Working Papers of the Cornell Phonetics Laboratory
1998, vol. 12, pp. 27-51

Temporal Properties of Madurese Consonants: A Preliminary Report*

Abigail C. Cohn and William H. Ham

Madurese, an Austronesian language of Indonesia, is unique among languages of the region in that stops exhibit a five-way contrast in place of articulation which is cross-cut by a three-way contrast in phonation on the one hand, and a singleton-geminate contrast in medial position on the other (Kiliaan 1897; Stevens 1966, 1968). The singleton-geminate contrast extends to all other consonants in the inventory except glottal stop, but is phonologically neutralized following the vowel /ə/. This paper reports the results of the first systematic acoustic investigation to be carried out on Madurese consonants. We focus in particular on the acoustic correlates of the three-way phonation contrast in the stops, the magnitude of the singleton-geminate contrast in all consonants, and the question of whether phonological neutralization of this contrast following /ə/ triggers phonetic shortening. Findings show that geminates are on average approximately 60% longer than their singleton counterparts, and that vowels preceding geminates are only half as long as those preceding singletons (i.e., in open syllables). Closed syllable shortening of this magnitude, along with the fact that geminates do not shorten preceding /ə/, suggests that Madurese spaces syllables at roughly equal temporal intervals. With specific reference to stops, results show that phonologically aspirated stops are only lightly aspirated, and that the durational difference between singleton and geminate voiced stops is (unexpectedly) the largest, despite the fact that they are always fully voiced.

1 Introduction

Madurese is a West Austronesian language spoken in Indonesia, on Madura and parts of East Java. It has a complex consonant inventory, richer than most of its neighbors in range of places of articulation, in the number of phonation contrasts, and in exhibiting a singleton-geminate contrast. The consonant inventory is presented in (1). In word-final position, the inventory is considerably smaller; only those consonants in boxes may appear word-finally.

(1) Consonant inventory of Madurese

		labial	dental	retroflex/ alveolar	palatal	velar	glottal
stops	voiced	b	d̪	ɖ	j	g	
	voiceless	p	t̪	ʈ	c	k	
	aspirated	p^h	$t̪^h$	$ʈ^h$	c^h	k^h	
fricative				s			
nasals		m	n̪		ɲ	ŋ	
approximants		(w)		l, r	y		ʔ h

* This paper was presented at 8-ICAL, Taiwan in December 1997. We'd like to thank the 8-ICAL audience, members of the Cornell Phonetics Lab, Allard Jongman and Lisa Lavoie for helpful comments. Thanks also to Eric Evans for technical assistance and Rob Podesva for assistance in data analysis. This work was supported in part by NSF grant no. SBR-9511185.

As shown in (1), Madurese exhibits five supralaryngeal places of articulation in word-initial and word-medial stops: labial, dental, alveolar retroflex, palatal and velar; most neighboring languages show only a four-way place contrast. The five-way contrast in Madurese is exhibited for word-initial voiceless stops in (2).

(2) Five-way place contrast

papa	t̪apa	ʈaʈat̪	capak	kapa
'humble'	'asceticism'	'parrot'	'statement'	'saddle'

It also has a three-way phonation contrast among voiced, voiceless and aspirated in word-initial and word-medial positions, as illustrated for initial bilabial stops in (3):

(3) Three-way phonation contrast

bɤt̪a	pakaʔ	pʰɤt̪a
'lift'	'type of fish'	'brick'

None of the neighboring languages shows such a three-way contrast. Javanese, Sundanese and Indonesian, for example, all have two-way phonation contrasts. As discussed by Stevens (1966), aspirated stops in Madurese derive historically from voiced stops, and what are now the voiced stops derive principally from glide hardening and borrowings. It should be noted that some researchers transcribe the aspirated series as voiced, but impressionistically they do not appear to be voiced. We will return to this question in our acoustic study, below.

Another important point about Madurese exemplified in (3) is that voiceless stops are followed by a distinct set of vowels from the voiced and aspirated stops. Specifically, only non-high vowels follow voiceless stops, while only high vowels follow voiced and aspirated stops (see Cohn 1993 and Cohn and Lockwood 1994 for a full account). The vowel inventory in (4) illustrates this split. Note that there is no phonemic distinction in vowel length in Madurese:

(4) Interaction between stop phonation and following vowel height

	front		back	
high	i	ɤ ɨ	u	[+high] after voiced and aspirated stops
mid	ɛ	ə	ɔ	[-high] after voiceless stops
low		a		

In addition, all consonants in (1) except glottal stop can be geminated in intervocalic position, as exemplified in (5).

(5) Singleton-geminate contrast

papa	pappa	kanaʔ	kannaʔ
'humble'	'stem of palm leaf'	'small child'	'come here'
paron	parron	kasaʔ	kassaʔ
'identical half'	'metallic base'	'voice'	'go over there'

None of the neighboring languages in Java feature contrastive gemination, although the languages of Southern Sulawesi do. According to Stevens (1966), principal historical sources for gemination include total assimilation in nasal-stop clusters, gemination after schwa and borrowings. All vowels except schwa can appear before either a singleton or a geminate consonant. The appearance of schwa is restricted to the environment preceding geminates or nasal-stop clusters. All consonants — again with the exception of glottal stop — appear as geminates following schwa. We have yet to find an example of a geminate voiced stop *not* preceded by schwa.

As far as previous work on Madurese is concerned, both Kiliaan (1897) and Stevens (1968) provide excellent descriptions, though these are based solely on impressionistic observations. The first acoustic work on Madurese was carried out by Cohn (1993) and Cohn and Lockwood (1994), focusing only on the interaction between phonation and vowel quality illustrated in (5). We report here on the results of a systematic acoustic study of the consonant system of Madurese — the first such study to be carried out. This paper is organized as follows: §2 outlines the methodology employed in the current study; §3 reports on the acoustic correlates of the three-way phonation contrast in the stops, as well as the durational characteristics of fricatives and sonorants; §4 describes the acoustic realization of the singleton-geminate contrast in stops, fricatives and sonorants, and discusses preceding vowel duration; finally §5 concludes with a summary of the findings of the present study and lays out directions for further research.

2 Methodology

This section briefly outlines the procedures followed here for data recording, digitizing and analysis. The word list developed for the present study contains at least one example of every Madurese consonant in every prosodic position in which it can occur within the word (recall that geminates occur only in intervocalic position, and that the inventory of phonemes which can appear in word-final position is considerably reduced). The structure of the list is summarized in Table 1, where targeted consonant positions are underscored, and 'T' represents bilabial, dental or velar voiceless stops:

prosodic shape	segmental content	target
a. CVCV(C)	_[a / ɤ]Ta(C)	all segments
b. CVCV(C)	Ta_[a / ɤ](C)	all segments
c. CVC:V(C)	Ta_[a / ɤ](C)	all geminates except voiced stops
d. CəC:VC	Tə_[a / ɤ](C)	geminate stops, fricatives, nasals and liquids
e. CVCVC	(C)aTa_	all segments appearing word-finally

Table 1. Structural overview of the wordlist.

Two native speakers of Madurese — one female (speaker E) and one male (speaker S) — were recorded in a sound-treated room in East Java, Indonesia, using a Sony portable cassette recorder and an Audio-Technica AT813 microphone. Both subjects are speakers of the standard dialect of East Madura, and neither had any known record of speech or hearing disorders at the time of recording. Each token from the wordlist schematized in Table 1 was read by the speakers in the frame sentence [bɤca _____ sapaɾo] ('read _____ partway'). Analysis of 100 words with three repetitions per speaker is presented here, encompassing a total of 600 tokens.

The 600 carrier phrases containing the target words were digitized at a sampling rate of 11.025 kHz on a SUN SparcStation LX and analyzed by generating waveforms and wide-band spectrograms using Entropic Signal Processing Systems (ESPS) Waves+® speech analysis software package. Durational measurements were derived from pairs of labels placed within these audio files, using the X-Label® attachment for Waves+® and a shellscript written in C-Shell and AWK scripting languages in the Cornell Phonetics Laboratory. For stops, measurements included closure duration, closure voicing, VOT, burst duration and preceding vowel duration. For the other consonants, measurements included closure duration and preceding vowel duration.

Based on both wideband spectrograms and waveforms (which can be viewed simultaneously in the Waves+® environment), the following segmentation criteria were employed in labeling the data: Vowel onset was taken to coincide with the onset of F2 energy, and vowel offset was taken to coincide with the offset of F2 energy. For stops, the closure phase was marked to span from the offset of the preceding vowel to the left edge of the release burst; positive VOT was labeled as the interval between the left edge of the release burst and the onset of F2 in the following vowel (VOT was not measured word-finally). Fricative onset was taken to correspond with the beginning of aperiodicity in the waveform, and fricative offset was marked at the onset of F2 in the following vowel (except word-finally — this case will be discussed in §3.2). Tap [ɾ] (and its geminate counterpart [r]) were labeled as the interval between offset of F2 in the preceding vowel and onset of F2 in the following vowel (or the end of the word). For sonorants (nasals, liquids, glides), onset was marked to coincide with a clear downward shift in formant energy (or in nasals, with the appearance of the antiformant), and offset was marked to coincide with a clear upward shift in formant energy at the onset of the following vowel (or with the end of the word). These segmentation criteria essentially follow those set forth in, e.g., Klatt (1976) and Fujimura and Erickson (1997).

3 Acoustic overview of singleton consonants

This section provides an overview of the durational properties of Madurese singleton consonants. We begin with a comparison of voiced, voiceless and aspirated stops across word positions in §3.1, and continue with a brief discussion of all consonant types in §3.2.

3.1 Stops

As outlined in §1, Madurese distinguishes three phonation types in the stops: voiced, voiceless and voiceless aspirated (following Stevens 1980). This subsection, which concentrates exclusively on the singleton stops, is concerned primarily with the question of how these three phonation types are differentiated acoustically. Representative wideband spectrograms of word-medial voiced, voiceless and aspirated singleton stops are provided in Figure 1:

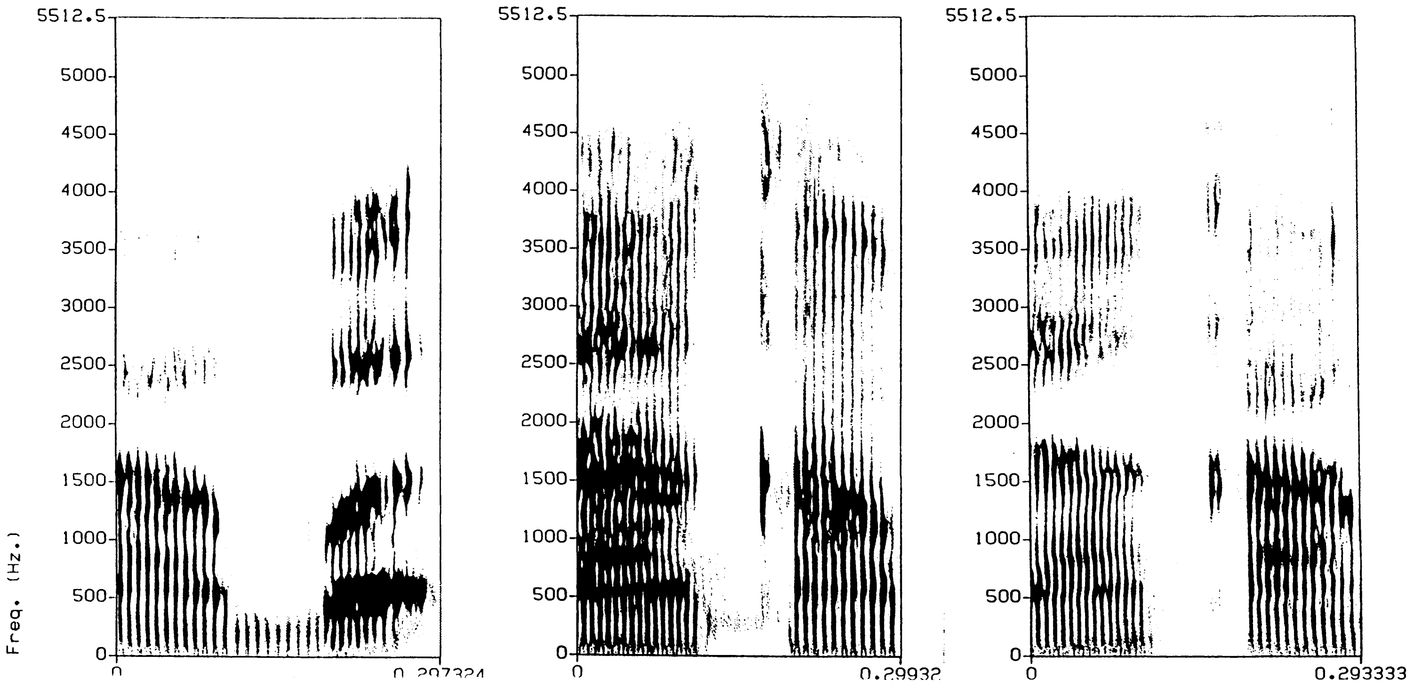

Figure 1. Wideband spectrograms of (left to right) word-medial bilabial voiced stop [b] in [kabɤʔ] ('wine'), bilabial voiceless stop [p] in [kapal] ('ship') and bilabial aspirated stop [pʰ] in [kapʰɤr] ('news'), produced by speaker S.

We begin with the contrast between the voiced and voiceless stops. In languages with a voicing distinction in the stops, there is a common tendency for the voiced series to be shorter in duration than the voiceless series due to aerodynamic considerations (Lisker 1957, Peterson and Lehiste 1960, Ohala 1983). In order to maintain vocal fold vibration throughout the closure, voiced stops are typically produced with considerably shorter closure durations due to the rapid equalization of subglottal and intraoral pressures.

Figures 2 and 3 summarize closure duration data for voiced, voiceless and aspirated stops in both word-initial and word-final positions (all five places of articulation combined). The first interesting point to be made here is that there is no strong tendency for voiced stops to be shorter than their voiceless counterparts. The second point of interest is the general tendency on the part of both speakers for stop closure to be longer in word-medially than word-initially. Acoustic studies of English consonant duration have consistently revealed that word-initial segments tend to be longer than word-medial segments (Oller 1973, Klatt 1976, Umeda 1977), even in phrase-medial position (like the Madurese data); Byrd *et al.* (1997) report similar effects in Tamil. The absence of consistent initial lengthening in Madurese suggests that languages may differ this regard, just as they differ with regard to the magnitude of other durational effects, such as the implementation of quantity distinctions for consonants (Ladefoged and Maddieson 1996:

91-93) and vowels (Hubbard 1994). Error bars indicate one standard deviation from the mean for all stops within each category; p-values are from two-sample t-tests performed on pairing of initial and medial stops:

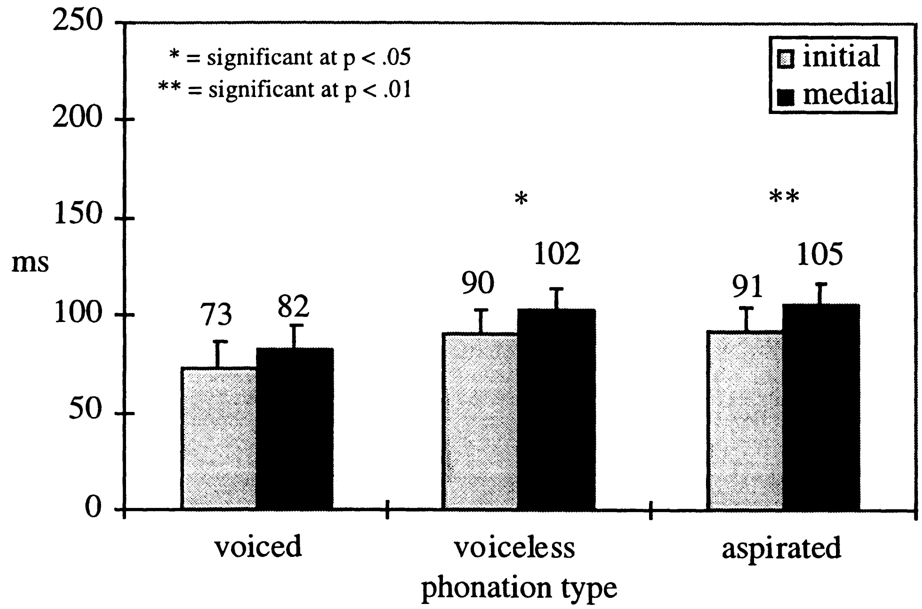

Figure 2. Mean closure duration for voiced, voiceless and aspirated (singleton) stops in word-initial and word-medial positions; five places of articulation combined (speaker E); n = 15 per category.

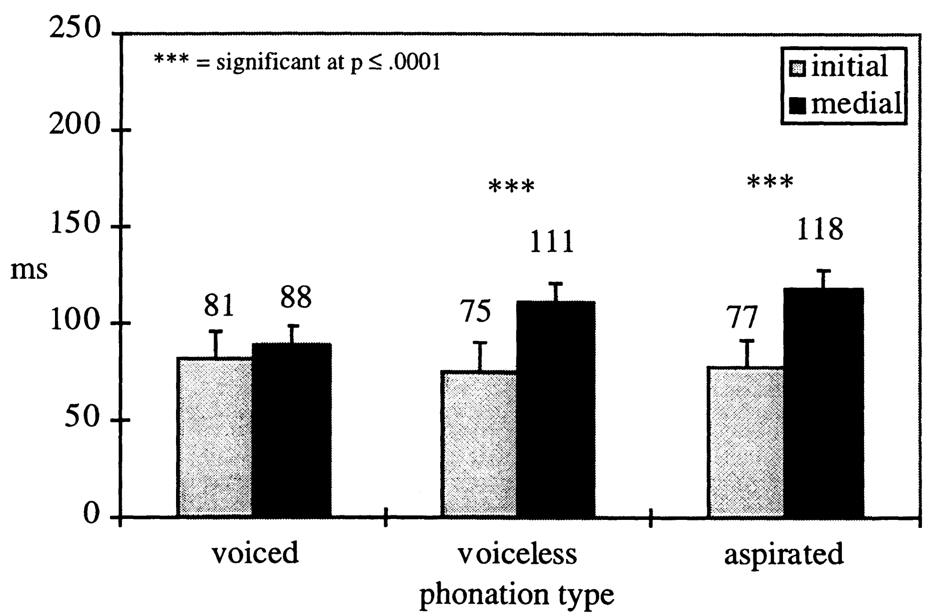

Figure 3. Mean closure duration for voiced, voiceless and aspirated (singleton) stops in word-initial and word-medial positions; five places of articulation combined (speaker S); n = 15 per category.

For speaker E, voiced stops are on average 23% shorter than their voiceless counterparts word-initially, and 24% shorter word-medially. For speaker S, the voiced stops are on average 8% *longer* word-initially, but 26% shorter word-medially. The lack of a robust difference in initial stops for speaker S may relate to the observation that initial position serves as the locus for articulatory strengthening, which may or may not be accompanied by some degree of lengthening (see, e.g., Fougeron and Keating 1997, Byrd *et al.* 1997).

Because the voiced stops in our sample are, with only a handful of exceptions, fully voiced throughout the closure, the absence of consistently shorter voiced stops may indicate that vocal fold vibration is sustained with the help of one or more articulatory strategies aimed at the enhancement and prolongation of obstruent voicing (note that even the partially voiced cases were voiced throughout at least 70% of the closure). One possibility along these lines is egressive leakage of air through the velum (also referred to in the literature as "hypervoicing" — see Lisker and Abramson 1971, Ohala and Ohala 1991, Henton, *et al.* 1992, Iverson and Salmons 1996). Beyond velic leakage, Ladefoged and Maddieson (1996: 50) review various other strategies which facilitate obstruent voicing via an expansion of the oropharyngeal cavity behind the point of closure. These include anteriorization of the closure during voicing, advancing the tongue root, and lowering of the larynx or the jaw. In particular, lowering of the larynx appears to be a distinct possibility in Madurese, given evidence that the systematic differences in vowel height (Cohn 1993) and pitch (Cohn and Lockwood 1994) conditioned by the phonation type of the preceding stop (see §1) are related to larynx height during stop production.

These methods of enhancement all involve an articulatory adjustment of some sort during obstruent production, and each could conceivably be accompanied by relatively longer closure duration in order to allow time to accomplish the necessary adjustments. Given the association between initial position and the intensification of articulatory gestures drawn by numerous studies (e.g. Browman and Goldstein 1992, Krakow 1989, Byrd 1996, Fougeron and Keating 1997), it is equally conceivable that such strategies, if employed, are exaggerated word-initially. For speaker E, this effect — whatever its cause — manifests itself as a relatively smaller durational difference between voiced and voiceless stops in initial position, while for speaker S, the difference essentially disappears.

We turn next to the contrast between voiceless and aspirated stops. It should first be noted that aspirated stops were consistently observed to be voiceless for both speakers (see Figure 1). As first discussed by Lisker and Abramson (1964), aspirated stops are distinguished from their unaspirated counterparts primarily in terms of the relative timing of stop release and the onset of voicing in the following vowel. Our expectation, therefore, is to find that aspirated stops in Madurese have higher positive VOT values than the voiceless stops across the board. As shown in Figures 4 and 5, this is precisely what we find for both speakers:

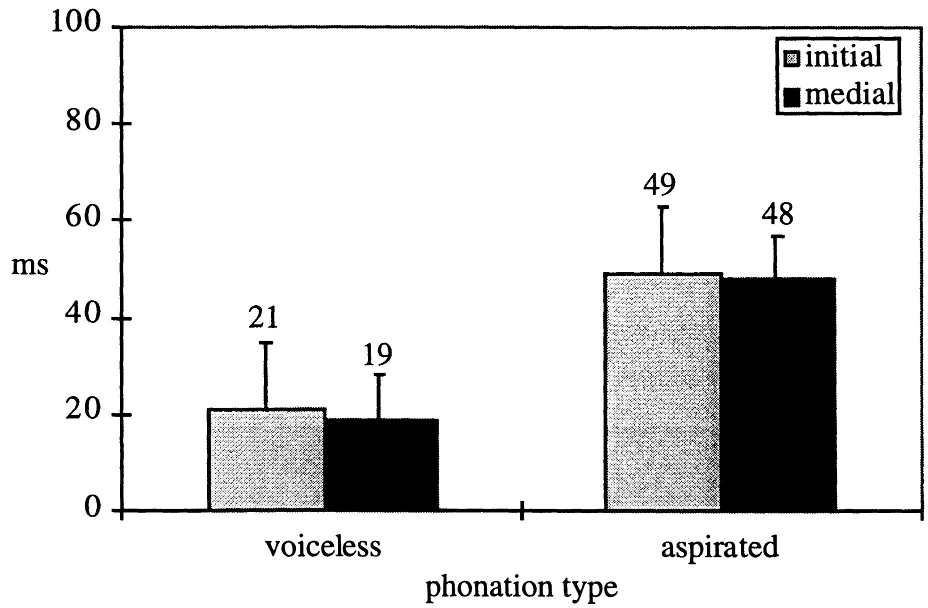

Figure 4. Mean VOT in voiceless and aspirated stops in word-initial and word-medial positions; five places of articulation combined (speaker E); n = 15 per category.

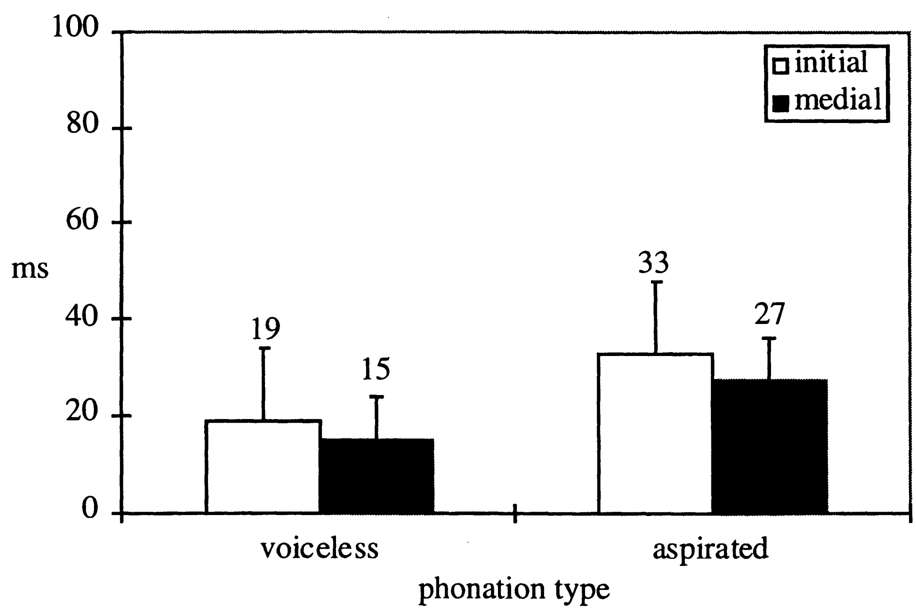

Figure 5. Mean VOT in voiceless and aspirated stops in word-initial and word-medial positions; five places of articulation combined (speaker S); n = 15 per category.

For speaker E, mean VOT in initial aspirated stops is 133% longer than in initial voiceless stops, and VOT in medial aspirated stops is 152% longer than in medial aspirated stops. Differences for speaker S are only about half as large: VOT in initial aspirated stops is 73% longer than in initial voiceless stops, and VOT in medial aspirated stops is 80% longer than in initial aspirated stops. Although the differences in VOT between voiceless and aspirated stops are statistically robust ($p < .01$ by two-sample t-test) in each case, the mean raw durational gap between the two stop types is relatively small, ranging between 15 and 39 ms for speaker E, and between 0 and 17 ms for speaker S.

This result is not particularly surprising, given that voiceless stops are always followed by non-high vowels, while aspirated stops (along with voiced stops) are always followed by high vowels. The two voiceless stop types can therefore be distinguished solely on the basis of the quality of the following vowel. Adding weight to the hypothesis that the contrastive role of VOT in the stops is being minimized is the observation that stops preceding high (tense) vowels tend to show significantly longer intrinsic VOT values than their non-high counterparts cross-linguistically (see Klatt 1975, Port and Rotunno 1979, and Nearey and Rochet 1994). We can therefore speculate that the small absolute differences in VOT in the Madurese data reflect a change in progress, whereby the contrast between voiceless and aspirated stops is being subsumed by a difference in following vowel quality. Historical shifts of this type have been documented in other language groups, perhaps most notably for the emergence of register systems in Mon-Khmer languages (see, e.g., Huffman 1976, Pinnow 1979).

A correlation between initial position and enhanced glottal articulation — including longer VOT — has been observed for English by Pierrehumbert and Talkin (1992), and for Korean by Jun (1993). The present data do not show a consistent trend in this direction for Madurese. For speaker E, a two-factor ANOVA performed on VOT using the category variables "phonation type" and "position" revealed a significant main effect for phonation [$F(1, 56) = 91.600$, $p \leq .0001$], but not for position [$F(1, 56) = .10876$, $p = .7429$] and no interaction between the two [$F(1, 56) = .02537$, $p = .8740$]. For speaker S, the same test showed no significant main effect for phonation [$F(1, 56) = .94016$, $p = .3364$], no main effect for position [$F(1, 56) = 1.7125$, $p = .7429$] and no interaction between the two [$F(1, 56) = 1.0735$, $p = .3046$].

Lastly, we turn briefly to the duration of word-final stops. As discussed in §1, the three phonation types appear to be neutralized to plain voiceless in word-final position; in addition, only bilabial, dental and velar stops occur word-finally. As data gathered for the present study are not sufficient to address the question of whether final neutralization of phonation type is complete, we compare only word-final stops which appear orthographically as <p, t, k> to voiceless bilabial, dental and velar stops in word-initial and word-medial positions. As shown in Figure 6, word-final stops tend to be longer than those in other positions, in agreement with reports of word-final lengthening in English by, e.g., Klatt (1976) and Fougeron and Keating (1997). Regression analysis indicates that position-in-word exerts a larger influence on closure duration for speaker S [$F(1, 25) = 58.1$, $p \leq .0001$, $r^2 = 69.9$] than for speaker E [$F(1, 25) = 14.5$, $p = .0008$, $r^2 = 36.7$]. Two-sample t-tests show that while differences between both word-initial and word-medial and word-medial and word-final stops are significant at at least the $p < .05$ level for speaker S, only the difference between word-medial and word-final is significant for speaker E:

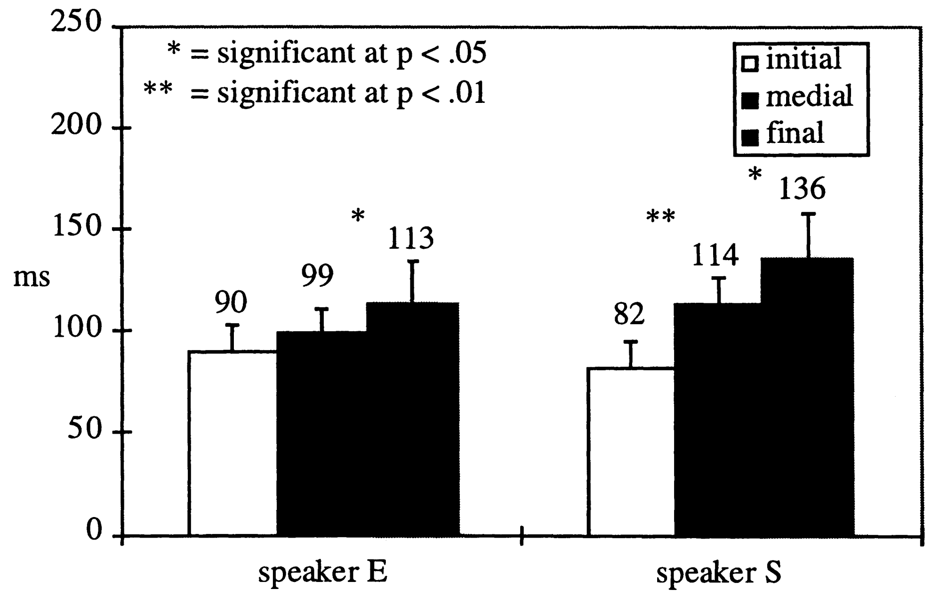

Figure 6. Mean closure durations for word-initial, word-medial and word-final voiceless (i.e., unaspirated) bilabial, dental and velar stops (places combined); n = 9 per category per speaker.

To briefly summarize, we have made a number of observations about the acoustic characteristics of Madurese stops. First, voiced stops are fully voiced, but are not produced with a dramatically shorter closure than voiceless stops, contrary to what we might expect *a priori*, given the aerodynamic difficulties with maintaining obstruent voicing. As discussed above, this may be the result of an articulatory compensation of some type aimed at accommodating vocal fold vibration in the presence of a complete oral closure. Second, while initial stops of all three phonation types are generally shorter than their word-medial counterparts, word-final (voiceless) stops are longest. Third, aspirated stops are not characterized by a larger amount of closure voicing than plain voiceless stops, but, as anticipated, do show longer VOT. Differences in VOT between the two types are rather small, suggesting that the height specification of the following vowel may be taking over the contrast in this case.

3.2 Other consonants

This subsection briefly outlines the durational properties of singleton fricative /s/, nasals, lateral liquid /l/, tap /ɾ/, and the glides in word-initial, word-medial and word-final positions. Recall from §1 that nasals occur in four places of articulation — bilabial, dental, palatal and velar (palatal nasals do not occur word-finally). Beyond this, the only segment type discussed here which needs further comment is tap /ɾ/, which is realized as trill [r] in word-initial position typically consisting of two to four linguopalatal ƒcontacts. That /ɾ/ might be realized as [r] in word-initial position is not entirely unanticipated, considering that subglottal pressure is presumably higher at word onset in most cases. This higher subglottal pressure gives rise to an increased airflow velocity from the lungs, and therefore to the relatively higher negative pressure between the tongue tip and alveolar ridge necessary for maintaining trill production (Ladefoged and Maddieson 1996: 217-218).

As shown in Figures 7 and 8, the realization of /ɾ/ as [r] in word-initial position has as its outcome for both speakers that segment duration is markedly longer word-initially for the rhotic only. Word-final segment duration is markedly shorter in word-final position than both word-initially and word-medially for [s].[1] Data for glides is limited to [j] here, since [w] is restricted to word-initial and word-medial positions.

[1] The mean duration for word-final [s], which is immediately followed by the initial [s] of the next word in the carrier phrase ([sapaɾo]), was calculated by subtracting the mean duration calculated for the initial [s] in [sapaɾo] from the mean duration of [s#s] sequences.

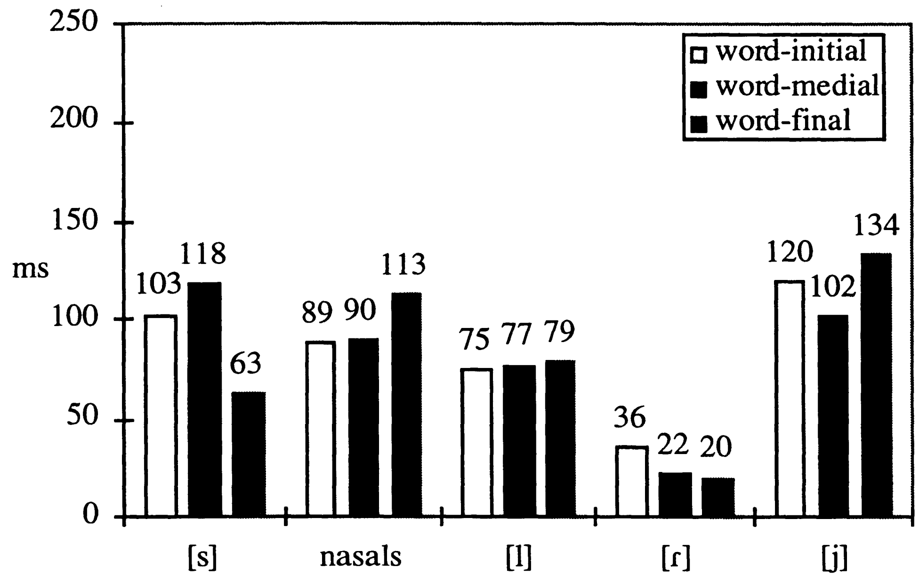

Figure 7. Mean durations for fricatives and sonorants in word-initial, word-medial and word-final positions; nasals averaged over place (speaker E); n = 3 per category

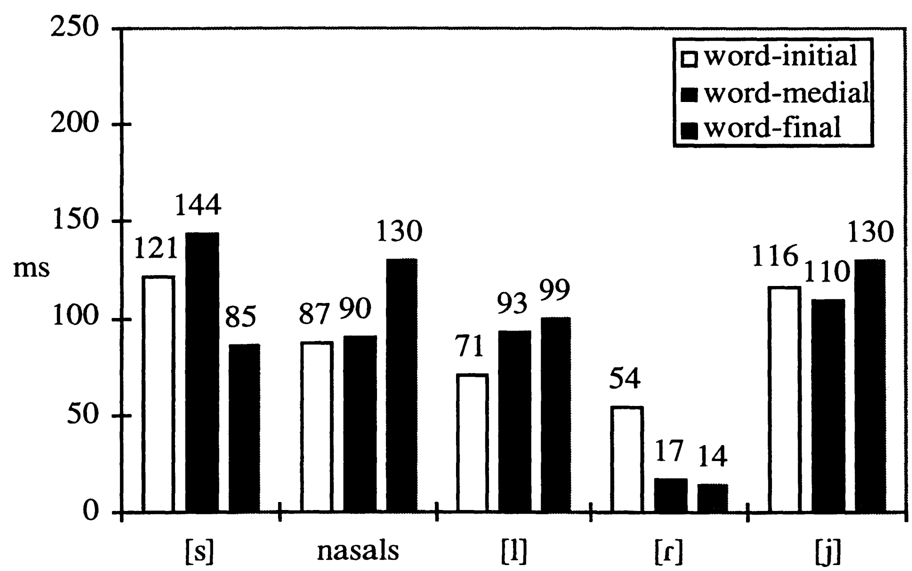

Figure 8. Mean durations for fricatives and sonorants in word-initial, word-medial and word-final positions; nasals averaged over place (speaker S); n = 3 per category

Final lengthening is in evidence for both speakers only for nasals, [l] and glides. As expected, initial trill [r] is considerably longer than medial and final tap [ɾ]. Statistical analysis was not carried out on the data in Figures 7 and 8 due to small sample sizes.

4 Acoustic overview of geminate consonants

This section examines the durational characteristics of the singleton-geminate distinction in word-medial consonants in Madurese. As in §3, we concentrate first on the stops (§4.1), and then move on to an examination of all other consonant types (§4.2).

4.1 Stops

Representative wideband spectrograms of geminate bilabial voiced, voiceless and aspirated stops are shown in Figure 9:

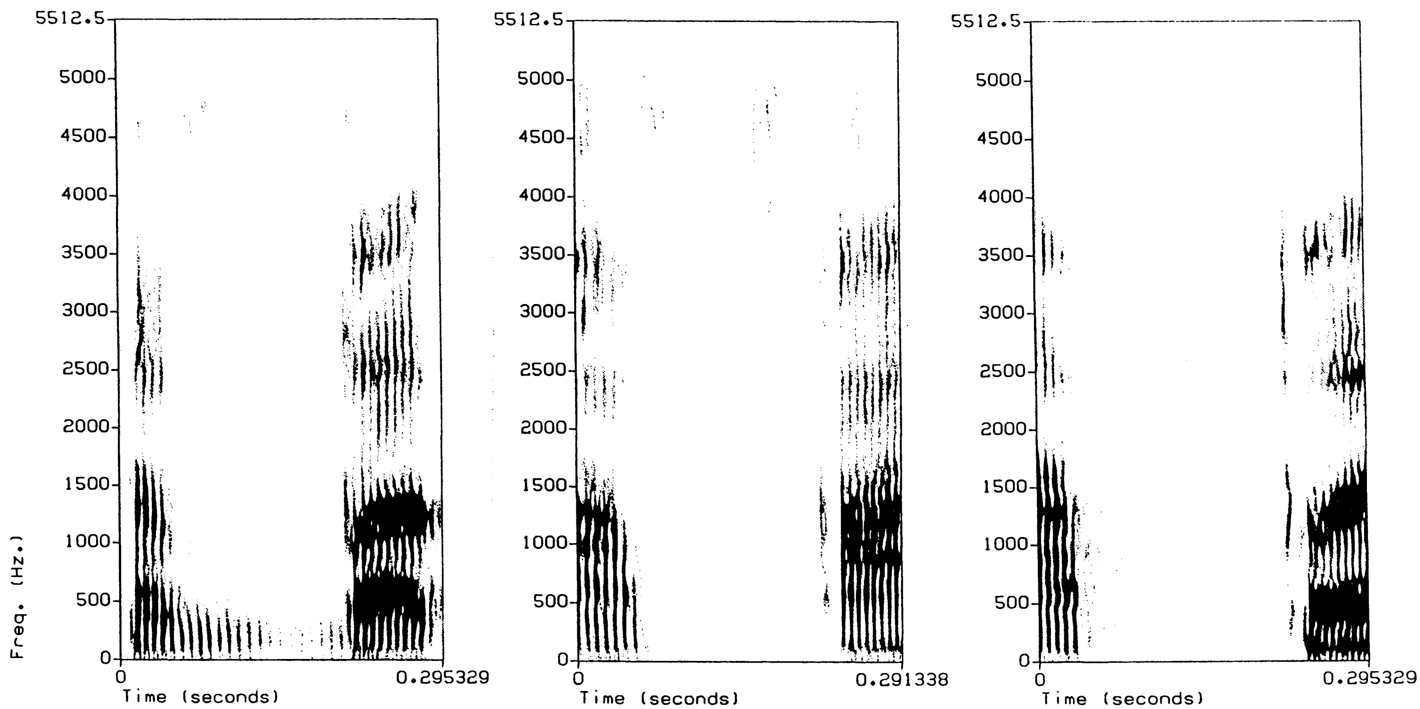

Figure 9. Wideband spectrograms of (left to right) word-medial geminate voiced bilabial stop [b:] in [səb:ɤk] ('torn'), geminate voiceless bilabial stop [p:] in [pap:a] ('type of stem'), and geminate voiceless aspirated stop [p^h:] in [sap^h:ɤɾ] ('patient'), produced by speaker S.

Figures 10 and 11 illustrate the mean closure durations of word-medial voiced, voiceless and aspirated singleton and geminate stops. Note that no data are available for geminate voiced dental stop [d̪:]; five places of articulation are therefore combined in all categories except geminate voiced stops, which includes only four places.

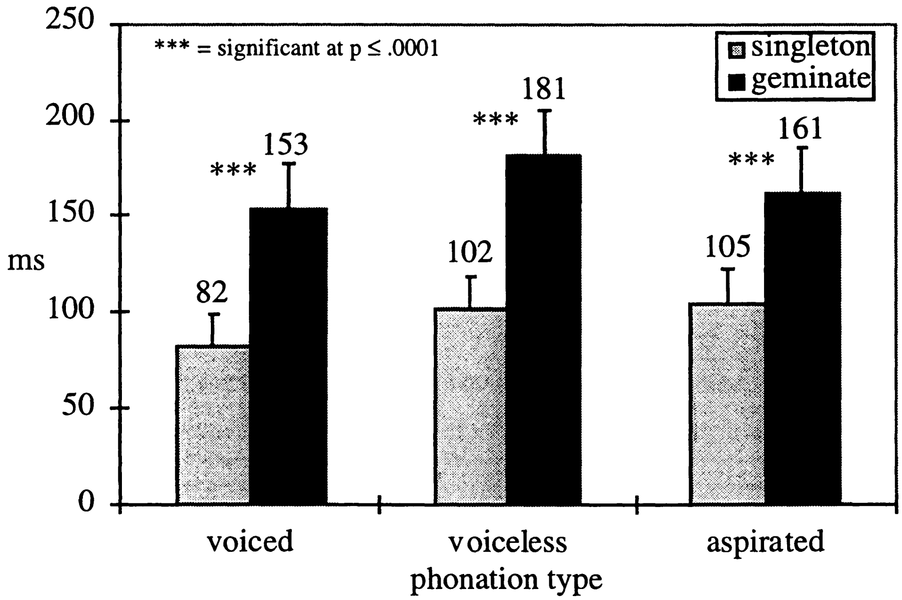

Figure 10. Mean closure durations of voiced, voiceless and aspirated singleton and geminate stops in word-medial position (speaker E); n = 15 for singletons and 12 for geminates

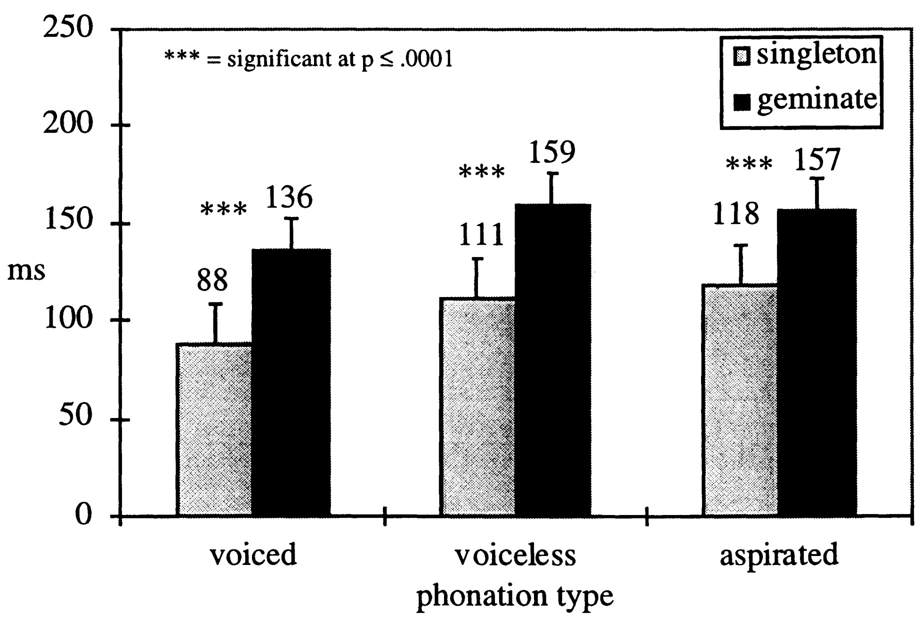

Figure 11. Mean closure durations of voiced, voiceless and aspirated singleton and geminate stops in word-medial position (speaker S); n = 15 for singletons and 12 for geminates

Two-factor ANOVAs performed for each speaker using the category variables "quantity (singleton/geminate)" and "phonation type" indicate that there are significant main effects on closure duration for consonant quantity [E: $F(1, 81) = 271.0$, $p \leq .0001$; S: $F(1, 81) = 159.53$, $p \leq .0001$] and for phonation type [E: $F(2, 81) = 11.75$, $p \leq .0001$; S: $F(2, 81) = 19.784$, $p \leq .0001$]. There is a marginally reliable interaction between the two factors [$F(2, 81) = 3.0235$, $p = .0542$] for speaker E, but not for speaker S [$F(2, 81) = .85336$, $p = .4289$].

With the exception of voiceless retroflex (and perhaps also velar) stops for speaker S, the differences in closure duration associated with the singleton-geminate contrast appear to be robust for all stop types in all places of articulation. It should be pointed out that voiced geminates are in fact fully voiced throughout the closure 94% of the time in the present sample, again suggesting that the maintenance of vocal fold vibration is facilitated by some method of articulatory enhancement, such as velic leakage. This hypothesis is further supported by the observation that the average difference in closure duration between voiced singletons and geminates is the largest, followed by voiceless and aspirated stops. Singleton-to-geminate closure ratios are provided for each speaker in Table 2:

	speaker E	speaker S
voiced	1:1.9	1:1.6
voiceless	1:1.7	1:1.4
aspirated	1:1.6	1:1.3

Table 2. Singleton-to-geminate closure ratios for word-medial stops.

It should be pointed out with regard to Figures 10 and 11 that voiced geminate stops in this sample and — to the best of our knowledge, in the language as a whole are preceded by the vowel /ə/, which for both speakers has an average duration of only about 15 ms. Voiceless and aspirated geminate stops may be preceded by any vowel (including /ə/), but intervocalic singletons of any type are never preceded by /ə/. This means that the singleton-geminate contrast is phonologically neutralized in the environment following /ə/, and we might therefore wonder if voiceless and aspirated geminates shorten in this environment as a result (since there is no pressure to contrast with singletons). Figures 10 and 11 include data for voiceless and aspirated geminates following vowels other than /ə/ only, but as shown in Figures 12 and 13, mean values for closure duration for bilabial and velar geminates following /ə/ are comparable to — and in some cases even longer than — their counterparts in Figures 10 and 11:

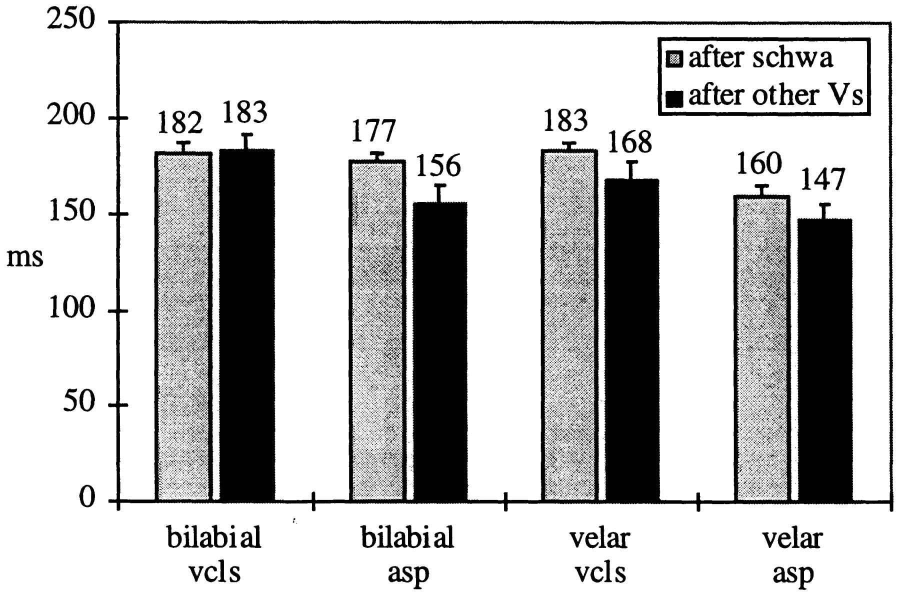

Figure 12. Mean closure durations for voiceless and aspirated bilabial and velar geminate stops following /ə/ and other vowels (speaker E); n = 3 per category

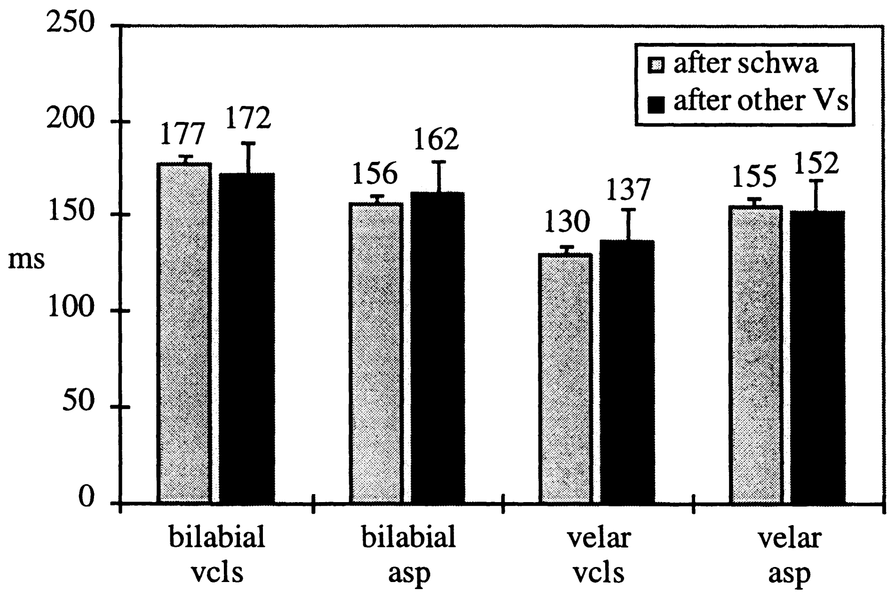

Figure 13. Mean closure durations for voiceless and aspirated bilabial and velar geminate stops following /ə/ and other vowels (speaker S); n = 3 per category

Once again, sample sizes were too small to permit reliable statistical comparison, but it appears that the present sample does not provide any evidence for a tendency to shorten following /ə/.

Lastly, we return now to the question of differences in VOT between voiceless and aspirated singletons and geminates. As illustrated in Figures 14 and 15, neither voiceless nor aspirated geminates show longer mean VOT values than their singleton counterparts:

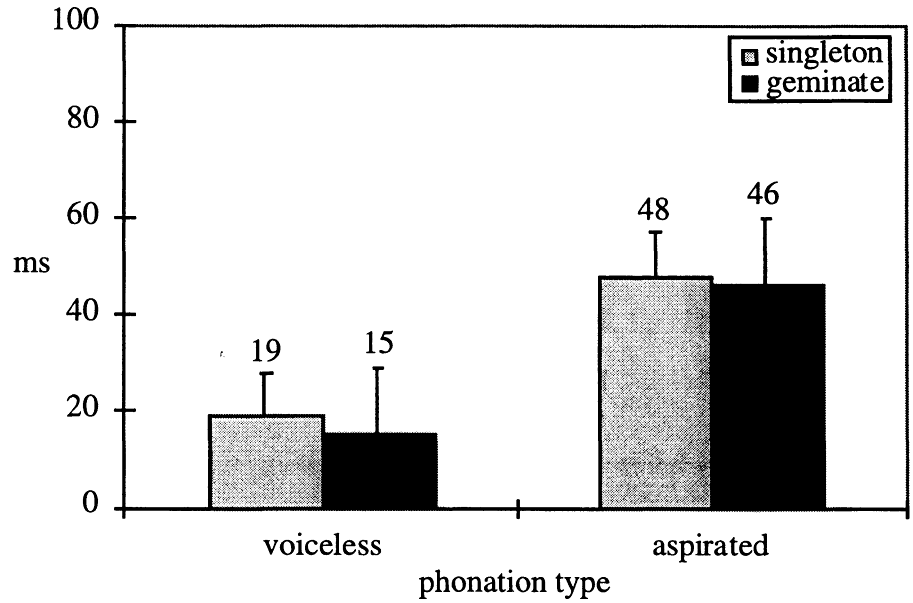

Figure 14. Mean VOT for voiceless and aspirated singleton and geminate stops in word-medial position; five places combined (speaker E); n = 15 per category

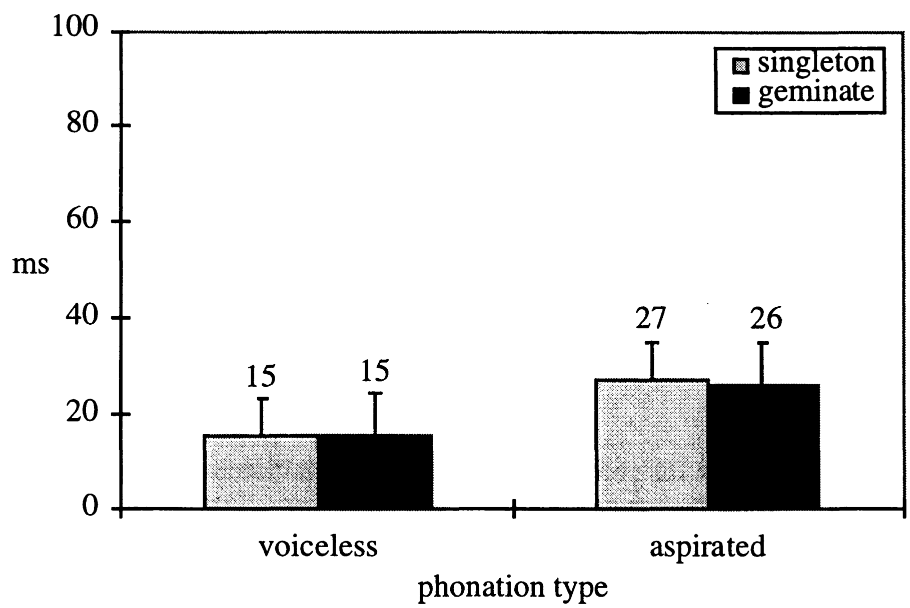

Figure 15. Mean VOT for voiceless and aspirated singleton and geminate stops in word-medial position; five places combined (speaker S); n = 15 per category

None of the differences with voiceless and aspirated pairs in Figures 14 and 15 approached statistical significance by two-sample t-test. It appears, then, that closure duration differences alone account for the singleton-geminate distinction in the stops.

To summarize, geminate stops of all three phonation types are consistently between 30% and 90% longer than their word-medial singleton counterparts. Surprisingly, the difference in closure duration is the largest in voiced stops for both speakers, providing a further indication that obstruent voicing may be achieved in Madurese with the help of an articulatory enhancement strategy. Geminate voiceless and aspirated stops do not appear to shorten in the neutralizing environment following /ə/; in fact, mean closure durations for geminates following the shorter vowel are somewhat longer in many cases. Finally, VOT in word-medial voiceless and aspirated stops does not vary as a function of phonemic consonant length, indicating that the singleton-geminate contrast is driven primarily by differences in closure duration.

4.2 Other consonants

This subsection provides a brief overview of the durational characteristics of the singleton-geminate contrast in all word-medial consonants besides the stops. As was the case of the stops, all geminate consonants except the glides (which appear word-medially only in recent borrowings) may appear following /ə/ as well as other vowels; also like the stops, word-medial singletons do not occur following /ə/, again creating a neutralizing environment for consonant length. Tap /ɾ/ again requires special comment here. Word-medial singleton /ɾ/ is invariably realized acoustically as [ɾ], produced with a single rapid contact. Geminate /ɾː/, however, is realized as trill [r], which typically includes four to six contacts (and is therefore on average somewhat longer than initial [r], discussed in §3.2). This is typical of the situation for tap-trill alternations in geminating languages (Italian, Hausa, Kurdish, Basque and Malayalam, to name a few) described by Banner-Inouye (1995: 155-235) and Ladefoged and Maddieson (1996: 217-232).

Figures 16 and 17 provide mean durations for word-medial singleton and geminate (following /ə/ vs. other vowels) fricative /s/, nasals, lateral liquid /l/, rhotics [ɾ] ~ [r], and glides (which do not occur after /ə/); values for nasals are averaged across four places of articulation:

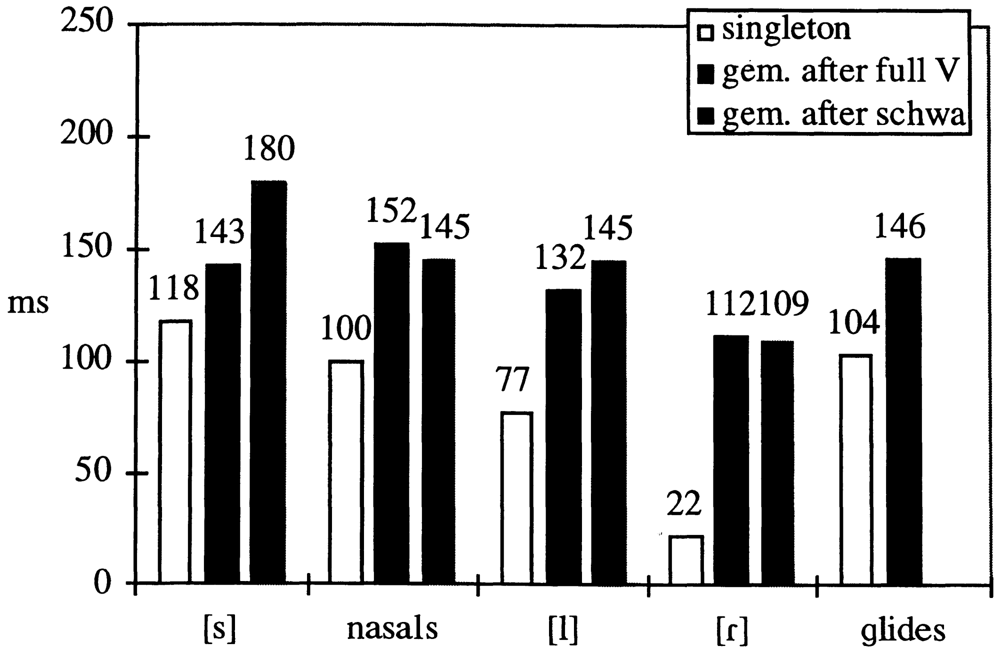

Figure 16. Mean durations for singleton vs. geminate fricatives and sonorants in word-medial position (speaker E)

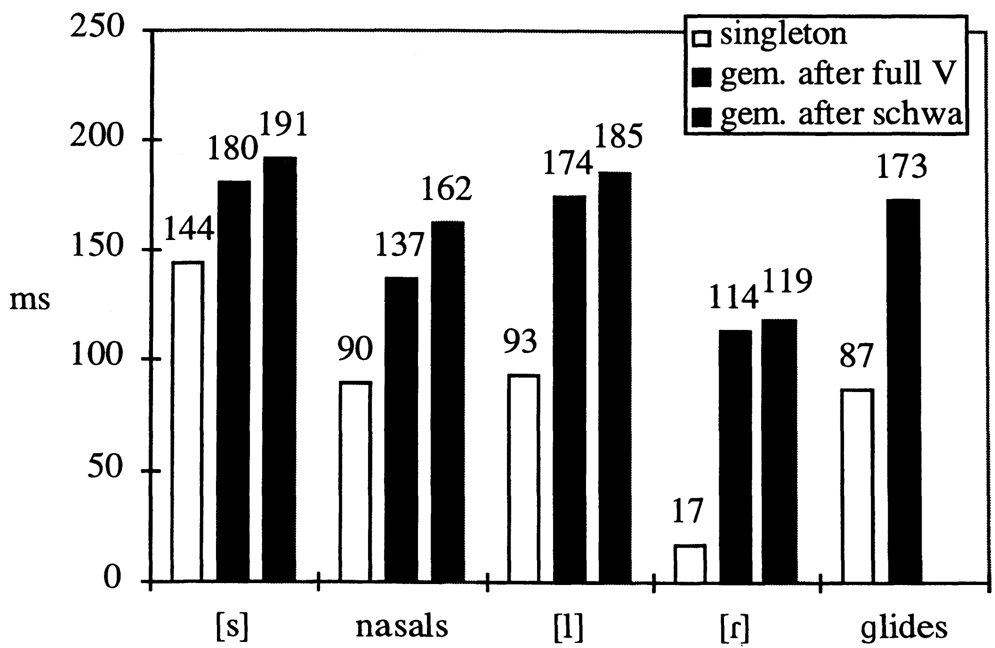

Figure 17. Mean durations for singleton vs. geminate fricatives and sonorants in word-medial position (speaker S)

For both speakers, the durational difference for fricative /s/ is the smallest, while that for tap [ɾ] vs. trill [r] is by far the largest. Here again, reliable statistical comparison was hampered by small sample sizes (n = 3 in all categories except nasals). Singleton-to-geminate durational ratios for all categories in Figures 16 and 17 are provided for each speaker in Table 3:

	speaker E	speaker S
V+[s]/[s:]	1:1.28	1:1.33
[ə]+[s]/[s:]	1:1.43	1:1.57
V+nasal	1:1.35	1:1.45
[ə]+nasal	1:1.46	1:1.45
V+[l]/[l:]	1:1.44	1:1.47
[ə]+[l]/[l:]	1:1.48	1:1.50
V+[ɾ]/[r]	1:5.09	1:6.71
[ə]+ ɾ]/[r]	1:4.95	1:7.00
V+glide	1:1.29	1:1.33

Table 3. Singleton-to-geminate closure ratios for word-medial fricatives and sonorants.

Here again, we see no evidence suggesting that geminates are shortened following /ə/ — in fact, as with the stops, mean durations tend to be longer in this environment. This suggests that there may be a tendency in Madurese for vowel and consonant durations to compensate for one another, such that vowel onsets are spaced at roughly equal intervals. Equal temporal spacing of this sort is typical of syllable-timed languages (see, e.g., Smith 1992 and Ham 1998 for overviews, and Ham 1998 for a detailed study of Madurese in this regard), which are also characterized by closed syllable shortening (see Maddieson 1985). As shown in Figure 18 for the vowel [a], vowels in closed syllables are in fact shortened in Madurese by an average of about 45%:

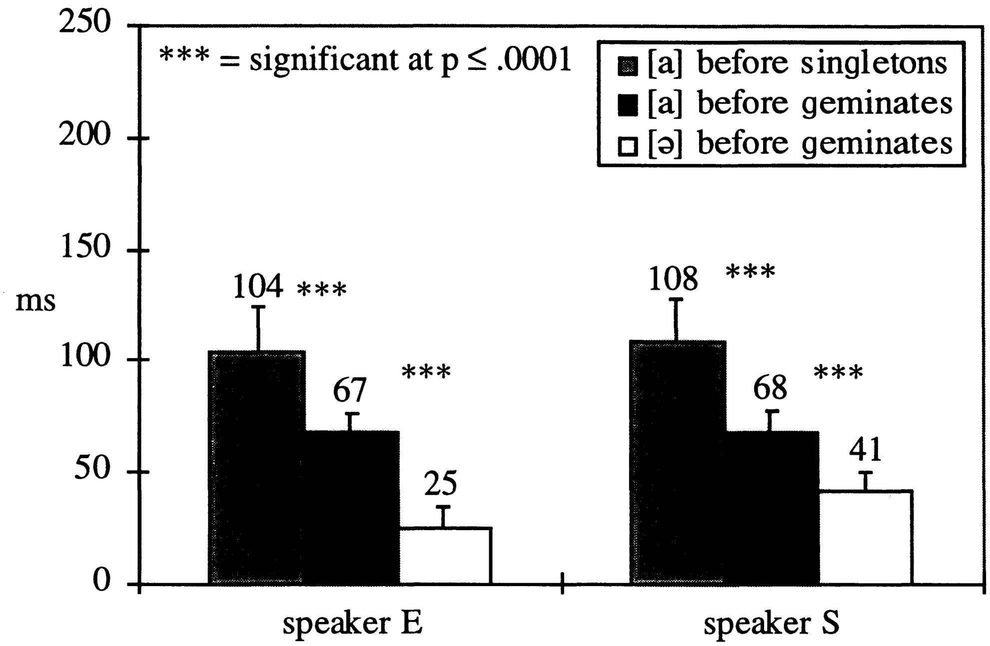

Figure 18. Mean durations for [a] in open (n = 108 per speaker) and closed syllables (n = 68 per speaker) and for [ə] in closed syllables (n = 36 per speaker)

Averaging over speakers and consonant types, geminates in the Madurese corpus are approximately 60% longer than their singleton counterparts. Taken together with the closed syllable shortening effects observed here, this figure indicates that the duration of VC and VC: sequences is kept more or less constant. Given such a timing strategy, we seem to have a plausible explanation for why geminates do not shorten preceding /ə/ (where, it will be recalled, they do not contrast with singletons): Because /ə/ is so short, reducing the duration of the closure phase in this environment would disrupt the overall syllable-timed pattern.

5 Conclusion

The present study has provided an overview of the durational characteristics of Madurese consonants across all places and manners of articulation, as well as across all positions in the word. However, this work only represents a first step in the acoustic study of Madurese consonants, and has highlighted a number of areas for future research. First is the question of why voiced stops are as long as they are, and why they consistently show the largest durational gap between singleton and geminate of any of the three stop types. We have hypothesized here that an articulatory enhancement strategy of some type

— perhaps velic leakage — is involved in prolonging stop voicing, but this can of course be confirmed only through further acoustic and articulatory studies. A second issue briefly touched upon here is the fact that voiced geminates appear to be preceded only by the short vowel /ə/, which might lead us to wonder whether this, too, represents a durational enhancement of voiced geminate stop closure relative to the duration of the preceding vowel. This, as well as the durational boundary between singletons and geminates of all types, can be tested by perceptual studies centered around both identification and discrimination tasks. A third question which could not be addressed in the present study for lack of a sufficiently large data corpus is whether the neutralization of the three stop phonation types to voiceless in word-final position is in fact complete, or continues to show small but systematic differences. And finally, we might also wonder whether the unexpected tendency observed here for word-initial consonants to be shortest carries over to a variety of prosodic environments (e.g. word-initial as well as phrase initial).

6 References

Banner-Inouye, S. (1995) *Trills, Taps and Stops in Contrast and Variation.* PhD dissertation, University of California at Los Angeles.

Browman, C. and L. Goldstein (1992) Articulatory phonology: an overview. *Phonetica* 49, 155-180.

Byrd, D. (1996) Influences on articulatory timing in consonant sequences. *Journal of Phonetics* 24, 206-244.

Byrd, D., S. Narayanan, A. Kaun and E. Saltzmann (1997) Phrasal signatures in articulation. To appear in *Proceedings of Laboratory Phonology V*. Cambridge (UK): Cambridge University Press.

Cohn, A. (1993) Consonant-vowel interactions in Madurese: the feature lowered larynx. *CLS* 29, 105-119.

Cohn, A. and C. Lockwood (1994) A phonetic description of Madurese and its phonological implications. *Working Papers of the Cornell Phonetics Laboratory* 9, 67-92.

Fougeron, C. and P. Keating (1997) Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America* 101, 3728-3740.

Fujimura, O. and D. Erickson (1997) Acoustic phonetics. In W. Hardcastle and J. Laver (eds.) *Handbook of the Phonetic Sciences.* Oxford: Blackwell, 65-115.

Henton, C., P. Ladefoged and I. Maddieson (1992) Stops in the world's languages. *Phonetica* 49, 65-101.

Ham, W. (1998) Phonetic and Phonological Aspects of Geminate Timing. PhD dissertation, Cornell University.

Hubbard, K. (1994) *Duration in Moraic Theory*. PhD dissertation, University of California at Berkeley.

Huffman, F. (1976) The register problem in fifteen Mon-Khmer languages. *Austronesian Studies Part I: Oceanic Linguistic Serial Publication* 13, 575-589.

Iverson, G. and J. Salmons (1996) Mixtec prenasalization as hypervoicing. *International Journal of American Linguistics* 62, 165-175.

Jun, J. (1993) *The Phonetics and Phonology of Korean Prosody*. PhD dissertation, Ohio State University.

Kiliaan, H. (1897) *Madoereesche Spraakkunst, Eerste Stuk: Inleiding en Klankleer.* Batavia: Landsdrukkerij.

Klatt, D. (1975) Voice onset time, frication, and aspiration in word-initial consonant clusters. *Journal of Speech and Hearing Research* 18, 685-706.

Klatt, D. (1976) Linguistic uses of segmental duration in English. *Journal of the Acoustical Society of America* 59, 1208-1221.

Krakow, R. (1989) *The Articulatory Organization of Syllables: A Kinematic Analysis of Labial and Velic Gestures*. PhD dissertation, Yale University.

Ladefoged, P. and I. Maddieson (1996) *The Sounds of the World's Languages*. Oxford: Blackwell.

Lisker, L. (1957) Closure duration and the intervocalic voiced-voiceless distinction in English. *Language* 33, 42-49.

Lisker, L. and A. Abramson (1964) A cross-linguistic study of voicing in initial stops: acoustical measurements. *Word* 20, 384-422.

Lisker, L. and A. Abramson (1971) Distinctive features and laryngeal control. *Language* 47, 767-785.

Maddieson, I. (1985) Phonetic cues to syllabification. In V. Fromkin (ed.) *Phonetic Linguistics: Essays in Honor of Peter Ladefoged*. New York: Academic Press, 203-221.

Nearey, T. and B. Rochet (1994) Effects of place of articulation and vowel context on VOT production and perception for French and English stops. *Journal of the International Phonetic Association* 24, 1-19.

Ohala, J. (1983) The origin of sound patterns in vocal tract constraints. In J. MacNeilage (ed.) *The Production of Speech*. New York: Springer Verlag, 189-216.

Ohala, M. and J. Ohala (1991) Nasal epenthesis in Hindi. *Phonetica* 48, 207-220.

Peterson, G. and I. Lehiste (1960) Duration of syllable nuclei in English. *Journal of the Acoustical Society of America* 32, 693-703.

Pierrehumbert, J., D. Talkin, O. Fujimura, L. Goldstein and I. Vogel (1992) Lenition of /h/ and glottal stop. In G. Dochtery and R. Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, Segment and Prosody*. Cambridge (UK): Cambridge University Press, 90-127.

Pinnow, H-J. (1979) Reflections on the history of the Khmer phonemic system. *Mon-Khmer Studies* 8, 103-130.

Port, R. and R. Rotunno (1979) Relation between voice-onset time and vowel duration. *Journal of the Acoustical Society of America* 66, 654-662.

Smith, C. (1992) *The Timing of Vowel and Consonant Gestures*. PhD dissertation, Yale University.

Stevens, A. (1966) The Madurese reflexes of Proto-Malayo Polynesian. *Journal of the American Oriental Society* 86, 147-156.

Stevens, A. (1968) *Madurese Phonology and Morphology*. American Oriental Series Vol. 52. New Haven: American Oriental Society.

Stevens, A. (1980) Formative boundary in phonological rules. In M. Aronoff and M.-L. Kean (eds.) *Junctures*. Saratoga: Anma Libri, 135-141.