

The consonant length contrast in Persian: Production and perception

Benjamin B. Hansen

The University of Texas at Austin
benhansen@utexas.edu

Scott Myers

The University of Texas at Austin
s.myers@austin.utexas.edu

Across languages, there is a tendency to avoid length contrasts in the most vowel-like consonant classes, such as glides or laryngeals. Such gaps could arise from the difficulty of determining where the boundary between vowel and consonant lies when the transition between them is gradual. This claim is tested in Persian (Farsi), which has length contrasts in all classes of consonants, including glides and laryngeals. Persian geminates were compared to singletons in three different speaking rates and seven different consonant classes. Geminates were found to have longer constriction intervals than singletons, and this length effect interacted with both speaking rate and manner of articulation. In one of two perception experiments, Persian speakers identified consonants as geminate or singleton in stimuli in which the constriction duration was systematically varied. The perceptual boundary between geminates and singletons was most sharply defined for obstruents and least so for laryngeals, as reflected by the breadth of the changeover region in the identification curve. In the other perception experiment, subjects identified the length class of glides differing in constriction duration and formant transition duration. Longer formant transitions led to more geminate responses and to a broader changeover interval.

Persian (Farsi) has a contrast in consonant length, as illustrated by the minimal pair bæŋd ‘building’ – bæŋnda ‘builder’. This contrast extends to all the consonant classes in the language, listed in [Table 1](#) (Samareh 1977, 1985; Mahootian 1997; Majidi & Ternes 1999; Deyhime 2000; Bijankhan & Nourbakhsh 2009; Hansen 2004, 2012). Samareh (1977) does not include geminate [ʒʒ] or [GG] in his inventory of Persian clusters, and we know of no instances of these geminates in Persian words. Yet we find that Persian speakers freely produce them in reading nonsense words with the relevant consonants marked with the length mark. This suggests that these are accidental gaps, and that these two consonants are not systematically avoided.

With that caveat, the consonants in [Table 1](#) all occur as singletons or as geminates, where the geminate differs from the corresponding singleton in having a longer constriction interval. The geminate has the same feature specification as the corresponding singleton, but

Table 1 Consonants of Persian.

	Bilabial	Labio-dental	Dental/Alveolar	Post-alveolar	Palatal	Uvular	Glottal
Plosives	p b		t d	tʃ dʒ	c ɟ	ɢ	ʔ
Fricative		f v	s z	ʃ ʒ		χ	h
Nasal stop	m		n				
Alveolar trill			r				
Approximant			l		j		

that feature specification is associated with two timing units in the geminate and one in the singleton (Leben 1980, Prince 1984, Hayes 1989).

The range of consonants with length contrasts in Persian is remarkable because in many languages the most vowel-like consonants, such as glides, liquids, and laryngeals, are excluded from the contrast (Podesva 2000, Aoyama & Reid 2006, Kawahara 2007, Maddieson 2008, Kawahara & Pangilinan to appear). For example, Icelandic (Garnes 1976) and Classical Nahuatl (Andrews 1975) have length contrasts for all classes of consonants except for glides. Koya (Tyler 1969: 34) and Telugu (Kostić, Mitter & Krishnamurti 1977) have length contrasts in all consonants except for laryngeals. Length contrasts in Luganda (Ashton et al. 1954) include all consonants except for glides and liquids, and those in Toba Batak (Nababan 1981) and Finnish (Lehtonen 1970) include all consonants except for glides and laryngeals. Glides, liquids and laryngeals are excluded from length contrasts in Punjabi (Malik 1995) and Pengo (Burrow & Bhattacharya 1970). Other languages combine such restrictions with other restrictions on what classes of consonants can be long, e.g. Japanese (no length contrasts in glides, liquids, laryngeals or voiced obstruents: Kawahara 2007) or Tamazight Berber (no length contrasts in glides or voiced obstruents: Penchoen 1973).

Podesva (2000) and Kawahara (2007) express these cross-linguistic tendencies within Optimality Theory, with markedness constraints which penalize long consonants (geminate) with high sonority. Both authors suggest that these phonological constraints have a phonetic basis in the difficulty of fixing the acoustic boundary between a sonorant consonant and a neighboring vowel. The transition between such a consonant and a vowel tends to be gradual, so that it is difficult to determine where the first sound in the sequence ends and the second begins (Peterson & Lehiste 1960, Myers & Hansen 2005).

These segmentation issues are illustrated in Figure 1 below in three spectrograms of Persian [æCɑ] sequences. The steady state interval of the medial consonant is marked off by the arrows at the bottom of the window.

Figure 1a shows an example of a medial nasal stop /n/, which displays the abrupt transitions that clearly delimit the closure interval of the nasal stop. It is easy to determine where the preceding vowel ends and where the following vowel begins, so it is easy to determine the duration of the consonant. Figure 1b gives an example with a medial palatal glide /j/, where the brief steady state interval marked off by the ticks is preceded and followed by long, gradual formant transitions. The formant values in the transition reflect both the consonant and the neighboring vowel (Öhman 1967), which makes it challenging to determine where the first vowel ends and the glide begins, or where the glide ends and the second vowel begins (Peterson & Lehiste 1960). Figure 1c gives an example with a medial laryngeal /h/. The formants for the preceding and following vowel extend through the laryngeal, so there is no formant transition marking the edge of the consonant (Keating 1988). The noise of the consonant interval builds up gradually, spreading from the higher frequencies to the lower ones. Reduced realizations of the laryngeals /h/ and /ʔ/ frequently consist just of an interval of breathy or creaky voicing (respectively) in the formants (Pierrehumbert & Talkin 1992). It

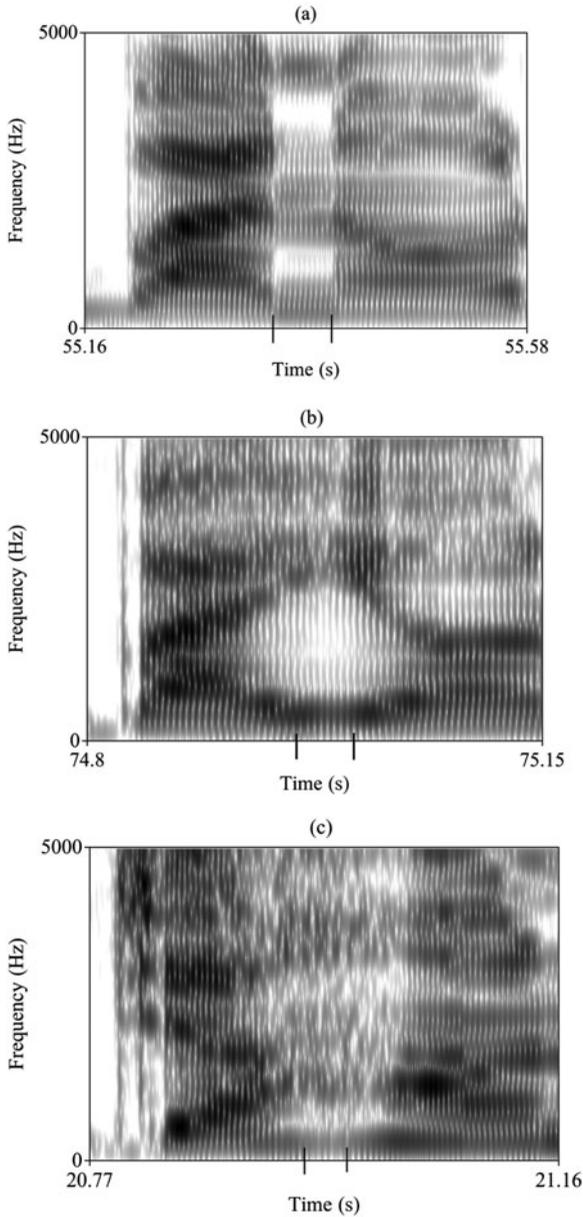


Figure 1 Sample spectrograms, with arrows marking the steady state of the test consonant: (a) */bæna/* 'building', with medial */n/*; (b) */bæjan/* 'explosion', with medial */j/*; (c) */dʒæhan/* 'world', with medial */h/*.

is challenging to determine a boundary between such a reduced laryngeal and a neighboring vowel, since the transition in voice quality and intensity is gradual (Umeda 1977: 846).

Kawahara (2007) demonstrated the perceptual challenge of identifying consonant length in more sonorous consonant classes, in an experiment with native speakers of Egyptian Arabic, a language in which the length contrast extends to all consonant classes. Participants heard stimuli in which the duration of the constriction interval of an intervocalic consonant was varied step-wise, and they had to identify the medial consonant as long or short. Kawahara

found that the participants' reaction times were significantly longer when the test consonant was a glide than when it was a liquid, and longer for sonorant consonants than for obstruent ones. He also found that the slope of the transition region of the identification curve was steeper for obstruents than for sonorants, and for liquids compared to glides. He concluded that the length contrast in Egyptian Arabic is less perceptible in glides than in liquids, and less perceptible in liquids than in nasals or obstruents.

Kawahara & Pangilinan (to appear) have explored the bases of these perceptibility differences in non-speech analogues. They synthesized sine wave complexes with 50 sine wave components, and a medial interval of silence, noise or a lower-amplitude sine wave complex with the same spectral specifications as the surrounding intervals. They found that in sine waves with a medial low-intensity interval (analogous to a medial consonant), discrimination and identification of differences in the duration of that interval were more accurate when the difference in intensity between medial low-intensity interval and the surrounding high-intensity intervals was greater. They also found that differences in the duration of the medial 'consonant' interval were more accurately detected when that interval consisted of silence or noise than when it consisted of the low-amplitude periodic interval. The series of experiments taken together provides evidence that length contrasts are harder to perceive the more acoustically similar the contrasting element is to the surrounding elements.

Such perceptual difficulties could lead to the observed cross-linguistic avoidance of length contrasts in more vowel-like consonant classes. If a contrast is difficult to perceive, that increases the likelihood that a language learner will fail to master it. Such a language learner would then lack the contrast, and if enough members of a speech community follow that learner, the contrast would be lost in the language (i.e. the contrasting elements would merge). As suggested by Ohala (1981, 1983, 1993), the tendency for listeners to systematically misperceive one sound (or sound sequence) can be seen as the basis for sound change and for recurring gaps in segment inventories

The range of consonant length contrasts in Persian provides an opportunity to test this claim, and examine how the length contrast differs across different consonant types. We hypothesize that the identification of Persian consonant length categories is more difficult the more vowel-like the consonant is. The more gradual the acoustic transition is between the consonant and the neighboring vowels, in formants, intensity, voice quality and other acoustic parameters, the more indeterminate the boundary is between consonant and vowel, and thus the more indeterminate is the duration of that consonant. Glides are similar to vowels in having prominent formants and in being relatively intense, and the transition between a glide and a neighboring vowel is long. Laryngeals are similar to neighboring vowels in sharing the same formant values, while differing from them in intensity and voice quality. With other classes of consonants, such as nasals and obstruents, the acoustic boundary between consonant and neighboring vowel is more distinct, which eases the perception of length contrasts in those consonants. We therefore expect the identification of consonant length in Persian to be less precise in glides and laryngeals than in other consonant classes.

We first consider how the length contrast is acoustically realized across consonant classes and speaking rates in Persian, presenting results from a production study. We then present the results of two perception studies investigating how the identification of consonant length classes varies across consonant classes, and how the duration of the transition between the consonant and the neighboring vowel affects that identification.

1 Experiment 1: Acoustic differences between singleton and geminate consonants in Persian

A production experiment was run to determine how the acoustic differences between long and short consonants varies according to manner of articulation class in Persian. The consonant classes included plosives, fricatives, nasals, liquids, glides and laryngeals.

Table 2 Test words by consonant length and consonant manner. The test consonant is boldfaced.

Segment type	Singleton	Geminate
Voiced palatal glide	/bæjan/ 'exposition'	/χæjjam/ (surname)
Voiceless laryngeal continuant	/dʒæhan/ 'world'	/sæhhaf/ 'book-binder'
Voiceless glottal stop	/fæʔal/ (see below)	/fæʔʔal/ 'active thing'
Voiced alveolar lateral approximant	/bæla/ 'haughty one'	/dællal/ 'laborer'
Voiced alveolar nasal stop	/bæna/ 'building'	/bænna/ 'builder'
Voiced alveolar fricative	/Gæza/ 'food'	/bæzzaz/ 'rug merchant'
Voiced dental plosive	/fæda/ 'sacrifice'	/mæddah/ 'eulogist'

Like other duration measures such as vowel duration or VOT, the duration of a consonant is less at faster speaking rates than at slower ones (Arvaniti 1999, Pickett, Blumstein & Burton 1999, Hirata & Whiton 2005). Hansen (2004) found that the consonant length contrast in Persian is maintained across speaking rates, but the difference in duration between singleton and geminate consonants is smaller at faster rates than at slower rates. The duration of geminate consonants is more strongly affected by variation in speaking rate than the duration of singletons is. Due to this interaction of the length contrast with speaking rate, variation in speaking rate was included in the production experiment, to test whether the acoustic differences between geminates and singletons hold across different rate conditions.

1.1 Method

1.1.1 Participants

Four adult native speakers of Persian participated in the study. Two were female (F1, F2) and two male (M1, M2). At the time of the study, all participants were residing in the United States, and their ages ranged from 38 to 44 years. All grew up in Tehran, stayed there at least through secondary school, and continue to speak the language on a daily basis. All four were fluent speakers of English as well.

1.1.2 Materials

The test consonants were the long and short counterparts of the consonants /d z n l ʔ h j/. These consonants were chosen to represent the laryngeals, the glides, and the other major manner of articulation classes: plosives, fricatives, nasals and liquids. Each consonant occurred in medial position in a two-syllable Persian noun between a preceding low front vowel /æ/ and a following low back vowel /a/, with main word stress on the second of those vowels. The 14 test words are listed in Table 2 above.

The test words were produced in the constant carrier frame /minu ___-ra did/ 'Minu saw ___'. In some cases, the object of the sentence was not a visible object (e.g. 'sacrifice'). In these cases, speakers were instructed to imagine that it was the orthographic representation of the test word that Minu saw.

The item /fæʔal/ is a conventional reference form for the Arabic *CaCaaC* verb template. This item was not familiar to all speakers, and so was treated by at least some of them as a nonsense word. They did not have any trouble reconstructing from the orthography how the word was supposed to be pronounced.

1.1.3 Procedure

The sentences were presented to the subjects in Persian script on sheets of paper, with filler items at the beginning and end of each page. There were 12 sheets, each with all 14 sentences in a different randomized order. Participants were asked to read the whole set of 168 sentences (12 sheets × 14 sentences) without sentence-internal pauses, at a normal, relaxed rate. Then

they were asked to read the same list of items again at a faster rate, though still speaking clearly enough to be understood. Finally, they were asked to read the sentences a third time speaking as fast as possible, without worrying about whether they were understood. Each participant thus produced 168 sentences in three rate conditions (*baseline/faster/fastest*), for a total of 504 sentences per participant. The distance from the speaker's mouth to the microphone was maintained at 6–12 inches. The readings were recorded in a sound-treated recording booth on a solid-state digital recorder at a sampling rate of 22,050 Hz at 16-bit amplitude precision.

1.1.4 Measurements

The duration of the following intervals was measured, using the Kay Elemetrics MultiSpeech acoustic analysis program: the consonant constriction, the formant transition from the preceding vowel to the test consonant, and the sentence frame excluding the test word.

The onset of the constriction was the end of the decline in intensity from the vowel, at the point at which the waveform attained its local minimum of intensity and complexity. The offset of the constriction was the end of that interval of low intensity and low wave complexity. In the glide /j/, where there often was no such clear low-point in intensity, the onset of the constriction was placed at the point at which F2 reached its local maximum and F1 its local minimum (whichever event took place last), and the offset was the end of the interval with those F1 and F2 values.

The onset of the formant transition was the onset of the movement in F1 or F2 (whichever changed first) from the steady state values for the preceding vowel towards those for the test consonant. The offset of the formant transition was the onset of the constriction. No formant transition duration could be measured for the laryngeals /ʔ h/ since they share the formant values of the neighboring vowels, so that any formant movement in the neighbourhood of the laryngeal reflected a transition from the preceding vowel to the following one (Lehiste 1964, Keating 1988). These consonants are therefore excluded from the analysis for this measure.

The onset of the frame sentence was the onset of acoustic energy at the onset of the utterance, and the offset of the frame was the offset of acoustic energy. The frame duration was the duration of the frame sentence minus the duration of the test word. It was thus the duration of the invariant portion of each sentence: /minu ___-ra did/ 'Minu saw ___'.

In addition, RMS amplitude (dB SPL) was measured, using a 20 ms window centered on one of two measurement locations: the midpoint of the preceding vowel and the midpoint of the constriction interval. Amplitude drop was the difference between these two measurements.

1.1.5 Hypotheses

1.1.5.1 Effects of consonant length

Geminates were expected to have longer constriction durations than corresponding singletons. This is the primary acoustic distinction between geminates and singletons (Obrecht 1965, Lehtonen 1970, Lahiri & Hankamer 1988, Pickett et al. 1999, Aoyama & Reid 2006, Ridouane 2007).

It might also be expected that geminates would have longer formant transitions than singletons, reflecting a lower velocity of articulator movement (and lower gestural stiffness) in producing the geminate, as found in articulatory studies by Smith (1995) and Löfqvist (2005, 2007) for Japanese. Myers & Hansen (2005) found that long vowels have longer formant transitions than short vowels in Finnish, and it would be reasonable to expect a similar difference between long and short consonants.

The drop in RMS amplitude from the preceding vowel might also be greater in geminates than in singletons. In a VCV sequence, undershoot of the consonant target due to overlapping vowel gestures can lead to stops without complete closure or fricatives without noise (Byrd & Tan 1996, Van Son & Pols 1999, Mauk 2003, Warner & Tucker 2011). However, there is greater tongue displacement in the transition from vowel to following consonant when that consonant is a geminate than when it is a singleton (e.g. in Japanese: Löfqvist 2007). This suggests that there might be systematically more complete closure in geminates than in

singletons in Persian, which should be reflected in a greater amplitude drop from the preceding vowel.

1.1.5.2 *Effects of speaking rate*

When a speaker increases his or her speaking rate, all the measurable durations in his or her speech get shorter. It is thus expected that in the three rate conditions (baseline/faster/fastest), faster speaking rate conditions will be associated with shorter durations for the constriction interval, the formant transition, and the frame.

The frame duration provided a further, gradient measure of variation in speaking rate. The words in the frame did not vary from condition to condition, so the duration of the frame material would not be expected to vary according to the target consonant, but rather only according to speaking rate.

Constriction and formant transition duration measurements were expected to be longer when frame duration was longer, since longer frame durations reflect slower speaking rates.

Moreover, Hansen (2004) found an interaction between speaking rate and constriction duration, in that the difference in constriction duration between geminates and singletons in Persian was greatest at slower speaking rates. The effect of speaking rate on constriction duration was greater for geminates than for singletons, as found in other languages by Pickett et al. (1999) and Hirata & Whiton (2005).

Longer frame durations would also be expected to be associated with a greater amplitude drop from the preceding vowel to the test consonant. This is because the constriction undershoot discussed in the previous section would be expected to be greater at faster speaking rates (Gay 1978, 1981).

1.1.5.3 *Effects of manner of articulation*

The central focus of the production study was to investigate how the manner of articulation interacts with the consonant length contrast. Aoyama & Reid (2006) investigated this issue in Guinang Bontok, finding that the ratio of geminate to singleton consonant duration was less for the palatal glide /jj – j/ than for any other class of consonant. They also found the ratios for the fricative /ss – s/ and the approximant /ɹɹ – ɹ/ were significantly lower than for plosives, nasal stops, or lateral approximants. Differences in constriction duration according to manner of articulation reflect the inherent duration of the different manner classes (Klatt 1976), and the limits on how compressible and extendable they are.

It was expected in the present study that the formant transition would be longer for the glide /j/ than for any other category in Persian, since the gradual gliding transition is a distinguishing characteristic of glides (Lieberman et al. 1956). It was also expected that the fricative /z/ would have a longer transition than the plosive /d/, since articulator movements have been found to be slower in fricatives than in plosives (Kuehn & Moll 1976).

The manner of articulation classes under consideration also differ in their inherent amplitude. All else being equal, sonorants are expected to have a higher RMS amplitude than obstruents, and sounds with a greater degree of stricture are expected to have a lower amplitude than those with a lesser degree of stricture (Lavoie 2000, Parker 2008). Such differences might be relevant to the perception of the length contrast in that a smaller amplitude drop from vowel to following consonant could make it more difficult to locate the edges of the consonant, and so increase the difficulty of perceiving length contrasts.

1.1.6 *Statistical analysis*

All statistical tests for experiment 1 were conducted with a mixed-model regression analysis, using the *lme4* package in *R* (Bates, Maechler & Bolker 2012), with degrees of freedom and *p*-values calculated using the *lmerTest* package (Kuznetzova, Brockhoff & Bojesen 2014). The random effects were intercepts for participant and item, and slopes for the interaction of participant with each categorical fixed effect. The fixed effects were Frame duration (a continuous variable, in ms), Length (short/long), Manner (sonorant nonglide/obstruent/laryngeal/glide), and Stop (nonstop/stop). In each categorical independent variable, the

Table 3 Classification of test consonants.

	/d/	/z/	/n/	/l/	/ʔ/	/h/	/j/
Sonorant nonglide	-	-	+	+	-	-	-
Obstruent	+	+	-	-	-	-	-
Laryngeal	-	-	-	-	+	+	-
Glide	-	-	-	-	-	-	+
Stop	+	-	+	-	+	-	-

unitalicized class (the first one) is the default reference level (coded as 0), and the italicized classes are the marked levels (coded as 1). All interactions among the fixed effects Frame duration, Length and Manner were included in the analysis. Interactions involving Stop were not included, since it does not fully cross-classify with Manner.

The classification of test consonants is given in [Table 3](#).

Follow-up analyses of subsets of the data have the same random effects structure, but are limited to one of the seven consonants, and have as fixed effects just Frame duration and Length.

1.2 Results

1.2.1 Constriction duration

Values of constriction duration for the test consonants are presented in [Figure 2](#) below as a function of frame duration for the four major manner of articulation classes. The solid regression line in each plot indicates the trend for geminates, and the dashed regression line marks the trend for singletons.

It can be seen that at any given frame duration (i.e. any given speaking rate) the geminates generally have a longer constriction duration than the singletons. Both geminates and singletons have longer constriction durations at longer frame durations, but the effect is greater for geminates than for singletons, as reflected in the steeper slope of the regression lines for the geminates. As found by Hansen (2004), the difference in constriction duration between singletons and geminates is greater at longer frame durations (i.e. slower speaking rates) than at shorter frame durations (i.e. faster speaking rates): 68 ms in the baseline rate condition, 56 ms in the faster rate condition, and 33 ms in the fastest rate condition.

The slope of the regression lines is very similar across the first three graphs, but is much flatter for the glides than for any of the other categories. The slope of the geminate regression line is .11 for the obstruents, .11 for the sonorant nonglides, .12 for the laryngeals, and .05 for the glides.

The mean constriction duration values for geminates and singletons of each consonant (averaged across speaking rate conditions) are given in [Table 4](#), with the ratio of the geminate mean to the singleton mean for each consonant. In every consonant, the mean constriction duration for the geminate was greater than that for the corresponding singleton. The greatest geminate–singleton ratios were in the (nasal, oral and glottal) stops /d n ʔ/. The smallest were in the continuants /z h j/.

Table 4 Mean constriction duration (ms) by Length and Consonant.

	/d/	/z/	/n/	/l/	/ʔ/	/h/	/j/
Geminate	114	113	113	110	116	131	63
Singleton	44	77	49	55	58	75	36
Ratio of geminate to singleton mean	2.59	1.47	2.31	2.00	2.00	1.75	1.75

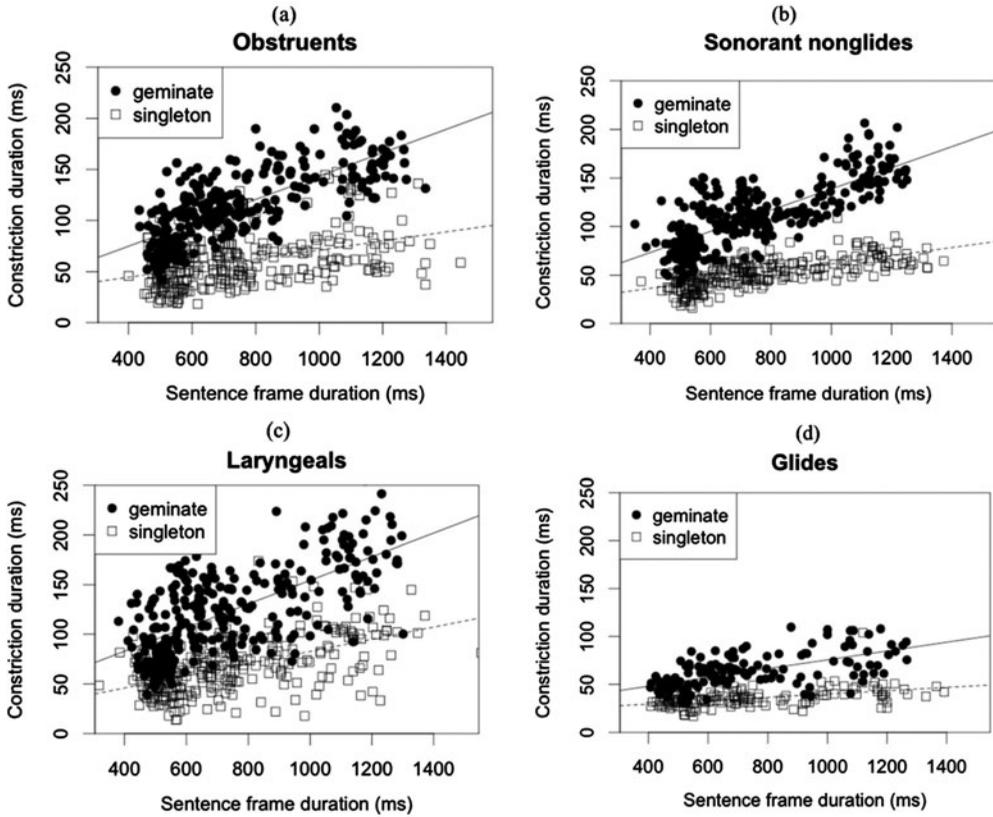


Figure 2 Constriction duration (ms) as a function of sentence frame duration for geminates and singletons in (a) obstruents, (b) sonorant nonglides, (c) laryngeals, and (d) glides.

The results of the mixed-model analysis are given in Table 5. Only significant effects ($p < .05$) are reported in this table. There were significant main effects of Frame duration and Stop. The duration of the constriction interval was significantly longer when the frame duration was longer, i.e. when the speaking rate was slower. The constriction interval was significantly shorter in stops than in continuants.

The interactions with Frame duration reflect the slopes of the lines in Figures 2a–d. The effect of increasing Frame duration on constriction duration was significantly steeper in geminates than in singletons (Frame duration \times Length), and in laryngeals compared to

Table 5 Significant fixed effects in a model of constriction duration.

Fixed effect	β	<i>d.f.</i>	<i>t</i>	<i>p</i>
Intercept	25.53	7.2	4.1	.004
Frame duration (ms)	0.04	1566.0	9.1	< .001
Stop	-12.17	3.2	-3.7	.031
Frame duration \times Length	0.08	1840.0	11.7	< .001
Frame duration \times Manner: Glide	-0.02	1627.0	-2.8	.006
Frame duration \times Manner: Laryngeal	0.02	1887.0	3.9	< .001
Frame duration \times Length \times Manner: Glide	-0.04	1916.0	-3.5	< .001

Table 6 Mean formant transition duration (ms) by length and consonant.

	/d/	/z/	/n/	/l/	/j/
Geminate	39	57	34	32	86
Singleton	38	50	32	33	72
Ratio of geminate to singleton mean	1.02	1.14	1.07	0.98	1.20

Table 7 Significant fixed effects in a model of formant transition duration.

Fixed effect	β	<i>d.f.</i>	<i>t</i>	<i>p</i>
Intercept	20.33	8.1	5.8	< .001
Frame duration (ms)	0.02	1277.0	7.8	< .001
Manner: Obstruent	20.75	1387.0	6.8	< .001
Manner: Glide	28.73	1378.0	8.0	< .001
Stop	-7.02	4.2	-5.7	.004
Frame duration \times Manner: Obstruent	-0.01	1336.0	-3.4	< .001
Frame duration \times Length \times Manner: Glide	0.03	1353.0	4.0	< .001

nonlaryngeals (Frame duration \times Manner: Laryngeal). On the other hand, that slope was flatter in glides than in nonglides (Frame duration \times Manner: Glide), and in particular in geminate glides (Frame duration \times Length \times Manner: Glide).

There was no significant main effect of Length. The effect of the length of the consonant was only reflected in the interactions of Frame duration with Length and Frame duration with Length and Manner: Glide. This reflects the fact that the difference between geminate and singleton was dependent on speaking rate, becoming greater at slower speaking rates.

Considering each manner of articulation class individually, mixed-model analyses were run with Frame duration and Length as the only fixed effects. In each of the manner classes, there was a significant main effect of Frame duration and a significant interaction of Frame duration and Length, as in the overall analysis. For /n/, there was also a significant main effect for Length, which was not the case for any of the other manner classes.

1.2.2 Formant transition duration

Table 6 gives the mean formant transition duration broken down by length class and consonant. The laryngeals /ʔ h/ are excluded since they had no formant transitions from the preceding vowel. The mean formant transition was not much longer in geminates than in singletons for any consonant, and in the case of /l/ the singletons had a slightly longer transition interval. The nonliquid continuants /z j/ had longer formant transitions and a greater geminate–singleton ratio than the stops /d n/.

The transition duration was greater at slower speaking rates than at faster ones. The mean transition duration was 50 ms for the baseline rate condition, 46 ms for the faster condition, and 38 ms for the fastest condition. As with the constriction duration, the difference in transition duration between geminates and singletons was greater at slower speaking rates: 6.8 ms in the baseline condition, 6.1 ms in the faster condition, and 1.6 ms in the fastest condition.

The results of a test of these effects are given in **Table 7**. Only significant effects are included in the table. There were significant main effects of Frame duration, Manner: Obstruent, Manner: Glide, and Stop. As with constriction duration, the formant transition duration was significantly longer when the frame duration was longer, i.e. when the speaking rate was slower. The formant transition was longer in obstruents than in sonorants, in glides compared to nonglides, and in continuants compared to stops.

There was no significant main effect of Length, and no interaction of Frame duration with Length. The only effect of Length was in the interaction Frame duration \times Length \times

Table 8 Mean RMS amplitude drop (dB) between vowel and following consonant by length and consonant.

	/d/	/z/	/n/	/l/	/ʔ/	/h/	/j/
Geminate	21	12	3	2	27	16	4
Singleton	10	11	2	2	10	8	2
Ratio of geminate to singleton mean	2.20	1.09	1.12	1.10	2.75	2.03	2.15

Table 9 Significant fixed effects in a model of the vowel–consonant amplitude drop.

Fixed effect	β	<i>d.f.</i>	<i>t</i>	<i>p</i>
Length	-3.92	39.1	-2.1	.041
Manner: Obstruent	6.25	10.1	2.7	.022
Stop	3.57	3.2	3.4	.040
Frame duration \times Length	0.01	1813.0	2.5	.012
Frame duration \times Manner: Glide	0.01	1896.0	2.6	.010
Frame duration \times Manner: Laryngeal	0.01	1762.0	6.0	< .001
Length \times Manner: Glide	6.81	1931.0	2.6	.010
Length \times Manner: Laryngeal	9.62	1938.0	4.4	< .001
Length \times Manner: Obstruent	4.95	1957.0	2.2	.028

Manner: Glide, which expressed a tendency for long glides to have longer formant transitions when the frame duration was greater, i.e. when speaking rate was slower.

Examining the effects of Frame duration and Length for each consonant, there was a significant interaction of Frame duration and Length for the obstruents /d z/ and the glide /j/, but not for the sonorant nonglides /n l/, which also had no main effect of Length.

1.2.3 Amplitude drop

Table 8 presents the mean RMS amplitude drop broken down by length and consonant. In all consonants, the mean amplitude drop was greater in the geminate than in the corresponding singleton, suggesting that the consonantal constriction was more complete when the constriction interval was longer. The amplitude drop was greater in consonants with lower inherent amplitude, so that the obstruent stop /d/ had a greater drop than the fricative /z/ and both obstruents had a greater drop than the sonorants /n l j/. The laryngeals /ʔ h/ pattern in this regard with the low-amplitude obstruents.

The amplitude drop was greater at slower speaking rates. The mean drop was 10.4 dB in the baseline rate condition, 9.5 dB in the faster rate condition, and 7.7 dB in the fastest rate condition. The difference in amplitude drop between geminates and singletons was also greater at the slower speaking rates, with a mean difference of 6.7 dB in the baseline rate condition, 7.2 dB in the faster rate, and 3.6 dB in the fastest rate.

The results of a test of these effects are given in Table 9. There were significant main effects for Manner: Obstruent and Stop, reflecting a significantly greater amplitude drop for the obstruents than for sonorants, and for stops compared to continuants.

The effects of consonant length on amplitude drop were complex. The interaction Frame duration \times Length indicates that the difference in amplitude drop between geminates and singletons gets greater with greater values of frame duration, i.e. at slower speaking rates. The interactions Length \times Manner: Glide, Length \times Manner: Laryngeal and Length \times Manner: Obstruent indicate that for all classes except for sonorant nonglides, geminates have a significantly greater amplitude drop than singletons. The main effect of Length, with its negative coefficient, reflects then the fact that in the default class of sonorant nonglides, there is no such difference between geminates and singletons.

The drop in amplitude was more rate-dependent in the laryngeals than in nonlaryngeals (Frame duration \times Manner: Laryngeal), and in glides compared to nonglides (Frame duration

Table 10 Significant effects by consonant type.

	/d/	/z/	/n/	/l/	/ʀ/	/h/	/j/
Frame duration			*		*	*	*
Length		*	*	*		*	*
Frame duration × Length	*		*	*	*		

× Manner: Glide). This suggests that these two categories showed more rate-dependent reduction than the other categories.

Follow-up analyses were made of the individual consonants, with only Frame duration and Length as fixed effects. The results were mixed, and are summarized in Table 10, which indicates which factors had significant effects for each consonant type. The main point to be drawn from these results is that Length had an effect in all consonant classes, either as a main effect or in interaction with Frame duration.

1.2.4 Summary of production study

As hypothesized, Persian geminates were found in this study to differ from singletons in constriction duration and in the intensity drop from the preceding vowel. In both cases, this effect was dependent on speaking rate, such that the difference between singleton and geminate was greater at slower speaking rates than at faster ones. The difference between geminates and singletons was not a simple additive effect, as it would be if the geminates had some fixed increment of duration above that of the singletons.

The hypothesized effect of length on the formant transition was not found. Geminates did not in general have longer formant transitions than singletons, either in absolute terms or relative to speaking rate. The one exception was in the glides, for which the effect of length was dependent on speaking rate.

Speaking rate had an effect on all three measurements, with longer frame duration intervals (i.e. slower speaking rates) associated with longer constriction duration, longer formant transition duration, and greater intensity drop from the preceding vowel. But the effect of speaking rate varied according to consonant length. There were stronger effects of speaking rate for geminates than for singletons, as reflected in the interaction of Frame duration and Length.

The manner of articulation classes differed in all three measurements. Continuants had significantly longer constriction intervals than stops. Laryngeals had significantly longer constriction intervals than nonlaryngeals, relative to speaking rate, and glides had significantly shorter constriction intervals compared to nonglides. Continuants had a significantly longer formant transition interval than stops, obstruents had a longer interval than non-obstruents, and glides had a longer interval than nonglides. Stops had a significantly greater intensity drop from the preceding vowel than continuants, and obstruents had a greater drop than sonorants.

Since manner of articulation affects constriction duration, formant transition duration and intensity drop, all of which also differ between geminates and singletons, one could expect manner of articulation to affect the perceptibility of consonant length. The next two experiments investigate whether this is the case in Persian.

2 Experiment 2: Effect of manner of articulation on length class identification

A perception experiment was designed to test how the differences among the consonant types that were seen in the production experiment are reflected in the identification of geminate consonants by Persian-speaking listeners. The overarching hypothesis is that the

perceptual boundaries for length identification will be less sharply defined for the more vowel-like consonant types (glides and laryngeals), since they have more gradual transitions to neighboring vowels.

2.1 Method

2.1.1 Participants

The participants in this study were 5 female and 5 male adult native speakers of Persian who had grown up in Tehran and stayed there through at least high school. All participants were residing in the U.S. at the time of the experiment, but continued to use Persian regularly in family interactions. Their ages ranged from 30 to 60 years at the time of the study. One male and one female participant in this experiment had also participated in the production study. One male participant failed to complete the experiment, so his responses were not included in the analysis.

2.1.2 Stimuli

The stimuli were nonsense words varying in constriction duration of a medial consonant, the primary acoustic correlate of the length contrast. A female speaker of Tehran Persian produced nonsense words of the form /gæC:ab/, where C: belonged to the set /dd zz nn ll jj ?? hh/. Each nonsense word was produced in the constant sentence frame /minu ___ra did/ 'Minu saw ___' (as in Experiment 1). A representative token of each nonsense word was excised from its sentence context, and resynthesized using the pitch-synchronous LPC synthesis tools in Multi-Speech Analysis-Synthesis Laboratory (Kay Pentax), with the sampling rate set at 11,025 Hz, filter order at 36, and pre-emphasis at 0.5. Cell intervals were set to the voiced-period marks where voicing was present, and to 15 ms elsewhere. The duration of the constriction interval of the medial consonant was varied by deleting or copying cells in the analysis window, and then resynthesizing the resulting file. Each continuum included 8 stimuli, with successive members of the same continuum differing by as close to 10 ms as was feasible through deletion or copying of whole glottal pulses. Each continuum covered a range of about 70 ms centered around the cross-over point in identification from singleton to geminate for each consonant, as determined in a series of preliminary perception tests. The ranges for each class are shown in Table 11.

Table 11 Ranges for stimulus constriction duration (ms).

	/d/	/z/	/n/	/l/	/ʔ/	/h/	/j/
Maximum constriction duration	120	152	150	140	148	156	102
Minimum constriction duration	51	82	81	71	71	80	28

In order to facilitate comparison across continua with different medial consonants, the duration of the preceding and following vowels were held constant. The preceding vowel was set in all tokens to a duration as close to 100 ms as was possible through the insertion and deletion of pulses in the analysis window, and the following vowel was set as close as possible to 170 ms. The vowel–consonant transition was kept as it was in the precursor recording in all tokens.

Each continuum was tested in a pilot study to insure that the perceptual boundary between geminate and singleton was included in the range of the continuum for each medial consonant. The perceptual boundary differed according to consonant, and the hypothesis is only concerned with the transition region in the identification function in which increases in duration are associated with increases in geminate responses. This is why the range of constriction durations differed from consonant to consonant.

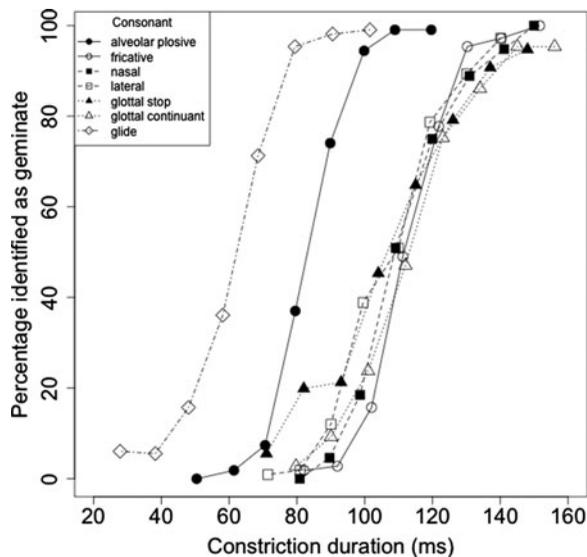


Figure 3 Percentage of geminate responses as a function of the medial consonant and its constriction duration (ms).

2.1.3 Procedure

The stimuli were blocked by consonant and presented in randomized order within block on a computer using SuperLab (Cedrus Corporation). Subjects were asked to press a button on a response pad to indicate whether they heard the medial consonant as the geminate, represented in Persian orthography with the length (*tashdid*) symbol, or as the singleton, marked by the same consonant symbol without the *tashdid* marker.

Seven test words with different medial consonants were included in the study, and each test word was the precursor for eight stimuli differing just in medial consonant constriction duration. Each item was heard by each subject 12 times. The total number of stimuli per subject was thus $7 \times 8 \times 12 = 672$.

2.2 Results

Figure 3 above represents the percentage of geminate responses for each stimulus as a function of the medial consonant and its constriction duration. In general, the percentage of geminate responses rises from close to 0% for short constriction durations to close to 100% for long constriction durations, confirming that constriction duration is an important perceptual cue for consonant length identification (Garnes 1976, Pind 1986, Hankamer, Lahiri & Koreman 1989, Esposito & Di Benedetto 1999).

The curves for the different consonants in Figure 3 differ in the constriction duration value at which they attain the 50% geminate identification level. The leftmost two curves represent the stimuli with medial /j/ (hollow diamonds) and /d/ (filled circles), the two consonants for which the 50% crossover point was well below 100 ms. These two consonants were the ones with the shortest overall mean constriction duration in the production study.

In order to model the relation between constriction duration and geminate response in each consonant class, each subject's set of responses for each stimulus continuum was modelled as a cumulative Gaussian distribution (McKee, Klein & Tiller 1985, Wichmann & Hill 2001a, b), using the *nls* package in R (Bates & DebRoy 2007). A psychometric function was fit to the proportion of geminate responses for each subject for each stimulus. One important value of these models is the THRESHOLD, which represents the crossover point in the constriction duration continuum from singleton judgements to geminate judgements.

Table 12 Mean threshold and breadth by consonant (Experiment 2).

	/d/	/z/	/n/	/l/	/ʔ/	/h/	/j/
Threshold	83	113	111	108	83	113	62
Breadth	7	6	10	12	21	17	9

Table 13 Significant fixed effects for a model of threshold (Experiment 2).

Fixed effect	β	<i>d.f.</i>	<i>t</i>	<i>p</i>
Intercept	114.55	9.2	32.2	< .001
Manner: Obstruent	-11.21	21.5	-3.7	.001
Manner: Glide	-52.62	20.7	-13.3	< .001
Stop	-10.89	45.2	-4.6	< .001

Another important model measure is the BREADTH (also known as the width or spread), which is the standard deviation of the function, and represents the steepness of the crossover portion of the identification function. Lower values of breadth represent a steeper slope, closer to a situation in which subjects came to a classification such that 100% of the stimuli with a constriction duration below the threshold were classified as singleton, and 100% of the remaining stimuli were classified as geminate. A higher breadth value, on the other hand, indicates a more gradual slope and a longer interval of the continuum in which the same stimulus was classified sometimes one way and sometimes the other way, by the same subject.

The mean values for threshold and breadth for each consonant class, pooled across participants, are given in Table 12. The threshold value of the model reflected the inherent duration of the different consonant classes. The highest threshold values were for /h/ and /z/, which were also the consonant types with the greatest mean constriction duration in the production study. The lowest threshold value was for /j/, which also had the lowest mean constriction duration of the segments in the production study.

A mixed-model analysis of threshold was constructed. The dependent variable was the threshold value for each subject for each consonant. The random effects were Subject and the interaction of Subject with each fixed effect. The fixed effects were the manner of articulation classes used in the analysis of Experiment 1: Manner and Stop. The results are given in Table 13, which includes just significant effects.

The value for the intercept, 103.34 ms, represents a baseline threshold across categories. The obstruents, glides and stops had values significantly below that threshold, distinguishing them from the laryngeals and sonorant nonglides. The differences indicate that listeners adjust the perceptual boundary between geminate and singleton to reflect the differences in inherent duration among the consonant classes.

Turning to the mean breadth values in Table 12, the highest values were for the laryngeals, representing a shallow slope to the identification function and a broad zone of equivocation in identification. The lowest values were for the obstruents. Surprisingly, the glide /j/ had a lower breadth value than /l/ or /n/, contrary to our expectation that length identification would be more difficult in the glide than in the nonglides.

Results of a mixed-model analysis, with the same structure as that for threshold, are given in Table 14. The laryngeals had a significantly greater mean breadth value than the nonlaryngeal consonants, while the obstruents had a significantly smaller mean breadth value than the sonorants. This indicates that listeners had a less sharply defined perceptual boundary between geminates and singletons for laryngeals, and a more sharply defined boundary for obstruents.

Unexpectedly, glides were not found to have a greater breadth than other consonants. One reason for this might be the important role of the formant transition duration for glides, which

Table 14 Significant fixed effects for a model of breadth (Experiment 2).

Fixed effect	β	<i>d.f.</i>	<i>t</i>	<i>p</i>
Intercept	10.42	8.0	5.6	< .001
Manner: Obstruent	-4.16	9.3	-2.3	.047
Manner: Laryngeal	7.83	8.0	2.8	.023

had by a considerable margin the longest mean transition durations of any of the consonant classes (Table 6 above), and significantly longer formant transitions in geminates compared to singletons. Formant transitions were held constant in the stimuli in this experiment, so if this is an important cue for the length contrast for glides, this might have led to less variation in response to glide stimuli than would be the case if the transitions were varying in a more natural manner. The next experiment was designed to specifically test the effect of such variation in formant duration on the perception of length in glides.

3 Experiment 3: Effect of formant transition duration on length identification in glides

Myers & Hansen (2005) presented Finnish-speaking listeners with stimuli including a glide–vowel sequence in which the vowel formant steady state and the glide–vowel formant transition varied in duration. Subjects had the task of identifying the vowel as short or long. Vowels with longer formant steady-state intervals were identified as long more often than those with shorter steady-state intervals, and vowels with longer transitions from the preceding glide were identified as long more often than those with shorter glide–vowel transition intervals. The authors concluded that listeners counted the formant transition as a weighted component of the vowel duration in determining whether the vowel was long or short.

The duration of the formant transition also played a role in Experiment 1 above, where it was found that glides had a longer formant transition in geminates than in singletons, with the difference being greater at slower speaking rates. One can wonder, then, if variation in the duration of the formant transition between a glide and a neighboring vowel in Persian would affect the perception of consonant length the same way it was found to affect the perception of vowel length in Finnish in the Myers & Hansen (2005) study cited above. If the two cases are indeed parallel, then we would expect that longer formant transitions between a glide and neighboring vowels would increase the likelihood that the glide will be identified as long. Moreover, if, in accordance with our initial hypotheses, longer transitions blur the boundary between neighboring segments, we should expect that longer transitions would be associated with greater uncertainty in length identification.

3.1 Method

3.1.1 Participants

Sixteen adult native speakers of Persian participated in the study: 8 female and 8 male. Their ages ranged from 30 to 65 years at the time of the study. All were literate in standard literary Persian as used in Tehran, and all were living in the U.S. at the time of the experiment. Eight of the subjects had previously participated in Experiment 2. One of the male subjects failed to complete the experiment due to an equipment failure. Data from two other male subjects were eliminated from the analysis: one because it turned out he had hearing deficits, and one because it became clear in the post-experiment debriefing that he had not understood the instructions. These exclusions left data from 13 subjects for analysis.

Table 15 Stimulus formant transition duration (ms) (Experiment 3).

Stimulus series	Duration of the [æ] to [j] transition (ms)	Duration of the [j] to [ɑ] transition (ms)
Short	57	88
Medium	78	108
Long	98	128

3.1.2 Stimuli

The stimuli were all based on a production of the nonsense word [gæjjab] drawn from the materials for Experiment 2. In all stimuli, the duration of the formant steady state for each vowel was maintained at the value in the precursor: 40 ms for [æ] and 209 ms for [ɑ].

The duration of three intervals was manipulated using the ASL synthesis package in Kay Elemetrics Multi-Speech: the steady state interval of the glide, the transition from the preceding [æ] to [j] and the transition from the glide to the following [ɑ]. The durations of both the VC and the CV transitions were manipulated in the same way, since it was observed in the production study materials that the durations of the two transitions were correlated, with longer VC transitions into the consonant tending to co-occur with longer CV transitions out of the same consonant.

The duration of the steady state interval of the glide was varied in 10 equally-spaced steps ranging from 20 to 110 ms. The durations of the interval between the vowel steady state and the glide steady state were set to the values in Table 15. There were three categories of stimuli with regard to duration of the transition between the vowel steady states and the glide steady state. The Medium formant transition stimuli preserved the formant transition duration of the original precursor recording. The Short transition stimuli had VC and CV transitions 20 ms shorter than in the precursor, and the Long transition stimuli had transitions 20 ms longer than in the precursor. The formant values between the vowel steady state and the glide steady state were filled in by straight-line interpolation.

3.1.3 Procedure

There were 30 stimuli in total (10 duration steps for the glide steady-state in three transition duration groups), and each stimulus was presented 12 times in a randomized order, for a total of 360 trials per subject. The items were presented over headphones on a computer using Superlab (Cedrus Corporation). Subjects were asked to judge whether the medial consonant was long or short, i.e. whether it should be written with the *tashdid* length marker or not. They indicated their response by pressing one of the two buttons on a response pad.

3.1.4 Hypotheses

If listeners take both constriction duration and formant transition duration into account in determining the length class of a glide, the perceptual boundary between singleton and geminate would be expected to be at a shorter constriction duration when the formant transition was longer (i.e. the model threshold would be lower).

If longer transitions lead to greater indeterminacy in length class identification, then the model breadth would be expected to be greater when the formant transition was longer.

3.2 Results

In Figure 4 below, the percentages of geminate responses (pooled across subjects) are presented as a function of the constriction duration and the transition duration. The regression lines for each transition duration condition represent the values predicted for each condition based on a cumulative Gaussian model for the data pooled across subjects.

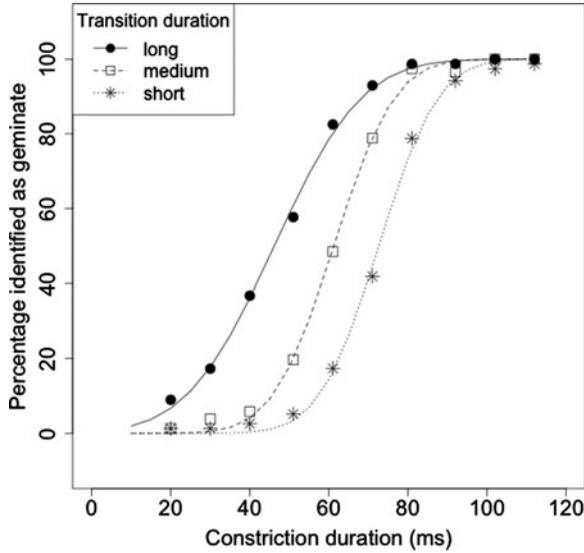


Figure 4 Percentage of geminate responses as a function of constriction and transition duration (ms).

It can be seen that stimuli with longer constriction durations are identified more often as geminates than those with shorter constriction durations. Moreover, stimuli with longer formant transition values were more often identified as geminates than those with shorter formant transitions, all else being equal. The overall proportion of geminate responses was greater for longer formant transition classes: .42 for Short, .53 for Medium, and .67 for Long.

As in Experiment 2, a cumulative Gaussian distribution model was fit to the responses, using the *nls* package. The pooled means for threshold and breadth are given in Table 16.

The threshold values reflect the crossover point in constriction duration from a singleton response to a geminate response. As seen in in Table 16 and Figure 4, the threshold was at higher constriction duration values for stimuli with shorter formant transitions. The results of a mixed-model analysis with subject as random effect and transition class as a fixed effect are given in Table 17. The medium transition class is treated as the default.

The threshold value was significantly higher for stimuli with short formant transitions compared to the other stimuli, and significantly lower for stimuli with long formant transitions

Table 16 Pooled means (in ms) for threshold and breadth by transition duration class (Experiment 3).

	Short transition	Medium transition	Long transition
Threshold	72	61	46
Breadth	9	10	14

Table 17 Significant fixed effects for a model of threshold (Experiment 3).

Fixed effect	β	<i>d.f.</i>	<i>t</i>	<i>p</i>
(Intercept)	61.1	18.1	28.5	< .001
Transition class: Short	11.2	24.0	6.8	< .001
Transition class: Long	-15.0	24.0	-9.1	< .001

compared to the others. In post-hoc pairwise comparisons, each pair of transition classes was significantly different in threshold. This indicates that there was a trade-off between the cues of constriction duration and transition duration, such that a stimulus with a longer formant transition did not require as long a constriction duration to be identified as geminate as a stimulus with a shorter formant transition. Both constriction duration and formant transition duration play a role in glide length classification in Persian.

The breadth values reflect the breadth of the transition zone in the identification curve. The mean breadth values in Table 16 above are greater for longer formant transition classes, and greater breadth values reflect greater equivocation in classification. These differences were tested in a mixed model analysis with the same structure as the last one, the results of which are given in Table 18.

Table 18 Significant fixed effects for a model of breadth (Experiment 3).

Fixed effect	β	<i>d.f.</i>	<i>t</i>	<i>p</i>
(Intercept)	10.4	28.8	10.2	< .001
Transition class: Long	3.5	24.0	3.0	.006

The breadth value for stimuli from the long transition class was significantly greater than those from the other two classes, but the short transition stimuli were not significantly different from the other classes. In a post-hoc pairwise comparison, the breadth was significantly greater for the Long class compared to either the Short or the Medium class, but there was no significant difference between the Short and the Medium class. These results indicate that the duration of the formant transition did affect the slope of the identification function, with the longest formant transitions associated with a shallower identification slope.

4 Conclusion

The production study has provided evidence concerning the acoustic differences between long and short consonants in Persian. Geminate had a longer constriction interval than singletons across speaking rates and consonant classes. Other acoustic differences between geminates and singletons depended on the consonant class. The formant transition was significantly longer in geminates than in singletons just in the glide /j/. The drop in intensity from the preceding vowel was significantly greater in geminates than in singletons for glides, obstruents and laryngeals, but not for sonorant nonglides.

The realization of the consonant length contrast was sensitive to variation in speaking rate. The differences between geminate and singleton in constriction duration, transition duration, and intensity drop were greater at longer frame duration values than at shorter ones. The effect of speaking rate on these acoustic variables was greater for geminates than for singletons.

The production study also provided measures of how similar the different consonant classes of Persian consonants were to the neighboring vowels. In terms of the intensity profile in time, the glides and nonglide sonorants showed a significantly smaller drop in intensity relative to the preceding vowel than the obstruents or laryngeals. In terms of spectral continuity, the longest formant transitions were those from the vowel to a following glide, and the smallest consonant effect on formant trajectories was in the laryngeals. Obstruents were the least vowel-like consonants in both dimensions.

Experiments 2 and 3 demonstrated that constriction duration is a perceptual cue for consonant length, with listeners identifying consonants with longer constriction durations as geminates. But while the acoustic differences between geminate and singleton are significant in all consonant classes, those classes differ in how easily listeners classify them into long and short categories. In Experiment 2, the breadth of the identification function for obstruents was

lower than for any of the other consonant types, and that for laryngeals was greater than for the other consonant types. This indicates that the perceptual boundary between geminates and singletons was most sharply defined for obstruents, and least sharply defined for laryngeals.

These results supported the general claim that more vowel-like consonant types would be harder to classify into length categories, although the fact that glides were not more difficult to classify than other sonorant consonants runs counter to this trend. It was suggested that this was due to the role of formant transition variation in the perception of consonant length, since glides had the longest formant transitions of any consonant class in the production study, and the greatest difference between geminate and singleton in formant transition duration.

Experiment 3 addressed the claim that formant transition duration contributes to the perception of glide length. It was found that intervocalic glides with longer formant transitions were more likely to be identified as long than glides with shorter formant transitions. This supported the hypothesis that listeners consider formant transition duration in making their decision as to whether a glide is geminate or singleton. It was also found that glides with longer formant transitions had a less steep identification curve than glides with shorter formant transitions, supporting the claim that longer transitions make length identification more difficult.

These results have provided evidence for the general claim that more vowel-like consonant types are more difficult to classify into length classes than less vowel-like ones, because the more vowel-like consonant types have more gradual transitions to neighboring vowels that blur the acoustic boundary between consonant and vowel (Kawahara 2007, Kawahara & Pangilinan to appear). In the case of glides, the relevant transition is a formant transition, while in the case of laryngeals the relevant transition is one in voice quality and intensity. These challenges could lead to a tendency for learners to fail to successfully learn length contrasts in glides and laryngeals, which in turn could account for the cross-linguistic avoidance of long glides and laryngeals noted by such authors as Podesva (2000), Kawahara (2007) and Maddieson (2008).

The ambiguity of these transitions also affects the perception of the length of the vowel. Kavitskaya (2002) suggests that the same kind of segmentation difficulty in vowel + glide and vowel + laryngeal sequences leads to lengthening of the vowel in classic compensatory lengthening. Myers & Hansen (2005) argue that such ambiguity in sequences of a vocoid and a vowel leads to a tendency for listeners to identify vowels in such a context as long, and a tendency for languages to require vowels to be long in that context.

Acknowledgements

The experiments reported on here were designed and conducted by the first author for his PhD dissertation (Hansen 2012), which was supervised by the second author. The authors would like to thank the following people for their help: Björn Lindblom, Harvey Sussman, Bob King, Megan Crowhurst, Golnaz Modarresi Ghavami, three anonymous reviewers for *JIPA*, and the very patient Persian-speaking participants in the studies.

References

- Andrews, J. Richard. 1975. *Introduction to Classical Nahuatl*. Austin, TX: University of Texas Press.
- Aoyama, Katsura & Lawrence Reid. 2006. Cross-linguistic tendencies and durational contrasts in geminate consonants: An examination of Guinang Bontok geminates. *Journal of the International Phonetic Association* 36(2), 145–157.
- Arvaniti, Amalia. 1999. Effects of speaking rate on the timing of single and geminate sonorants. *14th International Congress of Phonetic Sciences*, Berkeley, University of California (ICPhS XIV), vol. 1, 599–602.

- Ashton, E. O., E. M. K. Mulira, E. G. M. Ndawula & A. N. Tucker. 1954. *A Luganda grammar*. London: Longmans.
- Bates, Douglas & Saikat DebRoy. 2007. nls: Nonlinear Least Squares. <http://www.r-project.org/> (accessed 21 August 2008).
- Bates, Douglas, Martin Maechler & Ben Bolker. 2012. *lme4*: Linear mixed-effects models using Eigen and Eigen++ (version 0.999999-0). <http://CRAN.R-project.org/package=lme4> (accessed 9 July 2013).
- Bijankhan, Mahmood & Mandana Nourbakhsh. 2009. Voice onset time in Persian initial and intervocalic stop production. *Journal of the International Phonetic Association* 39(3), 335–364.
- Burrow, T. & S. Bhattacharya. 1970. *The Pengo language*. Oxford: Oxford University Press.
- Byrd, Dani & Cheng Cheng Tan. 1996. Saying consonant clusters quickly. *Journal of Phonetics* 24(2), 263–282.
- Deyhime, G. 2000. *Farhang-i Avay-i Farsi* [Dictionary of the sound of Persian]. Tehran: Farhang Moaser Publishers.
- Esposito, Anna & Maria Gabriella Di Benedetto. 1999. Acoustical and perceptual study of gemination in Italian stops. *The Journal of the Acoustical Society of America* 106(4), 2051–2062.
- Garnes, Sara. 1976. *Quantity in Icelandic: Production and Perception*. Hamburg: Helmut Buske Verlag.
- Gay, Thomas. 1978. Effect of speaking rate on vowel formant movements. *The Journal of the Acoustical Society of America* 63(1), 223–230.
- Gay, Thomas. 1981. Mechanisms in the control of speech rate. *Phonetica* 38(1–3), 148–158.
- Hankamer, Jorge, Aditi Lahiri & Jacques Koreman. 1989. Perception of consonant length: Voiceless stops in Turkish and Bengali. *Journal of Phonetics* 17(4), 283–298.
- Hansen, Benjamin B. 2004. Production of Persian geminate stops: Effects of varying speaking rate. In Augustine Agwuele, Willis Warren & Sang-Hoon Park (eds.), *Proceedings of the 2003 Texas Linguistics Society Conference: Coarticulation in Speech Production and Perception*, 86–95. Somerville, MA: Cascadilla Proceedings Project.
- Hansen, Benjamin B. 2012. *The perceptibility of duration in the phonetics and phonology of contrastive consonant length*. Ph.D. dissertation, The University of Texas at Austin.
- Hays, Bruce. 1989. Compensatory lengthening in moraic phonology. *Linguistic Inquiry* 20(2), 253–306.
- Hirata, Yukari & Jacob Whiton. 2005. Effects of speaking rate on the single/geminate stop distinction in Japanese. *The Journal of the Acoustical Society of America* 118(3), 1647–1660.
- Kavitskaya, Darya. 2002. *Compensatory lengthening: Phonetics, phonology, diachrony*. New York: Routledge.
- Kawahara, Shigeto. 2007. Sonorancy and geminacy. In Leah Bateman, Michael O’Keefe, Ehren Reilly & Adam Werle (eds.), *Papers in Optimality Theory III* (University of Massachusetts Occasional Papers in Linguistics 32), 145–186. Amherst, MA: Graduate Linguistics Student Association.
- Kawahara, Shigeto & Melanie Pangilinan. To appear. Spectral continuity, amplitude changes, and perception of length contrasts. In Haruo Kubozono (ed.), *Aspects of geminate consonants*. Oxford: Oxford University Press.
- Keating, Patricia A. 1988. Underspecification in phonetics. *Phonology* 5(2), 275–292.
- Klatt, Dennis. 1976. Linguistic uses of segment duration in English: Acoustic and perceptual evidence. *The Journal of the Acoustical Society of America* 59(5), 1208–1221.
- Kostić, Djodje, Alokandana Mitter & Bh. Krishnamurti. 1977. *A short outline of Telugu phonetics*. Calcutta: Indian Statistical Institute.
- Kuehn, David & Kenneth Moll. 1976. A cineradiographic study of VC and CV articulatory velocities. *Journal of Phonetics* 4(4), 303–320.
- Kuznetsova, A., P. B. Brockhoff & R. H. Bojesen. 2014. Tests in Linear Mixed Effects Models. <http://www.R-project.org> (accessed 18 December 2014).
- Lahiri, Aditi & Jorge Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics* 16(3), 327–338.
- Lavoie, Lisa. 2000. *Phonological patterns and phonetic manifestations of consonant weakening*. Ph.D. dissertation, Cornell University.
- Leben, William R. 1980. A metrical analysis of length. *Linguistic Inquiry* 11(3), 497–509.
- Lehiste, Ilse. 1964. *Acoustical characteristics of selected English consonants*. Supplement to *International Journal of American Linguistics* 30.

- Lehtonen, Jaakko. 1970. *Aspects of quantity in Standard Finnish*. Jyväskylä: Jyväskylä University Press.
- Liberman, Alvin M., Pierre C. Delattre, Louis J. Gerstman & Franklin S. Cooper. 1956. Tempo of frequency change as a cue for distinguishing classes of speech sounds. *Journal of Experimental Psychology* 52(2), 127–137.
- Löfqvist, Anders. 2005. Lip kinematics in long and short stop and fricative consonants. *The Journal of the Acoustical Society of America* 117, 858–878.
- Löfqvist, Anders. 2007. Tongue movement kinematics in long and short Japanese consonants. *The Journal of the Acoustical Society of America* 122(1), 512–518.
- Maddieson, Ian. 2008. Glides and gemination. *Lingua* 118(12), 1926–1936.
- Mahootian, Shahrzad. 1997. *Persian*. London: Routledge.
- Majidi, Mohammad-Reza & Elmar Ternes. 1999. Persian (Farsi). In IPA (ed.), *The handbook of the International Phonetic Association*, 124–125. Cambridge: Cambridge University Press.
- Malik, Amar Nath. 1995. *The phonology and morphology of Panjabi*. New Delhi: Munshiram Manoharlal Publishers.
- Mauk, Claude. 2003. *Undershoot in two modalities: Evidence from fast speech and fast signing*. Ph.D. dissertation, The University of Texas at Austin.
- McKee, Suzanne P., Stanley A. Klein & Davida Y. Teller. 1985. Statistical properties of forced-choice psychometric functions: Implications of probit analysis. *Perception & Psychophysics* 37(4), 286–298.
- Myers, Scott & Benjamin B. Hansen. 2005. The origin of vowel-length neutralization in vocoid sequences: Evidence from Finnish speakers. *Phonology* 22(3), 317–344.
- Nababan, P. W. J. 1981. *A grammar of Toba Batak* (Pacific Linguistics, Series D, No. 37). Canberra: Australian National University.
- Obrecht, Dean. 1965. Three experiments in the perception of geminate consonants in Arabic. *Language and Speech* 8(1), 31–41.
- Ohala, John J. 1981. The listener as a source of sound change. In Carrie S. Masek, Roberta A. Hendrick & Mary Frances Miller (eds.), *Papers from the Parasession on Language and Behavior* (Chicago Linguistics Society, 1–2 May 1981), 178–203. Chicago, IL: Chicago Linguistics Society.
- Ohala, John J. 1983. The origin of sound patterns in vocal tract constraints. In Peter MacNeilage (ed.), *The production of speech*, 189–216. New York: Springer.
- Ohala, John J. 1993. The phonetics of sound change. In Charles Jones (ed.), *Historical linguistics: Problems and perspectives*, 237–278. London: Longman.
- Öhman, Sven E. G. 1967. Numerical model of coarticulation. *The Journal of the Acoustical Society of America* 41, 310–320.
- Parker, Steve. 2008. Sound level protrusions as physical correlates of sonority. *Journal of Phonetics* 36(1), 55–90.
- Penchoen, Thomas G. 1973. *Tamazight of the Ayt Nahir*. Malibu: Undena Publications.
- Peterson, Gordon E. & Ilse Lehiste. 1960. Duration of syllable nuclei in English. *The Journal of the Acoustical Society of America* 32(6), 693–703.
- Pickett, Emily R., Sheila A. Blumstein & Martha W. Burton. 1999. Effects of speaking rate on the singleton/geminate consonant contrast in Italian. *Phonetica* 56(3–4), 135–157.
- Pierrehumbert, Janet & David Talkin. 1992. Lenition of /h/ and glottal stop. In Gerard J. Docherty & D. Robert Ladd (eds.), *Papers in Laboratory Phonology II: Gesture, segment, prosody*, 90–117. Cambridge: Cambridge University Press.
- Pind, Jörgen. 1986. The perception of quantity in Icelandic. *Phonetica* 43(1–3), 116–139.
- Podesva, Robert. 2000. Constraints on geminates in Buginese and Selayarese. *West Coast Conference on Formal Linguistics (WCCFL)* 19, 343–356. Somerville, MA: Cascadia Press.
- Prince, Alan. 1984. Phonology with tiers. In Mark Aronoff & Richard T. Oehrle (eds.), *Language sound structure*, 234–244. Cambridge, MA: MIT Press.
- Ridouane, Rachid. 2007. Gemination in Tashlyt Berber: An acoustic and articulatory study. *Journal of the International Phonetic Association* 37(2), 119–142.
- Samareh, Y. 1977. *The arrangement of segmental phonemes in Farsi*. Tehran: Faculty of Letters, University of Tehran.
- Samareh, Y. 1985. *Avāshināsī-i Zabān-i Fārsī*, [Phonetics of the Persian language]. Tehran: Markaz Nashr Dāneshgāhī.

- Smith, Caroline. 1995. Prosodic patterns in the coordination of vowel and consonant gestures. In Bruce Connell & Amalia Arvaniti (eds.), *Phonology and phonetic evidence: Papers in Laboratory Phonology IV*, 205–222. Cambridge: Cambridge University Press.
- Tyler, Stephan Albert. 1969. *Koya: An outline grammar*. Berkeley, CA: University of California Press.
- Umeda, Noriko. 1977. Consonant duration in American English. *The Journal of the Acoustical Society of America* 61(3), 846–858.
- Van Son, R. J. H. H. & Louis C. W. Pols. 1999. An acoustic description of consonant reduction. *Speech Communication* 28(2), 125–140.
- Warner, Natasha & Benjamin V. Tucker. 2011. Phonetic variability of stops and flaps in spontaneous and careful speech. *The Journal of the Acoustical Society of America* 130(3), 1606–1617.
- Wichmann, Felix A. & N. Jeremy Hill. 2001a. The psychometric function, I: Fitting, sampling, and goodness of fit. *Perception & Psychophysics* 63(8), 1293–1313.
- Wichmann, Felix A. & N. Jeremy Hill. 2001b. The psychometric function, II: Bootstrap-based confidence intervals and sampling. *Perception & Psychophysics* 63(8), 1314–1329.