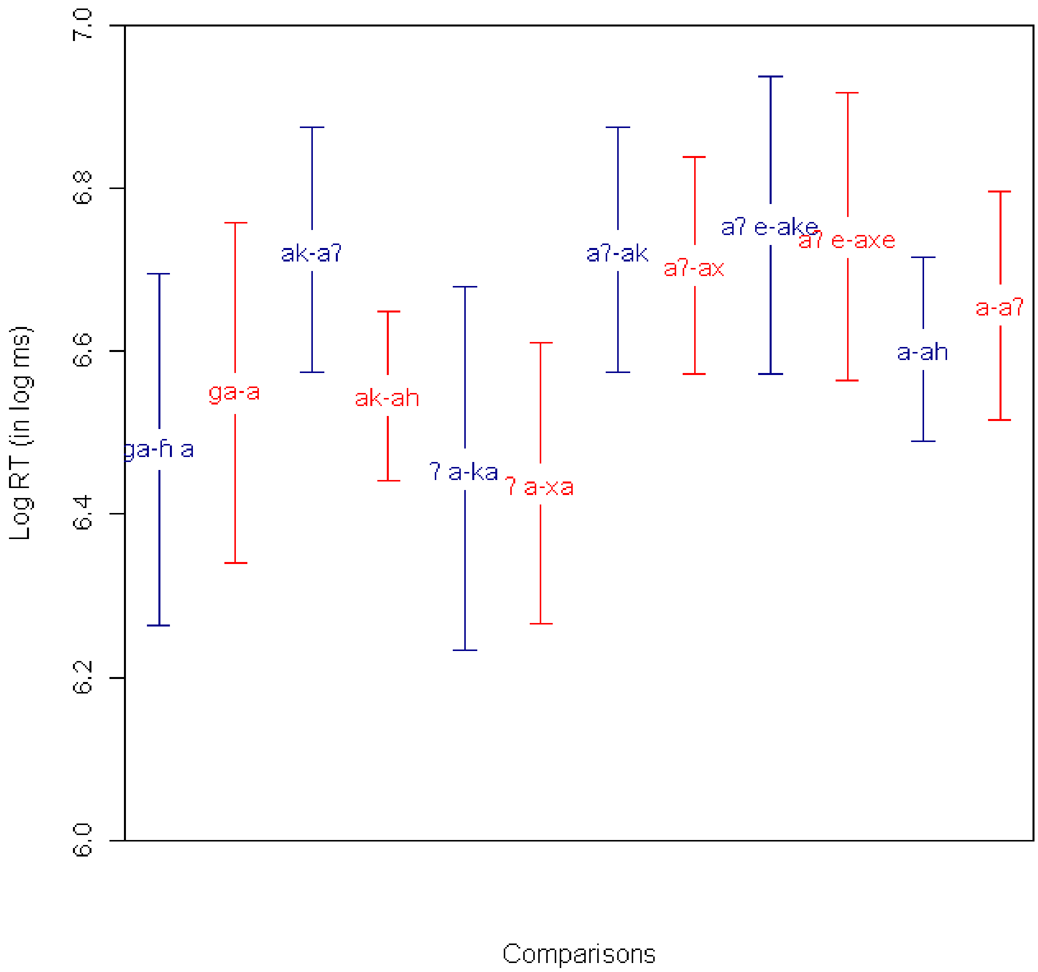

Figure 3.10: Log RT of paired sound comparisons from Table 3.11. 95% confidence intervals are shown, and the comparisons are given in the same order as the table. Note that the y-axis begins at 6.0 log ms, so differences in log RT may appear more dramatic than they are.

lation, this might also provide evidence for a perceptual faithfulness explanation for English glottal replacement. The other claims tested in the experiment were neither verified nor falsified by the results.

The response time comparison results are similar to the multidimensional scaling results. The predicted distances in Table 3.11 can be compared with the MDS visualizations from Figures 3.1 – 3.3. The only comparison that is obviously in line is the previously mentioned [ak aʔ] vs. [ak ah]. The other predictions are not obviously borne out in the MDS maps. There is some support, though, for the perceptual faithfulness view of the typology of debuccalization. Stop consonants will debuccalize to [h] because of articulatory similarity, and they will debuccalize to [ʔ] because of perceptual similarity (at least in the VC condition). Fricatives, on the other hand, only debuccalize to [h] in the typology. In each MDS map, [x] is no closer to [ʔ] than it is to [h]. Thus, all things being equal, fricatives would rather debuccalize to [h] because it is more articulatorily similar **and** because [ʔ] is no closer perceptually. The perceptual advantage of [ʔ] with respect to stops is not there for fricatives.

The dissimilation analysis claims that [a ah] is more similar than [a aʔ], and it appears like this could almost be the case in Figure 3.2. However, the only way to make it work is to weight the effect of Coordinate 2 as being more important than that of Coordinate 1. There is no external reason, though, to weight one dimension over another. In fact, the MDS function makes use of Euclidean distances—the algorithm responsible for reducing dimensionality measures distance between two sounds in a way that treats all resulting dimensions equally. Therefore, weighting one dimension more than another would lead to unsupported conclusions.

The dissimilation analysis of supplementary gestures could also be seen as making a rather strange typological prediction. If [h] is too vowel-like, then fricatives would also debuccalize to glottal stop, which is not attested in the typology. The perceptual results do not support this analysis—in all of the maps, [h] is no more vowel-like than [ʔ] is. Thus, the perceptual results are consistent with a dissimilation constraint, but they do not support dissimilation as a cause for supplementary gestures in Indonesian.

3.4 Conclusion

A speeded *same-different* discrimination task was performed to determine the perceptual similarity of various consonants in three prosodic positions, with the ultimate goal of finding evidence for or against the predictions made by the perceptual faithfulness and dissimilation analyses of supplementary gestures. The response times for correct *different* responses were used as an indication of similarity between segments. This data was then visualized using multidimensional scaling and hierarchical cluster analysis, and various t-tests were also performed. One of the predictions was verified by the results—that of the perceptual faithfulness analysis of Indonesian k-debuccalization.

The implications of this experiment are farther reaching than simply how they relate to supplementary gestures, though. The response time results were able to provide perceptual maps. Future research in perceptual phonetics can make use of these MDS maps, and additional claims about how perception relates to phonological systems can be substantiated by the maps, as long as they make use of a subset of the sounds utilized here. For example, if a researcher wanted to make a claim using the P-Map or some other perceptually-mediated phonological tool, and they were comparing [k g x ɣ], then the results here would be immediately applicable to such an analysis.

In the next chapter, I take a different experimental approach to evaluating analyses of supplementary gestures. Departing from perceptual similarity, the experiment in Chapter 4 focuses on a possible bias against neutralizing rules.

Chapter 4

Artificial grammar experiment

4.1 Introduction

Phonologists sometimes make claims about human language with reference to the idea of **neutralization avoidance**. The neutralization of two words into a single phonetic form, or more generally two phonemes into a single allophone, is sometimes argued to be avoided by the grammar. In particular, this type of push towards contrast preservation plays an important role in dispersion theory (as used in Flemming 1996, Padgett 2003b, Flemming 2004 and others). In the dispersion theory literature, constraints against neutralization are used alongside constraints that promote stronger contrastiveness and that maximize the number of contrasts possible. Neutralization avoidance was also outlined as a possible explanation of supplementary gestures in Chapter 2.

The goal of this chapter is to evaluate, by experiment, the claims of the neutralization avoidance analysis of supplementary gestures in debuccalization. In fact, we wish to find evidence for neutralization avoidance as a learning or perceptual bias in general, not just with respect to supplementary gestures. The experiment takes the form of an artificial grammar learning task. The question to be answered is: What effect does neutralization avoidance have on how learners acquire debuccalization processes? In short, if two languages are exactly the same except for whether or not a debuccalization rule is neutralizing, then the neutralization avoidance analysis predicts that the neutralizing rule should be harder to learn than the equivalent non-neutralizing rule.

The experiment described in this chapter was conducted to answer just this question. A small artificial language (called *Kasi*) was developed for the experiment. In this language, /k/ debuccalizes to [h] intervocalically (as it does in Florentine Italian fast speech). In order to incorporate neutralization avoidance, two dialects were necessary. Both dialects have the same debuccalization rule, but they differ in their lexical items. Dialect A does not have any morphemes containing /h/, so neutralization avoidance should not affect the debuccalization process in any way. Dialect B, on the other hand, has /h/ in its phoneme inventory, leading to our ability to probe the effect of neutralization avoidance.[1] If neutralization avoidance is an active part of grammar acquisition, then it stands to reason that, all things being equal, participants learning Dialect A would acquire the language more easily than those learning Dialect B.

The basic plan of the experiment is as follows. In the first part of the experiment, participants are auditorily presented with individual words (nouns and adjectives) in the artificial language. While they hear these words, they are visually presented with English glosses corresponding to the AG words. Here, the only difference between Dialect A and Dialect B is that Dialect A has no words containing /h/, while Dialect B does. Next, participants are presented with polymorphemic words which are the combinations of previously learned nouns and adjectives. Like before, these words are presented with English glosses. Finally, they learn five nouns that they have never encountered before. After the participants have been exposed to the language, they must decide whether new polymorphemic words (consisting of familiar adjectives and the newly-learned nouns) are part of the language or not. Because the two grammars are almost exactly alike except for the phonemic status of /h/, we expect any difference in accuracy or reaction

[1] Approximately 17% of the lexical items in Dialect B begin with /h/.

time to be solely due to the neutralizing vs. non-neutralizing nature of the rule.

4.2 Background on the artificial grammar paradigm

The artificial grammar paradigm has been in use for a long time, but there has been more renewed interest in it in the last decade or so. Perhaps the earliest artificial grammar example is Esper 1925, where participants attempt to learn the words for various shape/color combinations. Using a custom device like a slide projector, Esper trained the participants over the course of two months. He found that the participants learned more easily when the morphemes for color and shape correspond to syllables (so that syllable boundaries and morpheme boundaries coincide) and are in the English-like order of color + shape.

Pycha et al. 2003 look at the difference between formal simplicity and phonetic naturalness in rule learning. Two rules can be equally simple, but differ in terms of phonetic naturalness. In their experiment, they use backness harmony (formally simple, phonetically natural) and backness disharmony (formally simple, phonetically unnatural). They also include an arbitrary backness alternation (formally complex, phonetically unnatural). The language includes CVC stems (singular nouns) and an alternating plural -VC suffix. The choice of suffix depends on the group. The harmony and disharmony groups learned better than the arbitrary group, and the harmony group trended as better than the disharmony one. This suggests that both formal simplicity and phonetic naturalness may play a role in rule acquisition.

Carpenter 2010 is also interested in the role phonetic naturalness plays in learning. Her experiment investigates the interaction of stress and vowel height. Under the assumption that learners have a bias to stress low vowels, a natural rule would attract stress to low vowels, while an unnatural rule would attract it to high vowels. In that study, learners did indeed acquire the natural rule better than the unnatural one, providing evidence for the bias.

In Wilson 2006, a slightly different experimental procedure is used to find evidence for phonetic naturalness biases. Wilson looks at velar palatalization ([k] → [tʃ]) using a language game with an artificial lexicon. Participants are exposed to stimuli of the form "I say [kinə], you say [tʃinə]." Some participants see palatalization before [i], and others see palatalization before [e], but crucially they do not have evidence for whether the other front vowel conditions palatalization. All participants see that palatalization fails to occur before [ɑ]. Because [ki] and [tʃi] are more similar than [ke] and [tʃe], a biased learner might extend the [ke] palatalization rule to [ki] more easily than the other way around. In comparing the results to a computational model of categorization, Wilson found that a model that uses this phonetically-based bias was more like the experimental results than an unbiased model.

Moreton 2008 also investigates learning biases. He distinguishes between analytic biases (cognitive, in the mind of the learner) and channel biases, which are caused by transmission error in individual utterances, and can build up in a lexicon over time. Moreton argues that the bias in Wilson 2006 might be construed as either an analytic or channel bias, so Moreton compares two rules whose phonetic precursors are approximately the same—vowel height harmony (height-height), and vowel height depending on the voicing of the following obstruent (height-voice). Moreton finds that the height-height pattern is easier to learn (and more common cross-linguistically), providing support for an analytic bias independent of phonetic precursors.

The experiment in this chapter most closely follows the logic of Pycha et al. 2003 and Carpenter 2010. Two rules are compared, and if one of the rules is more easily learned or more easily used, then that provides evidence for it being more natural. In other words, if artificial grammar learners disprefer neutralizing rules then perhaps there is an analytic bias against them. The difference between their research and the present investigation is that here the basis of comparison is not phonetic naturalness. Both rules in the experiment are equally natural, by the very fact that they do the same thing (/k/ → [h] intervocalically). The question is whether non-neutralizing rules are analytically simpler than their neutralizing counterparts—easier to learn, easier to process, easier to use, etc. Thus, we are not claiming that neutralization is completely unnatural, in a phonetic sense or otherwise. In fact, there are many examples of neutralization in the world's languages. What we suggest is that perhaps there is some cost in learning and using neutralizing rules that is not there for non-neutralizing ones.

4.3 Methods

4.3.1 Stimuli production

Table 4.1 shows a few examples from the artificial language *Kasi*. Both Dialect A and Dialect B have the same debuccalization rule, where k → h intervocalically. They also have the same morphological rule for creating adjective-noun compounds—the adjective precedes the noun, and no additional morphological changes are made. The *difference* is that Dialect B has words with the phoneme /h/ (like *hinu* 'rabbit'), while Dialect A does not. This is because, in Dialect B, /h/ and /k/ are a source of contrast (because both can appear at the beginning of a word, regardless of the following vowel). In Dialect A, on the other hand, [k] can be found word-initially, and [h] can be found intervocalically (in derived environments), but the two sounds can never be a source of contrast due to their mutually exclusive environments.[2]

The artificial language involves productive morphology precisely so the debuccalization rule can be learned. In other words, participants must learn that there is an alternation between [k] and [h], not just a static phonotactic constraint

[2]The lexica of Dialect A and Dialect B are identical, except Dialect B has /h/ wherever Dialect A has /r/. While in Dialect B [h] is an allophone of both /h/ and /k/, in Dialect A [r] is an allophone of only /r/ and [h] is an allophone of only /k/. The exception to this is the adjective 'big', which is /ma/ in Dialect A but /ha/ in Dialect B.

Dialect A k → h / V__V no neutralization possible		Dialect B k → h / V__V neutralization with /h/ possible	
tu	'little'	tu	'little'
si	'red'	si	'red'
nusa	'shoe'	nusa	'shoe'
rata	'cat'	**hata**	'cat'
kani	'spoon'	kani	'spoon'
tunusa	'little shoe'	tunusa	'little shoe'
sinusa	'red shoe'	sinusa	'red shoe'
turata	'little cat'	**tuhata**	'little cat'
sirata	'red cat'	**sihata**	'red cat'
tuhani	'little spoon'	tuhani	'little spoon'
sihani	'red spoon'	sihani	'red spoon'

Table 4.1: Example words from both dialects of Kasi (Bold forms are different between the dialects.)

against intervocalic [k]. By providing the participants with multiple allomorphs of the same morpheme, they have the ability to internalize the rule without being directly told what the rule does.

For each dialect, 4 adjectives and 19 nouns were created. In the familiarization part of the experiment, the participants heard all of these words. They also heard all of the combinations of the first 14 nouns and all the adjectives. The testing of novel compound forms consisted of the final 5 nouns combined with all of the adjectives. As such, there were 25 monomorphemic words and 76 polymorphemic words, resulting in 101 distinct words in Dialect A and 101 distinct words in Dialect B. Additional ungrammatical forms were created for testing purposes.

The stimuli were synthesized using the MBROLA speech synthesis software package (Dutoit et al. 1996). A computer program concatenated the morphemes according to the adjective + noun rule, and then used the transcription of each word (monomorphemic and polymorphemic) to create the input to MBROLA. Using the Classical Latin male voice (`la1`), MBROLA made sound files for all of the words in Kasi, including the ungrammatical testing words. Speech synthesis

Gram.	Correct application Correct non-application	Intervocalic /k/ $\rightarrow$ [h] Surface mapping of non-/k/ cons.
Ungram.	Overapplication Underapplication Fortition (Dialect B only)	Intervocalic non-/k/ C $\rightarrow$ [h] Intervocalic /k/ remains unchanged Intervocalic /h/ $\rightarrow$ [k]

Table 4.2: Rule-application types

was chosen over natural speech in order to be as consistent and regular as possible, both within dialects and between Dialect A and Dialect B. All of the consonant sounds were 95 ms in duration. The final vowel of each word (which was stressed) was 245 ms in duration and had a falling pitch contour. The other vowels were 130 ms in duration and were relatively flat in pitch. All adjectives were of the form CV, and all nouns were of the form CVCV, for both dialects.

The compound forms that were heard were of five rule-application types. The two grammatical rule-application types were correct application of the rule (intervocalic /k/ $\rightarrow$ [h]), and correct non-application of the rule (when there is no intervocalic /k/ for the rule to apply to). The ungrammatical rule-application types were overapplication (making a non-/k/ intervocalic consonant into [h]), underapplication (failing to change an intervocalic /k/), and fortition (turning an intervocalic /h/ into [k]). Because Dialect A did not have the phoneme /h/, fortition was not a possible rule-application type for those participants, and they only encountered ungrammatical forms as a result of overapplication and underapplication of the rule. These rule-application types are outlined in Table 4.2.

4.3.2 Experiment

The experiment was created and run using the E-Prime computerized experiment software suite (Psychology Software Tools, Pittsburgh, PA). It included eight blocks—four familiarization blocks and four testing blocks. The block types were interleaved, as shown in Table 4.3. The first familiarization block introduced

Block	Type	Description	
1	Fam	Adjectives	(4 adj x 3 trials each = 12 trials)
2	Test	Adjectives	(4 adj x 6 trials each = 24 trials)
3	Fam	Nouns	(14 nouns x 3 trials each = 42 trials)
4	Test	Nouns	(14 nouns x 6 trials each = 84 trials)
5	Fam	Compounds	(4 adj x 14 nouns x 1 trial each = 56 trials)
6	Test	Compounds	(4 adj x 14 nouns x 2 trials each = 112 trials)
7	Fam	Nouns	(5 nouns x 3 trials each = 15 trials)
8	Test	Compounds	(5 nouns x 4 adj x 2 trials each = 40 trials)

Table 4.3: Description of experimental blocks
(Fam = Familiarization, Test = Testing)

the participants to four adjectives of the language. The sound file corresponding to each adjective was played to them over headphones. At the same time, an English gloss was provided on the screen. They were instructed to listen to the words and to try to remember them. After hearing the word, the participants pressed a button on the button box to continue to the next word. The second block tested these adjectives. The participant saw an English gloss on the screen, and heard a (now somewhat familiar) adjective from the previous block. The task was to decide if the Kasi adjective and the English word on the screen matched. Blocks 3 and 4 were very similar, except they used nouns instead of adjectives, and they had more trials.

Block 5 provided participants with every combination of previously heard adjectives and nouns. It is in this phase that they encountered evidence for the morphological rule (adjective + noun) and the phonological rule (/k/ $\rightarrow$ [h] intervocalically). Each trial proceeded in the following way: First, the participant was shown an English gloss of an adjective on the left side of the screen, and the corresponding Kasi adjective was played over headphones. Then, the English gloss of a noun was displayed on the right side, and the corresponding Kasi noun was played auditorily. Finally, they were shown the combination of these English words in the lower center of the screen, and they heard how these words were combined in Kasi. Like before, the participants pressed a button to continue to the next trial.

Block 6 tested participants on the phonological rule. The task remained the same as earlier testing blocks (yes-no grammaticality), but instead of testing memorization of lexical items, it tested rule application. Everything was the same as block 5 (with the adjective, noun, and combination shown on the screen), except the Kasi compound that was played over headphones was not guaranteed to be correct. The participant responded as to the correctness of the Kasi compound. Block 7 was identical to the earlier noun familiarization block, but with five new nouns. Block 8 was similar to the earlier compound testing block, but with novel compound forms (using the adjectives from block 1 and the nouns from block 7).

Feedback was given on testing blocks 2, 4, and 6. Participants were told whether or not they were correct after each trial, and they were also shown their reaction times and their current overall accuracy.[3] No feedback was given in block 8, as this block was intended to test how well participants could extend their knowledge to novel forms, and feedback would have provided evidence for later trials within the block.

As Carpenter 2010 argues, artificial language learning with feedback is quite similar to second-language acquisition found in the classroom. More than simply being similar on the surface, Carpenter claims that the task may plausibly employ the same learning system that adults use for second-language acquisition—specifically, a combination of a general learning mechanism and a language particular mechanism. This is an advantage to the present study, as it makes the results more applicable to natural human language learning than they might otherwise be. While it would be even better to model the task on first-language acquisition, such a system would be impractical. It would involve exposing children to the artificial language at an age before the critical period ends, and it would be extremely difficult to teach the sound-meaning pairings of morphemes and words

[3]Blocks 2 and 4 were combined in calculating overall accuracy in participant feedback, but the value was reset for block 6 for experiment programming reasons.

to the participants in a timely fashion.[4]

The experiment lasted approximately 30 to 40 minutes. After the experiment, participants were asked to complete a language background survey and an exit survey about the study.

4.3.3 Participants

There were 28 participants, all undergraduate students at the University of California, Santa Cruz. They were assigned to a dialect randomly, with 14 participants learning each dialect. They received course credit for participation in the experiment.

The participants' ages ranged from 18 to 25 years. The overall mean age was 20.2 years, with Dialect A being slightly younger (19.9 years vs. 20.5 years, not significantly different by t-test). Of the participants, 18 were women and 10 were men, with each dialect having the same number (9 women, 5 men).

Most of the participants were native speakers of English (and sometimes other languages). All but one (A05) had some experience with non-English languages. Two participants (A04, A10) in the Dialect A group were non-native speakers of English. Participant A04 spoke Russian natively, and participant A10 spoke Punjabi natively, but they were both proficient in English.

4.4 Results

According to the neutralization avoidance analysis, Dialect A (which has a non-neutralizing debuccalization rule) should be easier to learn than Dialect B (which has a neutralizing debuccalization rule). Participants exposed to Dialect A should have a greater frequency of correct responses, higher d′ (sensitivity) scores, and they should also show quicker reaction times.

[4] Recall that the sound-meaning pairings are necessary to demonstrate the alternation, as opposed to just the phonotactic generalization.

These three metrics of learning (or application of learning) are discussed in the following subsections. Most of the results are uninformative with respect to the neutralization avoidance hypothesis, but some of the accuracy and sensitivity scores lend support to the hypothesis.

Below, we only analyze the responses for block 6 (test - familiar compounds) and block 8 (test - novel compounds). This is because block 2 (test - adjectives) and block 4 (test - nouns) are not predicted to be different between the two dialects at all—at that point in the experiment, the neutralizing or non-neutralizing rule is not even encountered. Accuracy and sensitivity scores for these blocks (by participant) are shown, but only for comparative purposes. Furthermore, the odd-number blocks were familiarization blocks, and so these responses are not analyzed or shown.

All of the statistical analyses were performed using the R statistics software package (R Development Core Team, 2011), and all of the visualizations were also created using this software.

4.4.1 Accuracy

Table 4.4 shows the accuracy (in percent correct) for all participants in all testing blocks. Table 4.5 gives the average accuracy for each dialect.

Figures 4.1 and 4.2 show the accuracy of the relevant blocks in graphical form. The first figure is for Dialect A, and the second is for Dialect B. The accuracy scores of each participant are provided, with block 6 on the left and block 8 on the right.

In terms of statistical analysis, a t-test was performed on the combined percent correct for blocks 6 and 8. This t-test failed to show a difference ($t = 0.5166$, $df = 24.249$, p-value $= 0.6101$). (See Section 4.4.3 for a modification of this test that demonstrates a significant difference.)

	Bl 2	Bl 4	Bl 6	Bl 8	All Blocks
A01	83	83	95	98	90
A02	79	77	96	98	87
A03*	79	76	83	52	73
A04	79	89	96	100	91
A05*	79	58	71	50	65
A06	92	75	95	100	90
A07	67	80	97	92	84
A08	71	74	94	100	85
A09#	62	63	53	55	58
A10	100	96	95	90	95
A11	92	88	92	95	92
A12*	71	50	93	50	66
A13*	71	68	77	65	70
A14	92	70	96	90	87
B01	58	94	88	80	80
B02*	33	68	88	50	60
B03	71	82	81	85	80
B04	42	71	88	95	74
B05#	54	86	59	50	62
B06*	83	93	90	50	79
B07	92	89	76	75	83
B08	96	96	95	100	97
B09	100	92	95	100	97
B10	75	67	89	98	82
B11	88	75	96	90	87
B12	79	85	80	88	83
B13	92	80	71	82	81
B14	75	75	87	100	84

Table 4.4: Values of percent correct for all participants in all testing blocks (Asterisk marks *memorizers*, pound sign marks *non-learners*, and all others are *learners*.)

	Bl 2	Bl 4	Bl 6	Bl 8	All Blocks
Avg for A	80	75	88	81	81
Avg for B	74	82	84	82	81

Table 4.5: Average percent correct for Dialects A and B in all testing blocks

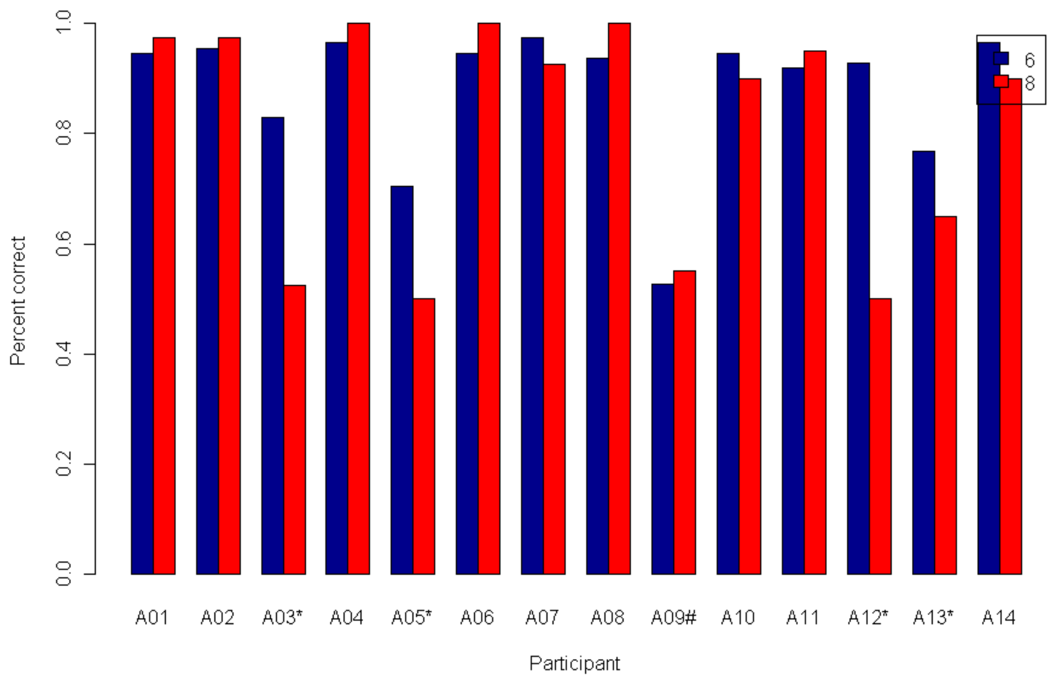

Figure 4.1: Accuracy by participant in blocks 6 and 8, Dialect A

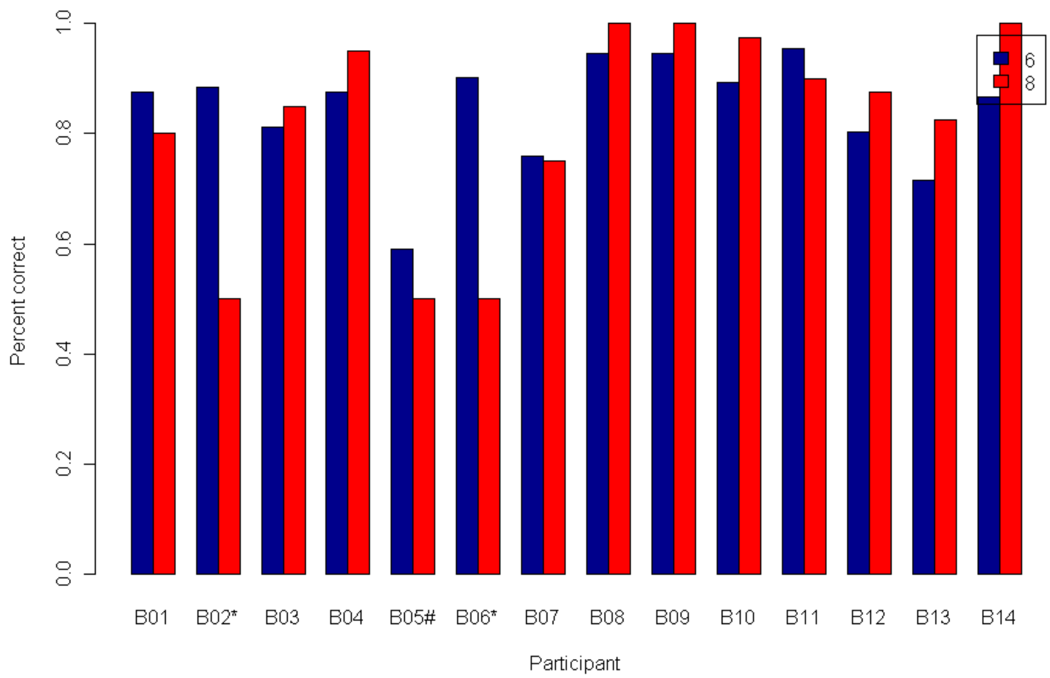

Figure 4.2: Accuracy by participant in blocks 6 and 8, Dialect B

4.4.2 Sensitivity

The sensitivity metric used here is d′. The d′ values were calculated using the number of correct responses on grammatical trials (hits), correct responses on ungrammatical trials (correct rejections), incorrect responses on grammatical trials (misses), and incorrect responses on ungrammatical trials (false alarms). These categories are schematized in Table 4.8. The values were modified slightly—if any of the categories were empty, then the count was changed from zero to one to assist in d′ calculations. The `sensR` package was then used to obtain d′ values.

The values of d′ obtained for all participants in all blocks are given in Table 4.6 and Table 4.7. Like for accuracy, a t-test was performed comparing the sensitivity (in d′) for Dialect A and Dialect B for blocks 6 and 8. This test failed to show a difference (t = 0.9882, df = 24.052, p-value = 0.3329).

4.4.3 Learner type: learners, memorizers, and non-learners

Using the results of d′ analysis, we can divide the participants into three categories of learning. There are those who learned the pattern in block 6 and extended it to novel forms in block 8. If participants received a d′ score of 1.00 or greater for both relevant blocks, then they were included in this category, called *learners*. Other participants were able to perform well in block 6, but were unable to extend this knowledge to novel forms. These *memorizers* received d′ scores of 1.00 or greater in block 6, but lesser than 1.00 in block 8. Finally, there are the *non-learners*, who did not perform well in either relevant block. If a participant received d′ scores of less than 1.00 in blocks 6 and 8, they were placed in this category. In the table of d′ scores (Table 4.6), memorizers are marked with an asterisk, non-learners are marked with a pound sign, and the learners are unmarked.

Although logically possible, there are no participants who performed poorly in block 6 (with a d′ below 1.00) and well in block 8 (with a d′ at 1.00 or above).

	Bl 2	Bl 4	Bl 6	Bl 8	All Blocks
A01	2.06	1.97	3.27	3.31	2.64
A02	1.64	1.51	3.45	3.31	2.43
A03*	1.81	1.55	2.08	0.02	1.57
A04	1.81	2.49	3.71	3.34	2.98
A05*	1.81	0.47	1.20	0.00	0.88
A06	2.77	1.43	3.27	3.34	2.46
A07	0.97	1.68	3.72	2.93	2.37
A08	1.11	1.31	3.26	3.34	2.24
A09#	0.95	0.76	0.14	0.25	0.39
A10	2.85	3.65	3.27	2.68	3.35
A11	2.77	2.38	2.85	2.95	2.71
A12*	1.11	0.00	2.93	0.00	1.11
A13*	1.11	0.93	1.66	0.80	1.16
A14	2.77	1.15	3.61	2.68	2.27
B01	0.46	3.13	2.30	1.68	2.14
B02*	-0.86	1.00	2.41	0.00	1.18
B03	1.11	1.89	1.91	2.12	1.84
B04	-0.42	1.14	2.47	3.29	1.65
B05#	0.29	2.19	0.55	0.00	0.94
B06*	2.06	3.16	2.59	0.00	2.25
B07	2.77	2.49	1.69	1.42	1.90
B08	2.81	3.65	3.22	3.34	3.54
B09	2.85	2.85	3.45	3.34	3.42
B10	1.59	0.88	2.50	3.31	1.85
B11	2.10	1.40	3.57	2.56	2.30
B12	1.64	2.18	2.03	2.49	2.07
B13	2.39	1.68	1.16	2.05	1.58
B14	1.59	1.48	2.53	3.34	2.18

Table 4.6: Values of d′ for all participants in all testing blocks (Asterisk marks *memorizers*, pound sign marks *non-learners*, and all others are *learners*.)

	Bl 2	Bl 4	Bl 6	Bl 8	All Blocks
Avg for A	1.82	1.52	2.74	2.07	2.04
Avg for B	1.46	2.08	2.31	2.07	2.06

Table 4.7: Values of d′ for Dialects A and B in all testing blocks

	Response: Gram	Response: Ungram
Stimulus: Gram	Hit	Miss
Stimulus: Ungram	False Alarm	Correct Rejection

Table 4.8: Signal detection applied to yes-no grammaticality task

This is in accordance with the assumption that novel forms will, in general, be more difficult than familiar ones (that is, training should never have a negative effect on performance).

The choice of $d' = 1.00$ as a cutoff point is somewhat arbitrary, but there are reasons to favor it. If a participant received a 69% correct score for trials with both grammatical and ungrammatical stimuli, then this would correspond to a d' of 1.00 (Keating 2005). A score lower than this for a yes-no grammaticality task shows minimal to non-existent learning.[5] It is also a useful number because none of the participants received d' scores for blocks 6 or 8 that were right on the cusp of this threshold—if the cutoff had been 0.90 or 1.10, the results would still look the same.

If we look at only the *learners*, then there is a significant difference of d' values from Dialect A and Dialect B ($t = 3.6648$, $df = 13.682$, p-value = 0.00264). This difference is in the direction predicted by the neutralization avoidance hypothesis—in other words, the mean of the d' values for Dialect A *learners* (discounting memorizers and non-learners) was higher than that of Dialect B (3.46 vs. 2.53). If we perform the same type of test for accuracy data, then there is also a difference ($t = 3.4909$, $df = 10.841$, p-value = 0.00516), also in the predicted direction (95% vs. 87%).[6]

Before determining this difference, multiple t-tests were performed. As reported above, the difference in d' for *all* speakers of Dialects A and B is not

[5]Because the stimuli were half grammatical and half ungrammatical, random responses should on average give 50% correct for both types of stimuli.

[6]This effect holds even when the two non-native English speakers in Dialect A are removed from the data set.

significant. Another test was done on just those d′ values above 1.0, which also failed to reach significance. The third test, which looked at the d′ values for those categorized as *learners*, did reach significance. Using Bonferroni correction, we can still be confident in this third t-test (p-value of 0.00516 x 3 = 0.01548, or even more conservatively for all six t-tests performed, p-value of 0.00516 x 6 = 0.03096).

In short, the only effect we found was when looking at *learners* of the language, while leaving out the *memorizers* and *non-learners*. This effect is found for both accuracy and sensitivity, and it is in the direction predicted by the neutralization avoidance hypothesis.

4.4.4 Reaction times

The reaction time results underwent pre-processing in the following manner: First, the results were pared down to include only correct responses to blocks 6 and 8 (testing on familiar compounds and novel compounds, respectively). This included 3614 responses. Looking at the logarithm of the reaction times in ms (logrt), responses that were two standard deviations away from the mean were removed. The upper limit was 7.91 logms (2723 ms), and the lower limit was 5.98 logms (397 ms), counting from onset of stimulus. There were 106 logrt values that were too high, and 38 that were too low. Thus, after preprocessing the data set consisted of 3470 responses. The resulting logrt values were somewhat normally distributed, as can be seen in the histogram in Figure 4.3.

The mean log reaction times for each participant are shown in Figure 4.4. Participants who were exposed to Dialect A (the non-neutralizing dialect) are shown on the left, while those who were exposed to Dialect B (the neutralizing one) are shown on the right. Both are sorted by mean log reaction time; bars indicate 95% confidence intervals.

Various mixed-effect linear regression models were fitted to the log reaction time results (using the `lme4` package). In each model, participant was a random ef-

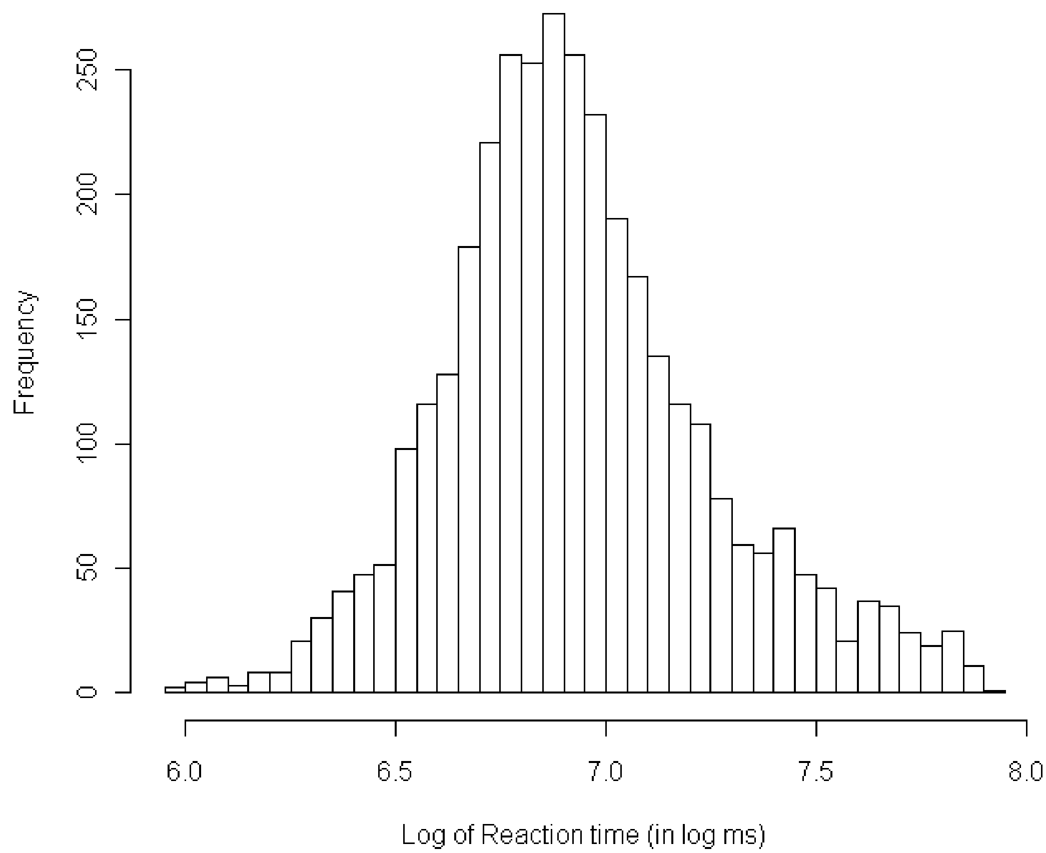

Figure 4.3: Histogram of log reaction times

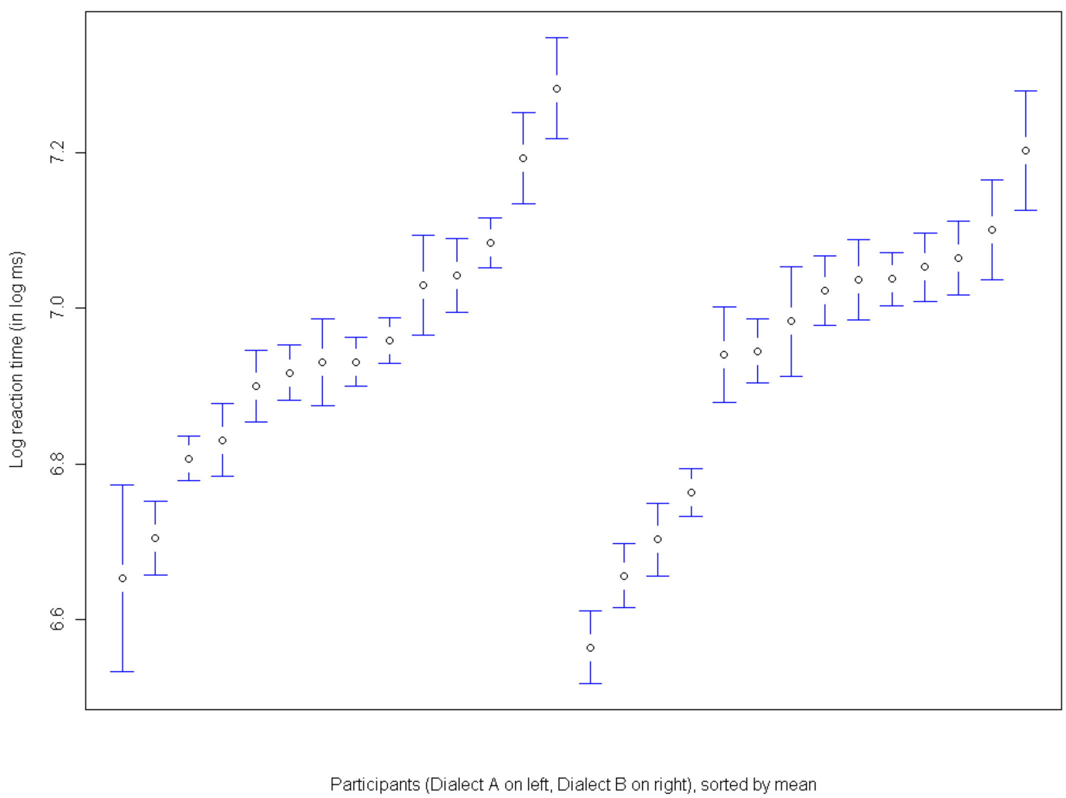

Figure 4.4: Log reaction times of responses in each dialect, by participant

fect. The fixed effects that were considered were: dialect type (neutralizing or non-neutralizing), block (familiar or novel compounds), learner type (learner, memorizer, or non-learner), and rule application type (whether the form resulted from overapplication, underapplication, correct application, correct non-application of the lenition rule, or whether the underlying /h/ underwent fortition). Because the log likelihood of model comparisons was important, and because the models had different fixed effects, the maximum likelihood estimation (MLE) method was used in place of restricted MLE (in `lme4`, `REML=FALSE`). Overall, dialect type was not found to be a good predictor of log reaction time, but block, learner type, and rule application type were.

The simple intercept-only model was compared to four models that each utilized one fixed effect. This comparison was performed at the same time using the `anova` command, and the measure of improvement was the chi-squared test on log likelihood values. The model with block as a fixed effect was an improvement over the intercept-only model, and the same was true for learner type and rule-application type. The model with dialect as the only fixed effect was not an improvement over the intercept-only model. Adding dialect to any of the other single-fixed-effect models did not improve them either, but those models were incrementally improved by adding the other fixed effects. At no point did the addition of dialect to the fixed effects improve any of the models that were considered. Thus, the best model is one that uses block, learner type, and rule application type, but does not use dialect type. In all of the models under consideration, participant was a random effect. The parameters of the final model are given in Table 4.9.

As can be seen in the coefficients of the model, responses to novel forms (block 8) tend to take longer than responses to familiar forms (block 6). This conforms to the intuition that familiarization can improve speed of correct judgments. In terms of learner type, reaction times by memorizers tend to be longer than those of learners. And for rule application type, correct non-application of the debuc-

Random effects:

Groups	Name	Variance	Std.Dev.
participant	(Intercept)	0.025445	0.15951
Residual		0.070740	0.26597

Number of obs: 3470, groups: participant, 28

Fixed effects:

	Estimate Std.	Error	t value
(Intercept)	6.91117	0.03820	180.94
block (novel forms)	0.08090	0.01082	7.48
learner type (memorizer)	0.16639	0.07517	2.21
learner type (non-learner)	-0.08675	0.12072	-0.72
rule type (correct non-appl)	-0.07145	0.01462	-4.89
rule type (overapplication)	0.02304	0.01513	1.52
rule type (fortition)	0.04254	0.02658	1.60
rule type (underapplication)	0.01494	0.01924	0.78

Table 4.9: Parameters of the best mixed-effect linear regression model considered, as given by `lme4`

calization rule tends to result in shorter reaction times than the default correct application of the rule. This supports the intuition that forms with rule application are more difficult than forms that simply combine the adjective and noun with no phonological change. Thus, the model confirms our general understandings of phonological experiments, but it does not lend support to the neutralization avoidance hypothesis, as adding coefficients related to neutralization fails to improve the model substantially. On the other hand, the model does not lend support to the opposite idea, that neutralization facilitates acquisition and use of rules.

Even if we remove the results from *memorizers* and *non-learners*, looking only at *learners* of Kasi, we reach the same conclusion. In the comparison of regression models, dialect type is not a good predictor of log RT, but block and rule application type are.

4.5 Possible mechanisms

The neutralization avoidance hypothesis makes the claim that neutralizing rules are more difficult to acquire and/or use than non-neutralizing rules, all things being equal. The rule in Dialect A is non-neutralizing, while the rule in Dialect B is neutralizing. Otherwise, the experiment was designed to minimize any other differences between the two dialects—to ensure that all things were indeed equal. Therefore, the experimental hypothesis is that Dialect A is easier to learn and/or use than Dialect B. There was no support for the hypothesis in the reaction time results, but the accuracy and sensitivity results provided evidence in favor of the neutralization avoidance hypothesis.

If the neutralization avoidance hypothesis is correct, then what is the precise mechanism for this slowdown in learning and/or processing of neutralizing rules? One possible proposal involves competition for phonemic categorization.

In Dialect A, the sounds [k] and [h] can always be categorized as the same phoneme /k/, regardless of the local environment of the allophone. Phonemic categorization in Dialect B is not so simple. There, initial [k] and [h] are categorized as their own phonemes (/k/ and /h/ respectively). Intervocalic [h], on the other hand, is in competition for the phonemes /k/ and /h/. When a listener comes upon an intervocalic [h], it is unclear whether it is underlyingly /h/ or whether it is /k/ having undergone the rule. Not knowing whether the rule applied or not might make acquisition of said rule more difficult. Additionally, in the computation of perception or production of an intervocalic [h], Dialect B might have a disadvantage when it comes to relating the output and input forms. The listener would have to rely on other information to determine its categorization, namely the allomorph of the morpheme that has either [k] or [h] initially.[7] The learner of Dialect A does not need to rely on this extra information, as all

[7]When /k/ and /h/ undergo neutralization in Dialect B, none of the forms result in lexical neutralization (homophony). That is, two lexical items are never pronounced the same, even though some phonemes may neutralize. Section 4.6 discusses the distinction between different types of neutralization in more detail.

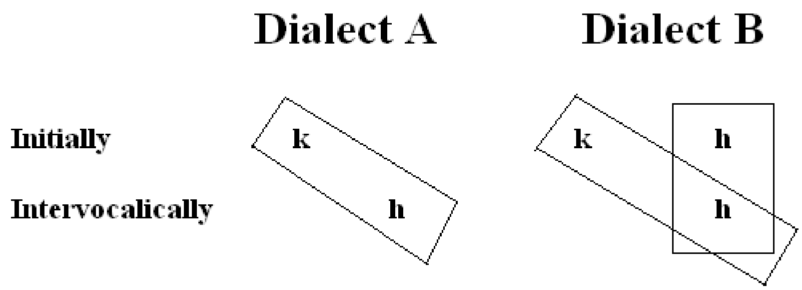

Figure 4.5: Categorization of allophones, in both dialects

instances of intervocalic [h] can simply be categorized as /k/. Thus, under this explanation, the competition of phonemic categorization hinders acquisition of the neutralizing rule, or it makes processing intervocalic [h] more difficult (or some combination of these two effects). The difference in categorization of these sounds is shown visually in Figure 4.5.

At this point, it is difficult to tell if this is the reason for the lower accuracy and sensitivity of Dialect B learners vs. Dialect A learners. When comparing the results for *learners* in trials with intervocalic [h], there is still no difference in log reaction time. There is a difference in accuracy, but this difference holds whether looking at trials with intervocalic [h] or trials without intervocalic [h]. Thus, there is no support for or against an explanation based on ambiguous categorization.

Another possible explanation involves perceptual warping. In Dialect B, /k/ and /h/ are separate phonemes. There is the possibility, then, for the perception of these sounds to be more distinct due to their phonemic status. Perceptual warping has been shown to occur as a result of categorization tasks, where members of the same category sometimes become more similar, and members of different categories sometimes become more distinct (Goldstone 1994, Guenther et al. 1999, among others).[8] Although the task in this experiment is not a categorization task

[8]Note that this effect is part of higher-level phonetic perception, the kind that the experiment in Chapter 3 was intended to avoid.

of the stimuli themselves, the relationship between the sounds that the participant learns is a type of categorization. If the phonemic status affects perception in this way, then the two sounds being related by the rule are more similar in Dialect A than in Dialect B. Furthermore, if there is a bias against rules that relate dissimilar sounds, then this may make the neutralizing rule more difficult to learn or use. Under this approach, then, perceptual warping coupled with a bias against relating distinct sounds is the source of the effect.

This explanation also does not have much evidence in its favor. In particular, I am unaware of any research showing that the phonetic similarity of inputs and outputs of rules affects their learnability. Further research will be necessary to better understand the source and nature of the neutralization avoidance bias.

There is also the question about how the neutralization avoidance bias affects languages. Specifically, how does the bias influence the way rules come about? It might be best to think of the situation in terms of reranking Optimality Theoretic constraints. When a rule first starts being applied, some speakers rerank their constraints some of the time in such a way that the rule applies. According to the analysis in Subsection 2.4.4, *Merge is the constraint involved in enforcing neutralization avoidance. If there is reranking, then sometimes the reranking will take *Merge into account due to it being highly ranked. Thus, one possible analysis of Indonesian k-debuccalization is that when the rule first started being used, *Merge was one of the constraints that was highly ranked, and /k/ became realized as [ʔ] instead of [h].[9] In other rule formation scenarios, *Merge plays less of a role, and the rule that is created has the possibility of being neutralizing.

4.6 Further issues

The division of participants into learner type was done on the basis of how well each individual participant performed in blocks 6 and 8, with no regard for the

[9]This specific proposal is not the only possible history of the k-debuccalization rule within this framework, though.

difference between Dialect A and Dialect B. Thus, it is likely that the difference in d′ and percent correct between the *learners* of A and the *learners* of B is due to the difference in neutralization. However, there is also a difference in the number of learners in each Dialect group. Of the 14 participants attempting to learn Dialect A, only 9 fully acquired the rule and demonstrated effective use on novel forms. For Dialect B, 11 of the 14 participants were able to do this. At this point, it is unclear if this is due to the random distribution of participants into the two Dialect categories, or if it reflects something more substantial about the difference between them.

One of the principle concerns about this experiment is whether or not the participants cared about the lexicon of the artificial language, and in turn its phoneme inventory. I am not familiar with previous artificial grammar experiments that suggest learners are sensitive to the phonemic status of a sound within the artificial language. There is, however, evidence that artificial grammar learners can become sensitive to phonotactic constraints rather quickly (Onishi et al. 2002, Pycha et al. 2003). The difference in accuracy and sensitivity lends support to the idea that the participants in this experiment were influenced by the lexicon and phoneme inventory. Moreover, the somewhat high accuracy and sensitivity scores for blocks 2 and 4 (as shown, for example, in Table 4.4) demonstrate that many of the participants memorized the sound-meaning mappings of the language. As such, they cared enough about the lexicon of the artificial language to succeed in the task, making the idea that they cared about the phoneme inventory more likely.

Another concern involves the type of neutralization at work. The type of neutralization in Dialect B is phonemic neutralization, because /k/ and /h/ are sometimes both realized as the same allophone [h]. However, there is no lexical neutralization (homophony). In other words, there are no minimal pairs in the lexicon, where one lexical item starts with /k/ and the other with /h/, with both of them being in the environment for k-debuccalization. Such an example would look like **ketu** 'apple' and **hetu** 'fish'. If these were both lexical items of the language, then the compounds for 'red apple' and 'red fish' would both be realized as **sihetu**. Just because there is no homophony, though, doesn't mean that the phonemes /k/ and /h/ can be identified solely by their allophones and the position of those allophones. There are still instances (when they are intervocalic) where we cannot identify the phonemic value of [h] from the surface form alone—we must also have information about related forms, where the phoneme surfaces word-initially. This is *possible homophony*—there could be homophony in the language, because of the phoneme inventory, morpheme structure constraints, and the debuccalization rule. The experiment never presents true homophony, but the language is set up so that it could occur if the lexicon were slightly different.[10]

The primary difference between Dialect A and Dialect B is the phoneme inventory, which in turn affects the neutralization status of the debuccalization rule. One other small difference between the dialects is how they were tested. In blocks 6 and 8, the participants responded as to whether a particular compound word was the correct way to combine the two component words in the artificial language. Both dialects had correct application and correct non-application of the rule. Furthermore, they both had ungrammatical forms created by overapplication and underapplication of the rule. The two dialects differed when it came to fortition (intervocalic /h/ $\rightarrow$ [k]). Dialect B had some ungrammatical forms that underwent fortition, while Dialect A did not (because that dialect does not have a phoneme /h/ to strengthen in the first place). The difference in mean accuracy as it varied by dialect and rule type is shown in Table 4.10.

4.7 Conclusion

To summarize, the artificial grammar experiment reported in this chapter was designed to test the neutralization avoidance analysis of supplementary gestures

[10]See Silverman 2010 for discussion on phonemic neutralization and homophony in Korean.

	Correct application	Correct non-appl	Over-application	Under-application	Fortition
Dialect A	0.92	0.98	0.96	0.89	NA
Dialect B	0.86	0.95	0.86	0.68	0.88

Table 4.10: Mean accuracy (in % correct) for *learners* of each dialect, by rule type

in debuccalization, and more generally, to look for evidence that neutralization avoidance is an active component of grammar acquisition and use. Participants were exposed to the artificial language *Kasi*, with adjective+noun morphological concatenation serving as the environment for the debuccalization rule (/k/ $\rightarrow$ [h] intervocalically). The language consisted of two dialects, which were largely the same except for the phonemic status of /h/. Half of the participants learned Dialect A (where the rule was not neutralizing), while the other half learned Dialect B (where the rule was neutralizing). The reaction time results showed no effect of neutralization avoidance, but the accuracy and sensitivity results did show an effect in the direction predicted, when the participants were divided into *learners*, *memorizers*, and *non-learners*.

Because an effect was shown for *learners* of the artificial language after only 30 minutes of exposure, support for neutralization avoidance as a phenomenon is somewhat encouraging. A more robust effect might be found in longer-term studies, where training takes place over the course of days or weeks rather than minutes. Extended training could provide time for the participants to internalize the basic sounds of the language, and also internalize which sounds are *not* basic sounds of the language, either because they are only derived by rule or because they are never used at all. Longer training sessions would also present the opportunity to use a larger number of lexical items, which could strengthen the effect of neutralization avoidance. Finally, future artificial grammar experiments on neutralization might benefit from using lexical neutralization (homophony), as opposed to the phonemic neutralization utilized here. Teasing apart the difference between these two types of neutralization would be difficult, but it would shed light on the motivation and mechanism behind neutralization avoidance in language.

Chapter 5

Conclusion

5.1 Which analysis is right?

Of the three competing analyses, no single analysis is clearly the "correct" approach for all examples of supplementary gestures in debuccalization. Specifically, none of the three are completely ruled out by the experiments. The analyses are all plausible, and they are all compatible with the principles of OT grammar

The perceptual faithfulness analysis of Indonesian k-debuccalization has some empirical evidence from the perceptual similarity experiment (Chapter 3). The pair [k ʔ] is more perceptually similar than the pair [k h] in coda position, as shown in the reaction time data. There is also evidence for the neutralization avoidance analysis, due to the results of the artificial grammar learning experiment (Chapter 4). When the participants are divided into *learners*, *memorizers*, and *non-learners*, the accuracy and sensitivity scores of the neutralizing dialect *learners* are lower than those of the non-neutralizing dialect *learners*. The dissimilation analysis does not have any evidence for or against it.

The perceptual faithfulness analysis also provides a basis for explanation of the typology of debuccalization processes. Under the assumption that glottal stop and oral stops are perceptually similar, leniting oral stops to glottal stop would serve a purpose under the perceptual faithfulness account, but leniting fricatives to glottal stop would not. Unfortunately, no evidence was found in Chapter 3 to support this particular claim, but future research might be able to find such evidence.

Thus, given the experimental evidence and potential explanatory power, I favor the perceptual faithfulness and neutralization avoidance analyses over the dissimilation analysis of supplementary gestures.

5.2 Is debuccalization a thing?

In Section 1.1, I defined debuccalization as any process that takes a consonant with oral constriction, and makes it a laryngeal consonant. I also categorized debuccalization as a subtype of lenition, using terms like *weakening* and *reduction* when referring to it.

If we assume that articulatory difficulty is the primary motivation for lenition, then, following Kirchner 2001, an overall analysis can be made with perceptual faithfulness constraints and constraints against articulatory effort. If debuccalization is always a subtype of lenition, then this analysis might be the right way to go.

According to our system, debuccalization processes are like other lenition processes, only they have a lower ranking of oral gesture faithfulness, causing those gestures to be unrealized. There is nothing else that separates debuccalization from lenition in general.

However, other constraints could be the primary motivation for some debuccalization processes. For instance, the process could be due to assimilation or dissimilation on the consonants themselves, or it could be the result of some perceptually-based sonority constraint. In other words, the assumption above might be wrong, and certain debuccalization processes might be non-lenition (non-weakening). An alternative definition of debuccalization could limit the scope of processes to just those that involve articulatory difficulty. It is not entirely clear what that alternative definition buys us, because it is hard to tell which motivation is at work when there are multiple possible motivations. This is especially the case when it comes to debuccalization, as opposed to other lenition processes (spirantization, voicing, degemination, etc.). In debuccalization, the resulting sound usually has a subset of the articulatory gestures that the non-debuccalized sound has. For instance, the fricative [s] has a glottal spreading gesture and a tongue-tip gesture at the alveolar ridge. If it debuccalizes to [h], then the glottal spreading gesture remains but the tongue-tip gesture is lost. The resulting gestural score for [h] is a subset of the one for [t]. Due to this subset-superset relationship, it is relatively straightforward to argue that [h] is less articulatorily effortful than [s]. Even so, we are not necessarily certain that the motivation for a voiceless obstruent becoming [h] in every case is articulatory ease. It is extremely difficult to show that articulatory ease constraints are the correct motivating force for a putative lenition process, because it would require demonstrating that all other possible forces are not motivating the process.

5.3 Why not simply delete?

Kirchner 2001 and Gess 2009 believe that debuccalization (and lenition in general) happens due to constraints in the OT grammar that wish to conserve effort—Kirchner calls this family of constraints Lazy, while Gess calls the constraint CAE (Conserve Articulatory Effort). For Kirchner, the Lazy constraints interact with other markedness constraints, with fortition constraints, and with faithfulness constraints. Through the constraint interaction, certain articulatorily easy (lenited) forms are chosen as the optimal candidates in particular phonological positions or registers of speech. Moreover, it is the faithfulness constraints that prevent full deletion. Some aspect of the underlying sound must remain, and debuccalization is one way to balance articulatory ease and faithfulness.

The neutralization avoidance analysis of debuccalization is also a possible explanation for why lenited forms do not simply delete. This is because, if the phoneme were to delete, then it would neutralize with zero. As such, even if the neutralization avoidance analysis is insufficient to account for all cases of supplementary gestures, it still might have a role to play in preventing deletion of the lenited consonant. The dissimilation analysis does not appear to be able to account for non-deletion. In fact, in the analysis of Indonesian k-debuccalization in Chapter 2, an extra constraint Max was needed to prevent deletion of the final /k/.

Supplementary gestures and non-deletion, while not always the same thing, are often closely related. In some circumstances, removing the oral gestures of a consonant would result in simple voicing. If this happens next to a vowel, then this would have the same effect as deletion. Thus, supplementary gestures in this case "save" the debuccalized consonant from the fate of deletion. In other circumstances, however, removing oral gestures would result in some laryngeal consonant, yet supplementary gestures are still added. These cases demonstrate that supplementary gestures and non-deletion are distinct ideas.

Bauer 2008 appears to take a less committed view to the cause of debuccalization. His definition is just as compatible with an Ohala-style innocent misperception view of things. Bauer argues that articulatory undershoot is the defining aspect of lenition. This undershoot may be grammatically controlled, as Kirchner 2001 would say, or it may originate as a production error. Even so, at some point the lenited form presumably gets a grammatical encoding, and it seems like Bauer is not committed to any particular encoding.

Widdison 1997 explains Spanish s-aspiration as driven by general tendencies to reduce phonetic material in coda position.[1] However, he doesn't argue that the [h] remains because of faithfulness constraints. Instead, Widdison stresses that glottal abduction occurs on vowels adjacent to a true [s] consonant, so s-aspiration can be seen as the continuation of this gesture without the [s]. The aspiration is a cue to the phoneme /s/, and is a part of both [s] and the debuccalized form. It

[1]Widdison argues that s-aspiration as a process then expanded to non-coda position by analogy.

might be the case that s-aspiration is a purely perceptually-driven change, or a combination of perception and articulatory ease, but Widdison 1997 supports the role of perception in some regard.

In sum, there are many possible reasons why a segment fails to simply delete, and many of these reasons are compatible with the ideas explored in previous chapters.

5.4 Implications

So what do we know from the investigation in this dissertation? We have a typology of debuccalization processes, given in Section 1.2. We know that debuccalization processes can be analyzed as an interaction between gestural faithfulness and articulatory ease constraints. In the case of supplementary gestures, some other constraint (perceptual faithfulness or neutralization avoidance) also comes into play, and the resulting sound is different from simple deletion of oral gestures. We have reasons to believe that both constraint types are possible, in the form of experimental results. The perceptual faithfulness account has perceptual similarity results that lend support to such an analysis for Indonesian k-debuccalization. The neutralization avoidance account also has support, due to the difference in the accuracy of artificial grammar learning, when dividing the participants into *learners*, *memorizers*, and *non-learners*.

The experiments also show things that are not directly connected to the evaluation of perceptual faithfulness, neutralization avoidance, and dissimilation. From Chapter 3, we have general perceptual maps of consonants to make further claims about faithfulness and markedness. In Chapter 4 some evidence was found in support of neutralization avoidance as a possible force in learning debuccalization processes. That evidence is not tied to debuccalization in particular, though. The effect of neutralization avoidance might be more general, as an effect in learning other types of phonological processes. At an even higher level, there might be an effect of neutralization avoidance in all manipulations of perceptual categorization. By repeating these types of tests with different varieties of rules and elements to manipulate, we can see how broad the scope is of this bias. For instance, there might be neutralization avoidance with rules that manipulate shapes or sizes of objects. If this were the case, it would show support for a view that neutralization avoidance is a general cognitive bias; if it were not, then perhaps this is a bias only in the linguistic or phonological realm.

Bibliography

Akinlabi, Akinbiyi M. (1992) Supraglottal deletion in Yoruba glides. *The proceedings of the Tenth West Coast Conference on Formal Linguistics* : 13–26.

Andrésen, B.S. (1968) *Pre-glottalization in English standard pronunciation*. Norwegian Universities Press.

Aronoff, Mark, Azhar Arsyad, Hasan Basri, & Ellen Broselow (1987) Tier configuration in Makassarese reduplication. *CLS 23: Parasession on autosegmental and metrical phonology* : 1–15.

Babel, Molly & Keith Johnson (2010) Accessing psycho-acoustic perception and language-specific perception with speech sounds. *Laboratory Phonology* **1**(1): 179–205.

Bandhu, C. M. & B. M. Dahal (1971) Nepali segmental phonology. *Summer Institute of Linguistics. Tribhuvan University: Kirtipur* .

Barker, M. A. R. (1964) *Klamath grammar*. University of California Press.

Bauer, Laurie (2008) Lenition revisited. *Journal of Linguistics* **44**(03): 605–624.

Beckman, Jill N. (1998) *Positional faithfulness*. Ph.D. thesis, University of Massachusetts Amherst.

Beckman, Mary E. & Janet B. Pierrehumbert (1986) Intonational structure in Japanese and English. *Phonology yearbook* **3**: 255–309.

Bessell, Nicola (1993) Phonological Wildcards: /?/, /h/. *ms* .

Blust, Robert (1990) Three recurrent changes in Oceanic languages. *Pacific island languages: Essays in honour of GB Milner* : 7–28.

Boersma, Paul & David Weenink (2010) Praat: doing phonetics by computer [Computer program]. URL `http://www.praat.org/`.

Browman, Catherine & Louis Goldstein (1986) Towards an articulatory phonology. *Phonology yearbook* **3**: 219–252.

Browman, Catherine & Louis Goldstein (1995) Gestural Syllable Position Effects in American English. *Producing speech: contemporary issues: for Katherine Safford Harris* : 19.

Buck, C. D. (1904) *A grammar of Oscan and Umbrian*. University of Washington Press.

Buckley, Eugene (1994) *Theoretical aspects of Kashaya phonology and morphology*. Center for the Study of Language and Inf.

Buckley, Eugene (1997) Explaining Kashaya infixation. In *23rd Annual Meeting of the Berkeley Linguistic Society*, Citeseer.

Campbell, L. (1985) *The Pipil Language of El Salvador*. Walter de Gruyter.

Carpenter, Angela C. (2010) A naturalness bias in learning stress. *Phonology* **27**(03): 345–392.

Clements, G. N. (1985) The Geometry of Phonological Features. *Phonology Yearbook* **2**: 225–252.

Coppenrath, Hubert & Paul Prévost (1975) *Grammaire approfondie de la langue tahitienne: ancienne et moderne*. Librairie Pureora.

Costa, D.J. (1991) The historical phonology of Miami-Illinois consonants. *International Journal of American Linguistics* **57**(3): 365–393.

Czaplicki, Bartłomiej (2006) Laryngeals and velars in Ukrainian. *The Slavonic and East European Review* **84**(1): 1–15.

de Lacy, Paul V. (2002) *The formal expression of markedness*. Ph.D. thesis, University of Massachusetts Amherst.

Dutoit, T., V. Pagel, N. Pierret, F. Bataille, & O. van der Vrecken (1996) The MBROLA project: Towards a set of high quality speech synthesizers free of use for non commercial purposes. In *Spoken Language, 1996. ICSLP 96. Proceedings., Fourth International Conference on*, vol. 3, IEEE, 1393–1396.

Edwards, Walter F. (1978) A preliminary sketch of Arekuna (Carib) phonology. *International Journal of American Linguistics* : 223–227.

Esper, Erwin A. (1925) A technique for the experimental investigation of associative interference in artificial linguistic material. *Language Monographs* .

Flemming, Edward (1996) Evidence for constraints on contrast: The dispersion theory of contrast. *UCLA Working Papers in Phonology* **1**: 86–106.

Flemming, Edward (2004) Contrast and perceptual distinctiveness. *Phonetically-based phonology* : 232–76.

Gerdel, F.L. (1985) Páez: Pitch and stress in the phonological word and phrase. *From Phonology to Discourse* : 31.

Gess, Randall (2009) Reductive sound change and the perception/production interface. *The Canadian Journal of Linguistics* **54**(2): 229–253.

Giannelli, L. & L.M. Savoia (1979) L'indebolimento consonantico in Toscana (II). *Rivista italiana di dialettologia* **3**(4): 38–101.

Gildea, Spike (1995) A comparative description of syllable reduction in the Cariban language family. *International Journal of American Linguistics* **61**(1): 62–102.

Goldsmith, John (1990) *Autosegmental and metrical phonology*. Blackwell.

Goldstone, Robert L. (1994) Influences of categorization on perceptual discrimination. *Journal of Experimental Psychology: General; Journal of Experimental Psychology: General* **123**(2): 178.

Guenther, Frank H., Fatima T. Husain, Michael A. Cohen, & Barbara G. Shinn-Cunningham (1999) Effects of categorization and discrimination training on auditory perceptual space. *The Journal of the Acoustical Society of America* **106**: 2900.

Gurevich, Naomi (2004) *Lenition and contrast: the functional consequences of certain phonetically conditioned sound changes*. Routledge.

Harris, John (1990) Segmental complexity and phonological government. *Phonology* **7**(2): 255–300.

Hayes, Bruce (1986) Inalterability in CV phonology. *Language* : 321–351.

Hendon, Rufus S. (1966) The phonology and morphology of Ulu Muar Malay (Kuala Pilah District, Negri Sembilan, Malaya) .

Hock, Hans Henrich (1991) *Principles of historical linguistics*. 34, Mouton de Gruyter.

Iverson, Gregory K. (1989) On the category supralaryngeal.

Kaneko, I. & S. Kawahara (2002) Positional faithfulness theory and the emergence of the unmarked: The case of Kagoshima Japanese. *ICU English Studies* **11**: 18–36.

Kari, J.M. (1976) *Navajo verb prefix phonology*. Dissertations-G.

Keating, Patricia A. (1988) *A survey of phonological features*. Indiana University Linguistics Club.

Keating, Patricia A. (2005) D-prime (signal detection) analysis. URL `http://www.linguistics.ucla.edu/faciliti/facilities/statistics/dprime.htm`.

Khalafallah, Abdelgahany A. (1969) A descriptive grammar of Sa̲ʕiːdi, Egyptian colloquial Arabic .

Kirchner, Robert M. (2001) *An effort based approach to consonant lenition*. Routledge.

Lapoliwa, Hans (1981) *A generative approach to the phonology of Bahasa Indonesia*. Dept. of Linguistics, Research School of Pacific Studies, Australian National University.

Lass, Roger (1976) *English phonology and phonological theory: synchronic and diachronic studies*. Cambridge University Press.

Lavoie, Lisa (1996) Consonant strength: Results of a data base development project. *Working Papers of the Cornell Phonetics Laboratory* **11**: 269–316.

Liljencrants, Johan & Bjorn Lindblom (1972) Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language* : 839–862.

Lindblom, Bjorn (1990) Explaining phonetic variation: a sketch of the H&H theory. *Speech production and speech modeling* **55**: 403–439.

Lipski, J. (1984) On the weakening of /s/ in Latin American Spanish. *Zeitschrift für Dialektologie und Linguistik* **51**(1).

Lombardi, Linda (1990) The non-linear organization of the affricate. *Natural Language and Linguistic Theory* **8**(3): 375–425.

Loving, R. & R. Loving (1966) Awa phonemes, tonemes, and tonally differentiated allomorphs. *Linguistic Circle of Canberra Publications A* **7**: 23–32.

MacDonald, Lorna (1990) *A grammar of Tauya*. Walter de Gruyter.

McCarthy, John J. (1988) Feature geometry and dependency: a review. *Phonetica* **43**(1988): 45.

Meira, Sérgio (2001) Linguistic theory and linguistic description: the case of Tiriyó [h]. *International Journal of American Linguistics* **67**(2): 119–135.

Mielke, Jeff (2005) Ambivalence and ambiguity in laterals and nasals. *Phonology* **22**(02): 169–203.

Milroy, James, Lesley Milroy, & Sue Hartley (1994) Local and supra-local change in British English the case of glottalisation. *English World-Wide* **15**(1): 1–33.

Moreton, Elliott (2008) Analytic bias and phonological typology. *Phonology* **25**(01): 83–127.

Morris, Richard E. (2000) Constraint interaction in Spanish/s/-aspiration: Three peninsular varieties. In *Hispanic Linguistics at the Turn of the Millennium: Papers from the 3rd Hispanic Linguistics Symposium*, 14–30.

Nivens, R. (1992) A lexical phonology of West Tarangan. *Phonological studies in four languages of Maluku* : 127–227.

O'Brien, Jeremy (2010) Perception and English t-glottalization, Ms., UCSC.

Onishi, Kristine H., Kyle E. Chambers, & Cynthia Fisher (2002) Learning phonotactic constraints from brief auditory experience. *Cognition* **83**(1): B13–B23.

Padgett, Jaye (1995) *Stricture in feature geometry*. CSLI Publications.

Padgett, Jaye (2003a) Contrast and post-velar fronting in Russian. *Natural Language & Linguistic Theory* **21**(1): 39–87.

Padgett, Jaye (2003b) Systemic contrast and Catalan rhotics .

Pisoni, David B. & Jeffrey Tash (1974) Reaction times to comparisons within and across phonetic categories. *Attention, Perception, & Psychophysics* **15**(2): 285–290.

Popjes, J. & J. Popjes (1986) Canela-Krahô. *Handbook of Amazonian languages* **1**: 128–199.

Poser, William J. (2001) A note on Sakhalin Ainu Morphophonemics .

Pugh, Stefan M. & Ian Press (1999) *Ukrainian: a comprehensive grammar*. Routledge.

Pulleyblank, E.G. (1984) *Middle Chinese: a study in historical phonology*. Univ of British Columbia Pr.

Pycha, Anne, Pawel Nowak, Eurie Shin, & Ryan Shosted (2003) Phonological rule-learning and its implications for a theory of vowel harmony. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, vol. 22, 101–114.

R Development Core Team (2011) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, URL `http://www.R-project.org/`, ISBN 3-900051-07-0.

Rice, Keren (1989) *A grammar of Slave*. Mouton de Gruyter.

Rose, Sharon (2000) Velar Lenition in Muher Gurage. *Lingua Posnaniensis* **42**: 107–116.

Rutgers, Roland (1998) *Yamphu: grammar, texts & lexicon*. Research CNWS, School of Asian, African, and Amerindian Studies.

Sapir, E., L.J. Frachtenberg, & W. Borgonas (1922) *Handbook of American Indian Languages*. Government Printing Office.

Schiffman, H. S. (1983) *A reference grammar of spoken Kannada*. Ginn and Co./The Athenium.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002) *E-Prime Reference Guide*. Psychology Software Tools, Inc, Pittsburgh.

Shepard, Roger N. (1980) Multidimensional scaling, tree-fitting, and clustering. *Science* **210**(4468): 390.

Shibatani, Masayoshi (1990) *The languages of Japan*. Cambridge University Press.

Silverman, Daniel (2010) Neutralization and anti-homophony in Korean. *Journal of Linguistics* **46**(02): 453–482.

Singh, Sadanand, David R. Woods, & Gordon M. Becker (1972) Perceptual structure of 22 prevocalic English consonants. *The Journal of the Acoustical Society of America* **52**: 1698.

Smith, Jennifer L. (2007) Markedness, faithfulness, positions, and contexts: Lenition and fortition in Optimality Theory .

Sommerstein, A.H. (1973) *The sound pattern of Ancient Greek*. Blackwell.

Sproat, Richard & Osamu Fujimura (1993) Allophonic variation in English/l/and its implications for phonetic implementation. *Journal of phonetics* **21**(291-311).

Steriade, Donca (2001) The phonology of perceptibility effects: the P-map and its consequences for constraint organization. *Ms., UCLA* .

Story, G.L. (1984) Babine and Carrier phonology: a historically oriented study. *Summer Institute of Linguistics Publications in Linguistics* **70**.

Tyler, S.A. (1975) Gondi/h/with Some Notes on Number and Gender in Proto-Central Dravidian. *Dravidian Phonological Systems. Seattle: University of Washington & Institue for Comparative and Area Studies* .

Vovin, Alexander (1993) *A reconstruction of Proto-Ainu*. Brill.

Watson, Kevin (2001) The realization of final/t/in Liverpool English. *Newcastle and Durham Working Papers in Linguistics* **5**.

Whitney, W.D. (1889) *Sanskrit grammar*. Harvard University Press.

Widdison, Kirk A. (1997) Phonetic explanations for sibilant patterns in Spanish. *Lingua* **102**(4): 253–264.

Wilson, Colin (2006) Learning phonology with substantive bias: An experimental and computational study of velar palatalization. *Cognitive Science: A Multidisciplinary Journal* **30**(5): 945–982.