Working Papers of the Cornell Phonetics Laboratory
1994, vol. 9, pp. 67-92.

A phonetic description of Madurese and its phonological implications*

Abigail C. Cohn
and
Katherine Lockwood
Eloquent Technology Inc.

Madurese, an Austronesian language of Indonesia, shows a systematic interaction of consonant phonation and vowel height. In this paper, we explore this interaction through an acoustic study of the consonants, vowels, and syllable structure of Madurese. A large effect of the consonant-vowel interaction is observed in terms of F1, where vowels following both the voiced and aspirated stops show systematically lower F1 values than those following voiceless stops and nasals. Systematic differences in F0 are also observed, with voiced and aspirated stops having a lowering effect on F0 and voiceless stops and nasals having a raising effect.

1. Introduction

Madurese (a Western Austronesian language, spoken in Indonesia, on Madura and parts of East Java) exhibits a number of interesting phonological properties, regarding both the inventory of sounds and phonological patterning observed. These include a three-way phonation contrast with direct effects on the quality of the following vowels, as well as a singleton-geminate contrast throughout the consonant inventory. These properties are of interest in and of themselves and because of differences with closely related neighboring languages. The present study is an acoustic analysis of both the consonants and vowels of Madurese including the following: (1) an examination of the acoustic properties of the Madurese consonants, in particular the realization of the three-way phonation contrast and its effect on the vowel system; (2) an investigation of the vowel system of Madurese; and (3) a consideration of the timing patterns of singleton vs. geminate consonants and their relationship to syllable structure.

The goal of this study is a comprehensive description of the sound properties of Madurese in order to better understand the unusual pattern of consonant phonation-vowel quality interactions observed. The structure of the paper is as follows. First in §2, phonological background on the sound system of Madurese is presented, based on the work of Kiliaan (1897), Stevens (1968), and Cohn (1993a, b). In §3 the structure and

* This is a revised version of a paper presented at the Seventh International Conference on Austronesian Linguistics, Leiden, Netherlands, August, 1994. Thanks to the 7ICAL audience and the members of Cornell phonetics-phonology group for useful comments and discussion, and specially to Allard Jongman and Becky Letterman for comments on an earlier draft. Thanks also to Sue Hertz and Eloquent Technology Inc. for technical support and to the American Council of Learned Societies for financial support for travel to 7ICAL. This research was supported in part by NSF grant # BNS-9111155 to the first author.

methodology of an acoustic study are outlined and in §4 the results are reported. We conclude in §5 with a discussion of these results and suggestions for further research.

2. The sound system of Madurese

The consonant and vowel inventories of Madurese are presented in Tables 1 and 2 respectively.

		labial	dental	alveolar	palatal	velar	glottal
stops	voiceless	p	t̪	ṭ	c	k	
	aspirated	pʰ	t̪ʰ	ṭʰ	cʰ	kʰ	
	voiced	b	d̪	ḍ	j	g	
nasals		m		n	ɲ	ŋ	
fricative				s			
approximants		(w)		r, l	y		ʔ, (h)

Table 1. Madurese consonant inventory (following Stevens 1980).

Several aspects of the consonant system are noteworthy. There are five oral places of articulation: labial, dental, alveolar (or retroflex), palatal (lightly affricated), velar, and [ʔ, h]. (/n/ is homorganic with both the dentals and alveolars.) This is a richer system of place contrasts than that observed in the closely related languages of Sundanese and Indonesian, but similar to the system observed in Javanese, a closely related and geographically proximate language.

Of considerable interest is the fact that all consonants, except glottal stop, can be geminated in intervocalic position. All vowels, except schwa, can appear before either a singleton or geminate consonant; thus medial gemination is contrastive, except after schwa. None of the closely related neighboring languages have geminates, although some of the languages of Sulawesi (e.g. Buginese and Macassarese) exhibit geminates, as do certain dialects of Malay. As argued by Stevens (1966), there are a number of historical sources for these medial geminates, including gemination after schwa, total assimilation of nasal-stop clusters, and gemination of some intervocalic consonants.

Also noteworthy is the three-way phonation contrast in the stops: voiceless, aspirated, and voiced. None of the neighboring languages show such a three-way contrast, e.g. Javanese, Sundanese, and Indonesian all exhibit two-way phonation contrasts (breathy-plain in the case of Javanese and voiced-voiceless in Sundanese and Indonesian). This three-way contrast results from a historical change whereby the series of voiced stops became aspirated stops; the current voiced stop series derives from glide hardening (/w/ > /b/ and /y/ > /j/) borrowings, and possibly a split in the labials (Stevens 1966). Some researchers transcribe the aspirated series as voiced aspirates, but impressionistically they do not appear to us to be voiced.

The basic vowel inventory of Madurese is shown in Table 2.

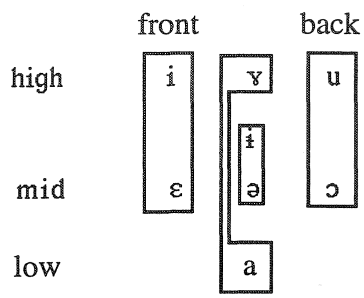

Table 2. Madurese vowel inventory (following Stevens 1980).

As described by Kiliaan (1897) and Stevens (1968, 1980) and shown in Table 2, the surface vowel system of Madurese consists of eight basic vowels. These constitute four alternating pairs, as shown in (1).

(1) Alternating pairs

[-high]	[+high]
ɛ	i
ɔ	u
a	ɤ
ə	ɨ

These two sets of vowels are in complementary distribution. Impressionistically the difference between each of these pairs is vowel height, so we refer to the one set as [-high] and the other as [+high], where [high] is taken as a neutral descriptive term. Observed patterns of alternation are exemplified in (2).

(2)

following voiceless stops/nasals/Ø		following voiced/aspirated stops	
-high		+high	
a. ɛ ~ i			
[pɛlɛ]	'choose'	[bilis]	'ant'
[mɛlɛ]	'choose, act.'	[pʰiris]	'whiskers'
[ɛpar]	'brother/sister-in-law'		
[lɛʔɛr]	'neck'		
b. ɔ ~ u			
[pɔla]	'perhaps'	[bulɤn]	'moon'
[kɔt̪ɔ]	'louse'	[kʰuʷɤ]	'cave'
[ɔbuʔ]	'hair'		
c. a ~ ɤ			
[paraʔ]	'almost'	[bɤrɤʔ]	'west'
[mala]	'misfortune'	[pʰɤlɤ]	'family'
[abu]	'dust'		
[rama]	'male term of address'		
d. ə ~ ɨ			
[pəllaŋ]	'endure'	[bɨlli]	'buy'
[əmpaʔ]	'four'		

In (2) we observe the phonological alternation between [ɛ~i], [ɔ~u], [a~ɤ], and [ə~ɨ]. Non-high vowels appear after voiceless stops, nasals, and word-initially. High vowels appear after voiced and aspirated stops. Word-initial liquids are followed by non-high variants, while word-medially, [r, l, ʔ] are transparent; in these cases a following vowel has the same height as a vowel preceding the medial consonant. (For a discussion of additional intricacies involving [s, h, y, w], see Cohn 1993b.) Thus the phonation of a preceding consonant conditions the height of a following vowel. Furthermore the consonantal effects may propagate over a long distance span: the quality of a consonant may affect a sequence of vowels, or a sequence of vowels with intervening liquids or a glottal stop. Systematic morphological alternations are also observed, as exemplified in (3).

(3) Morphological alternations

[palar]	'mean'	[malar]	'mean, act.'
		[N - palar]	
[bɤlɤ]	'speak'	[malaɛ]	'speak to, act.'
		/N - bɤlɤ - E/	
[bɤli]	'return'	[abɤliʔi]	'return for, act.'
		/A - bɤli - E/	

In (3) we see that the quality of a consonantal prefix conditions the quality of a root vowel. In a simple root, the vowel quality is determined by the voicing of the root initial consonant, e.g. [palar] 'mean' ~ [bɤlɤ] 'speak'. When this consonant becomes a nasal in the process of Nasal Substitution, marking the active form of a verb, the initial vowel of a voiced consonant initial root becomes [-high]. The quality of a vowel in a suffix is conditioned by the quality of the preceding vowel (as conditioned by the preceding consonant), as seen in the suffix /-E/, surfacing as [ɛ] after a root ending in a [-high] vowel and [i] after a [+high] vowel. Thus prefixes trigger alternations in roots and roots are responsible for alternations in suffixes.

As exemplified in (2) and (3), vowel height in Madurese is predictable, based on the quality of the preceding consonant or vowel. As argued by Stevens (1968), vowel height need not be represented in the underlying representation, as it can be accounted for by rule. We will informally refer to this rule as *Vowel Raising*. This leaves us with the following question: What features do the voiced and aspirated stops share (to the exclusion of the voiceless stops and nasals) that is spread to vowels to trigger this pattern of height assimilation?

Cohn (1993b) argues, based on phonological evidence, that the responsible feature is (1) a laryngeal feature, due to transparency resulting from coda neutralization, parallel to coda neutralization of voicing and aspiration; and (2) binary, due to the active behavior of both the "lowering" and "raising" consonants. Cohn (1993a, b) tentatively concludes that the observed effects are "register" differences, similar to the distinct vowel sets observed in the Mon-Khmer languages. Many of the Mon-Khmer languages show a contrast in the vowels between a breathy voice set and a clear set. These differences, referred to as *register*, have been shown to arise historically from the phonation of the preceding consonant (Huffman 1976; Pinnow 1979). Similar differences have been observed in Austronesian languages, including Cham (Edmondson and Gregerson 1993) and Javanese (Hayward 1993). Phonetic descriptions of Mon-Khmer register systems suggest that these two distinct sets are a result of larynx height and concomitant adjustments of other laryngeal muscles. The acoustic realization of register includes height (F1), pitch (F0) and voice quality differences in vowels.

In addition to the process responsible for these vowel height alternations, there are two other processes affecting vowel quality in Madurese. First, the vowel quality of [i, u] depends on syllable structure. This is shown in the following examples (Stevens 1968, p. 37):

(4) "Tense/Lax" in Madurese:

a.	íC/	—>	[ɪ]C/	e.g. ghínzhél 'kidney'	[k^hɪnc^{h}ɨl]
b.	other í	—>	[i]	e.g. ghíghí 'tooth'	[k^hik^{h}i]
c.	úC/	—>	[ʊ]C/	e.g. ghúndhúŋ 'bunch'	[k^hʊnt^{h}ʊŋ]
d.	other ú	—>	[u]	e.g. ghúrú 'teacher'	[k^huru]

Impressionistically, these vowels are "lax" in closed syllables and "tense" in open syllables. This suggests that the feature responsible for the height alternations is distinct from the feature Tense/Lax (or ATR, if these are assumed to be the same feature, see e.g. Kenstowicz 1994).

There is also a process of raising or tensing of the vowels /ɛ, ɔ/ after a nasal consonant as shown by the following examples:

(5)	pekang	[pɛkaŋ]	'to pinch'	mekat	[mekat]	'to trap'
	poka	[pɔka]	'to break'	moka	[moka]	'to break, act.'

Thus in addition to determining the feature responsible for Vowel Raising, we need to study its relationship to these other processes.

3. Acoustic Study

We turn now to the acoustic study undertaken to address these issues. In (6) we summarize the phonetic questions considered:

(6) a. Consonants:
voicing during closure
duration
aspiration
b. Vowels:
formant structure (F1, F2)
duration
vowel pitch (F0)
c. Role of syllable structure:
realization of geminates
open vs. closed syllables

3.1. Subjects

In the present study, the speech of two speakers (1 male (S), 1 female (E)) from E. Madura were investigated. They are both speakers of the "standard dialect" and their speech is very similar. Analysis was done from acoustic recordings made in E. Java, Indonesia, in a sound-treated room.

3.2. Stimuli

All tokens analyzed in this study are real words of Madurese of the shape CV́(C)CV(C), with stress on the first syllable. The first consonant included voiceless (P), aspirated (Ph), voiced stops (B), and nasals (M); in most cases labial, to control for differences due to place of articulation, since phonation and manner of articulation were of primary interest. The full system of eight surface vowels was studied: [+high] - [i, u, ɤ, ɨ] and [-high] - [ɛ, ɔ, a, ə]; the [+high] set after aspirated and voiced initial consonants and the [-high] set after voiceless stops, nasals, and word-initially. Each vowel was investigated in both open syllable position (followed by a syllable-initial single voiceless stop (Ø$P)), and closed position (followed by a nasal-stop cluster (M$P) or geminate (P$P)); except for [ɨ, ə] which only occur in closed syllables. The structure of the tokens is schematized in (7).

(7)	C1	V1	C2$C3	(V2 C4)
	{B, Ph}	{i, u, ɤ, ɨ}	{ØP, MP, P$P}	
	{Ø, P, M}	{ɛ, ɔ, a, ə}	{ØP, MP, P$P}	

The tokens were said in the frame [bɤca ______ saparo] 'Read ______ partway'. Analysis of two repetitions of each token for each speaker is presented here, for a total of 68 tokens per speaker. See the appendix for a complete word list.

3.3. Analysis

The analysis was conducted in two phases: acoustic analysis and database querying (see Hertz et al. 1994). The acoustic analysis was carried out using Entropics Waves+/ESPS on a Sun Workstation. The utterances were digitized at 11,025 Hz and spectrograms and wave forms were generated. These were labeled, indicating onsets and offsets of segments, voicing and aspiration. The acoustic information was extracted via customized software, developed at Eloquent Technology Inc. and entered into a database implemented with Borland International's Paradox software. This database was queried for acoustic information (duration, formant structure, F0) in conjunction with abstract linguistic information.

The results obtained are very consistent across the two speakers. Duration data have, in some cases, been pooled for the two speakers, but formant values and F0 values are reported separately, as there is a large difference in the fundamental frequencies of the two speakers.

4. Results

In discussing the results, we consider first the properties of consonants (§4.1), then vowels (§4.2), and then the role of syllable structure (§4.3).

4.1. Consonants

Of interest in the consonant system is the nature of the three-way phonation contrast between voiceless, aspirated, and voiced stops, with particular attention to what phonetic properties might be shared by the voiced and aspirated stops to the exclusion of the voiceless stops and nasals. First we consider the question of voicing during closure. In particular, are the aspirated stops voiced, indicating their historical source as voiced stops and identifying at least one property shared by voiced stops and aspirated stops? Representative wide band spectrograms of each consonant type are presented in Figure 1.

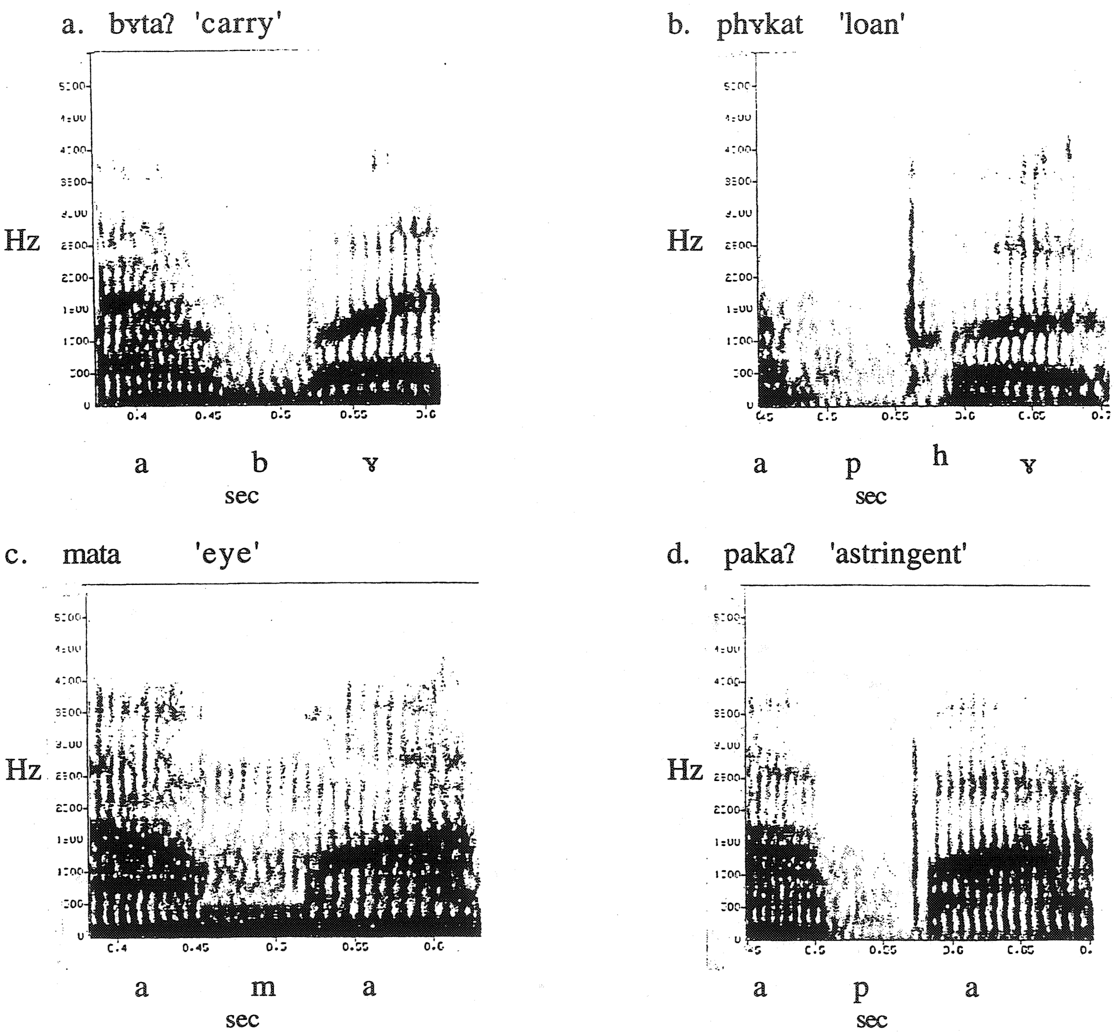

Figure 1. Wide band spectrograms of each Madurese consonant type, produced by S.

As illustrated in Figure 1a, the voiced stops are fully voiced, as are the nasal consonants (Figure 1c). On the other hand, there is no evidence of voicing during the aspirated stops. Both [p^h] and [p] show partial voicing at the onset of the consonant, as seen in Figure 1 b and d respectively; but these patterns are very similar and we conclude that this is due to the context of a preceding vowel. Both [p^h] and [p] are characterized by clear bursts, although impressionistically the bursts for [p^h] are of greater intensity.

Consider now consonant duration: how do the stops compare on closure duration and VOT? Average closure and burst durations and VOT values are presented by type in Table 3 and Figure 2.

Segment type	Closure duration	burst	VOT	Total
b	76	5	-	81
p	80	6	11	98
ph	90	7	25	121
m	77	-	-	77

Table 3. Consonant closure and VOT durations across speakers (in ms).

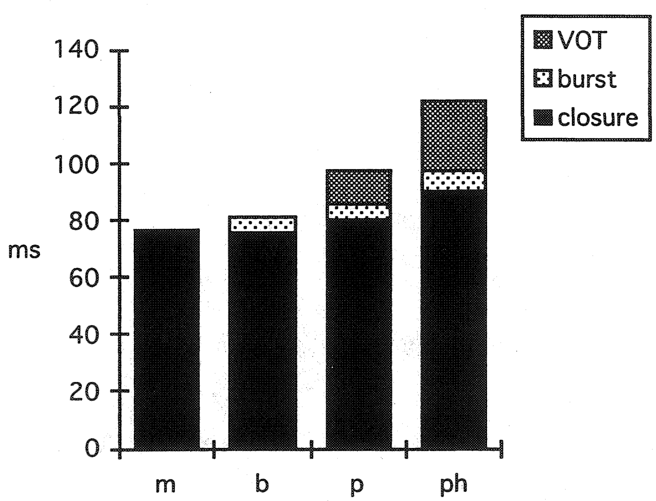

Figure 2. Closure and VOT durations for each consonant type across speakers (in ms).

Closure duration is quite consistent across the four types, except that [p^h] is slightly longer. Somewhat surprising is the relatively long duration of [b], considering the robustness of the voicing throughout. The burst durations are similar across the three stop types. The VOT difference between [p^h] and [p] is quite small, with [p^h] showing only moderate aspiration. The total average difference in duration between [p^h] and [p] is distributed almost equally between the closure and VOT differences. Overall the duration differences between [p^h] and [p] are quite small, but it is important to bear in mind that these two can always be distinguished by the quality of the following vowel[1].

[1] In using labials to study these consonant characteristics, we would expect to find the smallest VOT values and thus smaller differences between /p^h/ and /p/ and the greatest ability to maintain voicing during closure, thus long /b/ closures. In a fuller study of the Madurese consonant system now in progress, these characteristics for all places of articulation will be investigated.

4.2. Vowels

We turn now to the acoustic characteristics of the vowels, investigating duration, formant structure, and fundamental frequency. In terms of duration, cross-linguistically, it is observed that the lower the vowel height, the longer the vowel, thus considering intrinsic duration we would expect [a] > [ɛ, ɔ] > [i, u]. In order to understand the nature of the consonant-phonation interactions, it is interesting to see if there are duration effects other than these expected intrinsic differences. Also of interest is the duration of both [ə, ɨ], as their phonotactic patterning is different from the other vowels and in many languages schwa is very short. Average durations for each vowel and for all [-high] and all [+high] vowels are presented by speaker in Table 4 and Figure 3[2].

	V	S	E
[-high]	ɛ	77	66
	ɔ	66	87
	a	92	77
	ə	33	25
	all [-high]	69	66
[+high]	i	75	76
	u	67	73
	ɤ	83	87
	ɨ	29	36
	all [+high]	66	70

Table 4. Average duration by vowel by speaker (in ms).

2 The vowel durations of the [-high] vowels in word-initial position have been excluded from these averages, as they would skew the number of tokens of [-high] vowels in open syllables.

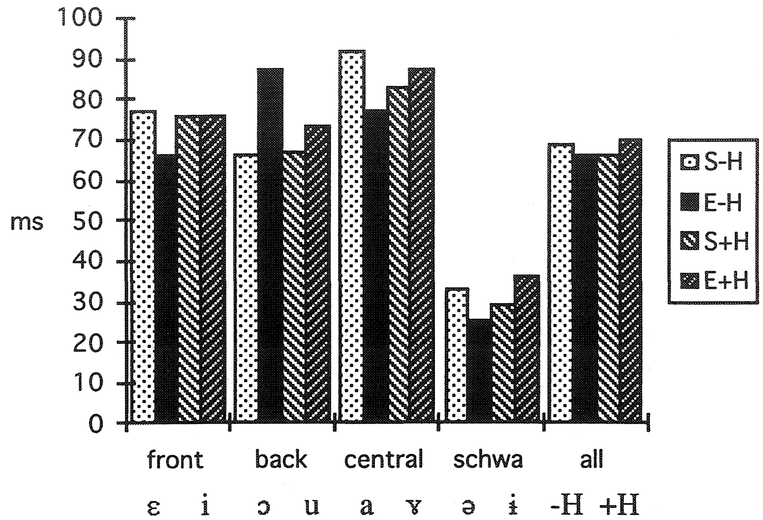

Figure 3. Average duration by vowel by speaker (in ms).

While [a] is longer than the non-low front and back vowels (except for [ɔ] for E) and all full vowels are considerably longer than the schwas [ə, ɨ], there are no significant differences within any of the alternating pairs—[ɛ, i], [ɔ, u], [a, ɤ], and [ə, ɨ]—for either speaker (based on independent t-tests). In fact, in each case the small differences observed go in the opposite directions for the two speakers. The lack of a significant difference between the [-high] and [+high] vowels may be due in part to the small number of tokens, but these results suggest that duration is not one of the primary factors distinguishing between the two sets. Both [ə] and [ɨ] are very short in duration, as expected. Surprising is the very long duration of [ɤ], a high (or mid) vowel; this may be in order to maintain the difference between it and [ə] and [ɨ], which impressionistically have similar qualities.

These duration results provide evidence against a tense/lax interpretation of the two vowel sets, since one of the defining properties of lax vowels is that they are shorter than tense vowels. Also, we do not see the sorts of duration differences which are sometimes observed in ATR systems, where [+ATR] vowels are longer than their [-ATR] counterparts (see e.g. Hess 1992).

We turn now to vowel quality. Measurements were made for F1, F2, and F3 at three points in the vowel. If the alternation is a register difference, differences in F1 (the primary acoustic correlate of vowel height) are predicted. Differences in F2 might also be expected. Also of interest are the formant patterns of the central vowels, as impressionistically there are three vowels with very similar vowel quality. Data presented here show F1 and F2 at the vowel midpoint. Average F1 and F2 values by speaker are presented in Table 5 and plotted in Figures 4 and 5 for S and E respectively.

	S F1	F2	E F1	F2
ɛ	422	2223	660	2282
ɔ	436	878	731	1119
a	870	1308	1036	1679
ə	512	1090	730	1370
all -H	567	1368	800	1629
i	310	2240	474	2690
u	304	754	456	1045
ɤ	412	1274	595	1622
ɨ	386	1088	488	1412
all +H	351	1356	505	1711

Table 5. Average F1, F2 values by speaker (in Hz).

In Figures 4 and 5, the inverse of F1 is plotted on the y axis and the inverse of F2 is plotted on the x axis. The average values for each vowel for each speaker are shown and the ellipses around these values indicate two standard deviations from the average. The vertical lines match up the four pairs of vowels. These values for the two speakers are plotted separately, as the formants for the two speakers are quite different; but the relationships between the vowel sets are very similar. Not only are the averages distinct, but as seen by the ellipses, the individual values for both speakers for each of the three pairs, [ɛ, i], [a, ɤ] and [ɔ, u] are largely distinct. However, the ranges of values for [ə, ɨ] are much greater, overlapping the acoustic space of neighboring vowels.

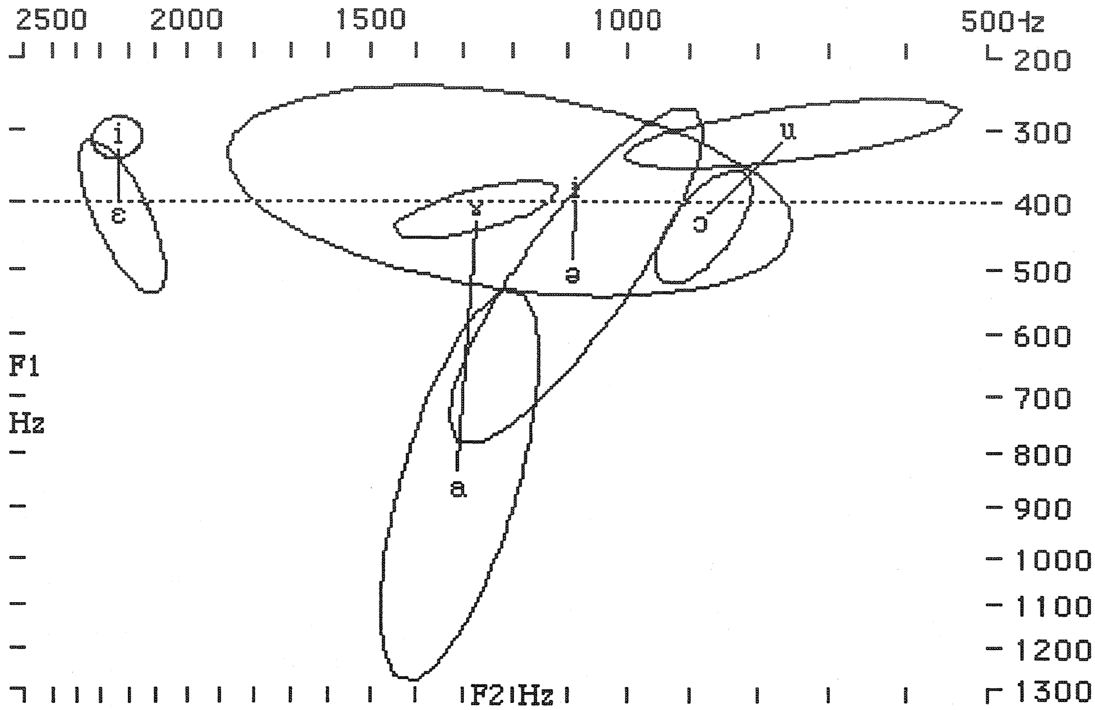

Figure 4. Formant plot of average values for F1 vs. F2 for S.

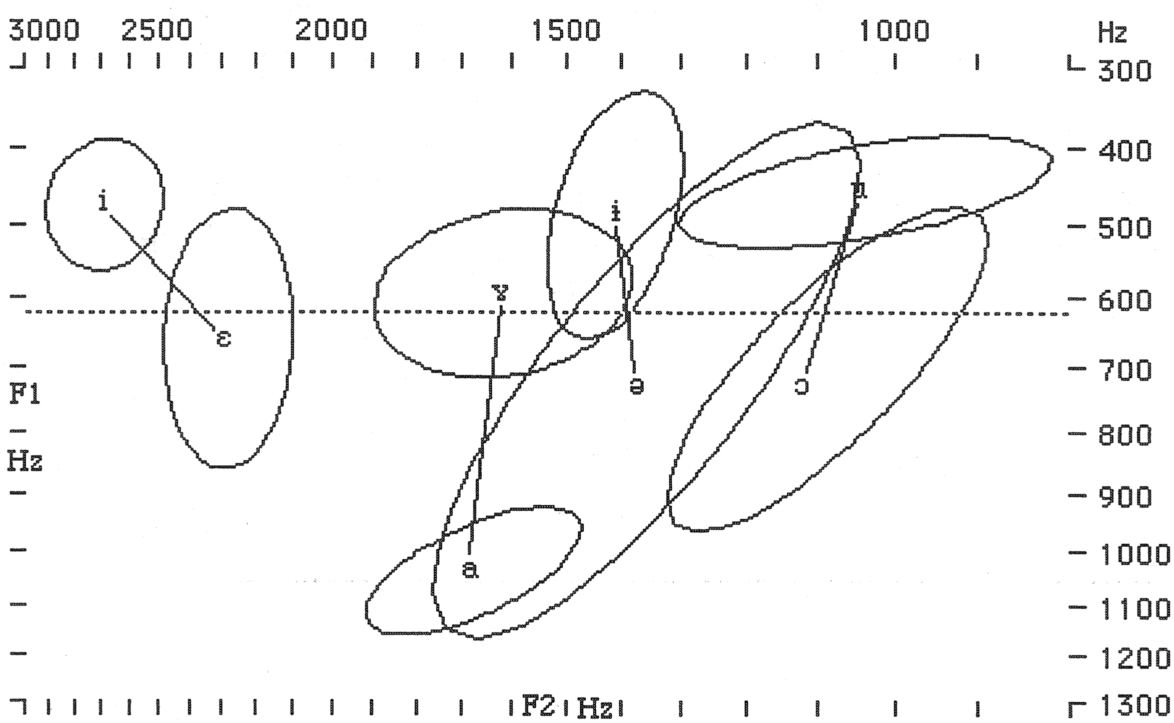

Figure 5. Formant plot of average values for F1 vs. F2 for E.

There are dramatic differences in F1 and very minimal differences in F2. As shown by the horizontal dashed line, the two vowel sets can basically be distinguished by F1 alone. These F1 differences are very large, much larger than those usually seen in either tense/lax or ATR systems. The magnitude of these differences can be seen in the average F1 values for all the [-high] and [+high] vowels pooled together as presented in Table 5. These differences are consistent with a register interpretation[3].

The less consistent formant patterns of [ə, ɨ] are not surprising, considering their very short duration, thus greater coarticulatory effects from the adjacent consonants would be expected. It might also be argued that these vowels are unspecified or less fully specified than the other vowels. For S, the acoustic space for [ɤ] is completely contained in the space for [ɨ], while for E, there is considerable overlap; this overlap may indeed account for the surprising duration of [ɤ] noted above.

Related to the question of vowel quality is the effect of a preceding nasal on vowel quality. As noted above, impressionistically [ɛ, ɔ] are raised or tensed after a nasal consonant. We can see the consequences of this in the acoustic signal by comparing vowel quality after [m] vs. [p]. This information is presented in Table 6 and plotted in Figures 6 and 7. (The vowels [ə, ɨ] are not included here, due to their variable character as discussed above.)

	S F1	S F2	E F1	E F2
pɛ	422	2186	689	2379
mɛ	338	2352	537	2228
Δ	84	-166	152	151
pɔ	434	862	745	1135
mɔ	378	923	576	941
Δ	54	-61	169	194
pa	852	1264	1029	1667
ma	845	1347	1087	1805
Δ	7	-83	-58	-138

Table 6. Average formant values after nasals vs. voiceless stops (in Hz).

Based on the average formant values shown in Table 6, there are clear differences in F1 for both [ɛ] and [ɔ] after [p] vs. [m] for both speakers, while the differences for [a] are not

[3] If register differences are due to larynx height, we would also expect F2 differences, due to the lengthening of the vocal tract. The absence of systematic F2 differences warrants further investigation.

systematic (with a small positive difference for S and a somewhat larger negative difference for E). There are also differences in F2, but these do not show a systematic pattern, with decreases for S and increases for E, except in the case of [a].

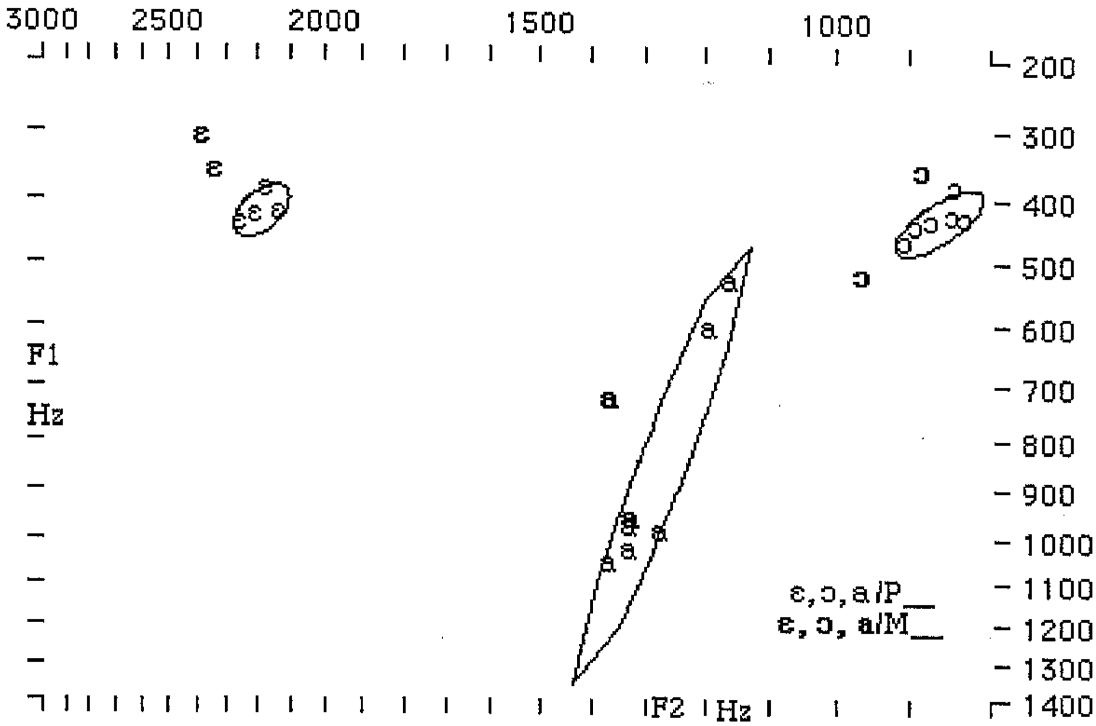

Figure 6. Formant plot of vowels after nasals vs. voiceless stops for S (in Hz).

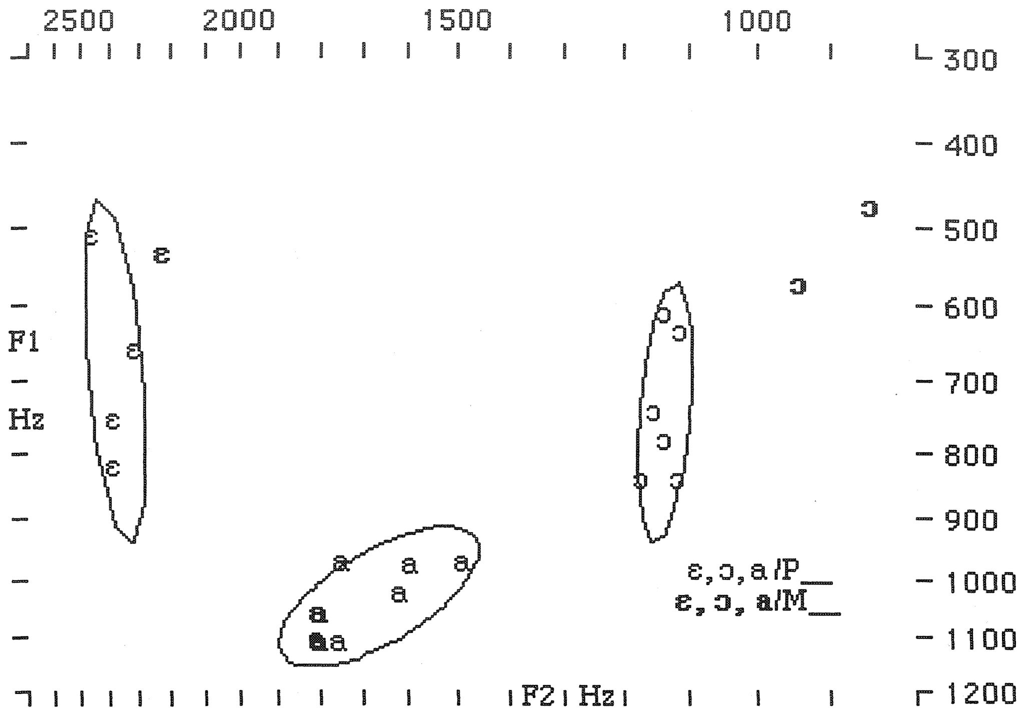

Figure 7. Formant plot of vowels after nasals vs. voiceless stops for E (in Hz).

The individual tokens are plotted in the Figures 6 and 7 for S and E respectively. The vowel tokens following [p] are in regular type enclosed by ellipses (indicating two standard deviations); the vowels following [m] are shown in bold (no ellipses could be draw as there was only one token with two repetitions for each speaker, with only one repetition in the case of [ɛ] for E). For both speakers, we see that the F1 values of [a] after nasals fall completely within the range of F1 values after voiceless stops. But most of the F1 values for [ɛ, ɔ] are lower after nasals that voiceless consonants. These results suggest differences in height for the mid vowels following a nasal consonant or voiceless stop. This appears to be a systematic pattern, but clearly more tokens of vowels after nasals are needed to reach a strong conclusion.

This might cause us to wonder whether nasal consonants were miscategorized as triggering following [-high] variants. Evidence from two difference sources argues that this is not the case, suggesting rather that this observed difference in vowel height is a subpattern within the [-high] vowels. First, this raising occurs only with the mid vowels [ɛ, ɔ] and does not affect [a]; secondly, long distance effects in forms such as /mɛlɛ/ 'choose, active' [melɛ] show that the nasals indeed trigger [-high] vowel variants.

We consider now the effects of a preceding consonant on fundamental frequency of the following vowel. Cross-linguistically, the higher the vowel, the higher the F0; the following intrinsic F0 differences would be expected: [i, u] > [ɛ, ɔ] > [a]. In addition it has been observed that F0 at vowel onset is affected by phonation of the preceding consonant (see e.g. Hombert 1978). F0 is lower after voiced stops or nasals than after voiceless stops. Obstruents and sonorants pattern together in this effect (see Maddieson's 1984 study of F0 effects in Burmese). F0 is higher after voiceless or aspirated stops, but lower after breathy voiced stops, as seen, for example, in Hindi. Finally, in register systems, F0 is lower in the breathy vowel set than in the clear vowel set.

Average F0 values for each vowel and for all [-high] and [+high] vowels are presented by speaker in Table 7 and are plotted separately for each speaker in Figures 8 and 9. Data are presented here for F0 at the vowel onset the vowel midpoint (mid).

	S	onset	mid	Δ	E	onset	mid	Δ
ɛ		133	123	-10		218	226	8
ɔ		137	125	-12		217	222	5
a		129	121	-8		223	218	-5
ə		130	139	9		228	232	4
all -H		132	126	-6		221	224	3
i		117	116	-1		191	213	22
u		112	118	6		194	199	5
ɤ		115	113	-2		179	193	14
ɨ		113	117	4		156	193	37
all +H		114	116	2		181	200	19

Table 7. Average F0 values by vowel by speaker (in Hz).

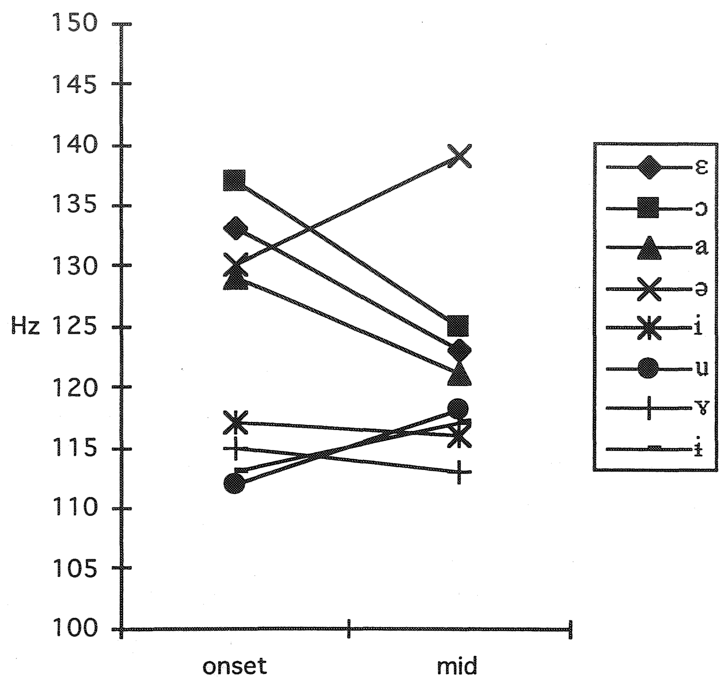

Figure 8. Average F0 values by vowel at the vowel onset and midpoint for S (in Hz).

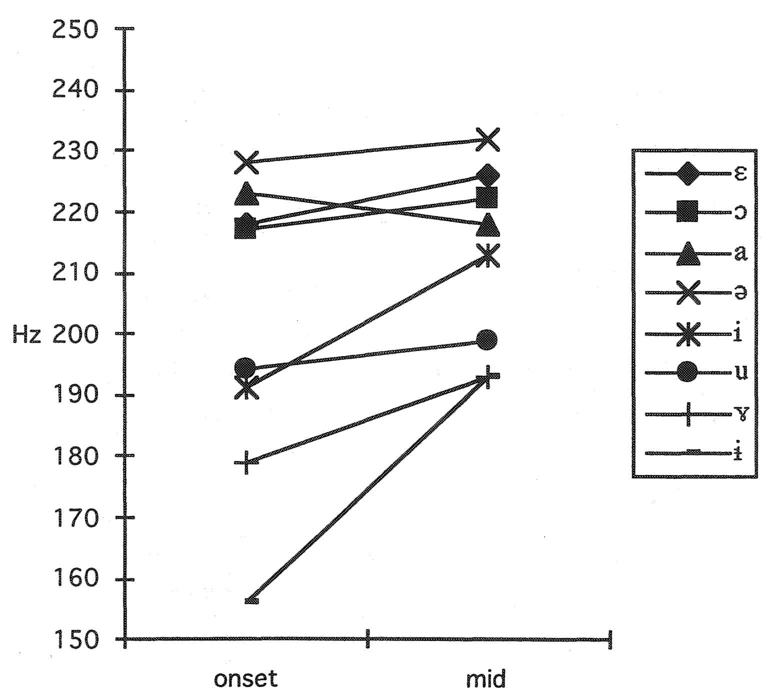

Figure 9. Average F0 values by vowel at the vowel onset and midpoint for E (in Hz).

For both speakers, we observe a large difference for F0 at the vowel onset, with all [+high] vowels showing much lower F0 values than all the [-high] vowels. Clearly these differences go in precisely the opposite direction of predicted intrinsic differences, that is [ɛ, ɔ] have higher, not lower, F0 values than [i, u]. At the vowel midpoint, the values are still distinct, but not as much as at the vowel onset. While these differences hold across both speakers, there is a systematic difference as well: for S, the [-high] values decrease and the [+high] values increase as they approach the vowel midpoint; while for E, both the [-high] and [+high] values increase. Overall the two classes of vowels are distinct in their F0 values, but we need to consider the relative contribution of the preceding consonants. Thus, for example, we want to be sure that the lower F0 values for the [+high] set of vowels is not due just to the effect of preceding voiced stops. The effect of the preceding consonant type is shown in Table 8. and Figure 10.

	S	onset	mid	E	onset	mid
p		134	128		223	226
m		129	124		221	219
b		114	116		184	196
ph		115	118		163	221

Table 8. Average F0 values by preceding consonant type by speaker (in Hz).

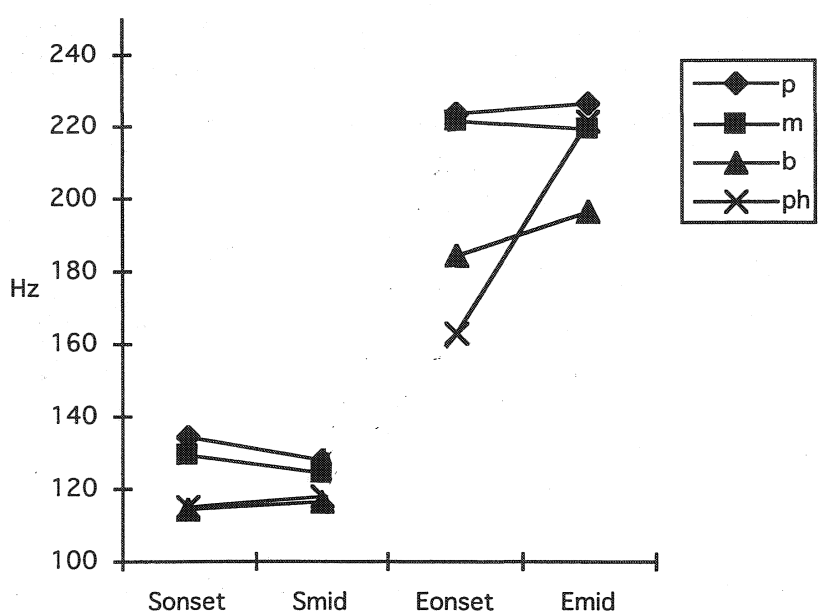

Figure 10. Average F0 values by preceding consonant type by speaker (in Hz).

The preceding consonant was shown to be a highly significant main factor for both speakers, as revealed by one-way ANOVA's (S: [$F(3, 54) = 20.0273$, $p = .0000$], E: [$F(3, 54) = 24.0337$, $p = .0000$]). Post-hoc Student-Newman-Keuls tests showed [p^h] and [b] to be significantly different from [m] and [p]. No significant difference between [m] and [p] was found; [p^h] and [b] were significantly different for E, but not for S. Thus we observe that F0 values are systematically lower after voiced and aspirated stops than after voiceless stops and nasals. At the vowel onset, the differences between the two types are quite large and the values within the groups of voiceless stops and nasals and voiced and aspirated stops are very similar. The averages of these two groups remain distinct at the vowel midpoint, except for the aspirated stops for E. The magnitude of the F0 differences after the two types of consonants are much larger than the differences typically seen from voicing; e.g., in the studies cited by Hombert (1978), the differences seen between voiced and voiceless consonants were typically on the order of 5-7 Hz, with the maximum difference observed being 10 Hz. Here we see differences within vowel pairs of at least 10 Hz for S and differences of 40 Hz or more for E. The differences in F0 are completely consistent with the differences in F1, again supporting a register interpretation.

4.3. The role of syllable structure

We turn now to the effects of syllable structure on vowels and consonants. We compared the structure of open syllables (CV) with closed syllables (CVC), where the coda consonant was either a nasal homorganic to the following stop, or half of a geminate.

First we consider the nature of geminates. As noted above, Madurese has a singleton-geminate contrast intervocalically. In Table 9, the average duration for a single intervocalic consonant (C3, taken to be the onset of the following syllable) is compared with average durations of nasal-stop clusters and geminates (C2 + C3).

	C2	C3	Total
ØP	Ø	120	120
MP	90	90	180
PP	84[4]	87	171

Table 9. Medial consonant duration (in ms).

[4] The geminates were divided in the middle of the closure to allow comparison with the MP cases, except in the few cases where there was some evidence of a slight release, in which case this was taken as the midpoint.

The duration of the nasal-stop clusters and geminates are very similar, about one and one half times longer than the singleton intervocalic consonants. This is quite consistent with cross-linguistic observations of geminate duration. As noted above, there is no geminate-singleton contrast after schwa. Schwas ([ə, ɨ]) only occur in closed syllables. We might thus wonder whether the geminates are truly geminate in this case, or whether their duration might be intermediate between the geminate and singletons in other contexts. In Table 10, geminate durations after schwa are compared with durations after all the other vowels and we see that the geminates after schwa are actually slightly longer.

PP/ ə, ɨ __	178ms
PP/V (except schwa)__	164ms

Table 10. Duration of geminates after schwas vs. other vowels (in ms).

We now turn to the effects of syllable structure on vowels. Impressionistically [i, u] are tense in open syllables and lax in closed ones. If this is borne out in the phonetics, we would expect the lax variants to be shorter and show more centralized formant values and the tense ones to be longer and show more extreme formant values. Durations for vowels and following consonants depending on syllable structure are shown in Table 11[5].

	V1	C2/C3	Total
VØ$P	88	120	208
VM$P	73	169	242
VP$P	62	164	226

Table 11. Duration of vowels and following consonants (full vowels only) (in ms).

Syllable structure was shown to be a highly significant factor as shown by one-way ANOVA's for both speakers (S: [$F(2, 24) = 13.3494$, $p = .0001$], E: [$F(2, 24) = 6.7623$, $p = .0047$]). Post-hoc Student-Newman-Keuls tests showed open syllables (Ø) to be significantly different from both types of closed syllables (M, P) for S; For E, open syllables were significantly different from syllables closed by geminates, but not from those closed by nasals. Vowels in open syllables are indeed longer than those in closed

[5] Only the full vowels have been included here, as the very short duration of [ə, ɨ] would skew the results, since these vowels occur only in closed syllables. The absence of the schwas results in different C2/C3 and Total values for closed syllables when compared with Table 9.

syllables. These vowel durations are consistent with a tense/lax difference or could be due to shortening in a closed syllable (or lengthening in an open one).

Let us consider now the effects of syllable structure on formant structure of the high vowels. In Table 12 and the formant plot in Figure 11, we present F1 and F2 values for [i, u] for each speaker before a single intervocalic consonant (Ø), before a nasal-stop cluster (M) and before a geminate (P). In open syllables, we expect lower F1 values for both [i, u] and higher F2 for [i] and lower F2 for [u] (that is, more peripheral values in the acoustic vowel space).

	S	F1	F2	E	F1	F2
iØ		301	2218		481	2717
iM		323	2237		427	2628
iP		314	2287		508	2698
uØ		302	778		470	1112
uM		330	751		402	945
uP		280	710		484	1009

Table 12. Average formant values of [i, u] based on following consonant type (in Hz).

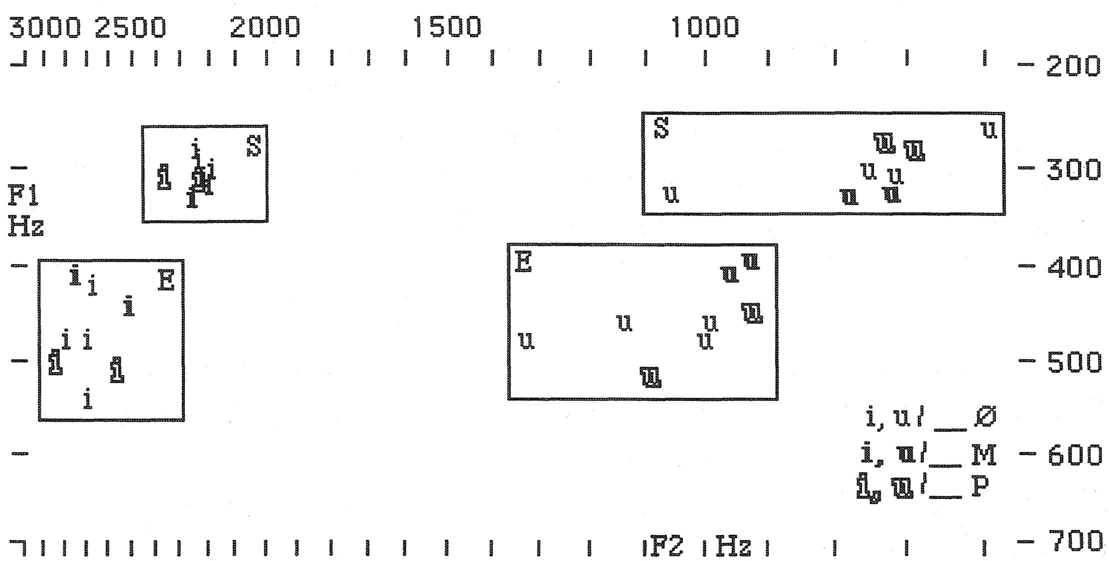

Figure 11. Formant plots of vowels in open syllables vs. preceding nasal-stop clusters or geminates for S and E (in Hz).

In Figure 11, formant values for the three syllable types are plotted for both speakers. (The boxes indicate the sets of values for each speaker.) For S, the values for [i] are very close;

while the small differences for [u] in closed syllables go in opposite directions. For E, the small differences for both [i, u] in closed syllables go in opposite directions. Thus we see no systematic difference for either speaker in formant values for [i, u] based on syllable structure. We conclude that there is not a systematic tensing of vowels in open syllables. The auditory impression of a tense/lax difference must be due to the duration differences.

5. Conclusions

Our acoustic investigation of the consonants and vowels of Madurese enables us to reach the following conclusions. In the case of consonants, no evidence of voicing during aspirated stops was found, leaving it an open question what properties are shared by the voiced and aspirated stops (synchronically). Investigation of voice quality and breathiness might offer insight into this relationship. Small duration differences between the aspirated stops and voiceless stops were observed both for closure and the period of aspiration. Thus synchronically these stops are quite similar, but can always be differentiated by the quality of the following vowel. Geminates show comparable durations to nasal-stop clusters for both full vowels and schwa.

In the case of vowels, three distinct effects on vowels were observed: height and pitch differences within the alternating pairs, height differences as a result of a preceding nasal, and duration differences as a result of syllable structure. The systematic differences in F1 (vowel height) and in F0 (pitch) between the [-high] and [+high] vowel sets support the conclusion that this is a register difference, similar to the patterns seen in the Mon-Khmer languages. An investigation of voice quality might lend further support to this conclusion and further investigation of the lack of F2 differences in warranted.

6. References

Cohn, A. (1993a) Voicing and vowel height in Madurese, a preliminary report. In J. A. Edmondson and K. J. Gregerson (eds.) *Tonality in Austronesian Languages* (pp. 107-121). *Oceanic Linguistics* publication no. 24.

Cohn, A. (1993b) Consonant-vowel interaction in Madurese: the feature Lowered Larynx. *CLS* 29, 105-119.

Edmondson, J., and K. Gregerson (1993) Western Cham as a register language. In J. A. Edmondson and K. J. Gregerson (eds.) *Tonality in Austronesian Languages* (pp. 61-74). *Oceanic Linguistics* publication no. 24.

Hayward, K. (1993) /p/ vs. /b/ in Javanese: some preliminary data. *SOAS Working Papers in Linguistics and Phonetics* 3, 1-33.

Hertz, S., E. Zsiga, K. de Jong, P. Gries, and K. Lockwood (1994) From database to speech: a multi-dialect relational database integrated with the Eloquence synthesis technology. *Proceedings of the 2nd ESCA/IEEE Workshop on Speech Synthesis.*

Hess, S. (1992) Assimilatory effects in a vowel harmony system: an acoustic analysis of advanced tongue root in Akan. *Journal of Phonetics* 20, 475-492.

Hombert, J.-M. (1978) Consonant types, vowel quality, and tone. In V. Fromkin (ed.) *Tone: A Linguistic Survey* (pp. 77-111). New York: Academic Press.

Huffman, F. E. (1976) The register problem in fifteen Mon-Khmer languages. *Austroasiatic Studies Part I, Oceanic Linguistics Special Publication* No. 13, 575-589.

Kenstowicz, M. (1994) *Phonology in Generative Grammar.* Cambridge (MA) and Oxford: Blackwells.

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

Maddieson, I. (1984) The effects on F0 of a voicing distinction in sonorants and their implications for a theory of tonogenesis. *Journal of Phonetics* 12, 9-15.

Pinnow, H.-J. (1979) Reflections on the history of the Khmer phonemic system. *Mon-Khmer Studies* 8, 103-30.

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 Volume 52. New Haven: American Oriental Society.

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

Appendix: word list

pattern	orthography	transcription	gloss
BɤP	bakkel	bɤkkel	vice
BɤM	bangkang	bɤŋkaŋ	naked
PhɤØ	bakat	phɤkat	loan
BɤØ	bata'	bɤtaʔ	carry (from the stove)
BɨP	bekkas	bɨkkas	former
BɨM	bengko	bɨŋko	house
PhɨP	bekkem	phɨkkəm	hit w/fist
PhiØ	bikang	phikaŋ	kind of cake
BiØ	bisa	bisa	to be able

BiM	bingka	biŋka	kind of cake
BiP	bittha'	biṭṭaʔ	to open, to force the lips open
BuP	bukka'	bukkaʔ	open
PhuØ	buka	phuka	to break the fast
BuØ	buta	buta	blind
BuM	bungko'	buŋkoʔ	hunchback
NaØ	mata	mata	eye
NəP	mekka	məkka	Mekka
NɛØ	mekat	mɛkat	to trap with glue
NɔØ	moka	mɔka	to break, to break the fast too early
ØaØ	apa	apa	what
ØɛØ	epar	ɛpar	brother/sister-in-law
ØəP	eppa'	əppa	father
ØɔØ	opa	ɔpa	to tip
PaØ	paka'	pakaʔ	astringent taste
PaM	pangkat	paŋkat	position
PaP	pappa'	pappaʔ	quid of tobacco (often mixed with betel)
PəP	pekka	pəkka	to split
PəM	pengko	pəŋkɔ	stubborn
PɛØ	pekang	pɛkaŋ	to pinch the thigh
PɛM	pengkot	pɛŋkot	to bind the hands
PɛP	keppar	kɛppar	to kick in martial arts
PɔØ	poka	pɔka	to break, to crack open
PɔM	pongka'	pɔŋkaʔ	to fall out, of a tooth
PɔP	poppo	pɔppɔ	to beat (a body)