The Dissertation Committee for Benjamin Bozzell Hansen Certifies that this is the approved version of the following dissertation:

The Perceptibility of Duration in the Phonetics and Phonology of Contrastive Consonant Length

Committee:

Scott Myers, Supervisor

Megan Crowhurst

Robert King

Björn Lindblom

Harvey Sussman

The Perceptibility of Duration in the Phonetics and Phonology of Contrastive Consonant Length

by

Benjamin Bozzell Hansen, B.S.; M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2012

Dedication

To Susan, Miranda, Safaa, Niku and Kent

Acknowledgements

This work came about because dedicated scholars consented to pass on to me a glimpse of the world they know. Each of my committee members, Myers, Crowhurst, King, Lindblom and Sussman, opened new windows for me to contemplate marvelous vistas in the landscape of language sounds. Other professors, Lisa Green, Marc Southern, Rajesh Bhatt, Bernhard Schwartz and Keith Walters helped me to understand human language more broadly and deepened my appreciation for Linguistics. The Departments of Linguistics and of Spanish and Portuguese at the University of Texas afforded me the opportunity to teach young students, which was a true honor and a joy for me. I am indebted to my experimental subjects-including one very close to me-who showed so much patience and indulgence. I am grateful for the help of students who preceded me in the department and who were kind enough to listen and discuss and share their travails, Golnaz Modarresi Ghavami, Augustine Agwuele and Willis Warren. Finally, the fellow students of my cohort who accompanied me in my first steps, particularly Fei Ren, B'alam Mateo-Toledo and Jia-Shiou Liao, who were wonderful companions and study-mates.

Lastly, I must thank my late parents Woodrow James Hansen and Dorothy Lee Hansen (née Stamps) who instilled in me a love of learning and who lent me the spiritual and material support that saw me through many rough seas.

The Perceptibility of Duration in the Phonetics and Phonology of Contrastive Consonant Length

Benjamin Bozzell Hansen, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Scott Myers

This dissertation investigates the hypothesis that the more vowel-like a consonant is, the more difficult it is for listeners to classify it as geminate or singleton. A perceptual account of this observation holds that more vowel-like consonants lack clear markers to signal the beginning and ending of the consonant, so listeners don't perceive the precise duration and consequently the phonological contrast may be neutralized in some languages. Three experiments were performed to address these questions using data from Persian speakers.

In Experiment I, four speakers produced singleton and geminate tokens of the voiced oral consonants [d,z,n,l,j] and the glottals [h] and glottal stop at three speaking rates. It was found that Persian speakers do distinguish geminate durations from singleton durations for all manners even at very fast speaking rates, and vowels preceding geminates are slightly longer than those preceding singletons. Speaking rate had more of an effect on geminates than on singletons for all segments studied: the durations of the geminates decreased more in fast speech than the durations of the singletons did.

In Experiment II, listeners heard manipulated continua of consonants ranging from singletons to geminates. Subjects' identification curves were modeled using the cumulative Gaussian model. The modeled standard deviation was interpreted as the

breadth of the perceptual threshold, and a broader threshold understood to indicate a less distinct perceptual boundary between the two categories. Obstruents [d,z] had smaller breadth values than the sonorants [n,l,j], and the glottals had the largest breadth values of all. This indicates that while sonorants were more difficult for listeners to categorize than obstruents, the glottals were the most difficult to categorize of the segments tested.

Experiment III tested whether the modification of a specific parameter, the formant transition duration, would affect the perceptibility of the geminate/singleton contrast. A single token containing the glide [j] was manipulated to produce three different continua, each having a distinctly different manipulated transition: short, normal or long. It was found that the longer the transition was, the broader the perceptual threshold, thus making the consonant harder to categorize.

Table of Contents

List of Tables

List of Figures

1.0 Introduction: Manner and Contrastive Consonant Length

A phonological contrast based on consonant duration such as Arabic [tabːala] "he spiced" vs. [tabala] "he consumed" (Madina 1973) is widespread geographically: Alabama (Hardy and Montler 1998), Finnish (Karlsson 1992), LuGanda (Clements 1986) and Bengali (Hankamer et al. 1989) are some of the languages that make the contrast.

In some of these languages, particular consonants do not occur as geminates, while in other languages such as Hungarian (Siptár and Törkenczy (2007), Arabic (Newman 2005) and Persian Deyhime (2000), no restrictions apply: any consonant segment may occur as a geminate.

1.1 GEMINATE SEGMENTAL RESTRICTIONS

A few of the language that have restrictions on particular segments are listed in Table 1.1 below:

Language	Non-ocurring Geminate Segment	Reference
Amharic	hː	Leslau, 1968
Madurese	ʔː	Cohn et. al. 1999
Buginese	wː, ʔː	Cohn et. al. 1999
Tiberian Hebrew	rː, ʔː, hː	Lowenstamm and Kaye 1986
Finnish	dː, vː, jː, hː	Karlsson, 1992
LuGanda	wː, jː, lː	Clements 1986
Classical Nahuatl	jː, wː	Andrews, 1975
Berber	wː	Basset 1929
Japanese	sː ("mimetic")	Kawahara 2006
Klamath	sː	Blevins 2004a
Alabama	bː	Hardy and Montler 1998
Kisar	All except nː, lː, rː	Christensen & Christensen 1992

Table 1.1 Languages having restrictions on geminate segments

Podesva (2000) and Kawahara (2007) have claimed that restrictions on geminate segments are selective, with higher sonority manner classes being more likely to be disallowed. Their claims of the higher markedness of sonorant geminates appeal to perceptual confusion, as in Podesva (2000): “the relatively small intensity difference between a geminate sonorant and the vowel that precedes it makes difficult the perception of the sonorant’s phonemic length.” By contrast, Blevins (2004b) has surveyed numerous languages from diverse language families and has identified several instances of languages that have geminate sonorants but no geminate voiceless stops[1]. She concludes that there are no “implicational relationships relating to sonority that affect geminate inventories,” attributing apparent gaps in the distributions to “a direct reflection of specific instances of geminate evolution, or a combination of them, nothing more.”

As for fricatives, Blevins (2004a) and Kawahara (2006) both attribute apparent restrictions on geminate [s] to confusion between the inherently long [s] and phonologically long [sː]. Kirschner (2000) claims that geminate fricatives in general are marked because he supposes that they require more articulatory effort to pronounce.

Ohala and Riordan’s (1979) account of aerodynamic constraints on the articulation of voiced geminate stops has been supported by subsequent investigation. For example, Hayes and Steriade (2004) provide extensive examples of languages whose geminate inventories are defined by this type of constraint.

Cross-linguistic tendencies are also suggested by inspection of surveys of phoneme inventories such as Maddieson (1984) and Ruhlen (1975). Geminate glottals and glides are of particularly low frequency, raising the possibility that there may be restrictions on geminates of this sort. The evidence for restrictions on sonorants and

[1] Gilbertese, Kisar, Palauan, Piro, Chuvash, Ocaina.

fricatives is weaker, but still suggestive: In Ruhlen's survey of 700 inventories, 103 languages are shown as having at least one geminate consonant.[2] Of the 103, only 18 have geminate glides and 16 have geminate glottals (81 have any glottals at all). 71 have geminate liquids, 82 have geminate nasals and 75 have geminate fricatives, while 93 have geminate stops. Usually, but not universally, the presence of higher-sonority or glottal geminates in a language implies the presence of lower-sonority geminates, particularly stops.

1.2 A PERCEPTUAL ACCOUNT

A basic claim of this dissertation is that the more vowel-like a consonant segment is, the more difficult it will be for listeners to perceive the precise boundary between it and adjacent vowels, thereby reducing the effectiveness of a phonological length contrast which is based primarily on duration. Such a segment would be predicted to be less likely it is to occur as a geminate, thereby accounting for the cross-linguistic scarcity of certain manners of geminates.

Perceptual ambiguity has often been given as the source of neutralizations and other historical changes (Ohala 1981). Myers and Hansen (2005) attributed vowel lengthening effects following a glide to ambiguous perceptual segmentation of the glide-vowel sequence. In the case of vowel-like consonants, if the primary cue for phonological length is ambiguous, the geminate becomes harder to distinguish from its singleton counterpart. The phonological distinction thus becomes less useful and so is either neutralized or an avoidance strategy is employed that preserves the lexical distinction while eliminating the phonological alternation. For example, in Tiberian

---

[2] Actually many more than 103 of the languages have geminates. Norwegian and Persian, for example, are not among the languages shown as having geminates in the survey, although they certainly do. It is likely that the primary sources differ in their treatment of geminates. Some sources consider phonological length to be suprasegmental and so do not note it in the inventory (Maddieson 1984).

Hebrew, the vowel preceding a prohibited geminate is lengthened and the lexical contrast is thereby preserved.

Kawahara (2007) tested the perceptibility of the geminate/singleton contrast by comparing the steepness of the geminate vs. singleton identification curves for Arabic consonants having different levels of sonority. He found that the perceptual threshold between geminates and singletons in Arabic is broader for [j] than for [l] and is broader for [l] than for [z] or [d]. He also found that subjects' reaction times were greater for approximants than for non-approximants, indicating that it is harder for them to judge the status of approximants. This supported his hypothesis that sonority is a factor in the avoidance of certain manners of geminates.

Sonority, if interpreted qualitatively as "amplitude," can be measured quantitatively as energy in decibels of sound pressure. By this definition, sonority is one possible dimension of similarity between consonants and vowels. Other dimensions include the relative openness of the vocal tract and the gradualness of the transition between the vowel and the consonant: Glottal consonants are vowel-like in that they have the open vocal tract configuration of vowels, and glides exhibit the gradual transitions characteristic of vowel-to-vowel sequences.

Hansen (2006) found that when the duration of the transition of the Persian glide [j] is varied, the breadth of the perceptual threshold between geminates and singletons is affected: the longer the transition, the broader the perceptual threshold as measured by the steepness of the curve-fitted identification curve model. This suggests that at least one factor apart from sonority can affect the perceptibility of the geminate/singleton contrast.

According to the central hypothesis of this dissertation, consonant segments having imprecisely demarcated interfaces with surrounding vowels are less likely to occur as geminates in the world's languages because listeners have a harder time judging the precise duration of the consonant gesture.

Using data from Persian, this dissertation will test the perceptibility of the geminate/singleton threshold for different manners of articulation: glottals, glides, liquids, nasals, fricatives and stops, using the Persian consonants [ʔ,h,j,l,n,z,d]. As in Kawahara (2007), the perceptual threshold between singleton and geminate sonorants is tested and compared to the threshold between singleton and geminate obstruents. The dissertation also compares the sonorant and obstruent thresholds to thresholds for glottal segments, which were not included in Kawahara's experiment.

The neutralization of a length contrast caused by the difficulty of precise identification of the consonant-vowel boundary was discussed in (Myers & Hansen 2005) in reference to the widely-observed pattern of the neutralization of all vowels to long following a glide. In perception experiments using Finnish, there was evidence that speakers classified part of the glide-vowel transition as glide and part as vowel. The longer the transition, the more likely they were to hear a long vowel. This indicates that the form of the consonant-vowel transition is important in the perceptual demarcation of the temporal boundary, and consequently in the perception of phonological length. The Finnish data provided an explanation of why all vowels following consonants with very long transitions (glides) would be classified in some languages as long: they *sound* like long vowels.

Kavitskaya (2002) in her discussion of CVC compensatory lengthening discusses the diachronic process of the phonologization of phonetic vowel lengthening followed by coda loss. Following Ohala (1981), she attributes the perception of a long vowel to listeners interpreting gradual vowel-consonant transitions as part of the vowel, leading to a long vowel classification which remains after the trigger consonant has deleted. But she does not predict compensatory lengthening to result from the deletion of every manner of consonants: "only the deletion of consonants whose transitions can be mistaken for a portion of the vowel or which can otherwise affect the phonetic duration, or the perception of the phonetic duration, of the preceding vowel should be able to cause CL (compensatory lengthening)." She specifically excludes stops from this category.

I will discuss two ways in which difficult-to-perceive length contrasts can be eliminated: neutralization, in which one of the two contrasting segments is phonologized as the other, and avoidance, by which one of the segments is somehow altered in quality so as to be more easily distinguished from the other.

1.3 NEUTRALIZATION IN CLASSICAL AZTEC

An example of the resolution of a length distinction in glides [j,w] by neutralization is provided by Classical Aztec, in which the juxtaposition of identical segments at morpheme boundaries results in segments that are phonologically long, unless the segments are glides, in which case the result is a singleton: "The /w/ and /y/ are special cases. When two /w/ or two /y/ sounds are thrown together, no long consonant is produced."[Andrews 1975 p.9][3]

Neutralization is also produced in the case of regressive assimilation at morpheme boundaries. For most consonants, when total assimilation occurs, the result is lengthening of what Andrews calls the "dominant" consonant. (Andrews does not provide glosses in his discussion of assimilation.)

---

[3] Andrews (1975) states that the /w/ phoneme has a voiceless articulation in syllable-final position. Female speakers do not produce a velar articulation in the phoneme, so the respective syllable-initial and syllable final allophones of /w/ for females are [β] and [ɸ].

sa**n** **s**eː	→	sa**sː**eː
naː**w- + -p**a	→	naː**pː**a

Table 1.2 Aztec Regressive Assimilation: Non-Glide

But when there is assimilation of a nasal coda preceding a morpheme beginning with a glide, there is no lengthening:

tonaː**nw**aːn	→	tonaː**w**aːn
a**m- + -j**aːskeh	→	a**j**aːskeh

Table 1.3 Aztec Regressive Assimilation: Glide

The model of phonologization model assumed by Kavitskaya (2002) and Myers and Hansen (2005) predicts such an outcome, whereby a variant assumes the value of the segment it sounds like. It is interesting to note that in this case, one must assume that long glides sound like short glides, and not the other way around. The assumption that a consonant that cannot be specified for length may be presumed to be phonologized as short (assuming there is a length contrast in the language) is supported by the absence, to my knowledge, of instances of consonant inventories in which the sole representative of a particular segment is a geminate, whereas opposite cases abound: all discussions of geminate consonant inventories cite languages in which a particular segment may only occur as a singleton. And it is universally attested that a language that has geminates also has singletons, not the other way around. These observations are classical tests for markedness, and may be thought of as expressing a principle of parsimony by which an added layer of complexity (such as phonological length) is not to be assumed where it need not be assumed. So wherever a prohibition exists against a length contrast for a particular segment, we expect that speakers consider all instances of the segment to belong to the short category.

1.4 AVOIDANCE IN TIBERIAN HEBREW AND AHAGGAR BERBER

Two examples of avoidance strategies are provided here: Tiberian Hebrew and Ahaggar Berber.

In Tiberian Hebrew (Lowenstamm and Kaye 1986)[4], vowel-final clitics trigger length in the initial consonant of the word they attach to:

mi +	po	mipːo
from	here	from here’
wa +	joːmar	wajːoːmar
[past]	‘he will say’	‘he said’

Table 1.4 Tiberian Hebrew Boundary Lengthening: Non-Gutturals

But if the word begins with one of the “gutturals” [r, h, ʔ, ʕ], the initial consonant remains short, and the preceding vowel is lengthened:

mi +	roʃ	meːroʃ[5]
from	beginning	‘from the beginning’
mi +	haːhaːr	meːhaːhaːr
from	the mountain	‘from the mountain’
wa +	ʔoːmar	waːʔoːmar
[past]	‘I will say’	‘I said’

Table 1.5 Tiberian Hebrew Boundary Lengthening: Gutturals

---

[4] I have converted the transcriptions to IPA

[5] A lengthened [i] occurs as [eː].

Gutturals also behave differently in verbal paradigms. In the templatic verbal paradigm for the root /ktb/ ‘to write’, the initial consonant in the imperfect is lengthened:

perfect	niktab
imperfect	jikːaːteːb

Table 1.6 Tiberian Hebrew Imperfect Lengthening: Non-Gutturals

But in the guttural-initial root /ʕmd/ ‘stand’ the preceding vowel is lengthened instead:

perfect	niʕemad
imperfect	jeːʕaːmeːd

Table 1.7 Tiberian Hebrew Imperfect Lengthening: Gutturals

It is to be noted here that the glide [j] is perfectly capable of lengthening to [jː] in *wajːoːmar*. Where compensatory lengthening exists, Kavitskaya’s (2002) theory predicts that it is the glide that would be affected in first instance. This is why she states that a phonetic explanation is not to be sought for the Tiberian Hebrew facts:

> Tiberian Hebrew raises a number of questions concerning both the validity of the account of CL through consonant loss presented so far and its applicability to CL in Hebrew. The inventory of segments which cause CL (compensatory lengthening) in Tiberian Hebrew cannot easily be accounted for through mechanisms of phonetic change. Here I will argue that it is not necessary to invoke phonetic explanation for CL in Tiberian Hebrew, since it is a clear case of a templatic, morphologically conditioned process.

I hold that a phonetic explanation is indeed called for. Since a templatic morphology is a pervasive and productive pattern signaling basic morphemic distinctions in the language, I reason that an overriding phonetic imperative would be required in order to dispense with the template. The disallowance of a particular derived form based on restrictions on the geminate that would be generated, would require a resolution that satisfies the requirements of communication while satisfying segmental well-formedness conditions. If the basis of the disallowance of a particular derived form is *perceptual* confusability, the solution must be one that *sounds like* the expected form, but whose cues are sufficient to signal the intended lexical contrast.

An alternative strategy for preserving the lexical contrast by avoiding a problematic segment class is exemplified by Ahaggar Berber (Basset 1929). In Ahaggar Berber, where the morphology generates a geminate according to the verbal template, certain segments are avoided and an obstruent is substituted. For example, a geminate [gː] occurs in place of the predicted *[wː] in alternation with [w], and the fricative [sː], not *[hː] alternates with [h]:

aorist	ihwg	
preterite	haggäg	"pour origine une relation w/ww"
aorist	iuhar	
preterite	wəssär	"pour origine une relation s/ss"

Table 1.8 Ahaggar Berber Length Alternations (author's transcriptions)

Kawahara (2007) describes such avoidance strategies within the framework of Optimality Theory that tend to disfavor marked (sonorant) geminates. He attributes the markedness of sonorant geminates to the fact that "the onsets and offsets of sonorants are perceptually hard to pin down, at least compared to obstruents." But he does not offer an explanation for why glottals should be avoided, simply stating in a footnote that "Explicating the markedness of geminate glottals is, however, beyond the scope of this paper."

Avoidance of perceptually problematic geminates can also be understood as an instance of "presupposition" coming into play within the framework of Lindblom's

Hypo-Hyper articulation theory (Lindblom 1990, also Lindblom et al. 1994): speakers heighten acoustic cues to preserve semantic contrasts.

How does the speaker know how to pronounce the word in order to satisfy the requirements of communication? One must refer to the concept of speaker-listener outlined by Lindblom et al. Part of the speaker's knowledge of his language is the knowledge of what will clearly sound like the word he wishes to express. By agreement among a community of speakers, an accepted solution can become customary and be available as long as the morphological context calls for it. Diachronically, once the morphological source of the conflict loses productivity, the segmental change becomes irretrievably phonologized. Conversely, as long as the morphological pattern remains effective, my approach predicts that the prohibited geminate can be regenerated by a different generation or community, for whom the subtle distinctions in duration are tolerable.

A possible example of the reinstatement of a lost geminate segment is to be found in the history of Arabic. In the view of classical Arabic authors, it was improper to distinguish the pronunciation of the plain and the geminate glottal stop. (Al-Nassir 1993, Cited in Blevins 2005). However, current speakers of Standard Arabic certainly do distinguish the two quantities of *hamza* (glottal stop) (Madina 1973).

I propose that in each of the cited cases of Classical Nahuatl, Tiberian Hebrew and Ahaggar Berber, alternative strategies are used to dispense with a contrast that is perceptually disfavored because of its high probability of confusion.

Persian is one of the languages that permits all classes of segments to contrast by length (Mahootian 1997)[6]. Therefore speakers are able to provide samples with which to characterize acoustic properties of all manners of geminates, and listeners are capable of judging the status of geminate vs. singleton tokens in perception tasks.

The consonantal phonemes of Persian are shown in Table 2.1 below.

	Bilabial	Labio-dental	Dental	Alveolar	Post-Alveolar	Palatal	Velar	Uvular	Glottal
Plosive	p,b		t,d			k,g		q	ʔ
Nasal	m			n					
Trill				r					
Fricative		f,v		s,z	ʃ,ʒ		x		h
Affricate					tʃ,dʒ				
Approximant						j			
Lateral Approximant				l					

Table 1.9 Consonantal phonemes of Tehrani Persian

The inventory follows descriptions by Mahootian (1997), Samareh (1985), Samareh (1977), Sepanta (1999) and Deyhime (2000). (I have used IPA [ʃ,ʒ,tʃ,ʔ] rather than [š,ž,č,']etc. used by some of the authors.) The "dental" consonants /t/ and /d/ are not pronounced interdentally as in the Romance languages, but "pressed to the back of the front teeth" (Samareh (1985)). Therefore, the blade of the tongue touches the alveolar ridge, and the effect is very much like an alveolar. I have heard Spanish speakers imitate a Persian accent in Spanish by pronouncing d's with a strong alveolar pronunciation, as

[6] Samareh's (1977) inventory of clusters has gaps for gg and ʒʒ. This is attributable to the fact that these sound are absent from Arabic loans, which are the source of nearly all of the geminates in Persian. I have tested the ability of Persian speakers to produce geminates gː and ʒː by asking them to read nonsense words with the geminate diacritic on [g] and [ʒ], and they have no problem in producing the sounds.

in English. Thus they could as well be described as alveolar. I have observed the articulation of /l/ to be light or apical, without a strong velar gesture. The phonemes /k/ and /g/ have a palatal articulation in all environments except before back vowels, though they are conventionally called “velars”. The uvular consonant is actually voiced except in voiceless clusters, but I follow the conventional usage and represent it as /q/. Samareh (1977) and Samareh (1985) contain useful discussions of the allophones of these sounds.

2.0 Experiment I: The Production of Geminate Consonants in Tehrani Persian

2.1 INTRODUCTION

There is a growing body of quantitative studies of geminate consonants: Lehtonen (1970), Bannert (1976), Homma (1981), Pind (1986, 1995), Lahiri & Hankamer (1988), Hankamer et al. (1989), Smith (1991), Han (1992), Abramson (1999), Arvaniti (1999), Pickett et al. (1999), Kawahara (1997). Every one of these studies contrasts the durations of the long and short segments, agreeing with Hankamer et al. that consonant duration is the primary cue for the geminate/singleton contrast. A smaller number of studies have compared other possible cues besides duration. Abramson (1999) measured the fundamental frequency in the vowel following word-initial geminates in Pattany Malay and found that there is a distinction in this parameter between geminates and singletons. Kawahara (1997) compared the amplitude difference between surrounding vowels and geminate and singleton consonants of various manners representing a range of sonority in Arabic and found that the more sonorous a segment is, the smaller the amplitude change is, but did not present a statistical comparison of geminates to singletons. However, in his figures it is observed that only for the stops [t] and [d] did the geminate seem to differ from the singleton, with the geminate showing a greater amplitude drop from V to C in each case.

Quantitative studies of Persian geminate consonants are few. Hansen (2003, 2004c) found that alveolar stop closure durations decreased with increased speaking rate, but geminate durations decreased more than singleton durations did. Mean singleton closure durations decreased from 63.6 ms at the normal rate to 51.7 ms at the fast rate to 44.1 ms at the fastest rate, while mean geminate durations decreased from 153.3 ms to 116.7 ms to 87.4 ms. There was some overlap between singletons and geminates:

although some geminates spoken at very fast rates were shorter than singletons produced at normal rates, no overlap occurred *within* a given speaking rate. This suggests that speakers adjust the category boundaries by taking speaking rate into account. The adjustment of temporal boundaries to the speaking rate has also been reported for consonant transitions (Miller 1981) and VOT (Kessinger & Blumstein 1997). In each case, the duration of the temporal parameter decreases as speaking rate increases. Pickett et al. (1999) likewise observed that geminate stop durations in Italian are more affected by speaking rate than singleton durations are (the Geminate*Rate interaction was highly significant, p<.0001). Hansen (2004a) observed that Persian speakers do distinguish between singleton and geminate categories for all manners of articulation, though the clusters are closer together at fast speaking rates, and that for manners other than stops (glottal, glide, nasal, liquid, fricative), geminate durations also adjusted to the speaking rate, with geminates being more affected by rate than singletons. These observations differed from Arvaniti (1999), who found no significant interaction between length and rate for Cypriot Greek sonorants, though the sonorant durations were affected by rate.

Hansen (2004b) performed a perception experiment to determine the singleton-geminate boundary durations for alveolar stops at different speaking rates. It showed that for every rate (normal, fast, very fast), and phonation type (voiced, voiceless), the perceptual boundary coincided with the gap between singletons and geminates produced by the speakers. The 50% identification crossover averaged 93 ms at the normal rate, 64 at the fast rate and 50 ms at the fastest rate, falling between the geminate and singleton averages reported in Hansen (2003). I conclude that listeners must be able to hear the relative duration of a segment as it is spoken in order to correctly categorize its phonological length under varying rate conditions, since it would be impossible otherwise to adjust the perceptual boundary to the speaking rate.

Experiment I characterizes the Persian consonant length contrast with regard to manner and rate. The consonants [ʔ,h,j,l,n,z,d] represent the range of sonority (glides, liquids, nasals, fricatives and stops), as well as gutturals (the glottals [ʔ] and [h]). Glottals have been grouped with glides with respect to compensatory lengthening (Kavitskaya 2002). Certainly as far as their apparent markedness as geminates goes, they appear to be more like glides than obstruents on the basis of their scarcity in segment inventories. Insofar as they lack oral specification as consonants, their formant structures have no discontinuity with surrounding vowels, so from this standpoint they are the most vowel-like of all consonants. Kawahara's (2007) study of the perception of Arabic geminate consonants did not study glottals, though Arabic does have glottal fricatives and stops (Madina 1973).

2.2 METHODS

2.2.1 Objectives

This production study has two main objectives:

- To confirm the fact that over a range of speaking rates Persian speakers do produce geminates differently than singletons, as observed in Hansen 2004a. This is important to establish because the basis for using Persian in the related perception studies is that speakers do distinguish the categories for all manners of articulations, and so may be expected to make linguistically significant judgments in a discrimination task.
- To characterize the Persian consonants using three measures relating to the sharpness of the vowel-consonant boundary: Amplitude Difference, Formant Transition Duration and the Preceding Vowel Offset F1. Each of these measures allows me to compare the seven consonants [ʔ,h,j,l,n,z,d] and determine which ones are more sharply distinguishable from the

preceding vowel. Since faster speech involves greater coarticulation between consonant and vowel, I also predict that all three measures will be more vowel-like at faster speaking rates.

2.2.2 Comparisons

The quantities measured to be measured in the production study were collected under four conditions. The data are classified according to the categories listed in Table 2.4, which are the independent variables used in the statistical analysis:

Condition	Description
Subject	Four subjects, two female (f1, f2) and two male (m1, m2)
Rate	Three self-selected speaking rates (Conversational, Fast, Very Fast)
Length	Phonological category (Geminate, Singleton)
Segment	Seven segments representing a range of vowellikeness [ʔ,h,j,l,n,z,d]

Table 2.1 Conditions

I recorded twelve tokens for each possible subcategory type representing a given Subject, Rate, Length and Segment combination. For example, f1 read twelve sentences at the fast rate containing singleton [l]. For the most part I was able to obtain twelve measured quantities for each variable, although there are instances in which only eleven quantities are logged because a measurement was not possible due to a poor quality or missing recording.

The following paragraphs describe the dependent variables to be measured for each of these conditions.

2.2.3 Frame Duration (ms)

Speakers were asked to provide data at three distinct self-selected speaking rates. The measure of speaking rate that I use in this study is the duration of the carrier sentence produced by the informant, minus the duration of the test word itself. This four-syllable sentence frame contains the same sequence of phonemes in all of the production data. I use the frame duration to determine whether the informants did use distinct speaking rates. Hansen (2003, 2004c) found that a threshold consonant duration, expressed as a proportion of the average syllable duration in the carrier sentence, could be used as a rate-independent discriminator between singleton and geminate stops in Persian.

2.2.4 Consonant Constriction (ms)

Consonant closure is considered to be the primary cue for the geminate/singleton phonological contrast (Lehiste 1970, Hankamer et al 1989, Pickett et. al. 1999). In this study Consonant Constriction is the measure of either closure duration, or the period of drastically reduced complexity and amplitude in the waveform within the consonant, corresponding to the consonant steady state. The hypothesis is that for all manners, geminate durations are significantly distinct from singleton durations at each speaking rate.

2.2.5 Preceding Vowel Duration (ms)

Pickett et al. (1999) and Pind (1995) have reported the ratio of the duration of a geminate consonant to the duration of the vowel preceding it to be a rate-independent discriminator for the consonant length contrast. However, in those languages for which this claim is made (Italian, Icelandic), the preceding vowel is shorter for geminates than it is for singletons. Hansen (2003, 2004c) found that in Persian the preceding vowel tends to be slightly longer for geminates than for singletons as in Japanese (Homa 1981) and Finnish (Lehtonen 1970).

2.2.6 Amplitude Difference (dB)

Amplitude is a measure of discriminability between vowels and consonants since vowels are produced with a relatively open vocal tract and consonants are produced by somehow blocking the free flow of air out of the mouth, thereby reducing the amplitude. So consonants that have higher amplitude may be more vowel-like and consequently it may be more difficult to tell when the vowel ceases and the consonant begins. Amplitude was the only correlate of sonority measured by Kawahara (2007). However, amplitude does not tell the whole story about sonority. A sibilant fricative may be loud, but is ranked lower in sonority than a nasal due to the fact that there is some obstruction of the vocal tract. Amplitude difference is therefore understood as one dimension of salience of the vowel-consonant boundary, but there are others to consider.

Generally speaking, vowels are louder than consonants. The purpose of comparing the amplitude of the consonants is to determine which of the seven segments analyzed was more vowel-like in terms of amplitude. Higher-sonority consonants such as the sonorants [j,l,n] are predicted to be louder than lower-sonority obstruents [ʔ,h,z,d] and to approach the amplitude of vowels. Besides the segments' amplitude relative to vowels, I wanted to determine whether they are even more vowel-like at faster speaking rates. Measurements of vowel amplitude indicated that some of the speakers tended to speak louder as they increased their tempo, so I used a difference measure of the amplitude drop between the vowel and the consonant so that I could observe whether consonants were more vowel-like at faster speeds, and not just louder. The parameter that I denote Amplitude Difference is the difference between the decibel amplitude measured at the midpoint of the vowel preceding the test consonant and the decibel amplitude measured at the midpoint of the consonant. A smaller Amplitude Difference indicates a more vowel-like consonant.

2.2.7 Formant Transition Duration (ms)

The duration of the formant transition (Bannert 1970, Recasens 1999) is a perceptually salient indication of the abruptness of the transition between vowel and consonant. It indicates the period of time during which articulators are adjusting from the vowel steady state to the consonant configuration. If the transition is long, there must be a period of indeterminacy which can make judgments of consonant duration more variable (Myers and Hansen 2005). Glides have characteristically long transitions. Referring to voiced fricatives, Stevens (1998 p. 481) states that "the rate of movement of F1 is somewhat slower than it is at the implosion or release of the voiced stop consonant." This slow movement implies that [z] would have longer transitions than [d]. The alveolar sonorants [l,n] have a similar point of articulation in Persian to [d], but there are differences in tongue configuration and velar setting, so I do not have prediction as to the relative lengths of the transitions into [l,n,d]. I do not anticipate formant movement preceding the glottals [ʔ,h] since these segments do not have an oral specification and so do not influence the formant contour. I consider a consonant with a larger value of Transition to be more vowel-like.

2.2.8 Preceding Vowel Offset F1 (Hz)

Finally, F1 in the offset of the preceding vowel relates to the amount of jaw opening at that point (Stevens 1998). The vowel preceding the test consonants in all of the data is the low vowel [æ] which has a high F1. Therefore, higher F1 preceding the oral consonants [j,l,n,z,d] indicates that rapid change in the formants occurs at the boundary between vowel and consonant, since each of the consonants is articulated with a relatively closed jaw position, and a corresponding low F1. For the oral consonants, a low value of Offset F1 indicates a more vowel-like consonant. In the case of the glottal consonants [ʔ,h] there is no change in F1 since these consonants do not have an oral specification. The F1 at the vowel offset preceding the glottal consonants is therefore expected to have the same high F1 value as [æ].

2.2.9 Subjects

The subjects used to record the materials for this study were four adult Texas-resident speakers of Tehrani Persian all of whom are also fluent English speakers. They were recruited through personal and family contacts and were not paid for their participation. Speakers were educated through at least secondary school in Tehran. Their time of residence outside of Iran ranges from 6 to 27 years, and all speakers continue to use Persian on a daily basis.

Female 1 (f1). Age 44, has resided in U.S. for 15 years and Venezuela for 12 years, also speaks Spanish. Attended high school in Tehran.

Female 2 (f2). Age 43, has resided in U.S. for 25 years, is a heritage speaker of Azari Turkish. Started, but did not complete high school in Tehran.

Male 1 (m1). Age 40, has resided in U.S for 6 years. Attended university in Tehran.

Male 2 (m2). Age 38, has resided in U.S for 20 years. Attended high school in Tehran.

Table 2.2 Experiment I (Production) Subjects

2.2.10 Materials

Fourteen test words were selected to represent five manner classes: glide, liquid, nasal, fricative and stop, and the glottals [h] and [ʔ]. For the oral consonants, place (coronal) and phonation (voiced) are controlled for. The glottal segments are voiceless in careful speech, but Hansen (2004a) observed that they may be voiced in the intervocalic environment. The words are common two-syllable nouns having stress on the second syllable, which is the normal stress pattern for Persian nouns. The direct object suffix -*ra* which attaches to the nouns is a de-stressed particle, resulting in an overall iambic sentence stress pattern [˘ ¯ ˘ ¯ ˘ ¯]. To avoid any effect of vowel height on the test consonant durations, the vowel context is controlled for: the all test consonants are preceded by a front low vowel and followed by a back low vowel. The low vowel context will provide for maximum F1 movement in consonant-vowel transitions, thereby facilitating measurement. Although geminate/singleton minimal pairs exist in Persian, they are rare enough that I was not able to use them for the materials except for the pairs bænɑ/bænːɑ and fæʔɑl/fæʔːɑl.

dʒæ**h**ɑn	world	sæ**hː**ɑf	book binder
fæ**ʔ**ɑl	verb	fæ**ʔː**ɑl	active thing
bæ**j**ɑn	exposition	xæ**jː**ɑm	Khayyam (surname)
bæ**l**ɑ	haughty one	dæ**lː**ɑl	laborer
bæ**n**ɑ	building	bæ**nː**ɑ	builder
qæ**z**ɑ	food	bæ**zː**ɑz	rug merchant
fæ**d**ɑ	sacrifice	mæ**dː**ɑh	eulogist

Table 2.3 Test Words

These words were presented within a four-syllable frame sentence to produce a six-syllable utterance:

Carrier: Minu _________ -rɑ did
Gloss: Minu (name) ____ -OBJ saw

Minu saw (the) ______

Table 2.4 Carrier Sentence

The carrier sentences were presented on sheets of paper in Persian script (14-point Naskh font). While it is customary in Persian printed text to omit the optional diacritic that marks geminate consonants, the diacritic is routinely used if necessary for disambiguation. I marked all geminates with the *tashdid* diacritic ـّ above the consonant for clarity. Subjects read twelve sheets, each of which contained all fourteen sentences arranged in a different random order on each sheet. Filler carrier sentences containing common two-syllable words were placed at the beginning and end of the list on each sheet. The twelve sheets were read once for each of the three speaking rates:

12 sheets x 14 words x 3 rates = 504 tokens per subject.

2.2.11 Recording

The recordings were made in the sound-proof recording booth in the Linguistics Laboratory at the University of Texas. Speakers sat 6-12 inches from a high-quality Shure microphone and read from the sheets placed on a table in front of them. The experimenter sat outside the booth and monitored the recordings with headphones. Subjects were first instructed to read the twelve sheets at a normal, relaxed speaking rate. They were instructed not to produce pauses between words within the sentences. They were then asked to read the twelve sheets faster, but were told that they were to make sure that they were still speaking clearly enough to be understood. Finally, they were told to read the twelve sheets very fast and not to worry about whether they could be clearly understood. The result was three self-selected rates (conversational, fast, very fast). The sentences were recorded onto a Marantz solid-state digital recorder at a 22,000 Hz sampling rate and saved as wav files.

2.2.12 Measurements

I used synchronized waveform and spectrogram displays in Kay Elemetrics Multispeech to measure quantities in the test sentences. Duration measurements were made in the waveform display with reference to the spectrogram display as noted below. Screen shots showing the coordinated waveform and spectrogram displays are shown in figures 2.1a through 2.1g on the following pages. Blue vertical lines represent the onset and offset of constriction.

- Sentence Duration: From the beginning of acoustic energy in the periodic wave marking the beginning of voicing of [m] until the end of acoustic energy of [d], whether burst release or modal voicing.
- Word Duration: Begins at the beginning of the constriction of the first consonant in the test word as signaled by reduced amplitude and complexity in the wave form (fricatives [f,s,x]), loss of formants signaling stop closure [q,b,d,dʒ], or non-complex nasal wave[m]. Ends at the beginning of the closure of the following trill [r] when test word end with [a], or the beginning of voicing visible in the spectrogram, or high-frequency frication of [r] visible in the spectrogram when test word ends in a consonant.
- Preceding Vowel Duration: Begins with periodic voicing visible in the spectrogram following the first consonant of the word. In the case of aspirated [t] and unaspirated [q,b,d] stops and affricates[dʒ], the VOT following the release burst is not included in the vowel duration. The vowel ends at the beginning of the consonant constriction.
- Consonant Constriction Duration:

  [d] From the point of abrupt energy decline in the waveform coinciding with loss of formants F2 and F3, until the beginning of the release burst.

  [z] The duration of abruptly decreased amplitude with reduced complexity in the waveform.

  [n] The duration of non-complex lower energy wave, coinciding with lighter shading indicating decreased intensity in the spectrogram.

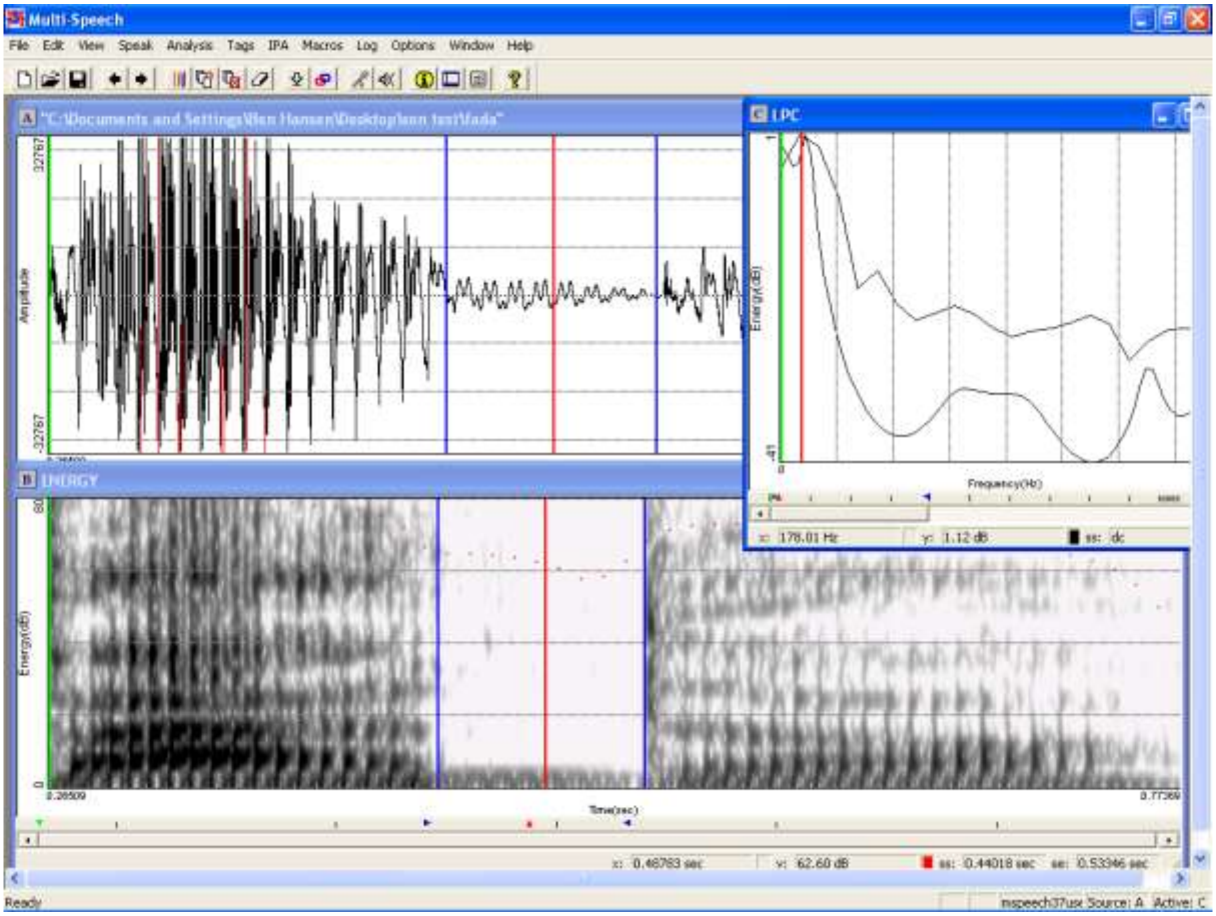

Figure 2.1a Sample Multispeech Display for Consonant [d]

Figure shows synchronized waveform and spectrogram displays from Multispeech. Onset and offset of constriction are shown by vertical blue lines. Red dots show the Decibel Amplitude (dB) track. Location of the consonant amplitude measurement is indicated by the vertical red line. Inset in upper right shown LPC spectrum at the amplitude measurement location.

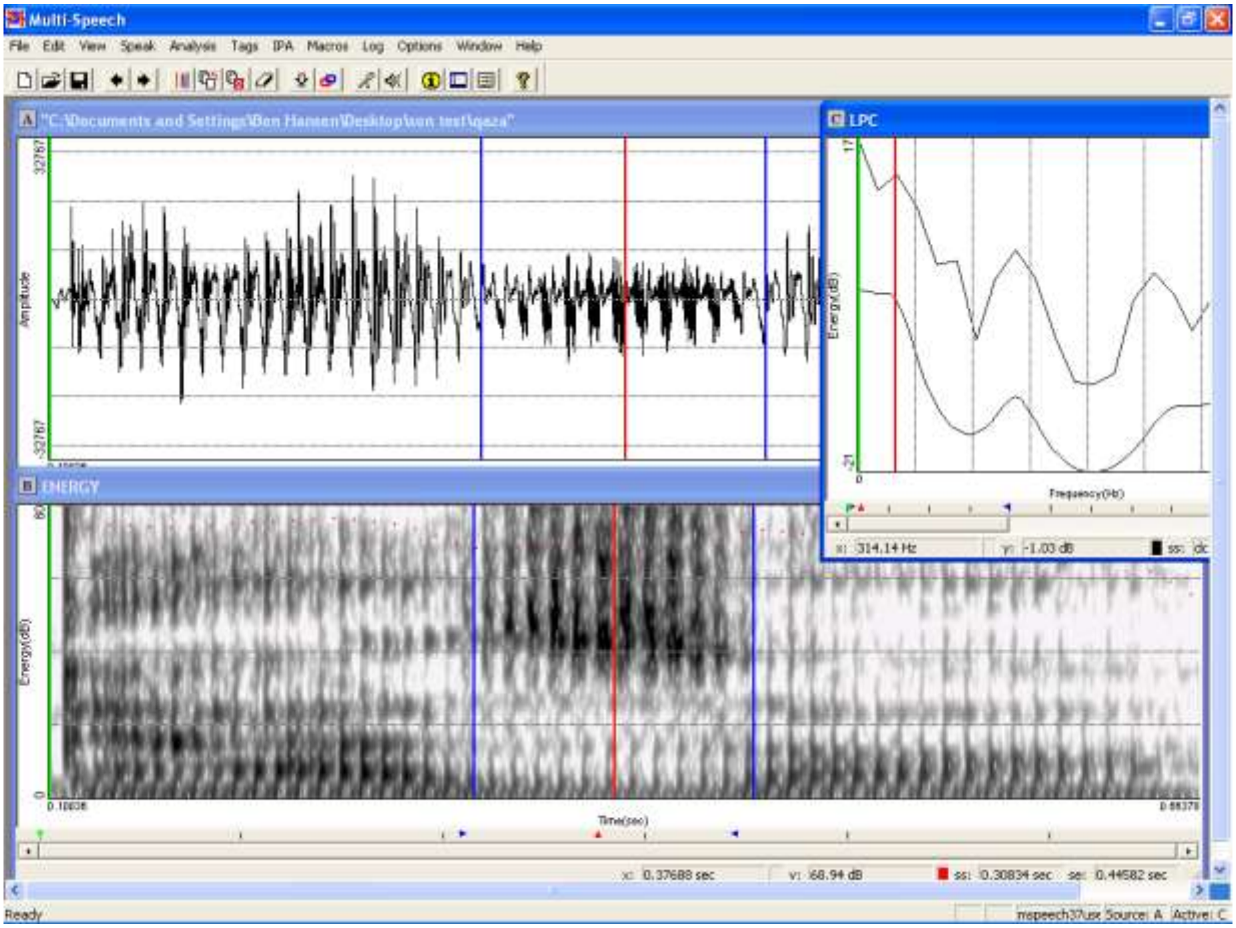

Figure 2.1b Sample Multispeech Display for Consonant [z]

Figure shows synchronized waveform and spectrogram displays from Multispeech. Onset and offset of constriction are shown by vertical blue lines. Red dots show the Decibel Amplitude (dB) track. Location of the consonant amplitude measurement is indicated by the vertical red line. Inset in upper right shown LPC spectrum at the amplitude measurement location.

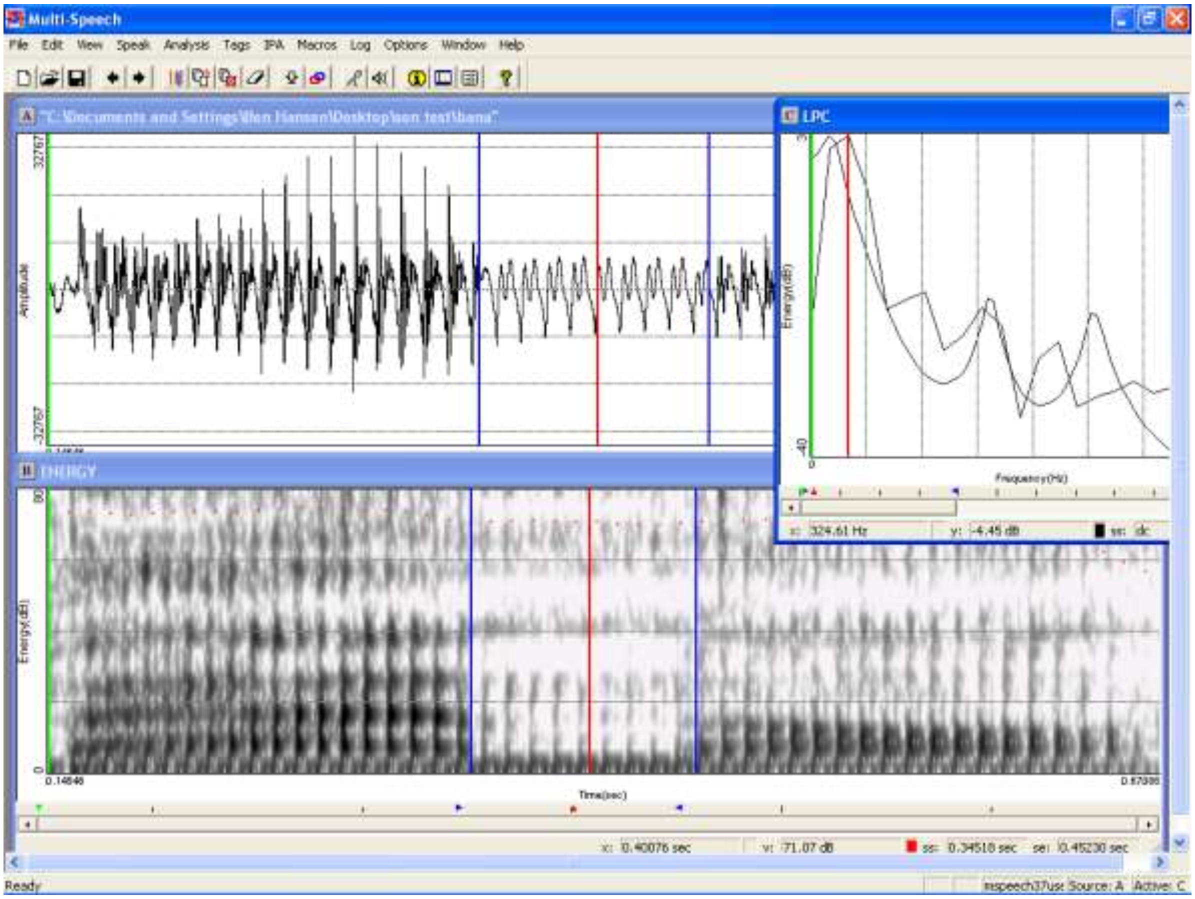

Figure 2.1c Sample Multispeech Display for Consonant [n]

Figure shows synchronized waveform and spectrogram displays from Multispeech. Onset and offset of constriction are shown by vertical blue lines. Red dots show the Decibel Amplitude (dB) track. Location of the consonant amplitude measurement is indicated by the vertical red line. Inset in upper right shown LPC spectrum at the amplitude measurement location.

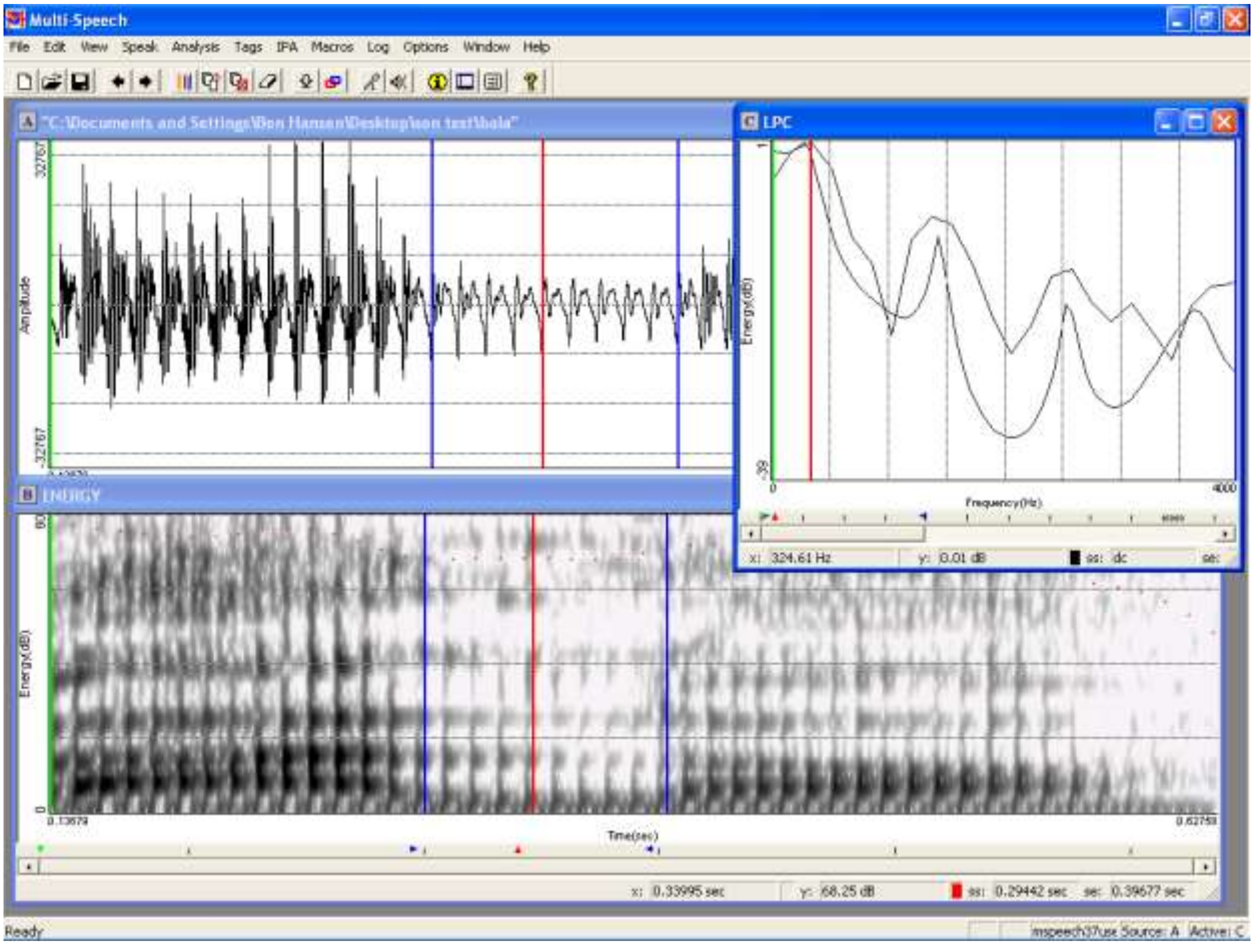

Figure 2.1d Sample Multispeech Display for Consonant [l]

Figure shows synchronized waveform and spectrogram displays from Multispeech. Onset and offset of constriction are shown by vertical blue lines. Red dots show the Decibel Amplitude (dB) track. Location of the consonant amplitude measurement is indicated by the vertical red line. Inset in upper right shown LPC spectrum at the amplitude measurement location.

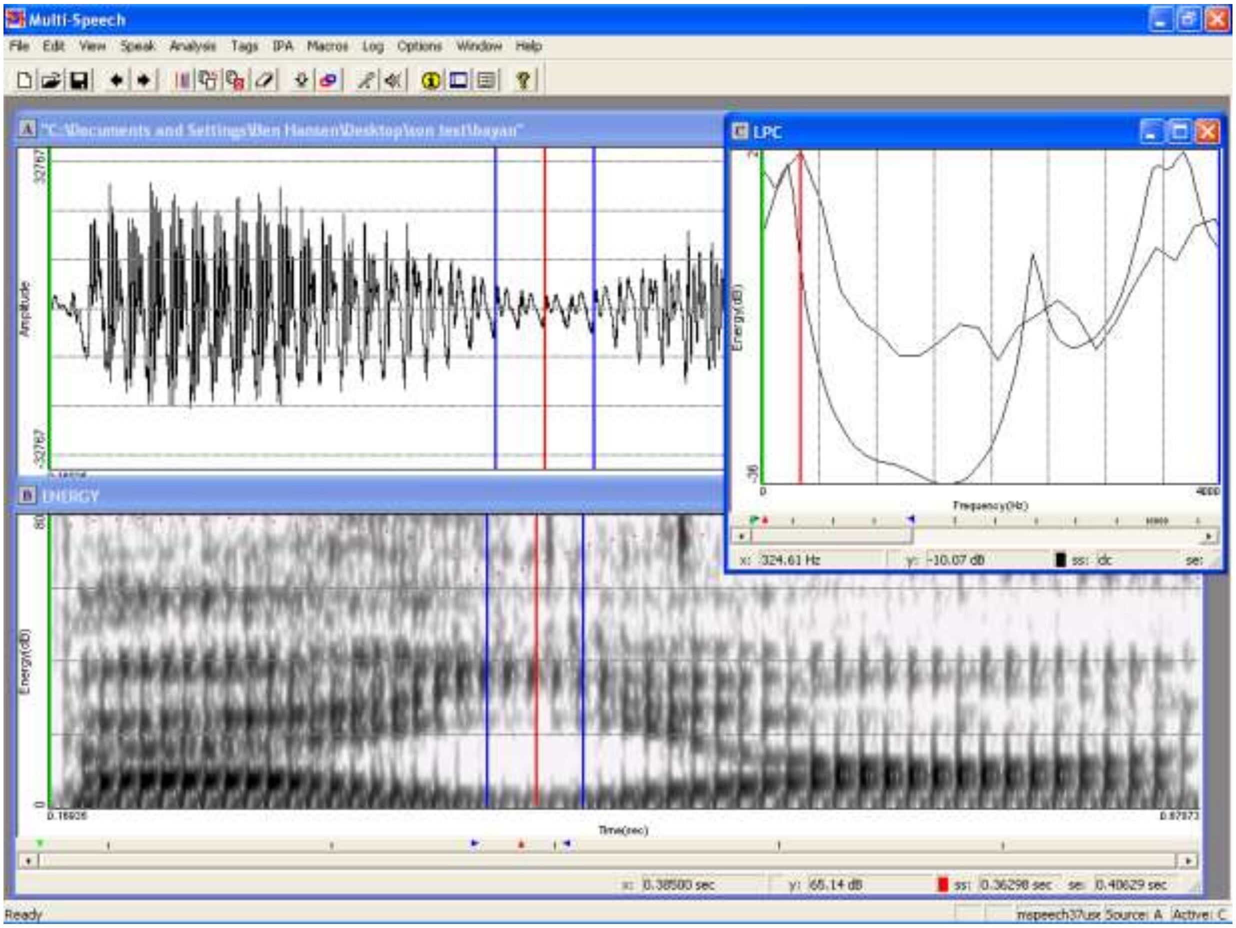

Figure 2.1e Sample Multispeech Display for Consonant [j]

Figure shows synchronized waveform and spectrogram displays from Multispeech. Onset and offset of constriction are shown by vertical blue lines. Red dots show the Decibel Amplitude (dB) track. Location of the consonant amplitude measurement is indicated by the vertical red line. Inset in upper right shown LPC spectrum at the amplitude measurement location.

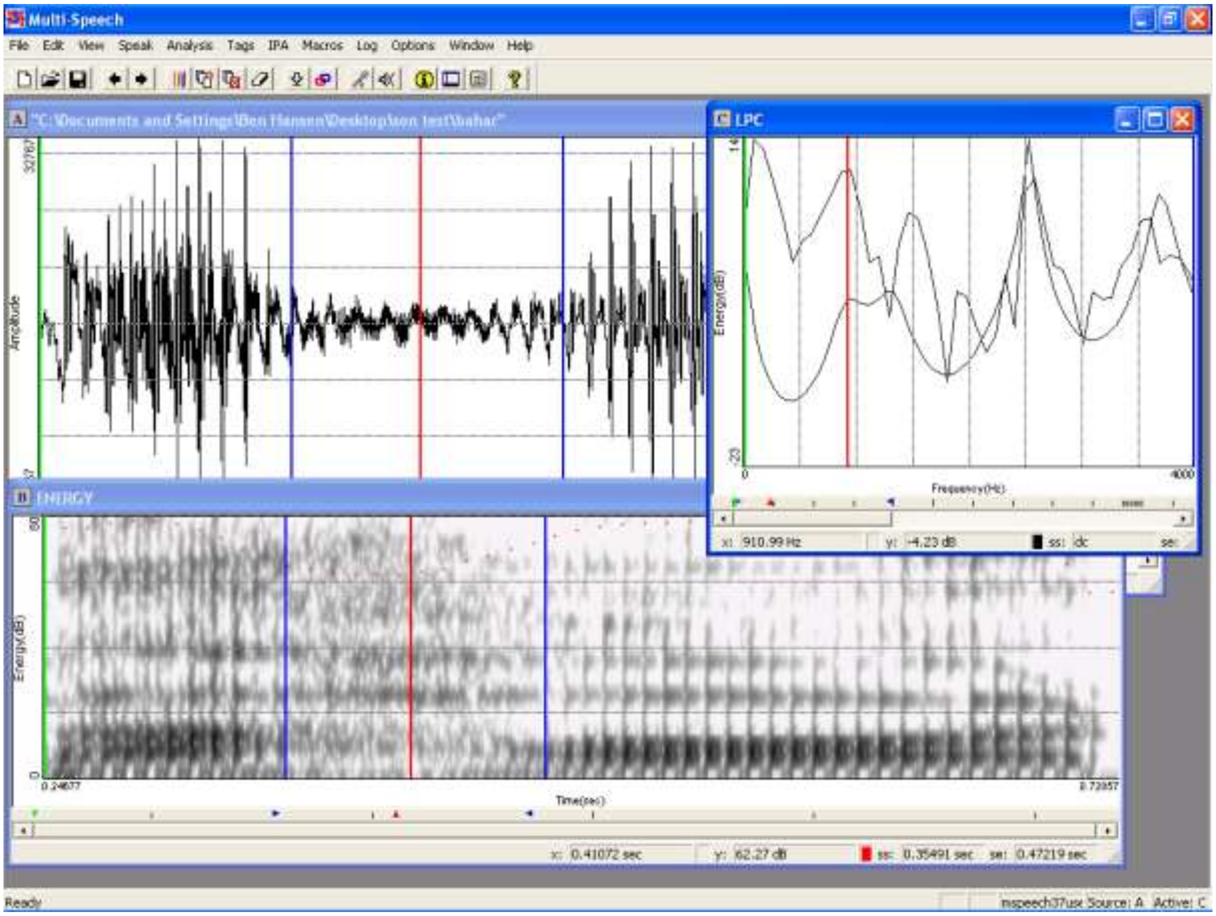

Figure 2.1f Sample Multispeech Display for Consonant [h]

Figure shows synchronized waveform and spectrogram displays from Multispeech. Onset and offset of constriction are shown by vertical blue lines. Red dots show the Decibel Amplitude (dB) track. Location of the consonant amplitude measurement is indicated by the vertical red line. Inset in upper right shown LPC spectrum at the amplitude measurement location.

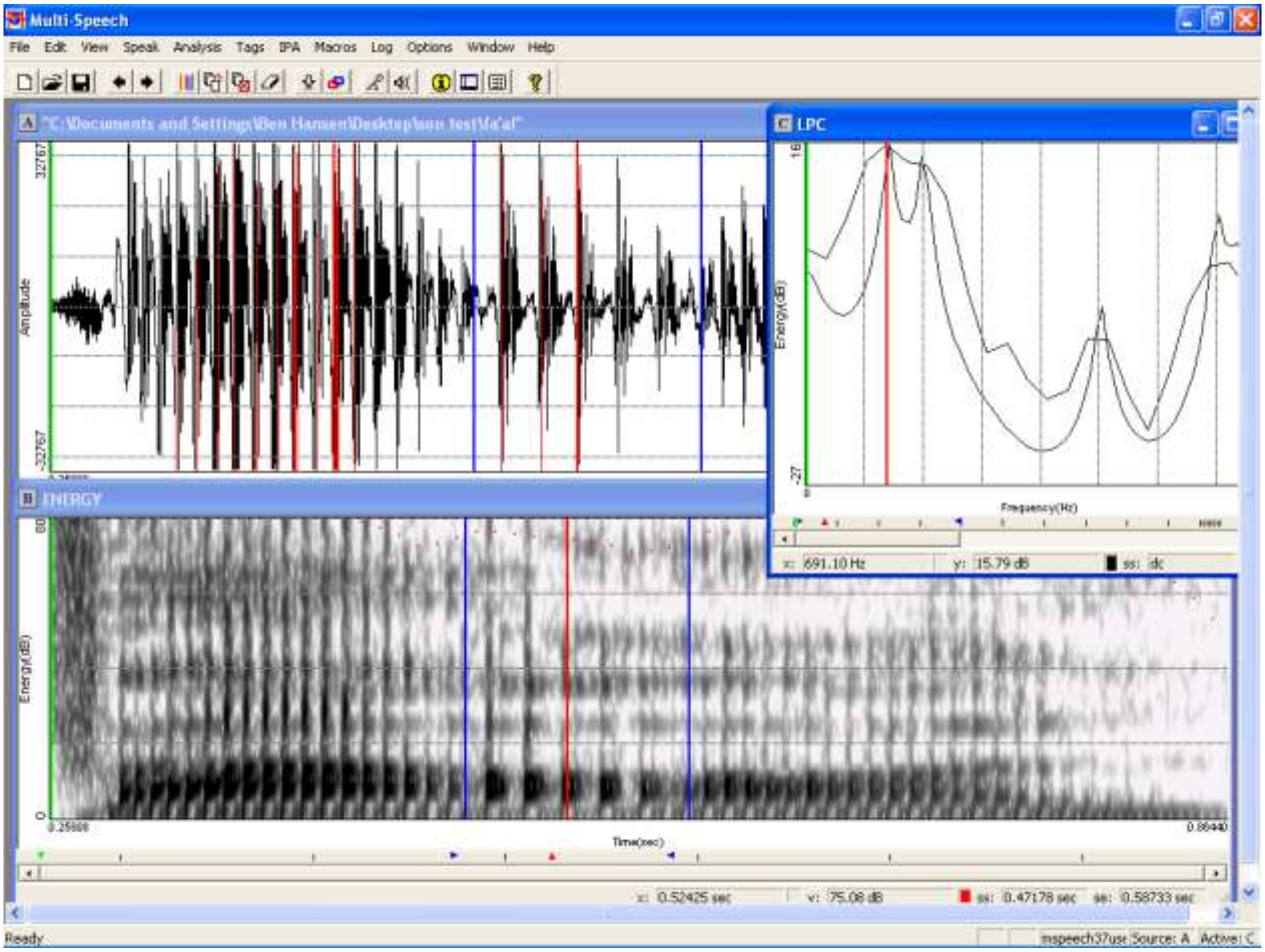

Figure 2.1g Sample Multispeech Display for Consonant [ʔ]

Figure shows synchronized waveform and spectrogram displays from Multispeech. Onset and offset of constriction are shown by vertical blue lines. Red dots show the Decibel Amplitude (dB) track. Location of the consonant amplitude measurement is indicated by the vertical red line. Inset in upper right shown LPC spectrum at the amplitude measurement location.

[l] The duration of lower energy wave, coinciding with steady state formant structure visible in the spectrogram.

[j] From the point in the spectrogram at which F1 and F2 movement from the low vowel ceases to move toward the target formants of [j], to the point at which F1 and F2 movement resumes toward the low vowel targets.

[h] The duration of decreased amplitude with reduced complexity in the waveform. This interval is taken to approximate the period of maximum abduction of the glottis.[7]

[ʔ] The interval measured from the onset of creakiness, a period of irregularly-spaced pulses of variable, generally decreasing amplitude, to the point at which amplitude sharply increases at the beginning of the first wave period of the following vowel.[8]

- Decibel Amplitude: I used the Multispeech Energy function to calculate the RMS average decibels of sound pressure (dB SPL) in 20-second windows. I logged the amplitude in dB at the midpoint of the preceding vowel and the midpoint of the test consonant constriction.
- Formant Transition Duration: In the spectrogram display, I measured the time from the beginning of either F1 or F2 formant movement in the vowel preceding the test consonant, up to the onset of the test consonant constriction. The glottal

---

[7] This is not "constriction" per se, but is the result of flexion of the posterior cricoarytenoid muscle responsible for abduction of the vocal folds.

[8] This interval corresponds only approximately to the period of maximum adduction, since the irregular airflow through the glottis is also a function of the pulmonic pressure, which could produce more or less creaky voicing while the glottis is adducted (Stevens 1998).

segments do not have a formant target different from the surrounding vowels, so there is no formant transiton interval measurement for these segments.

- Preceding Vowel Offset F1: I created a narrow band LPC analysis profile at the beginning of the last pulse of the vowel preceding the test consonant constriction and logged the frequency of the first peak following the fundamental frequency visible in the frequency analysis profile. The frame length for female speakers was 10ms and for male speakers, 15ms. In those instances in which it was impossible to distinguish F1 in the LPC analysis (often due to interference from the female speakers' high fundamental frequency), I obtained the F1 value from an FFT analysis.

2.3 RESULTS

2.3.1 Frame

Each speaker read the test materials at three distinct speaking rates. The continuous variable Frame is a measure of the duration of the carrier sentence, excluding the test word itself. Frame is thus a continuous measure of the speaking tempo: a four-syllable interval that is shortens as the speaking rate increases. Table 2.5 shows the mean values of Frame for each of the three rate categories in the Rate categorical parameter. The average speaking rate in syllables per second may be obtained by dividing the number four (which is the number of syllables within the frame sentence) by the value of Frame.

Speaker	Conversational	Fast	Very Fast
f1	1.20	0.86	0.56
f2	1.11	0.71	0.48
m1	0.99	0.72	0.51
m2	0.68	0.59	0.52

Table 2.5 Mean Frame Duration (seconds) by Speaker.

Mean duration of the four-syllable interval [minu___rɑ did] at three speaking rates.

Each individual speaker used a different speaking tempo at the conversational rate. Subjuct f1 had the longest average conversational frame duration at 1.20 seconds, corresponding to a relatively slow tempo of 3.3 syllables per second. Subject m2 had the shortest conversational frame duration, averaging 0.68 seconds, a much faster tempo of 5.9 syllables per second. Each speaker did reduce her or his average frame duration when instructed to speak at the fast rate, and reduced the frame even more at the very fast rate. All four of them spoke at about 8 syllables per second at this fastest rate (about a 0.5 second frame duration). I found no significant difference in the Frame variable between geminates and singletons or among the various segments. Frame differed significantly between each of the three rates. The three rate categories therefore do represent three different speaking rate populations.

2.3.2 Constriction

In line with my previous observations of Persian geminates, and consistent with observations of the geminate/singleton contrast in other languages, I consider the constriction duration to be the primary indicator of the Length variable, with singletons having significantly shorter durations than geminates. I also expect to find that all segments shorten as the speaking rate increases (Lehiste 1970). Thus the constriction at the conversational rate should be longer than the constriction at the fast rate, and the fast rate constriction should be longer than the very fast rate constriction. Although I do not have information on the inherent duration of Persian consonants, different inherent durations have been reported for different segments in other languages such as English (Klatt 1976). I expect that Persian consonants will also differ in duration.

Figures 2.2 through 2.8 below are scatter plots exhibiting the consonant constriction duration values for each consonant segment plotted against the average syllable duration, which is equal to one-fourth of the duration of the parameter Frame. Each figure represents one of the seven consonants studied, and is divided into four panels containing data from each of the four subjects. The ordinate represents Constriction in milliseconds, measured for each segment as described in section 2.2.12. The abscissa is the average syllable duration in milliseconds, which equals the value of Frame obtained for the token, divided by four. Filled diamonds represent instances of geminate consonants; unfilled diamonds are singletons. Tokens from all three speaking rates are plotted together. A least-squares linear regression model is plotted for each Length category. The regression equations and correlation coefficients appear on the plots next to the corresponding linear models.

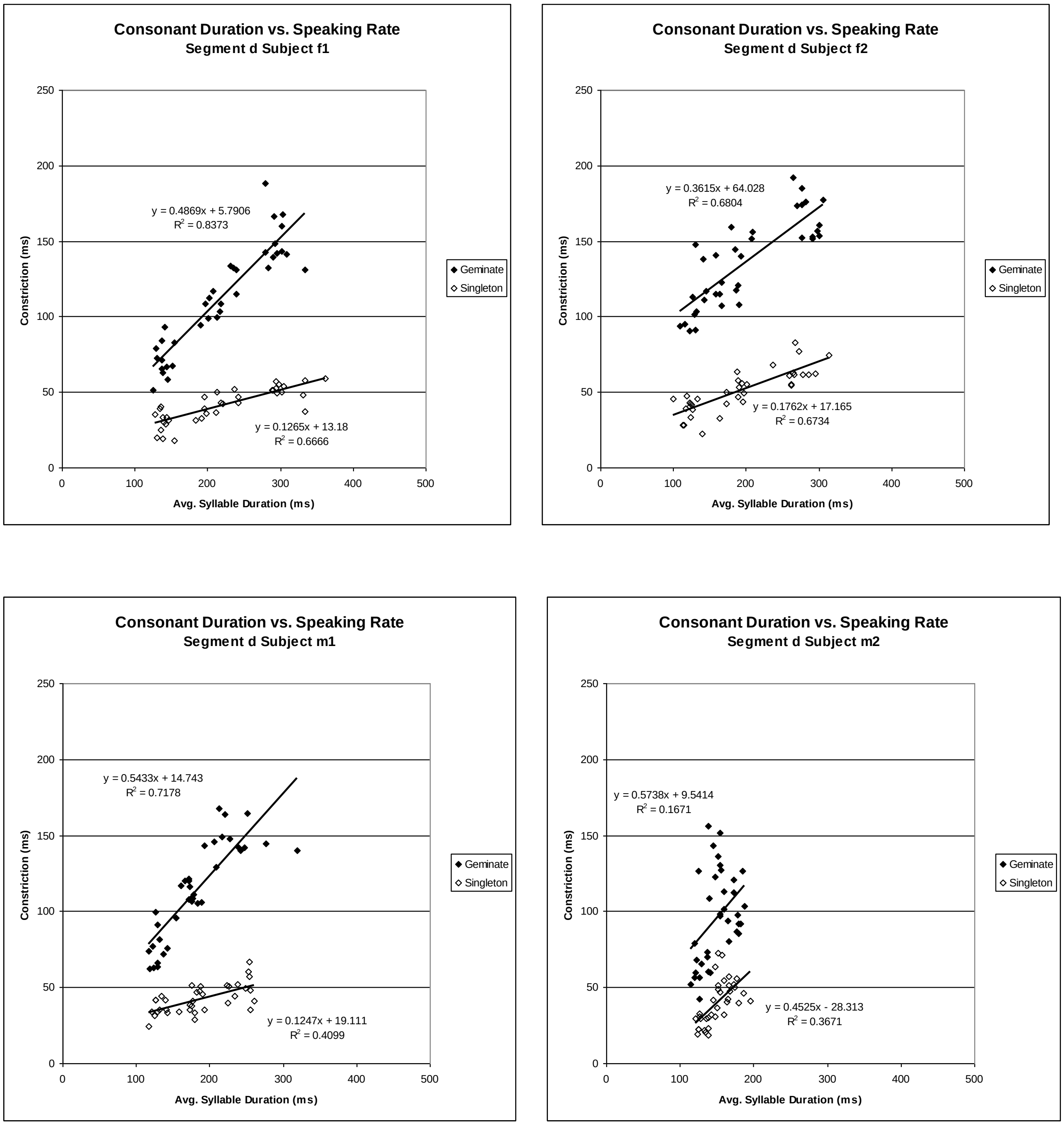

Figure 2.2 Consonant Duration vs. Speaking Rate Scatter Plots. Segment [d].

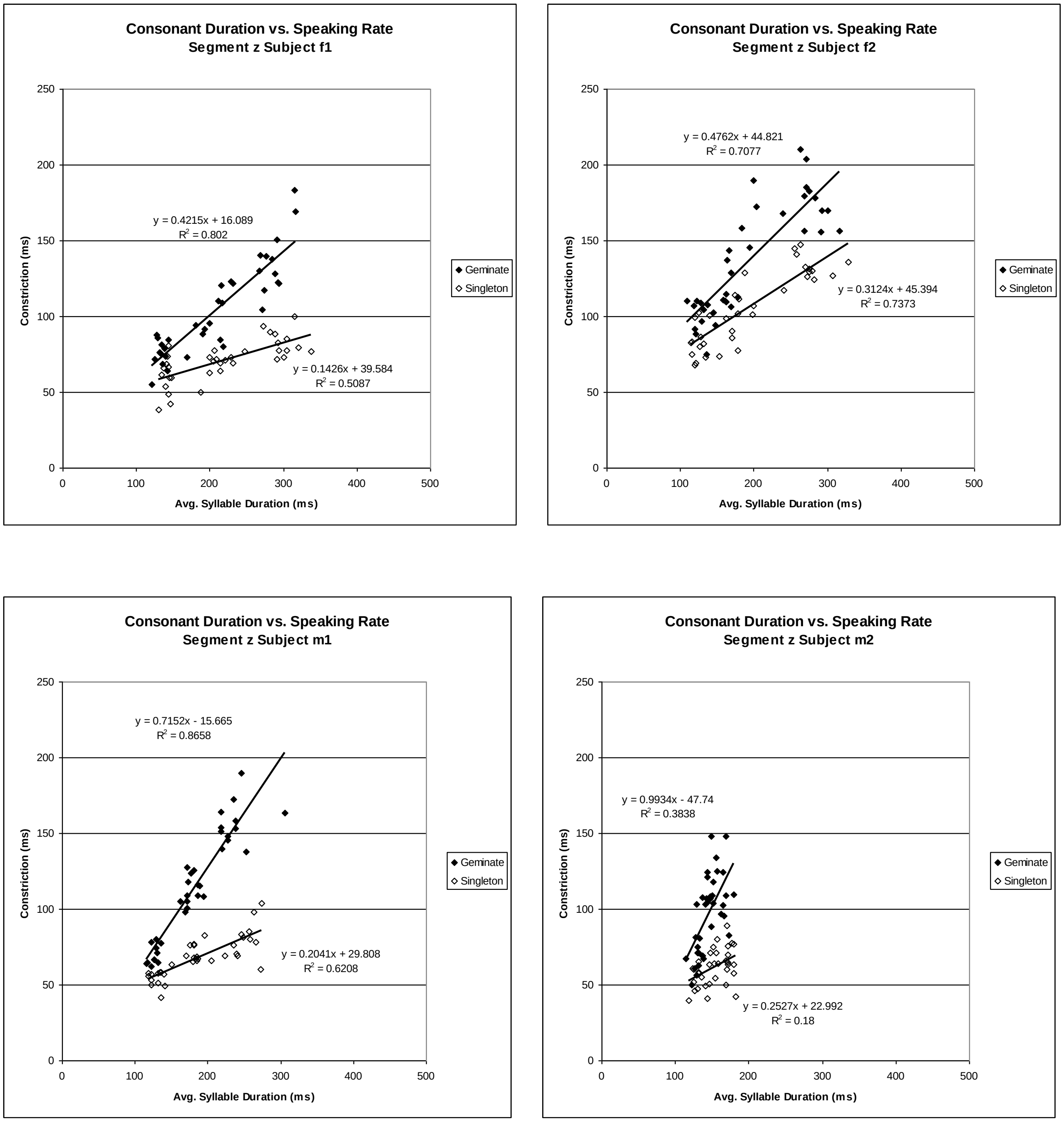

Figure 2.3 Consonant Duration vs. Speaking Rate Scatter Plots. Segment [z].

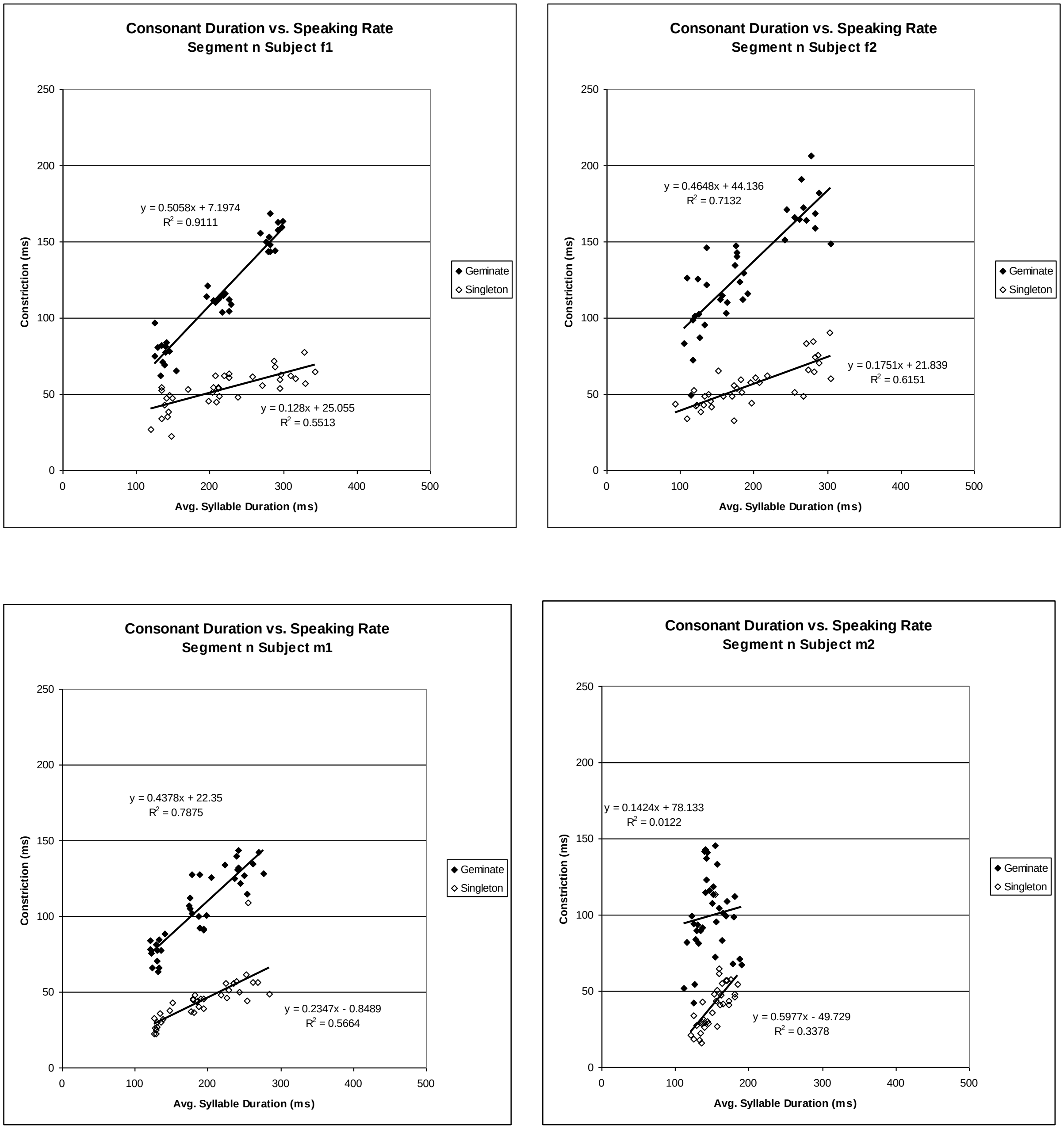

Figure 2.4 Consonant Duration vs. Speaking Rate Scatter Plots. Segment [n].

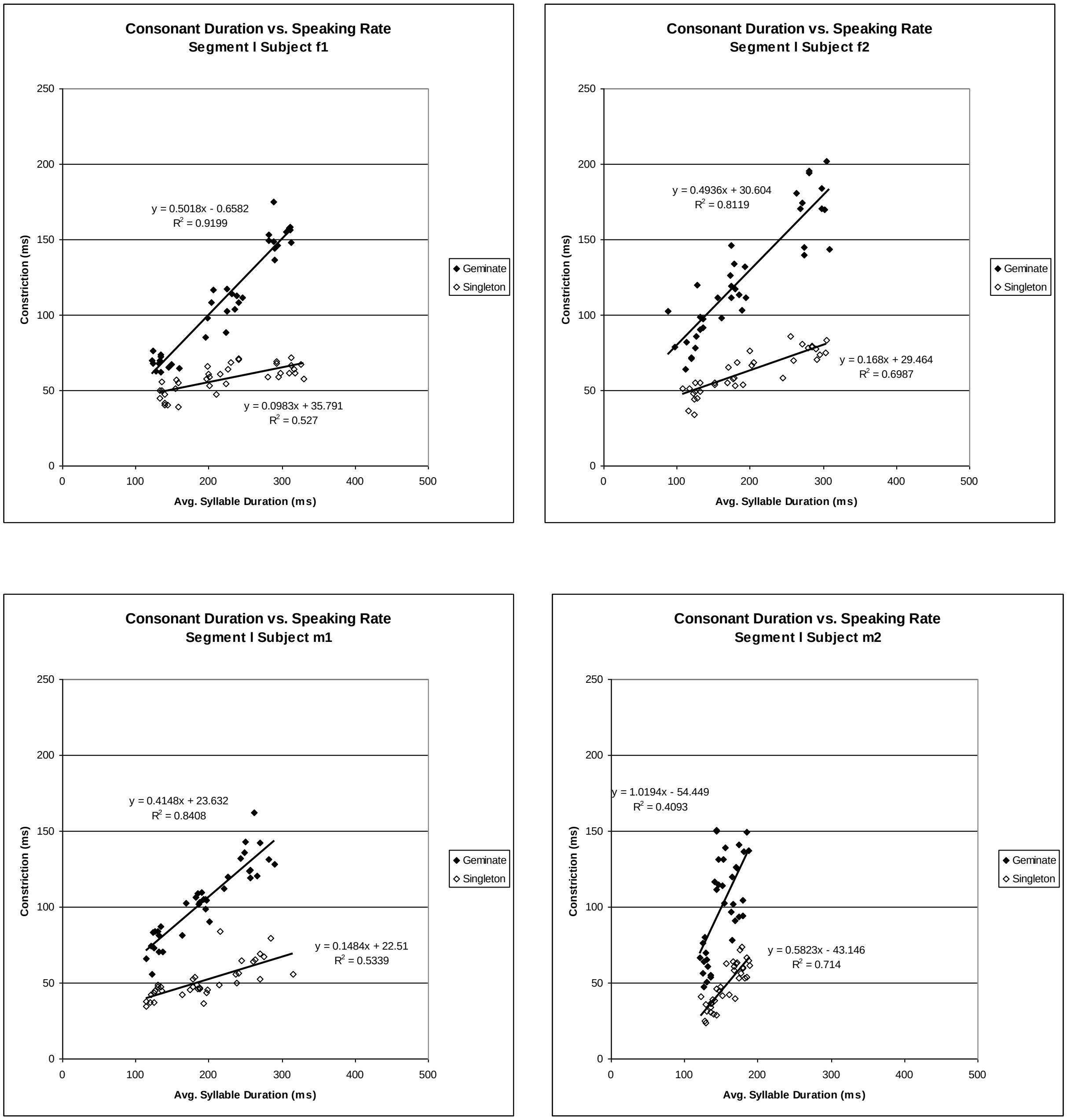

Figure 2.5 Consonant Duration vs. Speaking Rate Scatter Plots. Segment [l].

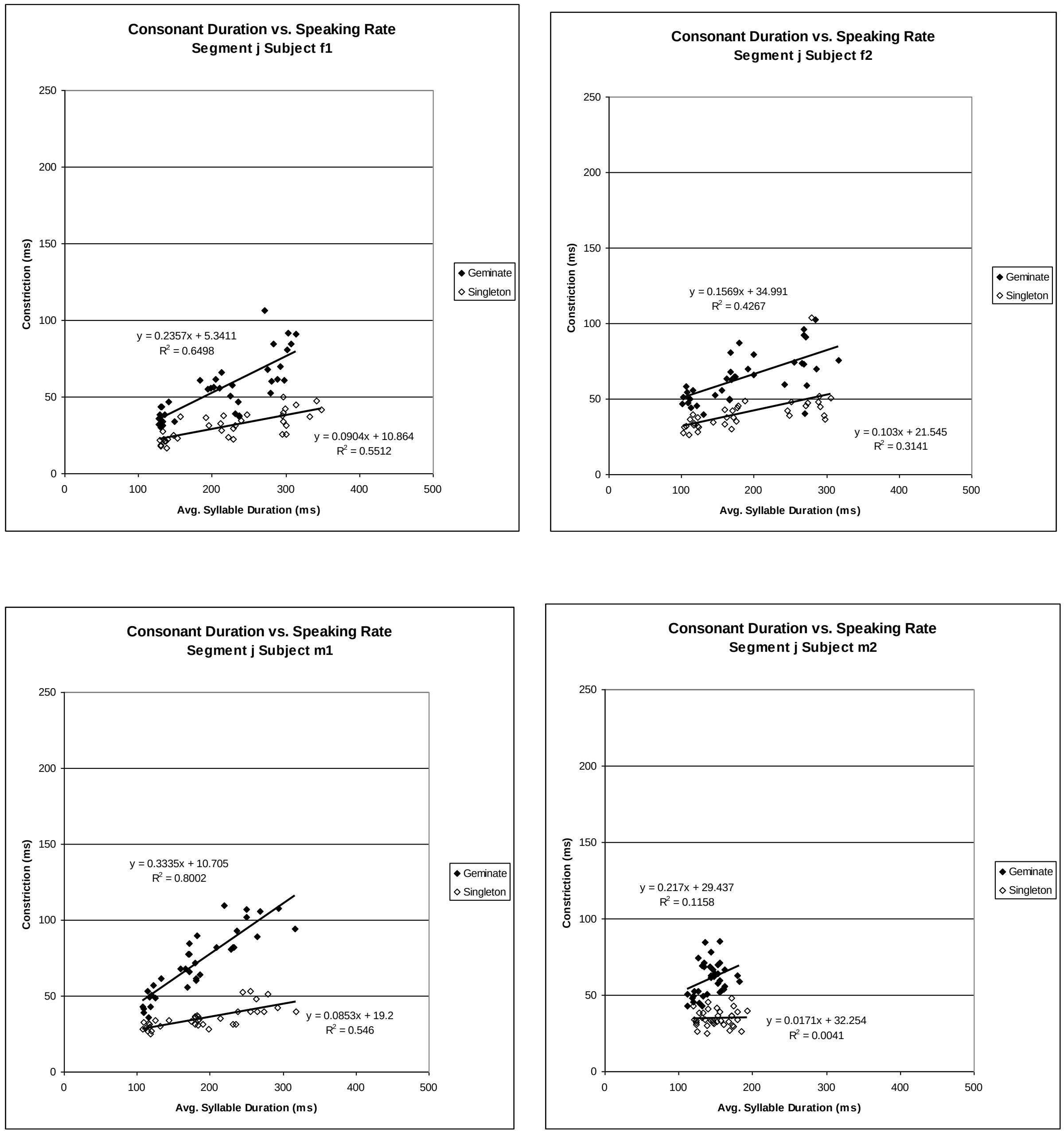

Figure 2.6 Consonant Duration vs. Speaking Rate Scatter Plots. Segment [j].

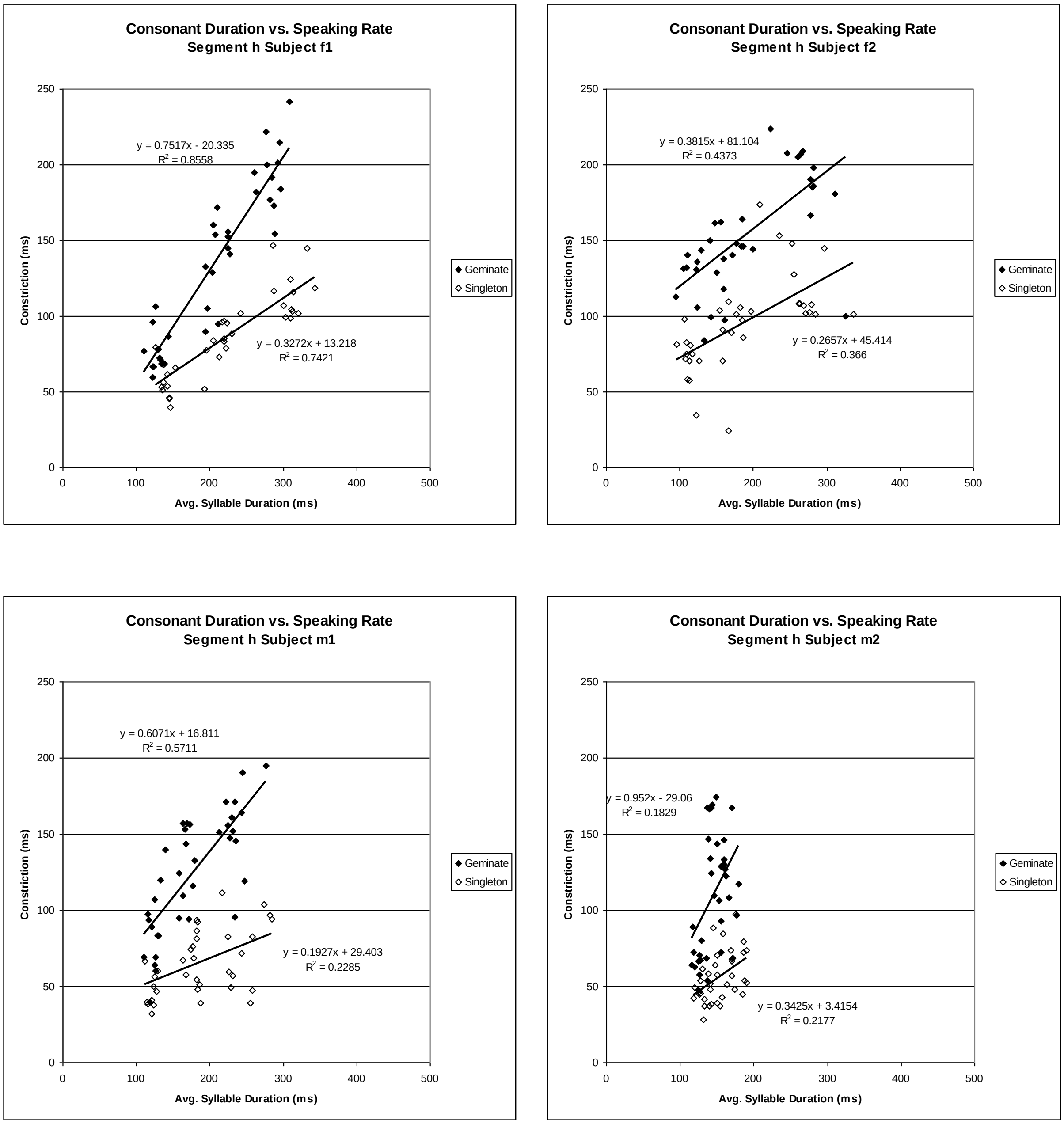

Figure 2.7 Consonant Duration vs. Speaking Rate Scatter Plots. Segment [h].

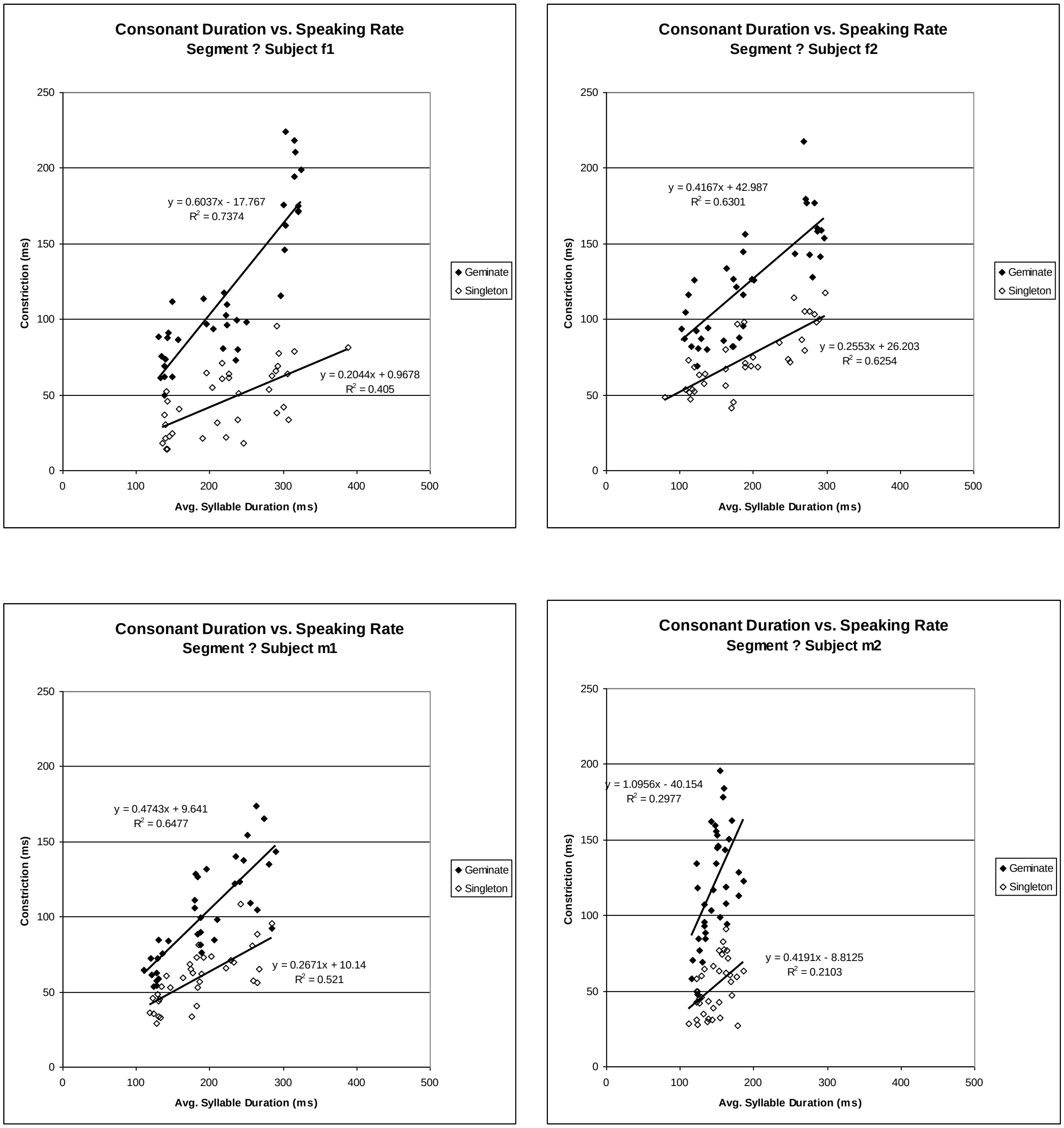

Figure 2.8 Consonant Duration vs. Speaking Rate Scatter Plots. Segment [ʔ].

Other authors (Pind 1995, Pickett et al. 1999, Hansen 2003) have plotted consonant duration against the duration of the preceding vowel, but I concluded in Hansen 2003 that the average syllable duration functions as a better metric of tempo for the purposes of discriminating between geminate and singleton clusters in Persian, due to the fact that, unlike Icelandic (Pind 1995) and Italian (Picket et al. 1999), the vowel preceding geminate consonants tends to be not shorter, but longer than the vowel preceding singletons. The tendency for vowels preceding consonants to longer than those preceding singletons has also been observed in Finnish (Lehtonen 1970) and Japanese (Smith 1991, Han 1994). By plotting Constriction against a non-correlated speaking rate parameter, it is possible to appreciate the fact that Persian speakers take speaking rate into account when producing geminates. That is, although a geminate produced at a fast speaking rate may have a shorter duration than that of a singleton spoken at a slower rate, geminates are largely longer that singletons at a *particular* speaking rate, as can be observed by the fact that geminate clusters are separate from singleton clusters in Figures 2.2 through 2.8. The geminate clusters begin to overlap only at the very fastest speaking rate. In general, the slopes of the regression lines are positive, indicating that the constriction duration increases as the syllable duration increases, i.e., as the tempo slows. The slopes of the geminate regression lines are steeper than the slopes of the singleton regression line, indicating that speaking rate has more effect on the geminate duration than it does on the singleton duration.[9]

[9] The lone exception is m2's production of segment [n]. This subject did not produce a wide range of speaking rates, so the correlation coefficients of the regressions were very low. In the case of geminate [n], the slope was flatter than the singleton, but the very low correlation ($R^2 = 0.0122$) leads us to discount the regression model for these data.

The alveolar stops [d] and [n] and the lateral [l] exhibit a clear separation between singletons and geminates at all speaking rates for all speakers except m2. In the case of the fricative [z] and the glide [j] there is some overlap between the categories as the regression lines appear closer together. In the case of [j] all durations are short in comparison to the other segments. The glottals [h] and [ʔ] exhibit quite a bit more variability than the other segments as evidenced by the generally lower values of R^2. This is undoubtedly at least partly due to the difficulty of segmentation of these segments.

Figures 2.9a and 2.9b presents each speaker's Constriction data in bar charts grouped by category. In Figure 2.9a, the data are grouped by rate. The white bars represent the average geminate constrictions for all 84 of the speaker's productions at each speaking rate. The gray bars show the mean singleton durations. The error bar shows the magnitude of the standard deviation for the category. This top panel indicates that the durations of the consonants decrease systematically as speaking rates increase, and that there is a greater decrease in the duration of singletons than in the duration of geminates.

Figure 2.9b shows mean Constriction grouped by segment-36 tokens per segment type-with the rates pooled for each speaker. It is apparent from the graph that there are inherent differences among the segment types. The fricatives [z] and [h] tend to have larger values of Constriction, and [j] much smaller ones. Singletons are shorter than geminates in every case as expected. Error bars on the geminate segments are longer than those on the singletons, but this may be a mere function of scale.

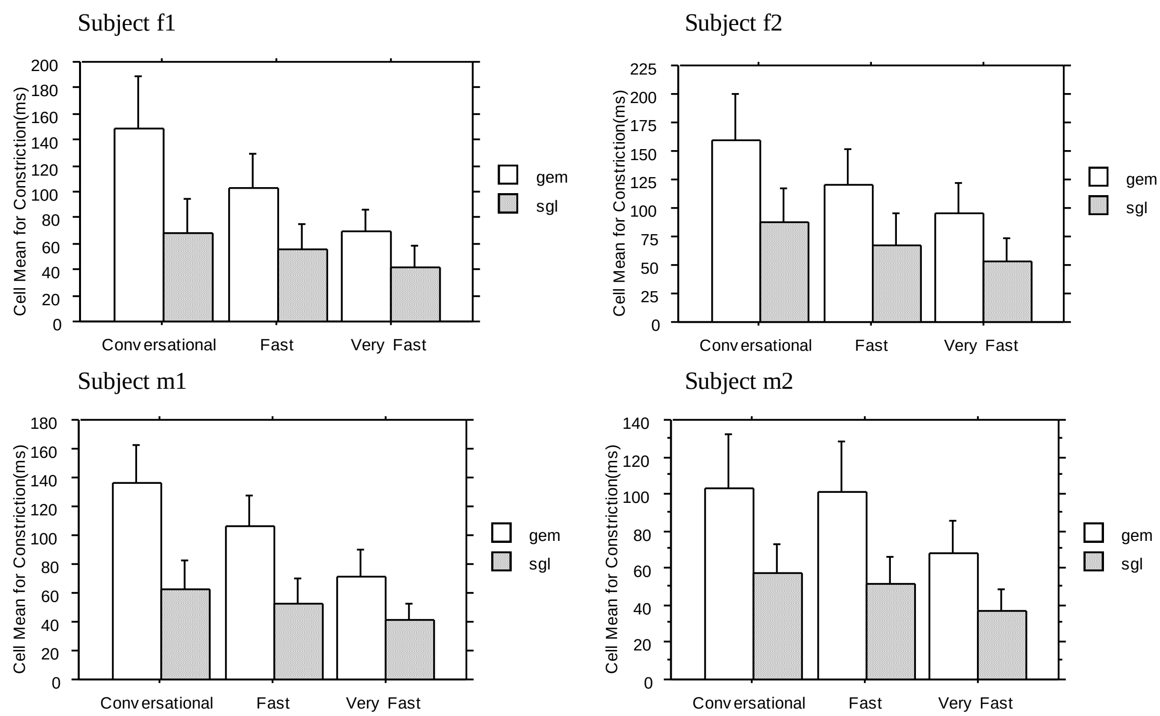

Figure 2.9a Consonant Constriction Duration (ms) by Rate
Error bars indicate standard deviation.

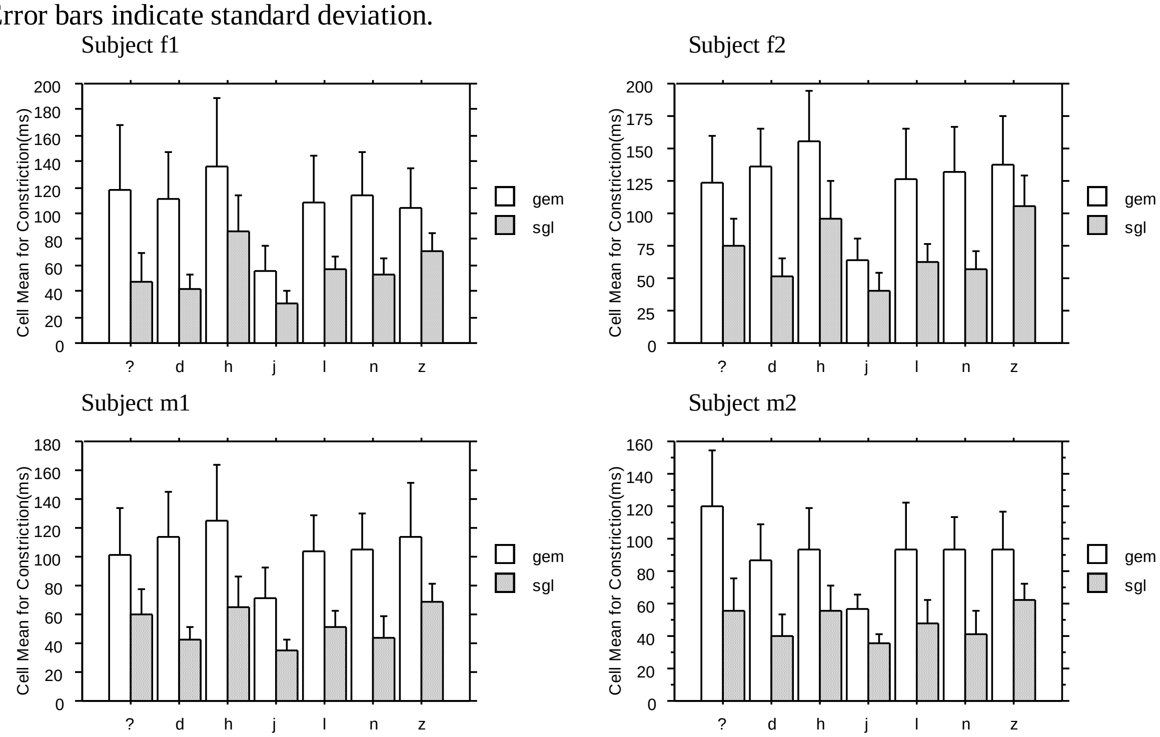

Figure 2.9b Consonant Constriction Duration (ms) by Segment
Error bars indicate standard deviation.

I performed individual 3-way ANOVA analyses of the dependent variable Constriction for each of the four subjects. The independent variables were Length, Rate and Segment. The analysis showed that the main effect of Length was highly significant for each speaker:

f1: F(1,461)=1872, p<.0001
f2: F(1,462)=1321, p<.0001
m1: F(1,461)=1952, p<.0001
m2: F(1,462)=1172, p<.0001

Table 2.6 ANOVA Results for Constriction Duration (ms): Main effect of Length

The results in Table 2.6 confirm what is apparent on inspection of Figure 2.9: geminates have a larger value of Constriction than singletons for all rates and segments.

In the 3-way ANOVAs for Constriction, Rate was also highly significant for each speaker, indicating that constriction durations did shorten as the subjects read faster.

f1: F(2,461)=676, p<.0001
f2: F(2,462)=339, p<.0001
m1: F(2,461)=445, p<.0001
m2: F(2,462)=198, p<.0001

Table 2.7 ANOVA Results for Constriction Duration (ms): Main effect of Rate

The mean geminate and singleton constriction durations produced at the three rates by each subject are depicted in Figure 2.9a. Singletons were shorter than geminates and both had shorter durations at faster rates in every case.

The factor Segment was also significant in the 3-way ANOVAS for Constriction.

f1: F(6,461)=160, p<.001
f2: F(6,462)=137, p<.001
m1: F(6,461)=74, p<.001
m2: F(6,462)=62, p<.001

Table 2.8 ANOVA Results for Constriction Duration (ms): Main effect of Segment

The Figure 2.9b shows the mean geminate and singleton constriction durations broken down by segment, averaged across all rates. The pattern observed here varies among speakers, but some general observations are possible. For all speakers the shortest constriction was produced in the segment [j]. For three of the four speakers, the longest constriction was in the segment [h], and for the other speaker (m2) [h] was next-to-longest after [ʔ]. The next longest segment after [h] for subjects f1, f2 and m1 was [z]. For m2, [z] was intermediate in duration between [n] and [d]. For all of the speakers, [n], [l] and [d] were not significantly different from each other, and their rank in terms of constriction duration varied among speakers. The duration of the glottal stop [ʔ] was extremely variable. As mentioned, for m2 it was the longest segment, while for m1 it was the shortest segment except for [j], and for f1 and f2 its duration fell in between the extremes.[10]

Fischer's PLSD post-hoc analysis resulted in the following constriction duration rankings for the seven consonants. In this analysis, all three speaking rates and both

[10] The variability of the glottal stop duration measurements must partly be attributed to the difficulty of segmenting this consonant. The precise point at which glottalization begins is sometimes not clear in the waveform. I observed that individual tokens showed variation in the acoustical reflex of the glottal adduction gesture. At times a full obstruction was produced, while at others a period of creakiness is present. At other times there is a period of stiff voice preceding the creak which lacks the irregular periodicity and amplitude in the wave form. The interval of stiff voice was not included in the constriction duration.

phonological length categories are pooled for a given segment. An asterisk indicates that there is a significant difference between the successive segment constriction durations:

f1	**[h *> z,n,ʔ,l,d *> j]**
f2	**[h,z *> ʔ,l,n,d *> j]**
m1	**[h,z *> d,l,n,ʔ *> j]**
m2	**[ʔ *> h,l,n,z,d *> j]**

Table 2.9 Segment Constriction Duration Rankings

Two-way interactions in the Length x Rate x Segment 3-way ANOVAs were all significant for each of the four subjects. The Length x Rate interaction indicates that geminates tend to shorten more than singletons do at faster rates. Hansen (2003) observed this interaction for stops, but Arvaniti (1999) did not observe Length x Rate interaction for some Cypriot Greek sonorants.

I performed post-hoc two-way ANOVAs on individual segments for each speaker and found that the Length x Rate interactions were significant at the .05 level for all consonants with the following exceptions: f2 and m2 [j] and f2 [h]. The [j] constriction is quite short to begin with and so although the geminate did shorten more than the singleton did for all speakers, the effect did not achieve significance for f2 and m2, possibly due to the fact that the geminate and singleton constriction duration values are close together relative to the measurement variance. In the case of [h] produced by f2, the large standard deviation in this segment (see Figure 2.9b) would require a bigger sample to demonstrate significance. Overall, the data are consistent with the observation that the Length x Rate interaction is not limited to particular manners of articulation, but applies to all of the segment types studied.

The Length x Segment interaction observed for all four subjects in the 3-way ANOVAs indicates that for some segments singletons are much shorter than geminates, while for other segments the difference is not as great. This phenomenon is apparent in the ratio of the geminate constriction durations to the singleton constriction durations. Arvaniti (1999) used this ratio to describe the geminate/singleton contrast, for example. The ratios of mean geminate constriction durations to mean singleton constriction durations are given in Table 2.11.

Subject	ʔ	h	j	l	n	z	d
f1	2.51	1.58	1.77	1.88	2.15	1.46	2.71
f2	1.65	1.62	1.57	2.04	2.33	1.30	2.68
m1	1.69	2.08	2.02	2.04	2.40	1.67	2.69
m2	2.14	1.66	1.58	1.94	2.27	1.52	2.18

Table 2.10 Geminate/Singleton ratios.
Ratios of geminate constriction durations to singleton constriction durations. Rates pooled.

It is notable that for all subjects, the fricative [z] has the smallest geminate/singleton ratio. [d] had the largest or second-largest ratio in every case.

A Rate x Segment interaction was observed in the 3-way ANOVAs. This Rate x Segment interaction means that the inherently longer segments such as [z] have a greater tendency to shorten at fast speaking rates than the shorter ones like [j] do.

Lastly, the three-way interaction Length x Rate x Segment was only significant for f1 and m2.

f1: F(12,461)=3.715, p<.0001
f2: F(12,462)=1.548, p=.104
m1: F(12,461)=1.431, p=.1481
m2: F(12,462)=1.983, p=.0241

Table 2.11 ANOVA Results for Constriction Duration (ms): Interaction of Length x Rate x Segment

The existence of a three-way interaction has several interpretations. One interpretation is that for these speakers, the Rate x Length interaction is stronger for some segments (such as the inherently long ones) than for others. Another interpretation is that the fact that some segments have much longer geminates than singletons is less evident at faster rates, since geminates shorten more than singletons do under fast speaking conditions.

2.3.3 Vowel

Vowels preceding geminates were longer than vowels preceding singletons. Overall, the vowel preceding a geminate averaged 105.1 ms (s.d.=32.1) vs. 99.0 ms (s.d.=25.7) preceding singletons. A three-way ANOVA analysis of the preceding vowel duration using the independent variables Length, Rate and Segment indicated the following:

The main effects of Length, Rate and Segment on the vowel duration are significant ($p<.02$) for all speakers. So the faster the rate the shorter the vowel duration is; and some of the consonant segments had shorter preceding vowels than others. All two-way interactions were significant at the .05 level, but the three-way interaction was only significant for m2. The latter result means that for most of the speakers, any extra shortening effects that may apply to vowels preceding geminates spoken at faster rates, seem to apply to all segment types.

These results indicate that in Persian, the duration of the preceding vowel may be a reinforcing cue to consonant length.

2.3.4 Amplitude Difference

Figures 2.10a and 2.10b show the comparison of geminate to singleton Amplitude Differences at the three speaking rates (2.10a) and by segment (2.10b). Positive values of Amplitude difference indicate that the vowel is louder than the consonant. Only for m1's [n] is the mean amplitude of the consonant greater than that of the vowel, producing a negative Amplitude Difference.

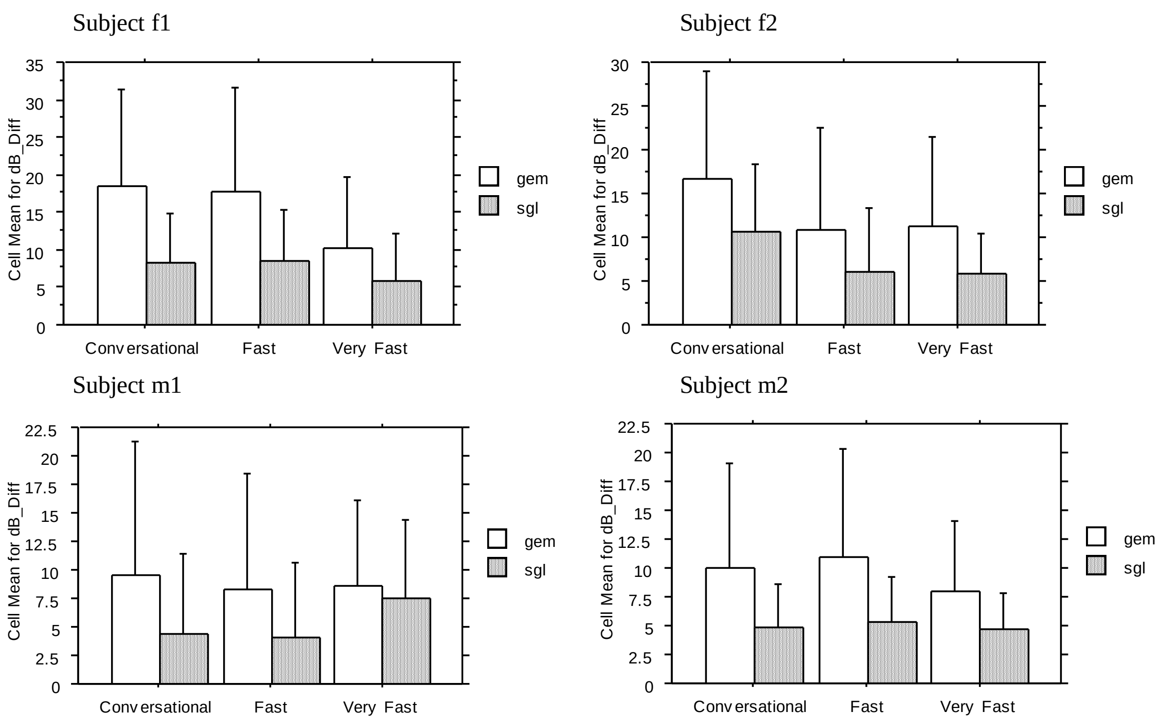

Figure 2.10a Amplitude Difference: V minus C (dB) by Rate
Error bars indicate standard deviation. Smaller numbers indicate greater consonant amplitude.

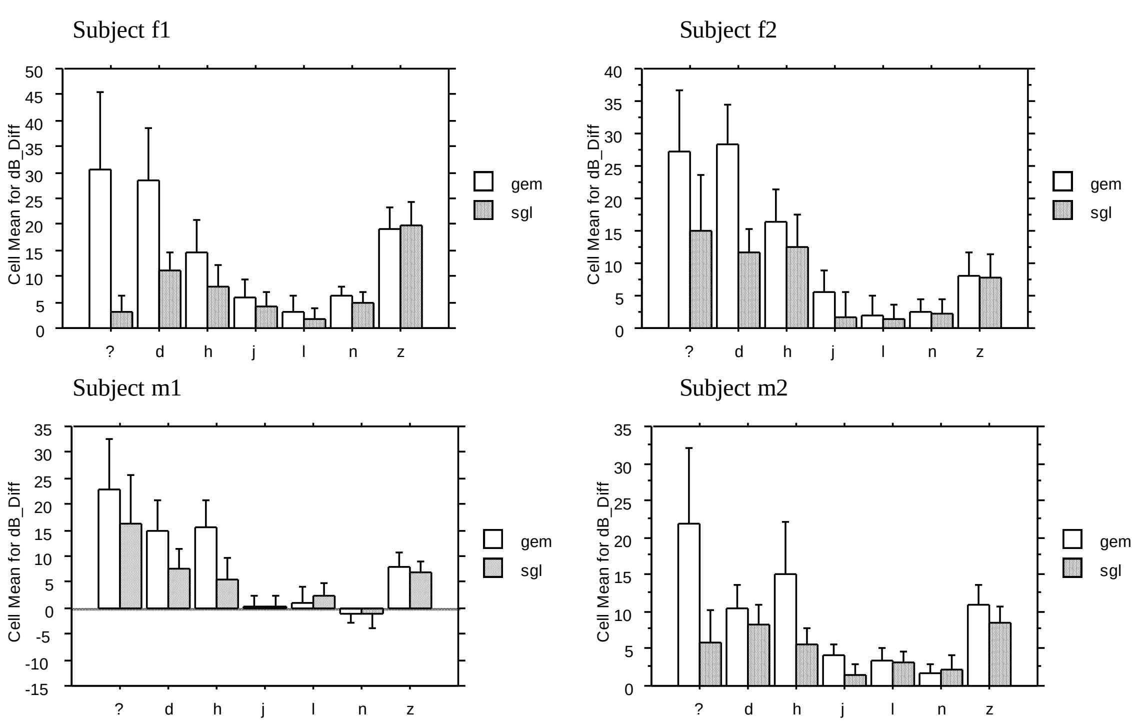

Figure 2.10b Amplitude Difference: V minus C (dB) by Segment
Error bars indicate standard deviation. Smaller numbers indicate greater consonant amplitude.

I performed a 3-way ANOVA on the variable Amplitude Difference using the independent variables Length, Rate and Segment. As expected, the main effect of segment was highly significant for all four subjects. The segments were expected to vary greatly in amplitude since the seven segments were selected to range in sonority from low to high.

f1: F(6,461)=191, p<.0001
f2: F(6,462)=368, p<.0001
m1: F(6,461)=229, p<.0001
m2: F(6,462)=123, p<.0001

Table 2.12 ANOVA Results for Amplitude Difference (dB): Main effect of Segment

Fischer's PLSD post-hoc analyses showed that sonorants [j,l,n] had the smallest amplitude difference among the seven consonants, and obstruents [z,d] consistently had relatively large amplitude difference. The glottals were variable, but generally had a greater amplitude difference than sonorants did, and so in this respect were less vowel-like than sonorants. In the tabulation below, speaking rates and length categories (geminates/singletons) are pooled. An asterisk indicates a significant difference between successive amplitude difference values.

f1	[d,z *> ʔ *> h *> n, j *> l]
f2	[ʔ,d *> z *> h *> j *> l,n]
m1	[ʔ *> d,h *> z *>l, j,n]
m2	[ʔ *> h,d,z *>l, j,n]

Table 2.13 Segment Amplitude Difference rankings

A striking pattern seen in Figures 2.10a and 2.10b is that singletons had less amplitude difference than geminates; i.e., singletons were louder than geminates. The main effect of Length for all four subjects had $p<.0001$. However, for each subject the Length x Segment interaction was also highly significant. In the case of segments [ʔ,h,d], the tendency for greater amplitude in singletons than geminates is observed for all four speakers. Singleton [j] was significantly louder than geminate [j] for all subjects except m1. In the case of [z], m1 and m2 produced significantly louder singletons than geminates, but f1 and f2 did not. On the other hand, except for f1, singleton sonorants [l,n] were not louder than their geminate counterparts at all. The geminate versions of m1's [l] and m2's [n] were actually louder on average than the singletons.

I expected that faster speaking rates would produce more coarticuated speech in which consonant amplitude approached the vowel amplitude. However, although a main effect of Rate on Amplitude Difference exists for all speakers (main effect of Rate, $p<.0001$), there is not a general tendency for faster segments to be more vowel-like, i.e. have a lower Amplitude Difference. Fisher's PLSD for Amplitude Difference did show that in the case of both of the female speakers, the mean Amplitude Difference decreased each time the speaking rate increased, but for m1, the fastest rate had a greater Amplitude Difference than either the conversational rate or the fast rate. For m2, the fast rate had a larger dB difference than the other two rates. In conclusion, contrary to my expectation, not all speakers's consonants became relatively louder at faster, presumably more coarticulated rates.

All interactions in the 3-way Length x Rate x Segment ANOVA for Amplitude Difference were significant ($p<.005$), with the following lone exception: for f2, the Length x Rate interaction was not significant ($F(462,2)=1.388$, $p=.2507$). For f2, the amplitude difference between geminates and singletons was about the same for all three rates.

As mentioned above, the Length x Segment interactions are highly significant ($p<.0001$ for all four subjects), a reflection of the fact that for some segment types geminates are not as loud as singletons, while for others, both are equally loud. The interactions involving rate were not interpretable because it was not clear what the effect of rate on amplitude difference was: not all speakers showed consistently reduced Amplitude Difference at faster rates. Likewise, the fact that the Length x Rate x Segment interaction was significant ($p<.002$ for all four subjects), seems to indicate that the strong impact of segment type on how much quieter geminates are than singletons is more apparent at some rates than others, but the meaning of this fact is obscure.

2.3.5 Formant Transition Duration

Because there is no target oral configuration for glottal consonants, formant transitions within the vowel preceding the test consonants are not seen for [ʔ,h]. Formant transitions are only observed for the oral consonants [j,l,n,z,d]. Therefore, no formant transition duration was logged for the glottal consonants [ʔ,h]. Figures 2.11a and 2.11b show the formant transition durations observed for the other segments [j,l,n,z,d].

The formant transition is longer before geminates than before singletons. In a 3-way ANOVA of Formant Transition Duration for the independent variables Length, Rate and Segment, the main effect of Length is significant.

f1: $F(1,330)=39.002$, $p<.0001$
f2: $F(1, 330)=5.419$, $p<.0205$
m1: $F(1, 330)=46.125$, $p<.0001$
m2: $F(1, 330)=5.662$, $p=.0179$

Table 2.14 ANOVA Results for Formant Transition Duration (ms): Main effect of Length

This result indicates that the longer geminate constriction gesture also involves a longer period of tongue movement before closure than is the case for singleton constrictions.

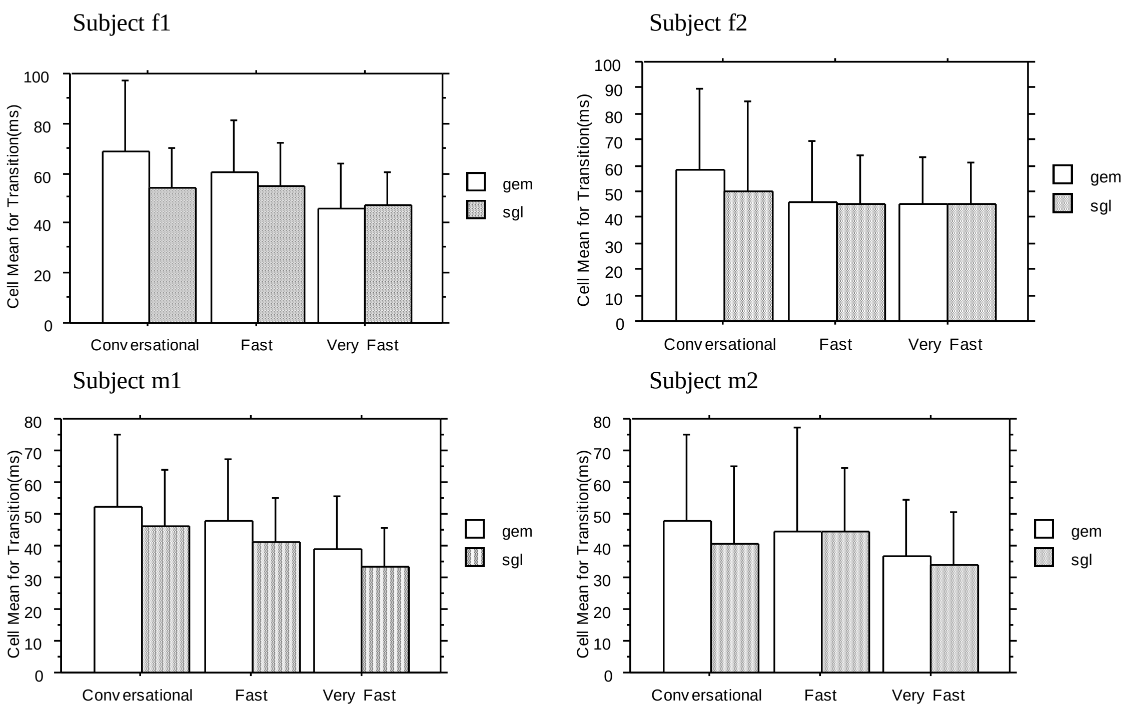

Figure 2.11a Formant Transition Duration Within Preceding Vowel (ms) by Rate

Error bars indicate standard deviation.

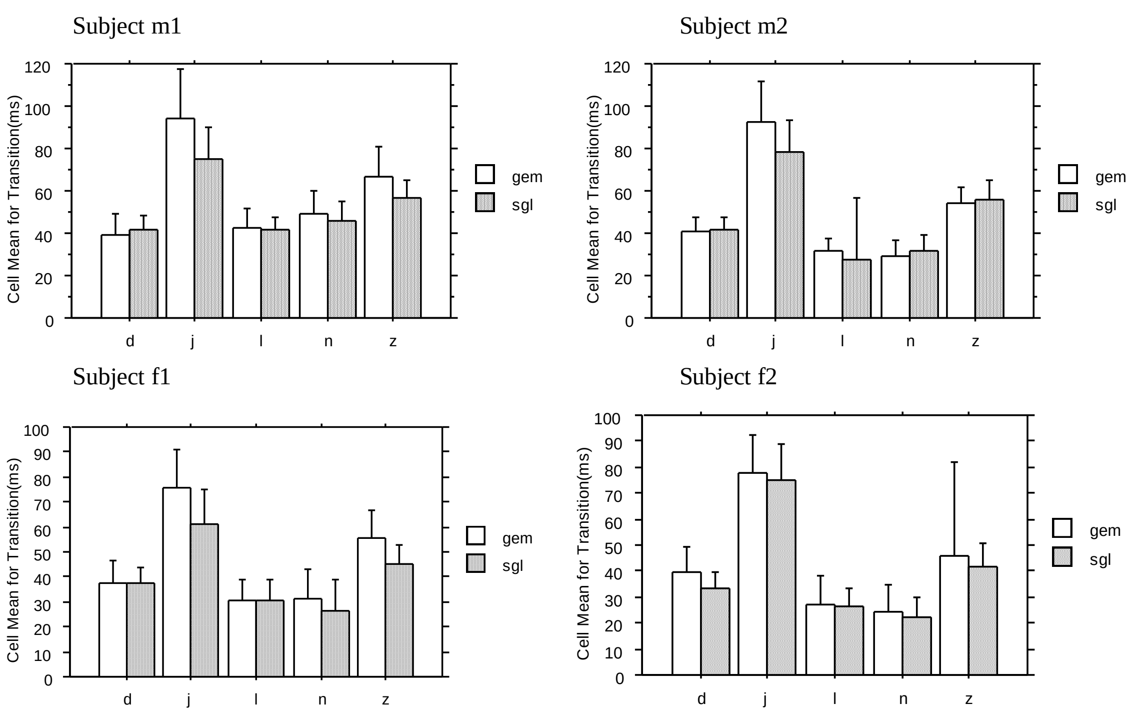

Figure 2.11b Formant Transition Duration Within Preceding Vowel (ms) by Segment

Error bars indicate standard deviation.

As with other temporal measures in speech, the formant transition becomes shorter as the speaking rate increases. The main effect of Rate is highly significant for all four subjects ($p<.0001$).

A crucial observation of this dissertation is that different segments are characterized by different formant transition durations. This observation is important since segments such as glides that have longer transitions are predicted less likely cross-linguistically to possess a length contrast. In effect, the main effect of Segment is highly significant for all four subjects ($p<.0001$), meaning that the type of segment involved strongly influences the duration of the transition. In Fisher's PLSD post-hoc tests, it was found that for all four subjects, the longest transitions were for [j], and the second longest transitions were for [z]. [d] had the next longest transitions for f2, m1 and m2, but [n] had a significantly longer transition than [d] or [l] for f1. [l] was not significantly different from [n] for f1, m1 and m2; and [l] was not significantly different than [d] for f1. The rankings are summarized below. Again, speaking rates and length categories are pooled and the asterisk indicates a significant difference between successive formant transition durations.

f2, m1, m2	[j *> z *>d *> l,n]
f1	[j *> z *>n *> l,d]

Table 2.15 Formant Transition Duration Rankings

By this measure, the glide [j] has the least distinct segment boundary with the preceding vowel. Somewhat surprisingly, [z] is next-to-least distinct. Although it is lower in sonority, the transition into it is gradual. [d,l,n] have shorter transitions without a clear order, though for 3 out of 4 subjects [d] had a longer transition than either [l] or

[n]. Certainly it is clear that higher sonority does not at all correlate with longer transitions, since the sonorant [l] is among the segments with the shortest transitions.

For all four subjects, the Rate x Segment interaction is significant ($p<.0001$), since the segments with the longer transitions tend to show more reduction in their transition durations at faster rates than those with shorter transitions do. On the other hand, Length x Segment interaction is significant only for f1, f2, and m1 ($p<.0001$), but not for m2 ($F(4,330)=.430$, $p=.7873$). This means that for m2, the difference in transition duration between geminates and singletons was about the same for all segments, while for the other subjects, the difference was not uniform. For these subjects, [j] had a particularly large difference between geminate and singleton transition durations. In a few instances mean transition durations were a bit larger for singletons. This was the case for f1's [d] and f2's [l,n,z] (see Figure 2.1lb).

The Length x Rate interaction was only significant for f1 and f2 ($p<.0001$), so for these speakers geminate transition durations were more affected by rate than singleton transition durations were; but for m1 and m2, the transition durations of geminates and singletons fared about the same, with all transitions shortening at faster rates.

Only for f1 was the Length x Rate x Segment interaction at all significant ($F(8,330)=2.196$, $p=.0274$). Post-hoc 2-way ANOVAs by segment showed that for this speaker the stronger effect of speaking rate on geminate (vs. singleton) transition durations was evident on all segments except for [n].

In summary, the most salient observation regarding formant transition durations is that there are strong differences among the segment types studied. Predictably, the glide [j] had quite long transitions, but somewhat surprisingly, so did the fricative [z]; much longer than any of [l,n,d]. For the most part, the transition durations behaved systematically: they were shorter at faster rates and longer for geminates than singletons.

2.3.6 Offset F1

The F1 value in the last glottal pulse of the vowel preceding the consonant constriction is strongly dependent on Segment, but not consistently so on Rate or Length.

In a 3-way ANOVA of Offset F1 for the independent variables Rate, Length and Segment the main effect of Segment was highly significant for all four subjects:

f1: F(6,461)=578, p<.0001
f2: F(6,462)=678, p<.0001
m1: F(6,461)=490, p<.0001
m2: F(6,462)=316, p<.0001

Table 2.16 ANOVA Results for Offset F1 (Hz): Main effect of Segment

Figures 2.12a and 2.12b show the Offset F1 values measured for these segments. Fisher's PLSD post-hoc tests yielded the following ranking among the segment types. Speaking rates and length categories are pooled and the asterisk indicates a significant difference between successive Offset F1 values.

f1	[h *> ʔ *> n,l *> d *> z,j]
f2	[h,ʔ *> n,l,d *> d *> z *>j]
m1	[h *> ʔ *> n *> l *> d *> j *>z]
m2	[ʔ,h *> l *> n *> d *> z,j]

Table 2.17 Offset F1 rankings

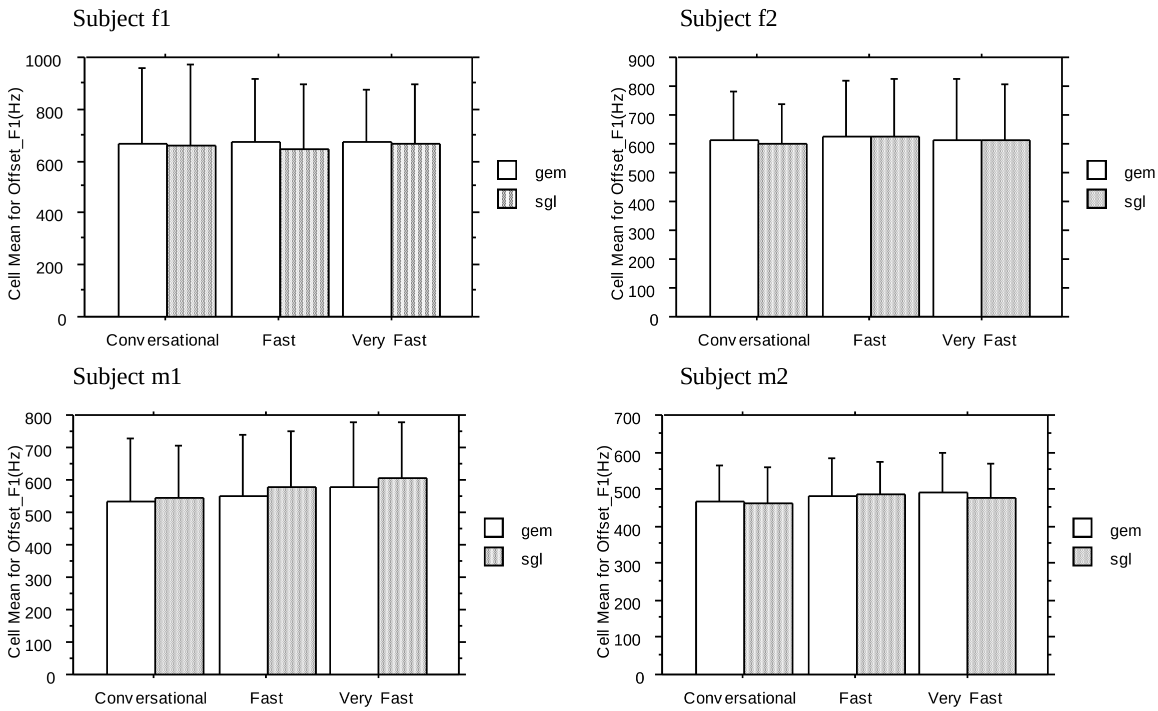

Figure 2.12a Preceding Vowel Offset F1(Hz) by Rate

Error bars indicate standard deviation. Larger number indicates less formant movement.

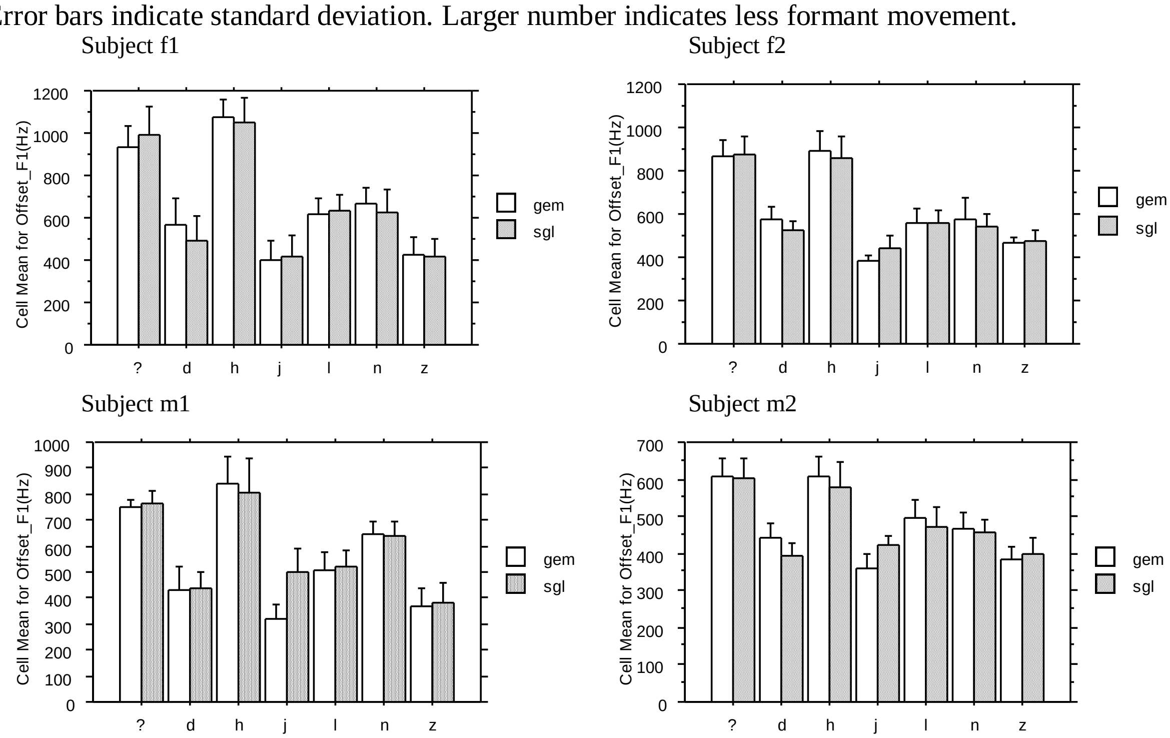

Figure 2.12b Preceding Vowel Offset F1(Hz) by Segment
Error bars indicate standard deviation. Larger number indicates less formant movement.

The glottal segments [ʔ,h] maintain the high F1of the preceding [æ] as expected. The lowest Offset F1 were shown by [j] and [z], indicating that a closed oral configuration is in effect before these consonants' constriction begins. The alveolars [n,l,d] had intermediate Offset F1 values. There is some lowering of F1 before the constriction, but not to the low level of the consonant's F1. This means that for these three consonants, an abrupt change in F1 occurs at the vowel-consonant boundary.

The main effect of Length was significant only for one of the informants, m1.

f1: $F(1,461)=1.110$, $p=.2926$
f2: $F(1,462)=.562$, $p=.4539$
m1: $F(1,461)=17.087$, $p<.0001$
m2: $F(1,462)=2.409$, $p=.1213$

Table 2.18 ANOVA Results for Offset F1 (Hz): Main effect of Length

So only for m1 were Offset F1 values higher for geminates compared to singletons. For the other three speakers, Offset F1 was not dependent on the Length category. A post-hoc 2-way Rate x Length ANOVA of Offset F1 split by segment type showed that m1 produced geminate [j] with a higher Offset F1 than singleton [j] ($F(1,66)=159$, $p<.0001$). His other consonants showed no Length effect at all.

The main effect of Rate is significant for 3 of the 4 informants, but Fisher's PLSD post-hoc analysis showed that only for m1 does the Offset F1 increase significantly for each Rate increase. In particular, for this speaker, singleton [j] and [z] showed systematic increases in Offset F1 as the speaking rate increased, an apparent undershoot phenomenon as these segments failed to reach their targets at faster rates. A consistent pattern was not observed from the other subjects.

Overall, neither Length nor Rate clearly affected the Offset F1 for all speakers, although m1 showed some effects that seemed to relate to coarticulation of [j] and [z].

Length x Rate interactions were not significant for any of the subjects, so whatever increase in Offset F1 was attributable to faster speaking rates, applied to both geminates and singletons.

Both Length x Segment and Rate x Segment interactions were significant for all four subjects ($p<.0001$). This means that any Length and Rate effects on Offset F1 were applicable to some of the segments but not others. Buth since the Length and Rate main effects don't show a clear pattern, these interactions are not informative. f2 and m2 had significant 3-way Rate x Length x Segment interactions ($p=.0023$ and $p=.0127$, respectively), but f1 and m1 did not. I do not attach importance to this fact for the reasons just stated.

2.4 CONCLUSIONS

Remembering that my objective is to identify characteristics of various segment types that could affect the perception of the vowel/consonant boundary, the results of this experiment show that segments are quite different from each other by each of the measures used, but the ranking of consonants in terms of the sharpness of the vowel/consonant boundary also depends on which measure is used. By the measure of Amplitude Difference, the sonorants [j,n,l] show the least change from the preceding vowel, and so should be harder for listeners to precisely segment for the purpose of distinguishing geminates from singletons. By the measure of Transition, [j,z] show the longest interval of oral adjustment preceding consonant constriction, providing a legthy indeterminate period which might be assigned by the listener to either vowel or consonant (Myers and Hansen 2005). By the measure of Offset F1, glottal segments [ʔ,h] had the same high F1 as the preceding vowel, and so there is no formant movement to cue a consonant onset. Likewise, [j,z] had the same low F1 as the ensuing constriction, and so there is no abrupt change at the vowel/consonant boundary.

Each of the segments can be said to be vowel-like in some dimension with the exception of [d]. Its amplitude is very low with respect to vowels, as evidenced by its high Amplitude Difference. It has a relatively short transition, so the indeterminate interval is short. And the Offset F1 of the preceding vowel is relatively high, leading to an abrupt interruption of the formant trajectory at the closure.

Segments [n,l] would lack salient boundary cues according to one of the measures employed: Amplitude Difference. This parameter is what Kawahara (2007) uses to account for the supposed markedness of geminate approximants, as supported by his perceptual experiments. But by the other parameters I studied, these sonorants do show a sharp boundary at the vowel/consonant interface: the formant transition is short and declines abruptly at closure.

Geminate fricatives have been claimed by researchers to be marked for reasons ranging from supposed excessive articulatory effort (Kirchner 2000) to perceptual accounts focusing on the inherent length of singleton fricatives (Blevins 2004a, Kawahara 2006). Indeed, in my data, the fricatives [h,z] were among the segments with the longest constriction durations (except for m2's [z], which was shorter than his [n]), and it is significant that for all speakers, [z] had the smallest geminate/singleton constriction ratio. Amplitude Difference has not been claimed to contribute to the lack of perceptibility of a phonological length contrast in fricatives, because fricatives, although sometimes quite loud, are considerd to be low in sonority.

In my data, [z] stands out in both Transition and Offset dB measues as lacking a sharp perceptual landmark at the vowel/consonant boundary. These measures relate to the rapidity with which the formant, in response to changes in the oral articulators, adjusts from the open vowel position to the closed consonant positon. Both of these measures would support the assertion that the precise duration of [z] is difficult to perceive because there is not a period of rapid formant change to serve as a landmark at the segment boundary.

If an ideal geminate is the least vowel-like, the glide [j] makes a poor geminate by every measure I studied. It is high in sonority, as evidenced by the low Amplitude Difference it exhibits; it has the longest transitions of any consonant I measured, and its Offset F1 is so low as to be indistinguishable from the F1 within the consonant's steady state constriction. It is not surprising that glides are among the least utilized segments in geminate/singleton contrasts in the world's languages.

The glottals [ʔ,h] present a bit of a paradox. While accounted as low in sonority by virtue of their classification as obstruents, they have the open vocal tract configuration of the vowel. In my measures of Amplitude Difference, their ranking was variable, but in no case did a glottal have as low an Amplitude Difference as the sonorants. On the other hand, they were not logged in the observation of formant transition duration, because there *was* no transition between the preceding vowel and the consonant. Likewise, the fact that the Offset F1 of the preceding vowel was highest in glottals does not indicate that there is a drastic fall in F1 at the consonant/vowel boundary; on the contrary, it indicates that the formant in the consonant is identical to that in the vowel. Therefore, formant movement can provide no cue for the consonant boundary between vowels and glottals. The fact that they are about as infrequent in the world's geminate inventories as glides is an indication that if the perceptibility of duration is a factor in the conformation of geminate inventories, a rapid, abrupt change in formants F1 and F2 can be required to provide a clear landmark for listeners to determine the phonological length category of a segment.

The observation that geminates may have lower amplitude than singletons has implications in the realm of dialectal variation and historical change. Where one dialect preserves an ancient geminate alternation, another may exhibit a manner alternation in which the singleton is lenited.

An example of how an amplitude difference can lead to neutralization is provided from the development of Latin singleton and geminate voiceless stops in Romance. Unlike Italian stops, Spanish stops underwent a process of lenition in the intervocalic position (Lewis 1998):

Latin	**Italian**	**Spanish**
SA**P**ERE	sa**p**ere	sa**b**er
RO**T**A	ruo**t**a	rue**d**a
AMI**C**A	ami**c**a	ami**g**a
CU**PP**A	co**pp**a	co**p**a
GU**TT**A	go**tt**a	go**t**a
VA**CC**A	va**cc**a	va**c**a

Table 2.19 Voiceless Stops in Romance (from Cravens 2002)

It is seen that while Italian preserves the singleton/geminate alternations of Latin: p/pp, t/tt, c/cc; Spanish has a manner alternation in these words: b/p, d/t, g/c. I posit that the geminate, having lower amplitude, is easily classified as a voiceless stop, while the louder singleton can be interpreted as voiced.

Experiment I has confirmed that a faster speaking rate acts to strongly decrease the duration of singletons more than it does the duration of geminates. I found this to be true of all segments. Geminates are on the whole of lower amplitude and have longer transitions than singletons. As hypothesized and supporting Kawahara (2007), I found that obstruents have a larger amplitude difference with the preceding vowel as compared to sonorants. As expected, the fricative [z] and the glide[j] had longer transitions than [d], though [n] and [l] did not, contrary to my prediction. The evidence from F1 Offset also only partially confirmed my hypothesis that more sonorant segments would have a larger decline in the first formant as it nears the consonant closure. The fricative and the glide did have a low F1 Offset, but the [n] and [l] did not. If [n] and [l] indeed are disfavored as geminates, it is due to other factors than the F1 and F2 transitions. Certainly their sonority could be an important factor, as could the possibility of coarticulation caused by a gradual velar gesture in either case.

3.0 Experiment II: The Effect of Manner on the Perceptibility of Phonological Length

3.1 INTRODUCTION

Various classes of consonants have been claimed to be marked with respect to their ability to contrast by length. Geminates belonging to the classes of glottals, represented by [h] and [ʔ] (Podesva 2002); sonorants, represented by [j], [l] and [n] (Podesva 2002, Kawahara 2007); and fricatives, represented by [s] (Podesva 2002, Kawahara 2007, Blevins 2004); have been claimed to be marked or disfavored by the authors. By contrast geminate stops are said to be well-formed in languages possessed of geminates, though there is agreement that voiced stops are marked compared to voiceless stops for articulatory reasons.[11] Experiment I found that on the basis of Amplitude Difference, sonorants are be more vowel-like than obstruents, so my hypothesis is that it is harder for listeners to distinguish between geminate and singleton sonorants than between geminate and singleton obstruents. The Amplitude Difference parameter also showed that sonorants are more vowel-like than glottals, but did not indicate that glottals are more vowel-like than obstruents, so any possible restrictions on glottal geminates are not to be attributed to differences in amplitude. The Amplitude Difference of [h,ʔ] showed a great deal of variability and depended on the individual speaker. I was unable to predict the status of glottals based on Amplitude Difference.

By another measure of the vowel-consonant interface, the Offset F1 (Hz), the fricative [z] and the glide [j] were the segments with the lowest F1 in the last pulse of the

[11] Articulatory factors favoring voiceless geminate obstruents over voiced geminate obstruents were first observed by Ohala and Riordan (1979), to wit: it is difficult to prolong voicing in obstruents due to the difficulty of maintaining a trans-glottal pressure differential when air is prevented from exiting the oral cavity. Hayes and Steriade (2006) provide a survey of the typology of geminate stop restrictions. In general, voiceless geminates are favored over voiced geminates, front geminates (where there is a larger oral cavity to fill) are favored over back geminates.

vowel preceding the test consonant. My conclusion was that since there was not a large change in the formant at the vowel-consonant interface, these segments would be difficult for listeners to categorize by length. The other oral consonants [d,l,n] would be easier to distinguish if Offset F1 were the only factor involved. The glottals, however, should be even more difficult to categorize than [z,j] because they showed no formant movement whatsoever at the vowel-consonant interface. My hypothesis that glottals [h,ʔ] are harder to categorize by Length than obstruents is based on the observation that F1 and F2 did not show movement at the vowel-consonant boundary.

The other parameter studied, Transition (ms), represented the duration of the F1 and F2 transitions in the vowel preceding the test consonant. The results for this parameter led to predictions consistent with those obtained for Offset F1, since the Transition in the oral consonants decreased systematically with increased Offset F1. That is, the longer the transition was, the lower the fundamental frequency of the last pulse of the vowel preceding the test consonant. Specifically, [z,j] had longer transitions than [d,l,n], indicating that it would be harder to categorize [z,j] by length. Transition was not measured for the glottal consonants, since there is no formant movement preceding [h,ʔ]. However, since the Transition parameter was selected as representing an interval of indeterminacy between the formant structure in the vowel and the formant structure in the consonant, I rank the glottals as least distinguishable because the period of indeterminacy involves the entire vowel.

Because of the differing predictions suggested by the results of the production study, it was not possible to make general predictions about whether [z,n,l] would be easy or hard to categorize. But [d] can be predicted to be easy to categorize according to both amplitude (low) and formant movement (short, abrupt) criteria. Conversely, [j] should be hard to categorize by both criteria: high amplitude and long formant transitions. [h,ʔ] can be predicted to be difficult to categorize due to the absence of any formant movement at the vowel-consonant interface, though the effect of amplitude is not clear.

3.2 APPROACH

Both perceptions experiments (Experiment II and Experiment III) require some quantifiable means of assessing the confusability of the length category for each of the different consonants. I used two different methods to accomplish this: the psychometric function and response times. Each of these techniques is further described below.

The psychometric function is a tool that has been used in Psychology and other behavioral sciences to describe the perceptual threshold between two possible stimuli categories that can be made to vary continually from one to the other along some quantifiable dimension, whether it is from category A to category B, or from a condition in which category B is absent to a condition in which B is fully present. The form of the identification curve is typically an s-shaped curve, where the percentage of "B" responses ranges from zero to 100 percent as the stimulus varies from A to B. The point at which the curve is steepest represents the stimulus level that is at the boundary between the two responses. If the response curve is modeled by fitting a mathematical curve-fitting model to the data points, differences in the regression-model abcissa reflect the lateral shift in the psychometric function, or bias. A difference in the slope of the phsychometric function indicates a difference in sensitivity. According to Wichman and Hill (2001), the slope of the function at the threshold is "a measure of the change in performance with changing stimulus intensity." For example, in describing an experiment measuring subjects' ability to distinguish between letters as the size decreased, Carkeet, Lee, Kerr and Keung (2001) state: "The level of optical defocus influenced the slope of the psychometric function."

In the context of an experiment in which listeners hear a continuum of stimuli ranging from singleton to geminate, the "stimulus intensity" is the closure duration. As one of my subjects told me, some of the tokens “sound like they have a heavy *tashdid*.” A steep slope thus indicates a large change in judgment in response to a small increase in duration: you can more easily hear the difference between the tokens. A flatter slope indicates that the judgment doesn’t change much as the duration is increased: you can’t hear the difference unless there is a large increase in duration.

I have made use of the cumulative gaussian curve fitting model as a replicable method of analysis for determining the slope or breadth of the geminate/singleton threshold. I would hope that other researchers may find this parameter to be informative.

Kawahara (2007) made use of cumulative identification curves to address the question of how easy it is for listeners to judge the phonological length category of consonants in a study of Arabic geminates. Although he did not mathematically model the psychometric function, he compared the slopes of the cumulative identification curves of continua of consonants whose constriction durations ranged from short to long. His “slope coefficient” was a measure of the steepness of the slope “between the last point above 80 and the last point above 10,” referring to the percent of singleton response to the stimuli. He predicted that smaller slope coefficients (i.e. a flatter slope) would be observed for sonorant stimuli than for obstruent stimuli, since there would be uncertain responses over a longer interval in the constriction duration continuum. He found that in terms of the slope of the identification curves, sonorants were harder to distinguish than obstruents, and that for particular segments, the order from hardest to easiest to distinguish was glides>liquids>nasals>obstruents.

The response time Response (ms) is another measure that can be used to compare the discriminability of phonological length for different consonants. Longer response times indicate that the subjects are struggling to decide on a response and indicate that the discrimination task is more difficult. A shorter response time indicates that the subjects can easily determine whether a given token is geminate or singleton. Kawahara (2007) also measured response time as a measure of discriminability between geminates and singletons. On average he found that sonorants had longer average reaction times than obstruents, glides>liquids>nasals, but the reaction times for nasals were not significantly different than those for obstruents. He concluded that the two measures of perceptual confusability gave consistent results.

I performed experiments using a language (Persian) that, like Arabic, allows all manners of consonants to be geminates. Differences between geminate and singleton glottals and glides are therefore linguistically significant. Other researchers pursuing this line of inquiry might investigate a language such as Hungarian for the same reason.

It was important to include glottals [ʔ,h] in my materials because there is striking evidence that despite being classified as "stops" or "fricatives," they readily participate in phonological processes such as deletion, compensatory lengthening (Kavitskaya 2002) and geminate avoidance (Podesva 2000) , which I suspect can be attributed to the lack of clear boundary cues at the glottal/vowel interface.

3.3 METHODS

3.3.1 Subjects

Five female and four male adult Persian speakers participated in the experiment. One of the females and one of the males had also recorded for the production study (Experiment I of this dissertation). The informants are all residents of the United States, and were selected from family, friends and University colleagues. They were not compensated for their effort, but were willing, often enthusiastic participants. Their ages range from 30 to 60 years. All are literate in standard literary Persian as used in Tehran, having attended at least high school in Iran. Though they reside in the United States and speak English regularly, they continue to use Persian in their family and social environments.

3.3.2 Materials

Experiment II is a forced-choice perceptual discrimination task in which listeners hear continua of words containing versions of the test consonants ranging from short to long as measured by the consonant constriction duration. The subjects listen to repeated instances of the words while recording their judgments on a computer. The percentage of responses at which the consonant is judged to be a geminate is plotted against the consonant constriction duration to create a series of identification curves. Similarly, the average response times recorded while subjects decided on their response were plotted against the consonant constriction duration. The response data are then analyzed to interpret whether specific classes of consonants such as obstruents, sonorants or glottals have significantly different identification curve slopes or response time averages.

I have used [d] as the representative stop since I wished to control for voicing and place by using voiced coronals for the oral consonants. Glottal consonants do not contrast for voicing in Persian and do not have an oral place specification. The seven consonants I have selected for these experiments may be grouped into three categories; obstruents [d],[z], sonorants [n],[l],[j] and glottals [h],[ʔ].

The recordings used to prepare the stimuli were produced by female informant f1 from Experiment 1. Prior to Experiment II she recorded nonsense words of the form [qæC(ː)ɑb] at a conversational speaking rate within the carrier sentence [minu ______ rɑ did]: "Minu saw the ______". A single representative token of each of the excised nonsense words [qædːab], [qæzːɑb], [qænːɑb], [qælːɑb], [qæjːɑb], [qæhːɑb], [qæʔːɑb] was used to provide precursors for all of the stimuli in this experiment. The sentence frames from which the tokens for this experiment were excised had durations ranging from 1.04 to 1.10 seconds, similar to the 0.99 second average value of Frame obtained for "Conversational" rate geminates in Experiment I. In Experiment I Frame is defined as the duration of the four syllables of the carrier sentence excluding the test word itself.

A continuum series of eight tokens was produced from each precursor word, with constriction durations ranging from short to long in approximately 10 ms increments. In order to avoid possible confounding effects due to differing segment durations surrounding the test consonants in the seven precursors, I controlled for vowel duration by setting the durations of the preceding vowels [æ] to approximately 100 ms and the following vowels [ɑ] to approximately 170 ms for all tokens. The duration of the final consonant of the test word [b], from closure to the end of acoustic noise, was set to 60 ms. The vowel durations include the formant transitions: vowels were lengthened or shortened by inserting or deleting pulses within the vowel steady state only so that in each series the transition durations were conserved as recorded. The ranges of durations of the resulting tokens are shown in Table 3.1 below:

d-dː	50.5 ms-119.6 ms
z-zː	82.1 ms-151.6 ms
n-nː	80.8 ms-149.9 ms
l-lː	71.4 ms-140.1 ms
j-jː	27.8 ms-101.5 ms
h-hː	79.7 ms-156.0 ms
ʔ-ʔː	71.0 ms-148.0 ms

Table 3.1 Consonant Constriction Series Continua Ranges

The ranges were selected to fall within the natural range of singleton and geminate consonants spoken at the conversational speaking rate. The eight 10 ms increments in the continua encompass approximately a 70 ms range in each case. Pilot experiments were performed to ensure that each continuum range straddled the threshold breadth for each consonant. Since the threshold was not the same for each consonant, the series beginning and endpoints are different.

I used Multispeech Analysis-Synthesis Laboratory to perform the manipulations. The excised tokens were analyzed at a sampling rate of 11025 Hz using filter order 36 and a preemphasis of 0.5. Cell intervals were set to the voiced-period marks where voicing was present, and to 15 ms elsewhere. I performed the pulse insertions and deletions in the resulting analysis file. The analysis file was then used to perform a synthesis, which generated the natural-sounding tokens for use in the perception experiment.

3.3.3 Response Observations

The experiment was carried out using the Superlab program installed on an IBM laptop computer in a quiet place convenient to the subject, such as a kitchen table or office. The perception materials were presented to the participants over earphones, while they looked at two alternative responses (the singleton and geminate forms) represented in Persian orthography on the screen. Subjects recorded their judgments as to whether they heard a geminate or a singleton consonant by clicking the appropriate key on an external response pad. One of the keys was marked with a simple dash, and the other with the Persian geminate diacritic (*tashdid* symbol). Subjects were instructed to respond as quickly as possible, and to guess if it was not clear which of the two categories the token belonged to. The response key selection and response time from the end of the token were logged in an ascii file on the computer.

The numerical probability of a geminate judgment for each of the tokens was determined by presenting each token multiple times. For each test consonant, the eight tokens in the continuum were presented in a single block in random order. Subjects heard the block twelve times, each time in a different random order, for a total of twelve responses per token. The number of possible judgments obtained per subject for this experiment was: 8 increments * 7 consonants * 12 repetitions = 672 judgments.

In a small number of cases (<1%), the subject failed to make a judgment. I was still able to determine the probability of a geminate response for a given token by calculating the proportion of geminate answers to the number of actual responses. In many more cases, I obtained a judgment, but failed to record a reliable reaction time. There were two reasons for failing to record a valid reaction time. First, the subject sometimes did not press the key firmly enough and had to press again, but the second response occurred after the two second window provided. Second, no response value was logged if the key was pressed before the token recording finished. Inspection of the histograms of the reaction times showed that the number of responses declines steeply above 800 ms and approaches zero around 1000 ms. Then a clearly separate population of responses could be observed above 1000 ms. where subjects repeated a response that didn't register on the first try. This second try response does not actually indicate how long it took for the subject to make a decision. I therefore discarded all reaction times above 1000 ms for all subjects. This resulted in the loss of 2-10% of the response time data for a given subject. However, since I was interested in the average response time for a token, and each token was heard twelve times, I was able to obtain an average response times for each token and speaker, even if some of these were based on fewer than twelve data points.

3.4 ANALYSIS

As I mentioned previously, besides the determination of the threshold duration, my analysis requires a measure of the breadth of the perceptual threshold between geminates and singletons. Several breadth measures can be devised, such as the inter-quartile distance or Kawahara's "slope coefficient", the inter-quartile range, or others. I sought a method that did not require any subjective decisions as to where to measure the breadth, and would be automatically repeatable for dozens of analyses. One mathematically simple means of reducing the breadth of a cumulative curve to a single parameter is by means of the Cumulative Gaussian model. This model is perhaps the simplest of several representations of the psychometric function that have been used to model responses to forced-choice two-category perceptual stimuli (Wichmann and Hill 2001a, 2001b).

In the Cumulative Gaussian model, two parameters are defined: the mean and the standard deviation. The mean corresponds to the 50% threshold, which I identify with the consonant duration at the geminate-singleton threshold. I call this variable Threshold, expressed in ms. The standard deviation is a measure of the breadth of the threshold and in the Gaussian Model is inversely proportional to the slope of the identification curve at the 50% identified level. That is, the flatter the curve, the smaller the slope value and the larger the standard deviation. I refer to the modeled standard deviation in this context as the variable Breadth, also measured in ms.

For each subject and consonant, I prepared an input matrix containing the proportion of geminate responses for the corresponding consonant constriction (C) value. An example input matrix is shown in Table 3.2.

cf1d	cf1z	cf1n	cf1l	cf1j	cf1h	cf17
0.000	0.000	0.000	0.000	0.000	0.000	0.000
0.000	0.083	0.000	0.000	0.083	0.000	0.083
0.167	0.250	0.500	0.000	0.167	0.143	0.167
0.583	0.750	0.917	0.417	0.250	0.571	0.500
0.917	1.000	0.917	0.500	0.500	0.857	0.500
1.000	0.917	1.000	0.750	0.917	1.000	0.708
1.000	1.000	1.000	0.958	0.917	1.000	0.917
1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3.2 Identification Curve Input Matrix Example
Proportion identified as geminates by Subject f1-Experiment II.

I used R statistical software to generate Cumulative Gaussian models of the identification curve data using least squares curve fitting.

The modeled Threshold and Breadth values obtained from the R analysis of each subject's responses formed the input data for the statistical analysis discussed below in the results section. This analysis provides a gauge of the significance of the differences observed among the consonants and natural classes studied.

3.5 RESULTS

3.5.1 Descriptives

Before addressing the statistical results of the comparisons of the various Gaussian parameters and what they say about the perceptibility of length for different classes of consonants, I will present descriptive data comparing characteristics of the geminate-singleton boundaries of the seven consonants studied.

A table listing the modeled Gaussian parameters and average response times for the seven test segments is shown in Table 3.3. The model parameters are based on pooled identification curves for all subjects, not the calculated averages of the individual subjects' gaussian models.

Segment	Natural Class	Threshold	Breadth	Reaction
d	Obstruent	83.26	9.87	244.0
z	Obstruent	112.31	11.24	257.0
n	Sonorant	110.30	14.22	268.1
l	Sonorant	107.42	16.65	266.9
j	Sonorant	61.57	12.49	274.2
h	Glottal	112.97	17.47	256.1
ʔ	Glottal	107.05	23.3	305.0

Table 3.3 Pooled Gaussian Model Parameters

Mean (Threshold), Standard Deviation (Breadth), and Average Response Times (Reaction) for Seven Test Consonants. All values are expressed in ms.

The values of Threshold shown in Table 3.3 are indicative of the inherent durations of the respective segments. Thus, the smallest value for Threshold is for [j] which is the consonant with the shortest average value of Constriction (See Table 2.9 for

Constriction duration rankings). Conversely, the fricatives [z] and [h] have the largest value of Threshold, and for three of the four speakers in Experiment I these were the two consonants with the largest value of Constriction. What we can conclude here is that listeners hear the threshold in the vicinity of where the sounds are actually produced. For example, if a [d] has a duration of 100ms, it is on the high side of its range, so it is heard as a geminate (Threshold = 83.26), while a [z] with a duration of 100ms is less than its Threshold of 112.31 and so is heard as a singleton. The identification of length is therefore not merely a matter of achieving a particular duration, but must take the segment identity and inherent duration into account.

The Breadth values in Table 3.3 reflect a pattern that is consistent with the hypothesis that obstruents [d,z] have the steepest identification curves (i.e. smallest Breadth values), and thus the sharpest perceptual boundary between geminates and singletons. It is interesting that the glottals [h, ʔ] have the broadest identification curves (largest Breadth values) indicating a blurry perceptual boundary, even though they did not have the small values of Amplitude Difference (Table 2.13) that would account for their confusability with vowels. Here is evidence that another dimension of consonant-vowel boundary blurring may be at work.

I had also predicted from the observation that [z] has a larger formant transition duration than [d], that [z] would have a larger Breadth value than [d], and the observations bear this out. However, [j] was predicted to have a larger Breadth value than the other sonorants, and this was not the case. I do not believe that this fact discounts the possibility that the duration of the transition can affect the breadth of the perceptual threshold, because since the observed constrictions of both singleton and geminate [j] are of short duration, there is a very limited range over which the Breadth variable can operate. Seen in this way, the observation that [d] has a smaller threshold breadth than [z] may also be attributed to a similar scaling effect, since [d] is inherently of shorter duration than [z]. I am therefore unable to state whether my hypothesis was borne out on the basis of the Threshold data. Experiment III of this dissertation deals with the issue of whether a change in the formant transition duration affects the perceptibility of the length contrast.

The Reaction values in Table 3.3 are a slightly inconsistent with the picture painted by the Breadth values in that the value for [h] is the second smallest, seeming to indicate that the geminate is [h] relatively easily distinguished from the singleton [h]. Among the sonorants, [j] had a very slightly larger Reaction value than [n] or [l], which would be consistent with my hypothesis that a longer transition would make it harder to distinguish phonological length if the difference were statistically significant. However, an ANOVA analysis comparing the mean Reaction values for each subject did not show that the main effect of Segment was significant for this variable ($F(6,56)=1.002$, $p=.4333$), and of the Fisher's PLSD post-hoc comparisons only [d] vs. [ʔ] was significant, which is consistent with my hypothesis that in a discrimination task an obstruent would have a shorter response time than a glottal. But overall I cannot say that the evidence from the subjects' response times provided strong support for my hypothesis that certain segments are more difficult to classify than others in terms of phonological length.

Identification curves and response time curves for the seven test segments are presented below in Figures 3.1 through 3.6. The data shown in the identification curves are pooled or averaged results for all nine subjects. That is, the individual data points shown in these plots represent the percentage of "geminate" responses given by all subjects for a particular token. The plots are grouped according to three natural classes

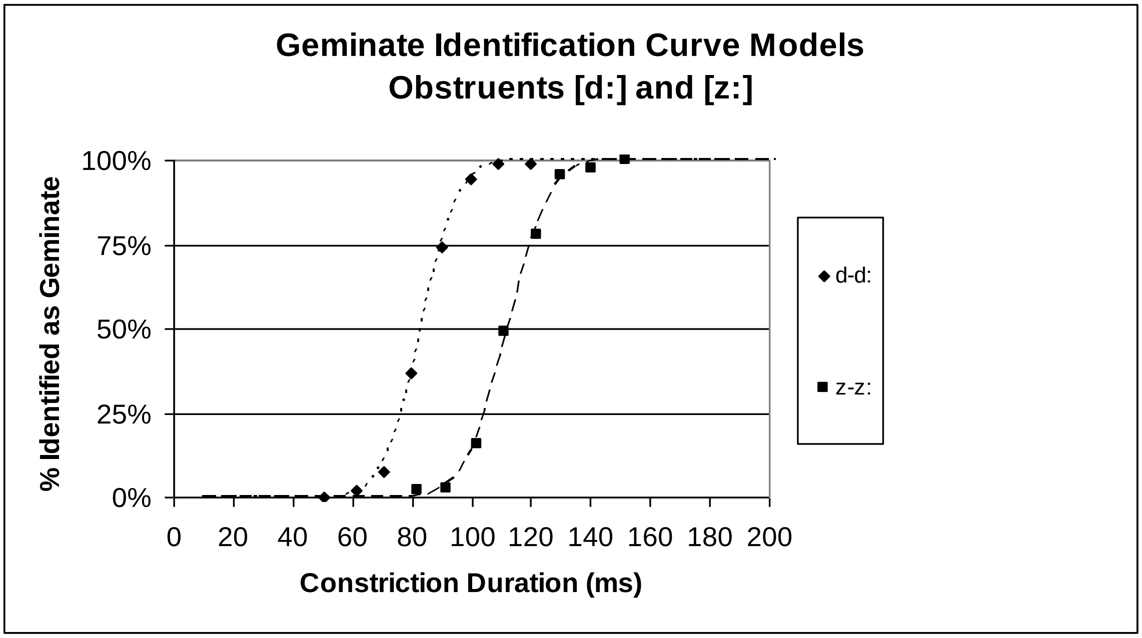

Figure 3.1 Geminate Obstruents Identification Curves.

Pooled data for 9 subjects. Curves show Cumulative Gaussian model for pooled data.

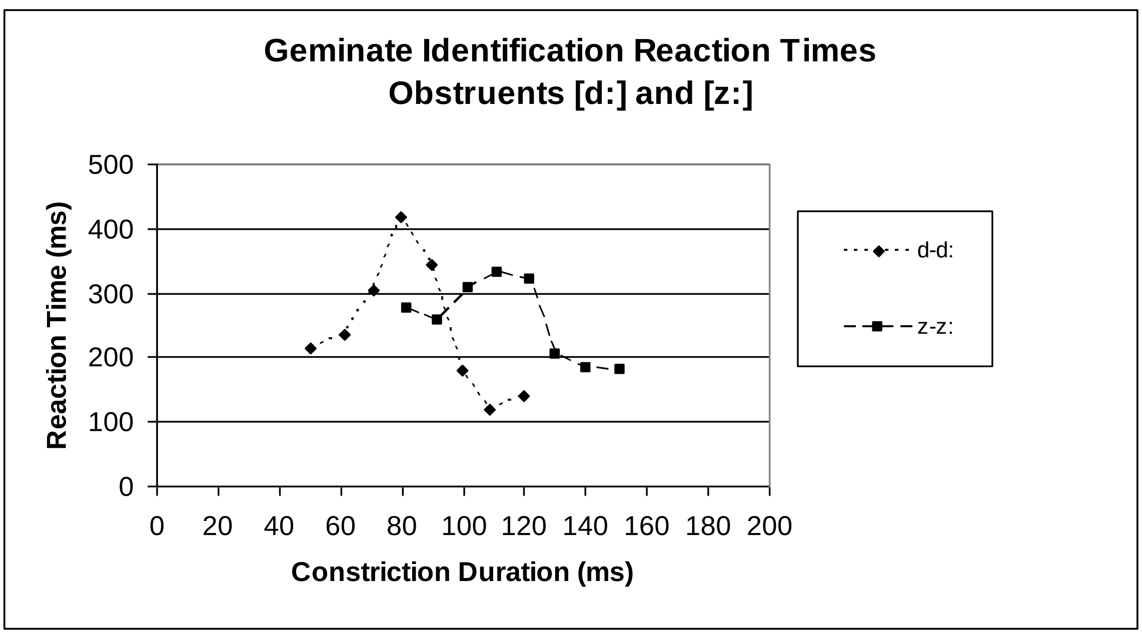

Figure 3.2 Geminate Obstruents Continua Reaction Times (ms).

Data averaged for 9 subjects. Curves show straight line interpolations between the pooled data.

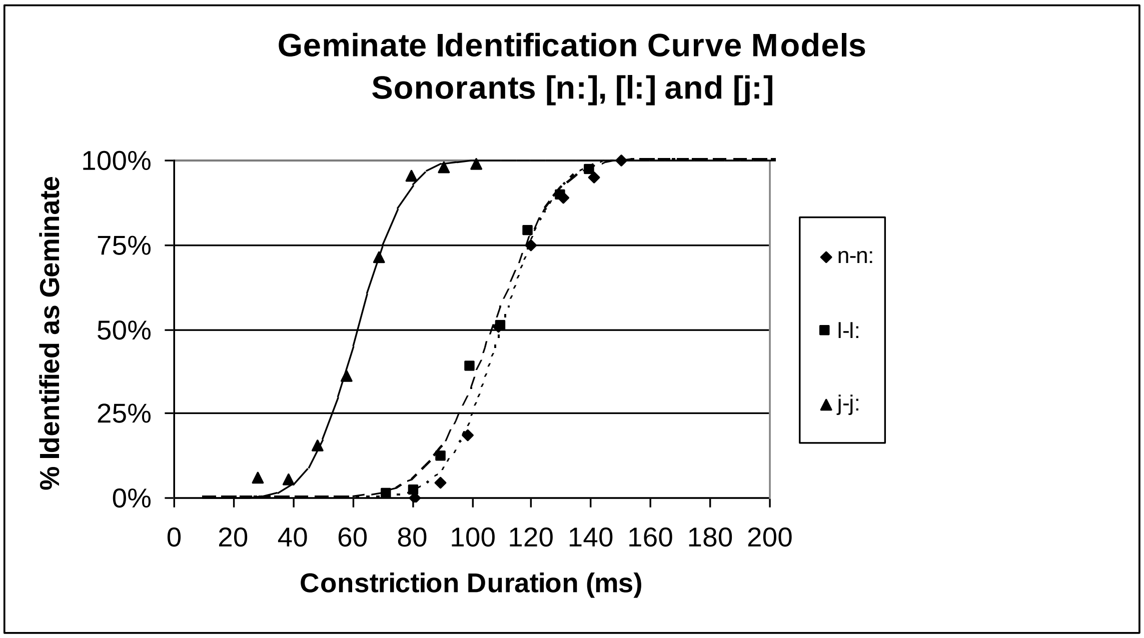

Figure 3.3 Geminate Sonorants Identification Curves.

Pooled data for 9 subjects. Curves show Cumulative Gaussian model for pooled data.

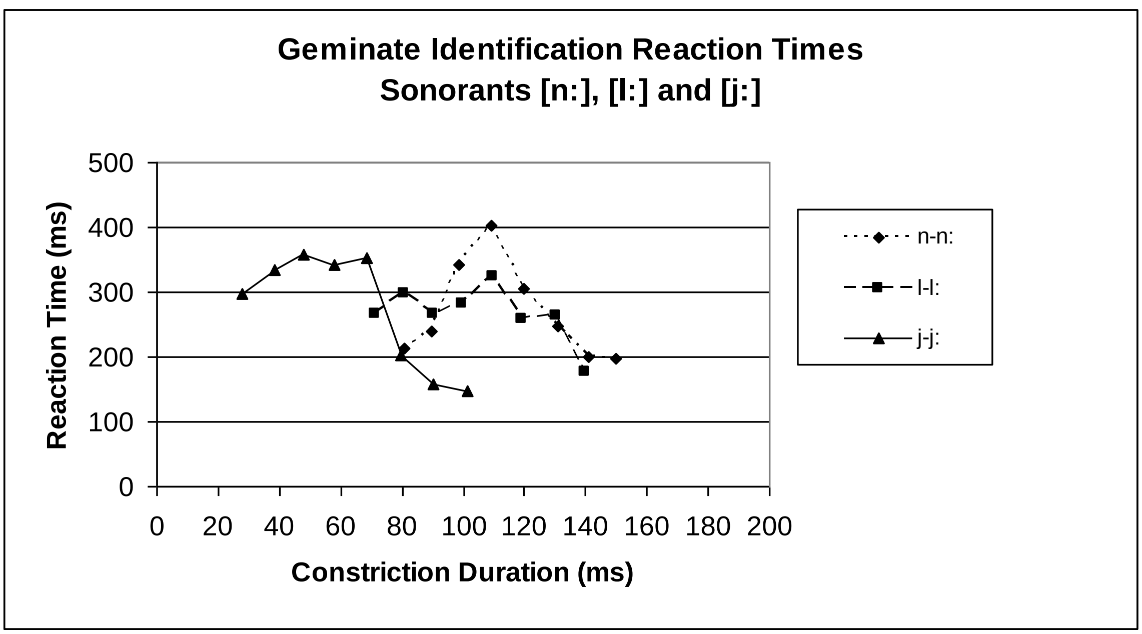

Figure 3.4 Geminate Sonorants Continua Reaction Times (ms).

Data averaged for 9 subjects. Curves show straight line interpolations between the pooled data.

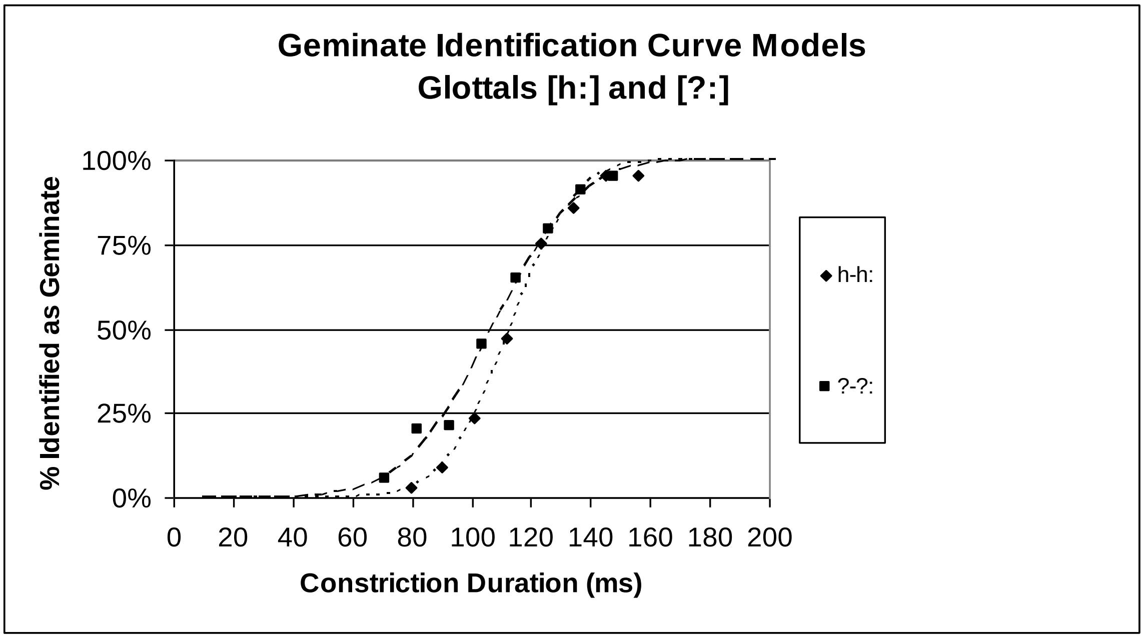

Figure 3.5 Geminate Glottals Identification Curves.

Pooled data for 9 subjects. Curves show Cumulative Gaussian model for pooled data.

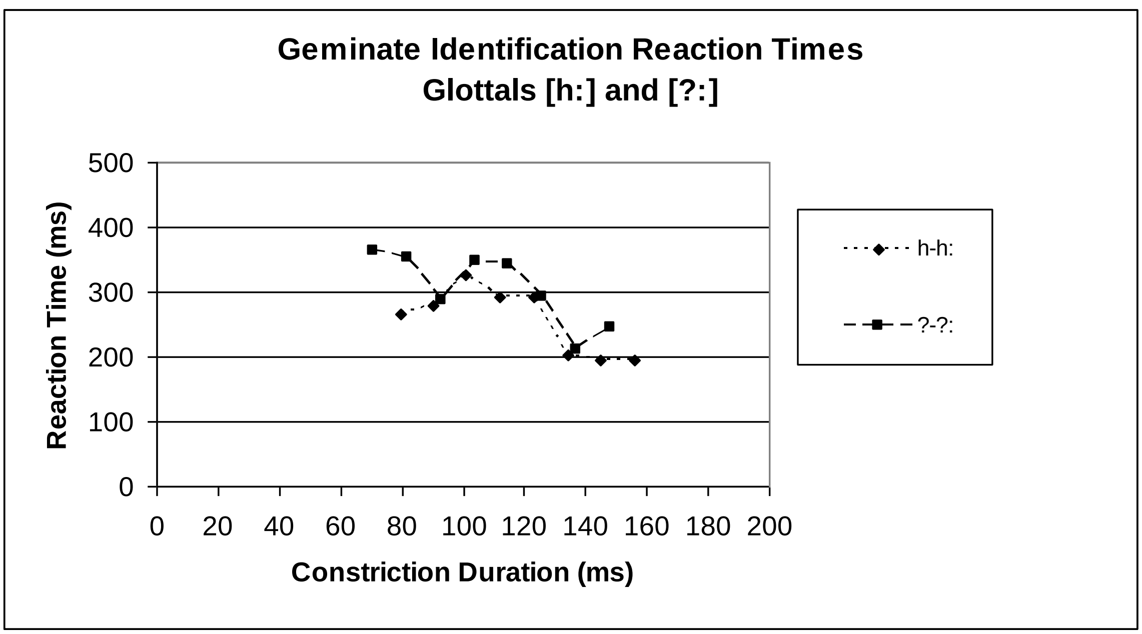

Figure 3.6 Geminate Glottals Continua Reaction Times (ms).

Data averaged for 9 subjects. Curves show straight line interpolations between the pooled data.

being evaluated in this experiment: obstruents, sonorants and glottals. For each class, a pair of figures is presented: the upper figure shows the percent geminate response curves along with the corresponding Gaussian model produced calculated using R. The lower figure shows the average reaction times observed for the judgments shown in the upper figure. The identification curve model parameters are those given in Table 3.3 above.

Since each subject heard a given token 12 times, each point is based on 12 x 9 subjects = 108 responses. For example, the fifth-longest token of the qædɑb-qædːɑb series in Figure 3.2 has a closure duration of 89.8 ms. Of the 108 times subjects heard this token, they selected "geminate" 80 times, or 74.1% of the time. Therefore, there is a filled diamond just below the 75% line at the abcissa value of 89.9 ms. The short dashed lines show the fitted model curve for the d-dː series represented by the filled diamonds. Similarly, modeled best-fit curves are shown passing through the points representing the other consonants.

In every case, there is a very good correspondence between the curves and the data points. Subjectively, the cumulative Gaussian model is a good representation of the psychometric function of forced category choice responses to ambiguous singleton/geminate stimuli.[12]

[12] Deviation from the model is most apparent at the extreme tail ends near the 0% and 100% response ordinates. Wichmann and Hill (2001) describe this kind of discrepancy as a *lapse rate* resulting from stimulus-independent errors: i.e. mistaken responses given even if the stimulus is not ambiguous. My cumlative Gaussian models did not account for the lapse rate.

The average reaction time curves in Figures 3.2, 3.4 and 3.6 above depict the averages of the mean response times produced by the subjects for a given token. Taking the same example of the fifth-longest token of the qædɑb-qædːɑb series discussed above, the nine subjects' mean response times were 271, 427, 350, 335, 155, 339, 333, 311 and 575 ms for this token, averaging 344.1 ms. As was discussed in Section 3.3 above, some response times were not recorded; in this instance, the nine subjects provided 12, 11,11,11,9,12,12,12, and 10 reliable response times respectively. So, for example, subject f2's 427 ms mean response time for this token was based on only 11 valid responses. The plotted point is the average of the nine means.

There is a distinct tendency for the peak of the average response to coincide generally with the threshold durations, defined as the steepest point in the model curves for the corresponding consonant series. In the d-dː series example, the peak average response value of 417.5 ms was observed for the fourth-longest token, having a consonant closure of 79.5 ms. This token was the closest token in the series to the modeled Threshold value of 83.26. Similar patterns are observed for the other consonants, as can be appreciated on the paired graphs shown in Figures 3.1-3.2, 3.3-3.4 and 3.5-3.6. What this indicates is that subjects do take longer to respond to the discrimination task when the stimulus is ambiguous.

3.5.2 Statistical Analysis by Segments

The Gaussian parameters listed in Table 3.3 above are not amenable to statistical analysis. Each parameter is a single value obtained by modeling the pooled response data from all of the subjects. As a means of obtaining a population of parameters that can be compared for the purpose of determining statistical significance, I created individual Gaussian models of the consonant identification curves for each of the nine subjects. As an example, Figure 3.7 shows the graphic output from R illustrating three of the Gaussian models produced for Subject f1. Each curve provides one value of Threshold and one value of Breadth for a particular consonant. The statistical analysis was performed on a population of nine parameters from the nine subject curves for each consonant.

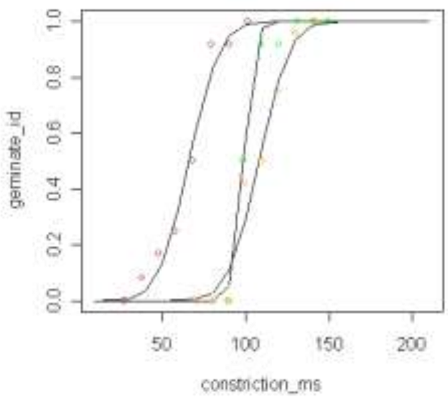

Figure 3.7 Subject f1 Gaussian Model Curves for [j], [n] and [l], from left to right.
Proportion identified as geminate against segment constriction for one subject.

Figure 3.8 presents the distribution of the nine such modeled values of Threshold in box plots.

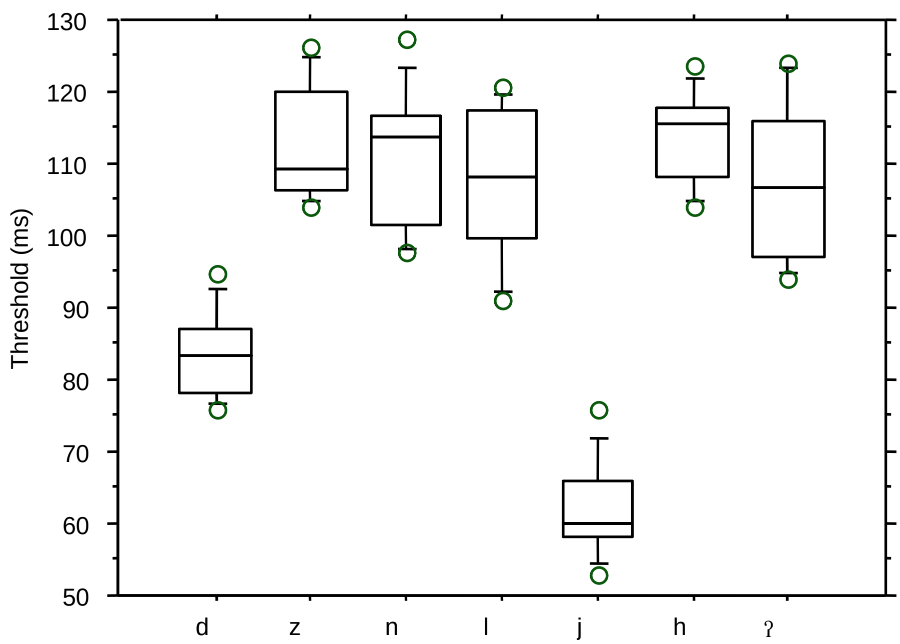

Figure 3.8 Threshold Box Plots.

Plots show Threshold values in ms for seven segment category conditions for n=9.

The one-way ANOVA analysis for the effect of Segment was significant ($F(6,56)=45.118$, $p<.0001$), indicating that the Threshold values differ among the segments studied. The means of the Threshold values obtained from the individual models are shown in Table 3.4. It is observable in particular that [j] has a very small Threshold and [d] an intermediate Threshold, consistent with the averages from the pooled data presented in Table 3.3.

Segment	Count	Mean	Std. Dev.
d	9	83.316	6.265
z	9	112.469	8.109
n	9	110.590	9.753
l	9	107.620	10.613
j	9	61.933	6.711
h	9	113.435	6.404
ʔ	9	106.930	11.48

Table 3.4 Means Table for Threshold (ms) by Segment. Nine Gaussian Models.

Post-hoc tests indicate that only the comparisons involving [j] and [d] are significant ($p<.0001$). None of the other segments have Threshold values that can be statistically differentiated from one other [j*<d*<ʔ,l,n,z,h]. This analysis indicates that the effect of Segment on Threshold was mainly the result of the fact that segments [d] and particularly [j], have much smaller values of Threshold than the other consonants.

Box plots of the nine modeled Breadth values for each of the seven segments tested are shown in Figure 3.9.

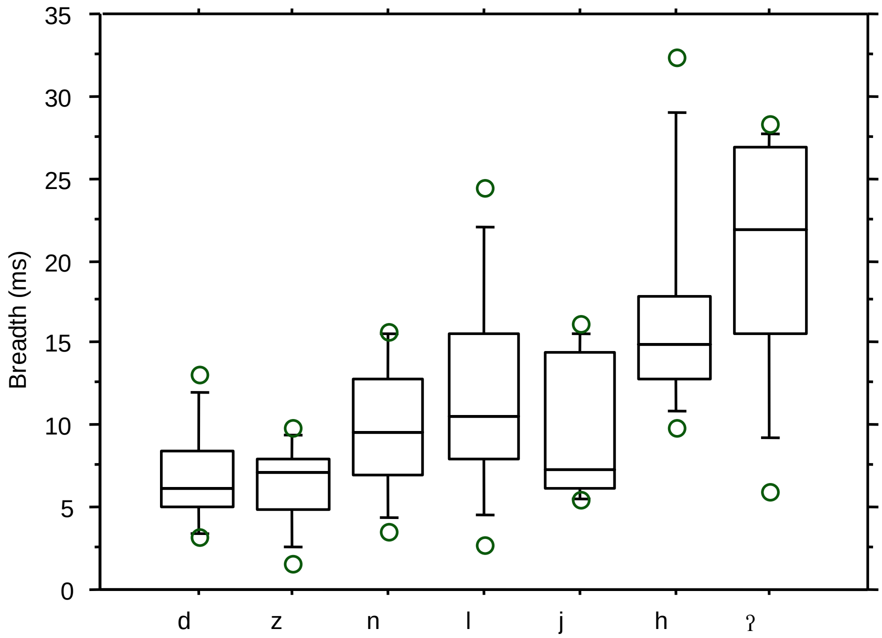

Figure 3.9 Breadth Box Plots.

Plots show Breadth values in ms for seven consonant category conditions for n=9.

Consistent with the main hypothesis of this dissertation, ANOVA analysis indicates that there is a highly significant effect of Consonant on Breadth (F(6,56)=8.52, p<.0001). Thus, some consonants have a broader—I claim less distinct—perceptual threshold than others do. Fisher's PLSD shows that each of the the glottal consonants that have significantly larger Breadth values than any of the oral consonants, but not a significant difference among themselves: [z,d,j,n,l*<ʔ,h]. However, while the obstruents [d,z] have smaller mean values of Breadth than the other consonants, the differences are not significant. As can be seen in Table 3.5, the standard deviation of the Breadth variable is large relative to the mean itself, so differences that may exist among the oral consonants cannot be demonstrated due to the relatively small sample size.

Segment	Count	Mean	Std. Dev.
d	9	6.831	3.149
z	9	6.407	2.523
n	9	9.740	4.135
l	9	11.810	6.501
j	9	9.428	4.344
h	9	16.717	7.047
ʔ	9	20.489	7.506

Table 3.5 Means Table for Breadth (ms) by Segment. Nine Gaussian Models.

So on the basis of the post-hoc analysis between individual segments I was not able to confirm my hypothesis that an obstruent has a smaller breadth value than a sonorant. However, the results do suggest that as a group the obstruents could have a smaller value of Breadth than the sonorants, and it is expected that the glottals should have a larger value of Breadth than either obstruents or sonorants. These tests were performed as described in the following section.

3.5.3 Statistical Analysis by Classes

Figure 3.10 shows box plots of the distribution of the Threshold parameter obtained from the subjects' Gaussian models grouped according to the three classes of consonants: obstruents, sonorants and glottals. The Gaussian models produced 18 values of Threshold for the obstruents [d,z], 27 values for the sonorants [n,l,j] and 18 for the glottals [h, ʔ].

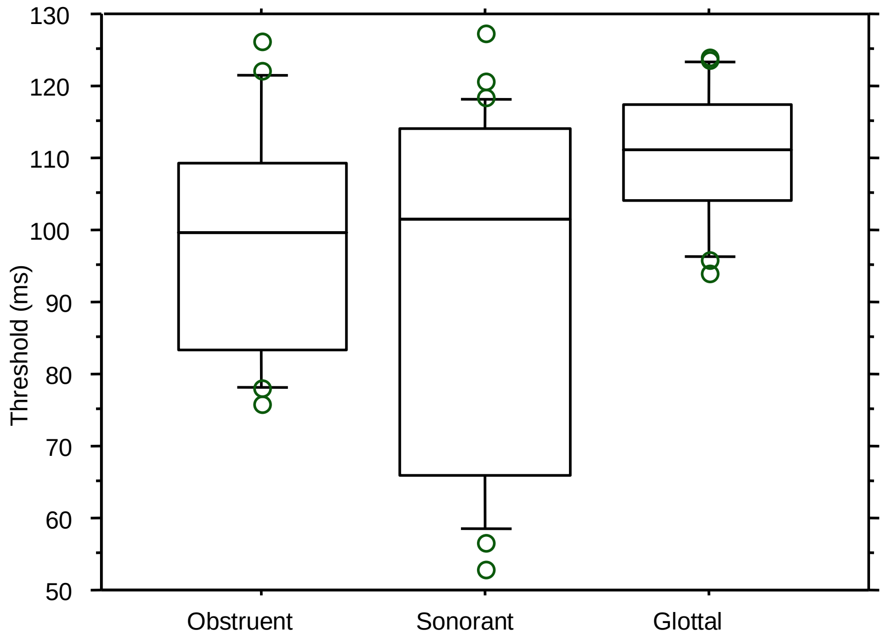

Figure 3.10 Threshold Values Grouped by Manner.
Box Plots show Threshold values in ms for the Obstruent, Sonorant and Glottal Category conditions.

The three classes were not expected to be systematically characterized by distinct Threshold values. Because each class is a heterogeneous set in that they contain inherently long-and short-duration segments, I did not have a prediction as to which group of consonants would have the highest average Threshold value. Nevertheless, in a the one-way ANOVA analysis of the parameter Threshold, the main effect of variable Manner (Obstruent, Sonorant or Glottal) was significant ($F(2,60)=4.318$, $p=.0177$) , indicating that not all of these groups can be assumed to have similar Threshold values. Table 3.6 Shows the means table for the variable Threshold according to the grouping by manner.

Class	Count	Mean	Std. Dev.
Obstruent	18	97.893	16.564
Sonorant	27	93.381	24.348
Glottal	18	110.194	9.615

Table 3.6 Means Table for Threshold (ms) by Manner Class from Gaussian Models.

Inspection of Fisher's PLSD for Threshold shows that the Glottals are significantly different from the Sonorants. This can be attributed to the fact that both components of the Glottal class have quite large thresholds while the Sonorants include [j], which has the smallest value of Threshold.

Figure 3.11 exhibits the comparison that I consider to be the crux of this dissertation. It presents box plots representing the distribution of the values of Breadth for the three manner classes.

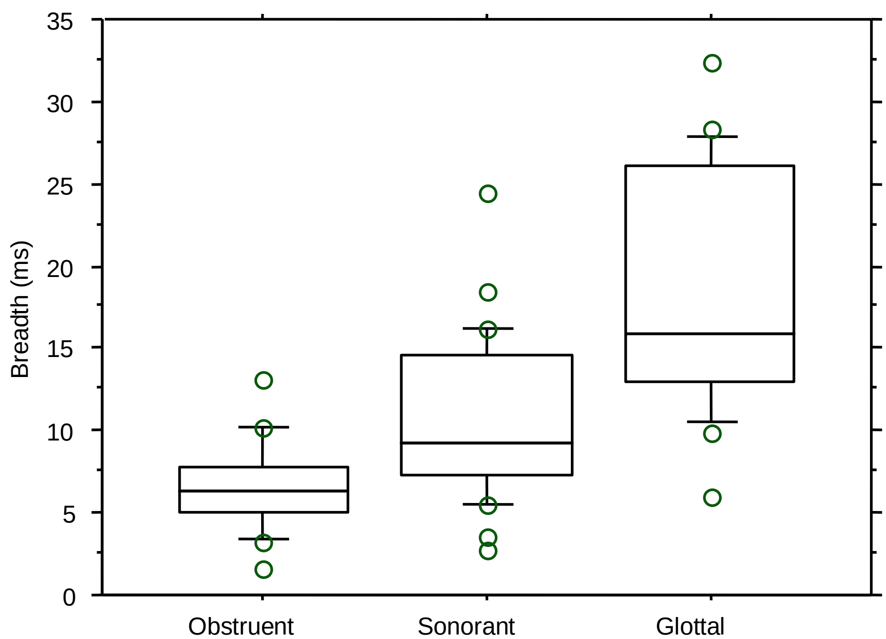

Figure 3.11 Breadth Statistics by Manner.
Box Plots show Breadth values in ms for the Obstruent, Sonorant and Glottal Category conditions.

Obstruents have smaller value of Breadth than sonorants and sonorants have a smaller value of Breadth than glottals as can also be seen in Table 3.7, the means table for Breadth.

Class	Count	Mean	Std. Dev.
Obstruent	18	6.619	2.776
Sonorant	27	10.326	5.023
Glottal	18	18.603	7.324

Table 3.7 Means Table for Breadth (ms) by Manner Class from Gaussian Models.

The one-way ANOVA analysis for of the Breadth shows that the main effect of Manner is highly significant ($F(2,60)=24.244$, $p<.0001$), indicating that Breadth is strongly influenced by whether the consonant is an obstruent, a sonorant or a glottal.

Furthermore, each of the comparisons in the Fisher's PLSD post-hoc analysis was also significant: Glottals were different from both sonorants and obstruents ($p<.0001$), and sonorants and obstruents were different from each other ($p=.0256$). This strongly confirms my hypothesis that it is harder for listeners to distinguish geminates from singletons if the consonant is a sonorant as compared to an obstruent, and it harder still if the consonant is a glottal.

On the other hand, there was not a significant effect of Manner on the response time. Figure 3.12 shows box plots of the Reaction variable grouped by manner class. The mean Reaction variable increases from 251 ms for Obstruents to 270 ms for Sonorants to 281 ms for Glottals—conforming to the expected pattern—but significance was not demonstrated in this case ($F(2,60)=1.26$, $p=.2909$). Nor were any of the paired comparisons between different manners significant in a Fisher's PLSD post hoc test. If indeed Manner does affect response time as hypothesized, a larger sample would be needed to demonstrate the fact. Of course this experiment does not disprove the notion that it takes listeners longer to categorize the phonological length of glottals than of obstruents, for example. The results are suggestive of it, but it cannot be asserted on the basis of this response time experiment.

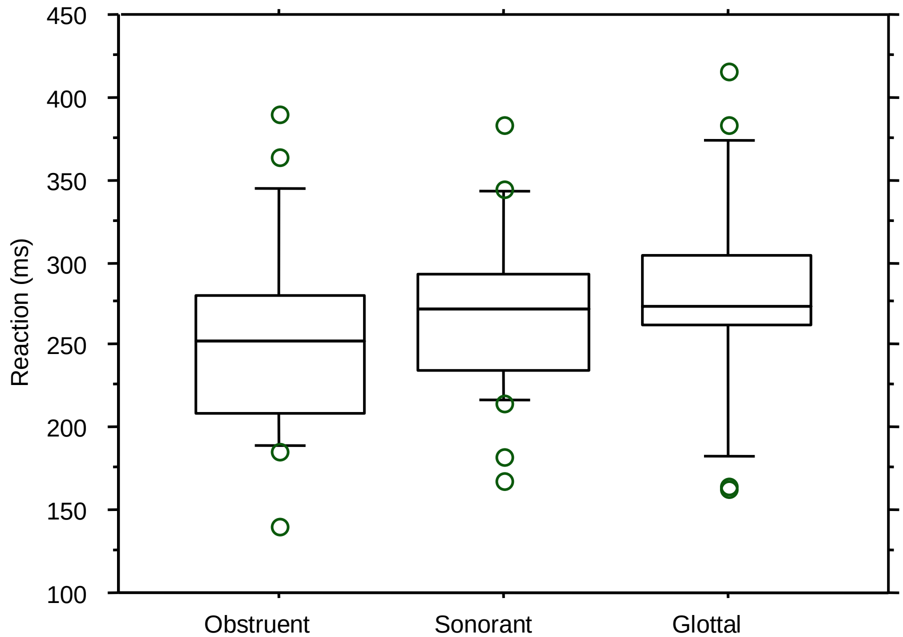

Figure 3.12 Response Time Statistics by Manner.
Box Plots show mean Reaction values in ms for the Obstruent, Sonorant and Glottal Category conditions.

3.6 DISCUSSION

When I set out to evaluate whether there are differences in the perceptibility of phonological length between various consonants in Persian, I defined certain parameters I thought would be relevant as perceptual markers for the vowel-consonant interface. I hypothesized that where there was a sharp change in the parameter between the vowel and the consonant, it would be easier for the listener to distinguish the duration of the consonant, and subsequently correctly categorize the consonant as “long” or “short.”

In Experiment I of this dissertation I tested three parameters at the vowel-consonant interface:

- Amplitude Difference: The loudness difference in decibels between the preceding vowel and the test consonant.
- Transition: The duration of the F1 and F2 formant transitions between the preceding vowel steady state and the test consonant formant steady state
- Offset F1: The frequency value of the first formant in the last pulse of the vowel preceding the test consonant formant steady state.

Each of the obstruents [d,z] was observed to have a significantly larger Amplitude Difference than any of the sonorants [n,l,j]. This was true for each of the four informants. On this basis, I predicted that sonorants would be harder for listeners to categorize for phonological length than obstruents. This hypothesis was confirmed by testing in Experiment II in which the slopes of the identification curves for the obstruents, as measured by the parameter Breadth, were significantly steeper than those for the sonorants, indicating that the identification task was easier for obstruents. I also hypothesized that the consonants that are easier to categorize would show faster response times. However, the response time test was inconclusive in that the smaller response time values observed for obstruents were not significantly different from those of the sonorants.

The Transition and Offset F1 parameters are inter-related measures of the abruptness of the vowel-consonant formant transitions. Since the oral consonants tested were all characterized by low first formants, those having longer transitions also had low Offset F1 values, which is unremarkable since the long transition interval provides ample time for the articulators to adjust to the constricted position. The consonants [z,j] are typical long Transition-low Offset F1 consonants of this type, as described in the discussion of Experiment I. By contrast [d,n,l] had shorter transitions and consequently the F1 value held relatively high at the moment of closure, indicating an abrupt formant transition at the vowel-consonant boundary. In terms of the phonological discriminability of long-transition consonants, Experiment II did not demonstrate that they were harder to categorize by phonological length than other oral consonants. Thus [z] did not have a larger Breadth value than [d], nor did [j] have a larger Breadth value than the other sonorants [n,l]. Therefore, on the basis of these perceptual tests on oral consonants, I cannot conclude that the abruptness of the formant transition is a factor in the perceptibility of phonological length, as hypothesized.

The behavior of the glottal consonants, on the other hand, does suggest that the lack of a formant movement cue could contribute to confusability between geminate and singleton glottals. Despite the fact that they had larger Amplitude Differences than sonorants did for all informants, they also produced much larger Breadth values in the perception tests. This suggests to me that Amplitude Difference is not the only determining factor in whether consonant length is confusable for a given segment type. It remains a possibility that the lack of an abrupt formant transition can contribute to the confusability of geminates and singletons. Experiment III in the following chapter explores this possibility.

4.0 Experiment III: Formant Transitions as Imprecise Segment Boundary Landmarks

4.1 INTRODUCTION

The objective of the present experiment is to determine whether it is possible to observe that listeners have greater difficulty distinguishing geminates from singletons if the boundary between them is artificially blurred. One way to support a claim that a parameter affects the perceptibility of duration in contrastive length judgments is to vary that parameter systematically, to intentionally blur or sharpen the vowel-consonant boundary, and observe whether listeners' ability to categorize segments varies as predicted.

Of the measures of vowel-to-consonant boundary sharpness studied in Experiment I, the F1 and F2 transition duration was the most convenient for the present purpose because I had experience producing natural-sounding tokens utilizing manipulated vocoid formant transitions: formant transitions were successfully varied synthetically to produce distinct conditions in perception tests in Myers and Hansen (2005).

There are also other acoustic transitions that occur as articulation proceeds from vowel to consonant that are typical of particular segment classes. For example, in nasals, there is a gradual lowering of the velum within the preceding vowel, independent of the oral gesture. This velar lowering increases airflow through the nasal cavity, thereby inducing antiformants that are only fully developed when the oral closure is complete. In some languages (though not Tehrani Persian) "dark" laterals also have a strong posterior articulation in addition to the anterior one. In these cases, an anticipitory raising of the tongue root can begin within the vowel preceding the anterior contact of the [l]. In glottals, the abduction or adduction gestures are gradual, introducing a gradual increase in breathiness or stiffness, respectively, in the vowel preceding the maximum extent of the glottal gesture (Stevens 1998). These and other measures of the vowel-consonant interface that relate to a limited class of segments are not suited to making general predictions about relative length discriminability of different manners of articulation. Still, my hypothesis would predict that to the extent that the velar, tongue root and glottal gestures overlap with the vowel articulation, listeners will find it difficult to categorize phonologically long and short versions of these segments on the basis of duration. These are open questions that could be addressed by future experimentation.

4.2 THIS EXPERIMENT

The present experiment investigates whether increasing the transition length makes it harder for listeners to make a categorical decision as to the phonological length of a glide segment. I have used three transition length conditions (Short, Medium,Long) ranging from the low end to the high end of the natural range of F1 transition durations in [æjɑ] sequences observed in the production data. I predicted that glides with longer transitions would be harder for listeners to categorize than glides with shorter transitions.

As in Experiment II and Kawahara (2007), I used a measure of the steepness of the slopes of the percent-identification curves to gauge the sharpness of the threshold. I also recorded response times as measures of the degree of difficulty of the geminate discrimination task.

A separate question regarding the segmentation of glides is, is the transition perceived to belong to the vowel or to the glide? In keeping with Myers and Hansen's (2005) observation that the glide-vowel transition is perceived partly as glide and partly as vowel, I expect that tokens with longer transitions will be more likely to be judged to contain more consonant material and so even tokens with a fairly short constriction durations will be heard as geminates by listeners if the transitions were long, resulting in lower threshold values for the Long transition condition. Conversely, tokens with short transitions are more likely to be heard as singletons and so listeners will not identify them as geminates unless the closure interval was much longer, and so the threshold for the Short transition condition should be quite high.

4.3 METHODS

4.3.1 Subjects

Eight female and five male adult Persian speakers participated in the experiment. The group of female subjects included all five of the female subjects from Experiment II. Three of the five male subjects were also subjects in Experiment II, including the one who also recorded for Experiment II. The informants are all residents of the United States, and were selected from family, friends and University colleagues. As in Experiment II, they were not compensated for their effort, but were willing, often enthusiastic participants. Their ages range from 30 to 65 years. All are literate in standard literary Persian as used in Tehran, having attended at least high school in Iran. Though they reside in the United States and speak English regularly, they continue to use Persian in their family and social environments.

4.3.2 Materials

The recordings used to prepare the stimuli were produced by female informant f1 from Experiment 1. Prior to Experiment II she recorded nonsense words of the form [qæC(ː)ɑb] at a conversational speaking rate within the carrier sentence [minu ______ ra did]: "Minu saw the ______". An excised nonsense word [qæjɑb] was the sole precursor for all of the stimuli in this experiment. The particular sentence frame from which the token for this experiment was excised had a duration of 1.04 seconds.[13]

---

[13] By comparison, the mean values of Frame for f1's repetitions of [pæyam] in the same context were 1.238 at the conversational rate and 0.887 at the fast rate, so the precursor for this experiment can be said to have been produced at a fast conversational rate for this speaker.

Three series of continua were produced from the precursor word: a Short transition series, a Medium transition series, and a Long transition series. The Medium series conserved the vowel and transition durations as recorded. The measured durations in the precursor were the following:

æ steady state	40 ms
æ-j transition	78 ms
j steady state	51 ms
j-ɑ transition	108 ms
ɑ steady state	209 ms

Table 4.1 Medium Transition Series Precursor Durations

I manipulated the recordings using the same synthesis and analysis procedures I used for Experiment II. In order to make the comparison of the three transition series comparable, I smoothed the [æ-j] and [j-ɑ] transition intervals by straight-line interpolation of formants F1 and F2 prior to performing the successive insertions and deletions of glottal pulses from within the steady state of [j]. The resulting synthesized series of 10 stimuli had [j] steady states ranging from 20 ms to 112 ms in approximately 10-11 ms increments. The other durations were as in Table 4.1 for all of the Medium series tokens. The tokens were natural-sounding and comparable to the short and long transition tokens in that they also had straight-line interpolated F1 and F2 transitions.

The Short transition series was produced by deleting four pulses totaling approximately 20 ms from near the center portions of both the [æ-j] and [j-ɑ] transitions of the precursor token. F1 and F2 transitions were smoothed by straight-line interpolation, resulting in a token having the following interval durations:

æ steady state	40 ms
æ-j transition	57 ms
j steady state	51 ms
j-ɑ transition	88 ms
ɑ steady state	208 ms

Table 4.2 Short Transition Series Precursor Durations

Ten steady state [j] durations ranged from 20 ms to 112 ms in approximately 10-11 ms increments, with the other interval durations retaining the constant values shown in Table 4.2.

The Long transition series was based on the same [qæjɑb] precursor as the Medium and Short series. [æ-j] and [j-ɑ] transition durations were increased by approximately 20 ms each by duplicating four glottal pulses near the centers of the transitions. Again the F1 and F2 transitions were smoothed by straight-line interpolation. The resulting token interval durations are the following:

æ steady state	40 ms
æ-j transition	98 ms
j steady state	51 ms
j-ɑ transition	128 ms
ɑ steady state	208 ms

Table 4.3 Long Transition Series Precursor Durations

The ten steady state [j] durations ranged from 20 ms to 113 ms in approximately 10-11 ms increments, and the other interval durations retained the constant values shown in Table 4.3.

For all 30 synthesized tokens the consonants and vowel steady states preceding and following the test consonant and its transitions, had approximately the same durations, thereby minimizing confounding factors relating to different token durations. The only difference between the three series was the duration of the transitions surrounding the test consonant. No distortion or unnatural-sounding noise was present in the tokens.

4.3.3 Procedure

The experiment was carried out using Superlab program installed on an IBM laptop computer as in Experiment II. The experiment was conducted in a quiet place convenient to the subject, such as a kitchen table or office. Subjects listened to the tokens over headphones and recorded their judgments as to whether they heard a geminate or a singleton glide by clicking the appropriate key on either the computer keyboard or an external response pad[14]. Subjects were instructed to respond as quickly as possible, and to guess if it was not clear which of the two categories the token belonged to. As in Experiment II, the response key selection and response time from the end of the token were logged in an ascii file on the computer.

All thirty tokens were presented in a single block in random order. Subjects heard the block six times, each time in a different random order. After a brief rest, they heard the block six more times for a total of twelve responses per token. In this way I was able to derive a numerical probability of a geminate judgment for each of the tokens. I thus obtained 10 increments * 3 transition conditions * 12 repetitions = 360 judgments per subject. As in Experiment II response times greater than 1000 ms were excluded from the data.

4.4 ANALYSIS

I used the same Cumulative Gaussian modeling procedure described in the discussion of Experiment II to obtain Threshold and Breadth values for each of the

[14] On the keyboard, I initially used 1 for singleton and 2 for geminate until I had some trouble with the keys sticking. I then used “s” for singleton and “d” for geminate. The Keytronic response pad had clear key covers that allowed me to mark one key with a dash for singleton and the other with a Persian tashdid diacritic for geminate.

thirteen subjects at each transition condition. The modeled Threshold and Breadth values obtained from the R analysis formed the input data for the statistical analysis discussed below in the results section.

I also pooled the responses from the subjects and modeled the Cumulative Gaussian curves for each transition condition.

4.5 RESULTS

There were marked differences between the three identification curves resulting from the three transition length conditions (Figure 4.1). The points in Figure 4.1 represent the pooled overall percentages identified as geminates at a particular constriction value, with different symbols representing the three transition conditions.

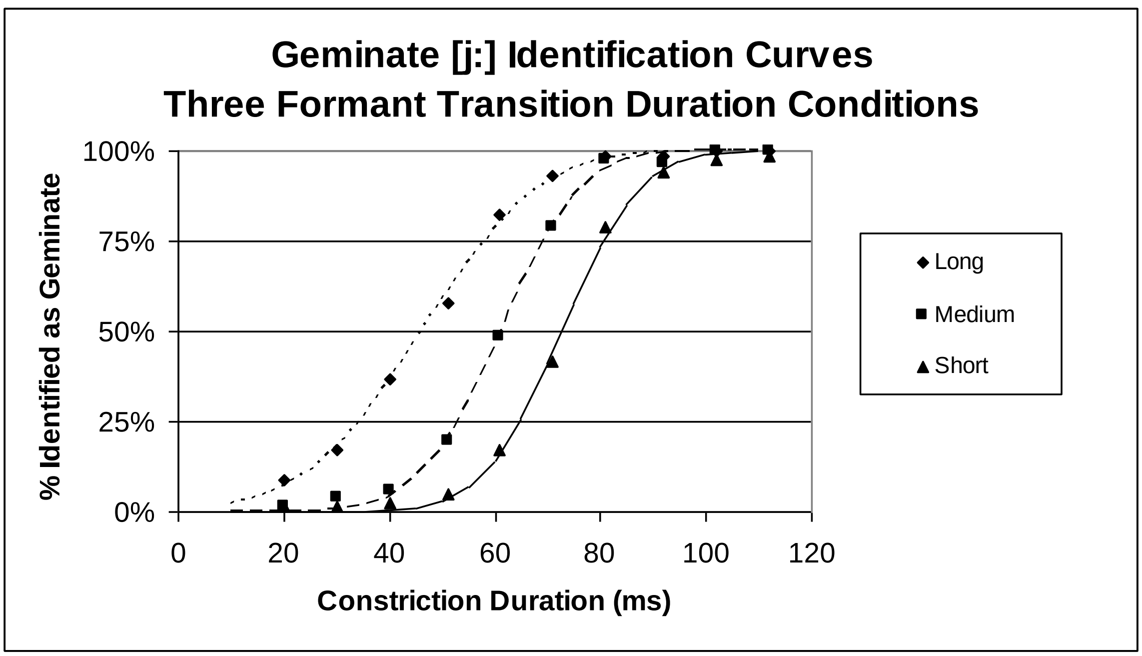

Figure 4.1 Geminate [j:] Identification Curves.

Three transition duration conditions: Long, Medium and Short. Pooled data for 13 subjects. Curves show Cumulative Gaussian model for pooled data.

The curves shown in figure 4.1 were modeled in R on the pooled data points. The glides with the longer transitions had the lower Threshold values, and their identification curves had the flatter slopes. Table 4.4 Shows the Threshold and Breadth parameters obtained from the Cumulative Gaussian R models of the pooled response data. Since the Breadth value increases along with the transition duration, the results support the hypothesis that it is harder for a listener to categorize the length category of a segment with a longer transition.

Consonant	Transition Category	Threshold	Breadth
j	Short	72.59	11.85
j	Medium	61.23	12.15
j	Long	46.18	17.45

Table 4.4 Pooled Modeled Gaussian Parameters

Mean (Threshold) and Standard Deviation (Breadth) for Three Transition Conditions. All values are expressed in ms.

The observation that Threshold is larger for shorter transition conditions validates the prediction that tokens with longer transitions would be more readily classified as geminates and supports the claim in Myers and Hansen (2005) that a portion of the vowel-glide transitions is perceived as part of the glide.

To test the significance of the qualitative observation that the glides with longer transitions had lower thresholds and steeper slopes, I modeled each subject's identification curves separately in R and analyzed the resulting parameters Threshold and Breadth as statistical populations of thirteen subjects. The one-way ANOVA of Threshold for the independent variable Trans. Cat. (transition duration category) was highly significant ($F(2,36)=37.508$, $p<.0001$). This indicates that the threshold increases systematically as the transition duration is reduced. Fisher's PLSD post hoc tests showed that the differences between the Threshold values was highly significant ($p<.001$) between all conditions. The box plot in Figure 4.2 represents the distribution of the modeled Threshold parameter for each of the three transition conditions. Table 4.5 is the means table for the same data.

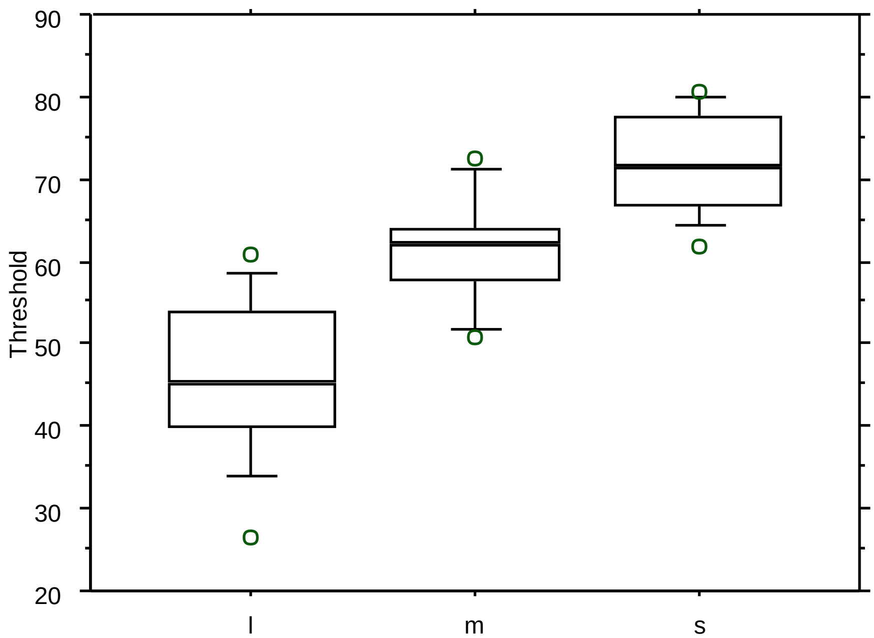

Figure 4.2 Glide Threshold Box Plot.

Box Plots (top) show Threshold values in ms for the Short (s), Medium (m) and Long (l) Transition Category conditions. n=13 for each category.

Transition	Count	Mean	Std. Dev.
Long	13	46.146	9.842
Medium	13	61.177	6.549
Short	13	72.338	6.314

Table 4.5 Means Table for Breadth (ms) by Segment. Thirteen Gaussian Models.

Figure 4.3 shows box plots representing the modeled Breadth parameter for the three transition conditions based on the thirteen subjects' Gaussian models. The trend of decreasing value of Breadth with shorter transition durations is evident, but there is a degree of overlap among the three categories. In the one-way ANOVA for Breadth, the main effect of Transition Category is significant ($F(2,36)=5.418$, $p=.0088$), but the

Fisher's PLSD post hoc tests were not significant between all transition conditions. The difference between the Medium and the Short conditions was not significant, though it was in the expected direction. Differences between the Long and the other two transition conditions were significant to the .05 level.

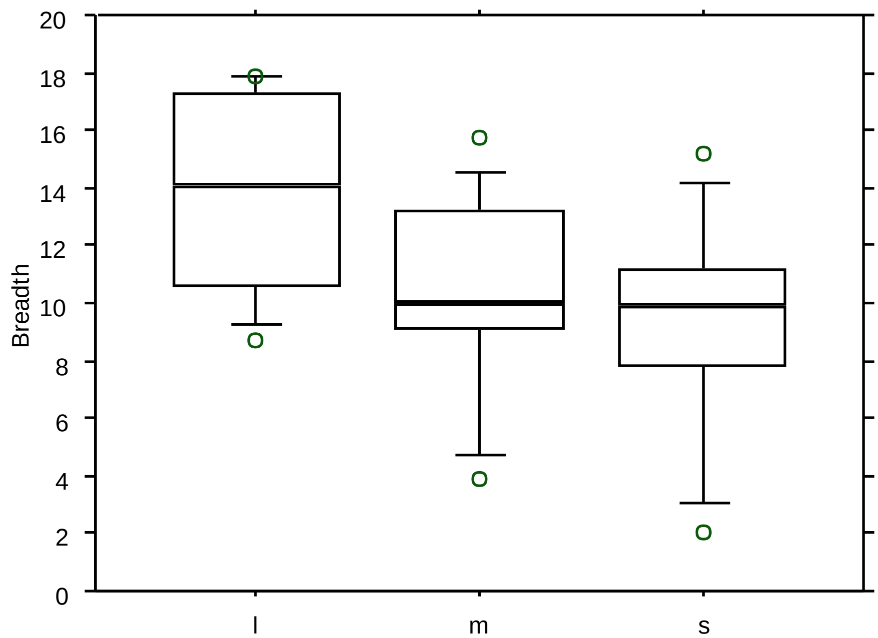

Figure 4.3 Glide Breadth Statistics.
Box Plots (top) show Breadth values in ms for the Short (s), Medium (m) and Long (l) Transition Category conditions. n=13 for each category.

Table 4.6 is the means table for the data presented in Figure 4.3. The means of the Threshold values of the 13 models differ somewhat from modeled values of Threshold for the pooled response curves given in Table 4.4 above due to the variability in the Breadth values obtained from the Gaussian models which are sensitive to small changes in the percent identified of individual subjects.

Transition	Count	Mean	Std. Dev.
Long	13	13.823	3.567
Medium	13	10.354	3.447
Short	13	9.323	3.925

Table 4.6 Means Table for Breadth (ms) by Segment. Thirteen Gaussian Models.

The other measure of the perceptual confusion used is the response time. Since the formant transition duration *following* the consonant constriction was altered for the Short and Long conditions, and since I wanted to be able to compare the response time from the time that the consonant is perceived, for Experiment III I compared the response times measured from the end of the consonant constriction, rather than from the end of the token as was done in Experiment II. To accomplish this I adjusted the response times by adding 304 ms to the Short condition response times, 324 ms to the Medium condition response times, and 344 ms to the Long condition response times. The results are displayed in Figure 4.4.

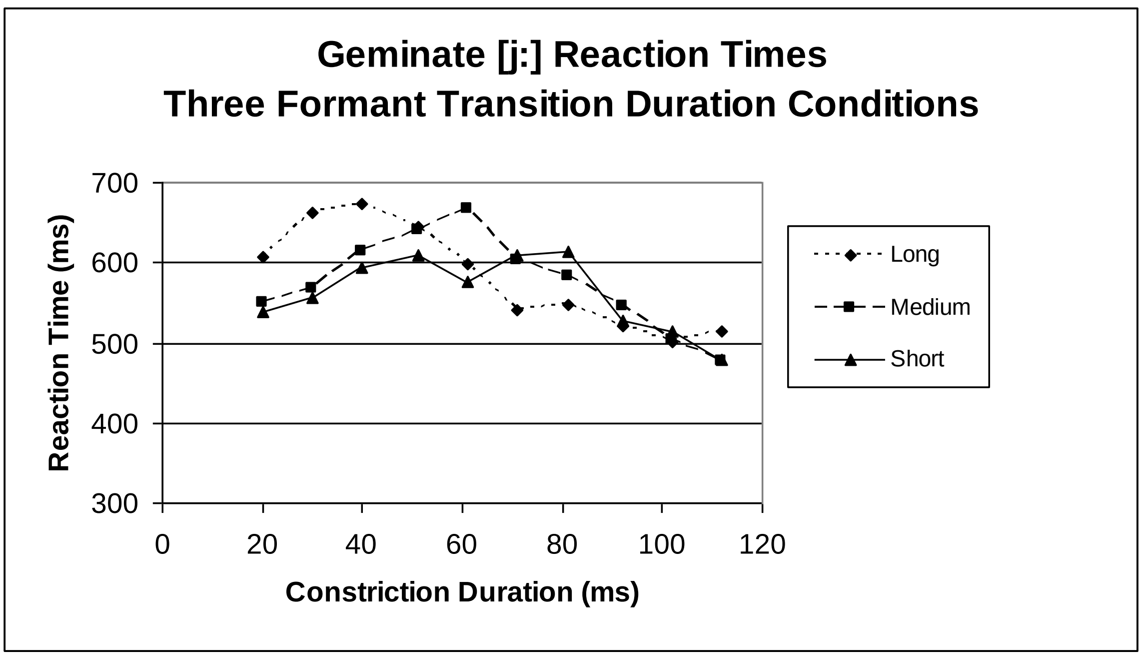

Figure 4.4 Geminate [j:] Reaction (Response) Times (ms).

Three transition duration conditions: Long, Medium and Short. Pooled data for 13 subjects. Curves show straight line interpolations between the pooled data.

The peaks of the three curves correspond well with the threshold values (See Figure 4.1 above), showing that subjects took longer to decide on tokens that were intermediate between geminates and singletons. A peak in response times was noted in perception experiments for tokens that are intermediate between two length categories by Myers and Hansen (2005) in relation to phonologically long and short vowels and by Kawahara (2007) in relation to singleton-geminate consonant continua.

A box plot representing the distribution of the average response times for the thirteen subjects is shown in Figure 4.5.

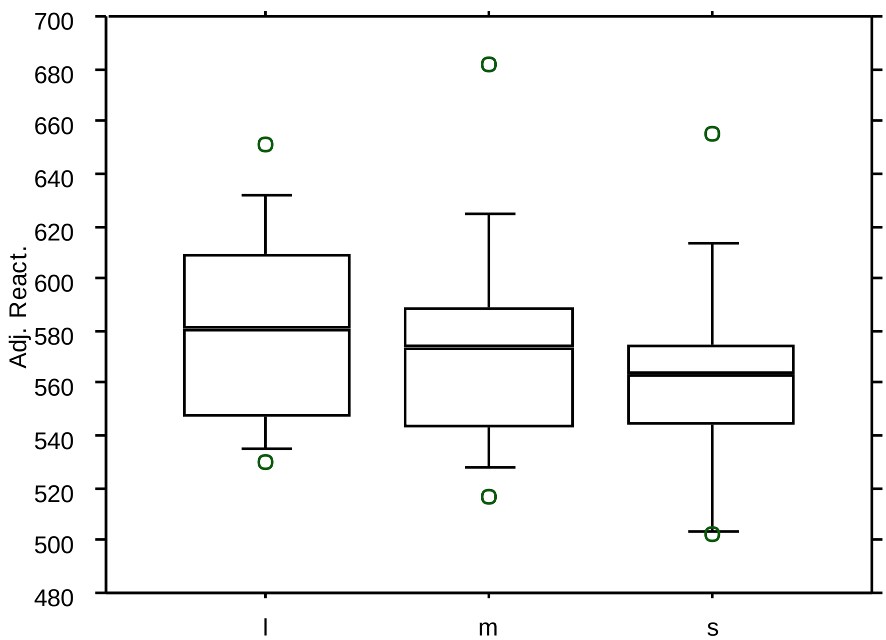

Figure 4.5 Glide Response Time

Box Plots show Adj. React. values in ms for the the Short (s), Medium (m) and Long (l) Transition Category conditions. n=13 for each category.

The response time, as adjusted to account for differences in transition durations following the test consonant, is labeled Adj. React. in Figure 4.5. Considerable overlap in the response time statistic is evident, yet the mean response time does decrease as the transition shortens. However, the one-way ANOVA of Adj. React. for the independent variable Trans. Cat. does not indicate that the pattern is significant ($F(2,36)=.778$, p=.4669). Neither are any of the Fisher's PLSD post hoc tests significant. Therefore we lack evidence for the proposition that it takes longer for people to categorize a glide with a longer transition, though the proposition is certainly not disproven.

4.6 DISCUSSION

Experiment III demonstrates that the length of the formant transition alone can alter the perception of phonological length. Not only does threshold shorten as the transition duration increases, but the breadth of the threshold increases, indicating that there is uncertainty over a wider range of constriction durations.

The shorter threshold observed for longer transitions was expected because previous work by Myers and Hansen (2005) showed that in vowel-glide sequences, the transition was perceived partly as vowel, partly as glide. My reasoning is that when an increase in the transition duration supplies additional material to be perceived as part of the glide, listeners register a geminate percept upon insertion of a rather small interval of steady state into the glide. Conversely, when there was a smaller than normal transition to work with, listeners needs quite a long addition of glide steady state before the stimulus would signal geminate to the subjects. The experiment showed this prediction to be correct, which supports the notion that the glide transition is partly perceived as glide, partly as vowel.

My central hypothesis that fuzzy, indeterminate boundaries between consonants and vowels make it difficult to gauge consonant duration and thus listeners cannot easily judge whether such a consonant is geminate or singleton, is supported by this experiment. I had predicted that the long-transition glides would be harder for listeners to categorize as geminate or singleton, and so would exhibit a flatter, broader threshold than glides with short transitions. This was shown to be the case.

My other prediction was that subjects' response times would lengthen along with the transition length. This was true on average in the data, but did not achieve the level of significance. A larger sample of responses could be obtained in an experiment specifically designed to test this point As it is, my response data neither support nor discount an effect of transition length on response time.

5.0 Conclusion

This dissertation has addressed questions about the nature of Persian singleton and geminate consonants and about the perception geminates by Persian speakers. It has also attempted to shed light on proposals regarding the existence of gaps in geminate inventories in other languages.

There has been very scant instrumental phonetic description of Persian. This is a first attempt to quantify such things as the duration and amplitude of its consonants and the effect different segment lengths have on an adjacent vowel. Lehtonen (1970) described the effects of consonant length on adjacent vowels in Finnish. Like Finnish, vowels preceding geminates in Persian appear to be longer than those preceding singletons. This is completely unlike other languages for which durations of vowels preceding geminate have been reported to be shorter (e.g. Pickett et al. 1999 on Italian).

This dissertation investigated geminate durations at three speaking rates. It was found that geminate consonants are more sensitive to the speaking rate than singletons. Arvaniti (1999) found this to be true for Cypriot Greek stops but not for sonorants. I found that in Persian geminate stops, fricatives, sonorants and glottals all had a more greatly reduced duration at faster speaking rates than their singleton counterparts did.

In the perception experiments of this dissertation, it was found that it is harder to discriminate between geminate and singleton sonorants than between geminate and singleton obstruents. A similar result was obtained by Kawahara (2007) with respect to Arabic geminates. My finding that geminate glottals are even harder to distinguish than sonorants is a new result.

In this dissertation I performed an experiment that tested another dimension of vowel-likeness apart from amplitude, which, it has been theorized, is the source of the perceptual confusion of geminate and singleton sonorants (Kawahara 2007). My experiment on the effect of the duration of vowel-glide transitions on the perceptibility of the geminate/singleton contrast was based on the idea that a loosely delimited transition between vowel and consonant would affect how easily the distinction is perceived. I found that the longer the transition, the broader the perceptual threshold between singleton and geminate, indicating a greater confusability.

One contribution to the analysis of linguistic data I believe was made by this dissertation was the use of a cumulative Gaussian curve to model the perception function. The advantage of this model, which has been used in Psychology to model the psychometric function (Wichmann and Hill 2001a and 2001b), is that it produces one parameter that represents the threshold of the perceptual boundary, and another that represents the breadth of that boundary. The breadth of the boundary can be used as a correlate of the relative fuzziness (broad) or sharpness (narrow) of the boundary.

6.0 References

Abramson, A. S. (1999). Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. In J. Ohala, Y. Hasegawa, M. Ohala, D. Granville and A. Bailey (eds.) *Proceedings of the XIVth International Congress of Phonetic Sciences*. Linguistics Department, University of California at Berkeley. 591-594.

Al-Nassir, A. (1993). *Sibawayh the Phonologist: A Critical Study of the Phonetic and Phonological Theory of Sibawayh as Presented in His Treatise Al-Kitab.* Library of Arabic Linguistics, Monograph no. 10. New York: Kegan Paul International.

Andrews, R. (1975). *Introduction to Classical Nahuatl* p. 9. Univ. of Texas Press, Austin.

Arvaniti, A. (1999). Effects of speaking rate on the timing of single and geminate sonorants. In *Proceedings of the XIVth International Congress of Phonetic Sciences, San Francisco, CA.* 599-602.

Bannert, R. (1976). *Mittelbairische Phonologie auf akustischer und perzeptoricher Grundlage*. Travaux de l'Institut de Linguistique de Lund 10. CWK Gleerup, Malmö.

Basset, A. (1929). *La Langue Berbère-Morphologie, Le Verbe* p.23. Paris: Ernest Leroux.

Blevins, J. (2004a). Klamath sibilant degemination: implications of a recent sound change. *International Journal of American Language.* 70 (3). 279-89.

Blevins, J. (2004b). *Evolutionary Phonology.* Cambridge University Press.

Blevins, J. (2005). *The typology of geminate inventories: Historical explanations for recurrent sound patterns*. Seoul Linguistics Forum 2005. Seoul: Language Education Institute, Seoul National University. 121-37.

Christensen, J. and Christensen, S. (1992). Kisar Phonology. In D. A. Burquest and W. D. Laidig (eds.) *Phonological studies in four languages of Maluku*. Summer Institute of Linguistics and the University of Texas at Arlington Publications in Linguistics 108. 33-65.

Clements, G. (1986). Compensatory lengthening and consonant gemination in LuGanda. In L. Wetzels and E. Sezer (eds.) *Studies in Compensatory Lengthening*. Dordrecht: Foris Publishers.

Cravens, T. (2002). *Comparative historical dialectology: Italo-Romance clues to Ibero-Romance sound change*. (Current Issues in Linguistic Theory 231.) Amsterdam & Philadelphia: John Benjamins Publishing Company. Pp. xi+163.

Deyhime, G. (2000). *Farhang-i Avayi-i Farsi* (Persian Pronunciation Dictionary) Tehran: Farhang Moaser Publishers.

Han, M. S. (1992). The timing control of geminate and single stop consonants in Japanese: a challenge for nonnative speakers, *Phonetica* 49. 102-127.

Hankamer, J., A. Lahiri & J. Koreman (1989). Perception of consonant length: voiceless stops in Turkish and Bengali. *Journal of Phonetics* 17. 283-298.

Hansen, B. B. (2003). Production of Persian Geminate Stops: Speaking rate effects, A Qualifying Paper for the Department of Linguistics, University of Texas at Austin (unpublished).

Hansen, B. B. (2004a). Measurements of geminate and singleton consonant durations: data from one speaker. Unpublished notes.

Hansen, B. B. (2004b). Production and perception of Persian geminate stops at three speaking rates. Poster presentation: Acoustical Society of America 75th Anniversary Meeting, New York, May 28, 2004.

Hansen, B. B. (2004c). Production of Persian geminate stops: Effects of varying speaking rate. In A. Agwuele et al. (eds.) *Proceedings of the 2003 Texas Linguistics Society Conference*. Somerville, MA: Cascadilla Proceedings Project. 86-95.

Hansen, B.B. (2005). Does the sharpness of the perceptual boundary between geminates and singletons impose restrictions on geminate segment inventories? Talk handout: Univ. of Texas at Austin Department of Linguistics April 22, 2005.

Hansen, B. B. (2006). The role of vowel-to-consonant transitions in the perception of phonological consonant length. Poster presentation: Acoustical Society of America 151st Meeting, Providence, June 5, 2006.

Hardy, H. and Montler, T. (1998). Imperfective gemination in Alabama. *International Journal of American Language* 54. 399-415.

Hayes, B. & Steriade, D. (2004). Introduction: the phonetic basis of phonological markedness. In B. Hayes, R. Kirchner & D. Steriade (eds) *Phonetically-Based Phonology*. Cambridge: Cambridge University Press.

Homma, Y. (1981). Durational relationship between Japanese stops and vowels, *Journal of Phonetics* 9. 273-281.

Karlsson, F. (1992). Finnish. Article in Bright, W. (ed.) *International Encyclopedia of Linguistics* v. 2. Oxford: Oxford University Press. 14-17.

Kato, H. and Tajima, K. (2004). Effects of speaking rate on the perception of phonemic length contrast in Japanese. Poster presentation. 147th Meeting of the Acoustical Society of America, New York City, May 24-28, 2004.

Kavitskaya, D. (2002). *Compensatory Lengthening: phonetics, phonology, diachrony*. New York: Garland.

Kawahara, Shigeto (2006). Mimetic gemination in Japanese: A challenge for Evolutionary Phonology. *Theoretical Linguistics* 32 (3). 411-424.

Kawahara, S. (2007). Sonorancy and Geminacy. *University of Massachusetts Occasional Papers in Linguistics 32: Papers in Optimality Theory*. 145-186.

Keer, E. (1999). Dissertation, Rutgers.*Geminates, the OCP and the nature of* CON. New Brunswick.

Kessinger, R. H. & S. E. Blumstein (1997). Effects of speaking rate on voice-onset time in Thai, French and English. *Journal of Phonetics* 25. 143-168.

Kirschner, R. (2000). Geminate inalterability and lenition. *Language* 76 (3). 509-545.

Klatt, D. (1973). Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America* 54 4. 1002-1004.

Klatt, D. (1976). Linguistic uses of segment duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59. 1208-1221.

Lahiri, A. & J. Hankamer (1988). The timing of geminate consonants, *Journal of Phonetics* 16. 327-338.

Lehiste, I. (1970). *Suprasegmental*s. Cambridge: MIT Press.

Lehtonen, J. (1970). *Aspects of Quantity in Standard Finnish*. Jyväskylässa, Finland.

Leslau, W. (1968). *Amharic Textbook*. Berkeley and Los Angeles: University of California Press.

Lewis, Anthony M. (2002). Contrast maintenance and intervocalic stop lenition in Spanish and Portuguese: When is it alright to lenite? In T. Satterfieldet al. (eds.) *Current Issues in Romance Languages*. 159–171.

Lindblom, B. (1990). Explaining phonetic variation: a sketch of the H&H theory. In W. J. Hardcastle and A. Marchal (eds.) *Speech Production and Speech Modelling*. The Netherlands: Kluwer Academic. 403-439.

Lindblom, B., S. Guion, S. Hura, S-J. Moon & R. Willerman (1995). Is sound change adaptive? *Rivista di Linguistica* 7. 5-36.

Lowenstamm, J. and J. Kaye (1986). Compensatory lengthening in Tiberian Hebrew. In L. Wetzels and E. Sezer (eds.) *Studies in Compensatory Lengthening*. Dordrecht: Foris Publishers.

Maddieson, I. (1984). *Patterns of Sounds.* Cambridge University Press.

Mahootian, S. (1997). *Persian*. London: Routledge.

Madina, M.Z. (1973). *Arabic-English Dictionary*. New York: Pocket Books.

Miller, J. L., (1981). Some effects of speaking rate on phonetic perception. *Phonetica* 38. 159-180.

Myers, S. and Hansen, B.B. (2005). The origin of vowel length neutralization in vocoid sequences. *Phonology* 22. 317-344.

Newman, D. L. (2005). Contrastive analysis of the segments of French and Arabic, in Elgibali, A. (ed.) *Investigating Arabic*. Brill.

Ohala, J. (1981). The listener as a source of sound change. In C. S. Masek et al. (eds.) *Papers from the parasession on language and behavior*. Chicago: Chicago Linguistic Society. 178-203.

Ohala, J. & C. Riordan (1979). Passive vocal tract enlargement during voiced stops. In J. J. Wolf & D. H. Klatt (eds.) Speech Communication Papers, Acoustical Society of America. [*JASA* 65 (S1): S23] New York. 89-92.

Pickett, E. R., S. E. Blumstein & M. W. Burton (1999). Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56. 135-157.

Pind, J. (1986). The perception of quantity in Icelandic. *Phonetica* 43. 116-139.

Pind, J. (1995). Speaking rate, voice-onset time, and quantity: The search for higher-order invariants for two Icelandic speech cues. *Perception & Psychophysics* 57. 291-304.

Podesva, R. J. (2000). Constraints on Geminates in Buginese and Selayarese. In Billerey and Lillehaugen (eds) *WCCFL 19 Proceedings*. Somerville, MA: Cascadilla Press.

Recasens, D. (1999). Lingual coarticulation. In W. J. Hardcastle and N. Hewlett (eds.) *Coarticulation*. Cambridge University Press.

Ruhlen, M. (1975). *A guide to the languages of the world*. Language Universals Project, Stanford University.

Samareh, Y. (1977). *The Arrangement of Segmental Phonemes in Farsi*. University of Tehran Faculty of Letters.

Samareh, Y. (1985). *Avāshināsī-i Zabān-i Fārsī*. Tehran: Markaz Nashr Dāneshgāhī.

Sepanta, S. (1999). *The sound Pattern of Persian*. Esfahan: Golha Publisher.

Siptár, Péter and Miklós, Törkenczy (2007). *The Phonology of Hungarian*. Oxford University Press, USA

Smith, C. (1991). The timing of vowel and consonant gestures in Italian and Japanese, *Proceedings of the 12th International Congress of the Phonetic Sciences* 4. 234-237.

Stevens, K. N. (1998). *Acoustic Phonetics.* MIT Press.

Wichmann, F. A. & Hill, N. J. (2001a). The psychometric function: I. Fitting, sampling and goodness-of-fit. *Perception and Psychophysics* 63 (8). 1293-1313.

Wichmann, F. A. & Hill, N. J. (2001b). The psychometric function: II. Bootstrap-based confidence intervals and sampling. *Perception and Psychophysics* 63 (8). 1314-1329.