

Degemination in Indonesian Phonology and Phonetics *

Niken Adisasmito-Smith

While Indonesian has no geminate consonants, identical consonants can become adjacent through morpheme concatenation, such as the concatenation of a root and a suffix. Lapoliwa (1981) suggests that in these cases, a rule of degemination takes place. The duration measurements in the current acoustic study provide an argument against Lapoliwa. While the differences between one and two identical consonants are not of the magnitude of a singleton-geminate contrast, identical consonant sequences tend to be longer in duration than a single consonant, suggesting that in Indonesian such sequences undergo partial neutralization. The result of duration measurements of the vowel preceding the contrasted consonants suggests that vowels are shorter when followed by two consonants than by one consonant, though only slightly. This study extends the investigation to cases where identical consonants become adjacent at word boundaries, and shows that while the distinction between one and two identical consonants at these boundaries is not neutralized, the magnitude of the duration difference is less than that of a singleton-geminate contrast.

1 Introduction

This study is an acoustic investigation of the duration of consonants in Indonesian, a Western Austronesian language. Of particular interest are the durational patterns of identical consonants at morpheme boundaries, namely root-suffix and word boundaries. These identical consonants have been analyzed as undergoing phonological degemination (Lapoliwa, 1981), which consequently would be reflected in the duration of these consonants. This acoustic study provides durational measurements and comparisons of one vs. two consonants at the root-suffix and word boundaries, presenting an argument for or against the claim of phonological degemination in Indonesian.

Due to the phonotactic constraints of Indonesian, only a subset of consonants in the phoneme inventory occurs at the relevant boundaries. The consonant inventory in Indonesian is shown in (1).

* This work owes improvement to comments and suggestions from Abigail Cohn, members of the Cornell Phonetics Lab, and the 8-ICAL audience. Thanks to Tobey Doeleman, Allard Jongman and Lisa Lavoie for comments on this draft, to Eric Evans for technical assistance in obtaining my measurements. Any errors are solely mine. This research is supported by NSF grant #SBR-9511185 to Abigail C. Cohn.

(1)	<i>bilabial</i>	<i>alveolar</i>	<i>palatal</i>	<i>velar</i>	<i>glottal</i>
stops	p b	t d	č ǰ	k g	ʔ
fricatives		s			
nasals	m	n	ɲ	ŋ	
liquids		l, r			
glides	w		j		

Most consonants in the phoneme inventory of Indonesian can occur in the onset position of a root, except /ŋ/ (Adisasmito-Smith 1998, and works cited therein). Palatal consonants are banned from coda position, as are voiced obstruents, except in borrowed words. There are no geminates in Indonesian. However, identical consonants become adjacent across morpheme boundaries, when the morpheme-final consonant is identical to the initial consonant of the following morpheme. Only a limited set of consonants can occur in suffix-initial position. Consequently, at the root-suffix boundary there is only a limited set of identical consonants that can become adjacent. Some examples are shown in (2), which cover all of the possible identical consonant sequences that occur at this boundary.

(2)	<i>Obstruents:</i> [p, k]	
	raŋka- <u>p</u> un	‘frame, emphatic’
	raŋka <u>p</u> - <u>p</u> un	‘doubled, emphatic’
	kata- <u>k</u> u	‘my saying/word’
	kata <u>k</u> - <u>k</u> u	‘my frog’
	<i>Sonorants:</i> [m, l]	
	raga- <u>m</u> u	‘your body’
	raga <u>m</u> - <u>m</u> u	‘your style’
	minta- <u>l</u> ah	‘ask, imperative’
	pinta <u>l</u> - <u>l</u> ah	‘weave, imperative’

Lapoliwa (1981) argues that a sequence of identical consonants in cases such as in (2) undergo degemination, in keeping with the constraint against geminates in Indonesian. The phonological rule of degemination is given in (3).

$$(3) C_1 + C_1 \rightarrow C_1$$

According to Lapoliwa (1981:85), “[...] a sequence of two identical consonants may become phonetically realised as a single consonant.” Thus, the output form of *ragamu* ‘your body’ and *ragammu* ‘your style’ would be neutralized to [ragamu], for example.

Identical consonants often become adjacent across a word boundary. The set of identical consonant sequences that can be observed at this boundary is relatively larger than that of consonants at the root-suffix boundary. In (4) are pairs where the first member has one consonant across the word boundary, whereas the second member has two identical consonants.

- (4) *Obstruents:* [p, t, k, s]
- | | |
|----------------------|----------------------|
| dara p <u>u</u> laŋ | ‘(a) girl goes home’ |
| harap p <u>u</u> laŋ | ‘please go home’ |
| suka t <u>u</u> laŋ | ‘like bone’ |
| sikat t <u>u</u> laŋ | ‘brush out of bone’ |
| bata k <u>a</u> sar | ‘rough brick’ |
| watak k <u>a</u> sar | ‘rough personality’ |
| bata s <u>a</u> ja | ‘brick only’ |
| batas s <u>a</u> ja | ‘border only’ |
- Sonorants:* [m, n, r]
- | | |
|----------------------|---------------------|
| kala m <u>a</u> nis | ‘sweet time’ |
| kalam m <u>a</u> nis | ‘sweet words’ |
| ĵala n <u>a</u> ŋka | ‘net for jackfruit’ |
| ĵalan n <u>a</u> ŋka | ‘Jackfruit Street’ |
| kasa r <u>a</u> mai | ‘busy cashier’ |
| pasar r <u>a</u> mai | ‘busy market’ |

Across a word boundary, the adjacent identical obstruents that occur are the stops /p/, /t/, and /k/, and the fricative /s/. The occurring sonorants are the nasals /m/ and /n/, and the liquid /r/. Even though adjacent identical /l/ occurs across word boundary, no good (near-) minimal pairs have been found so far.

Although a sequence of identical consonants at the root-suffix boundary may undergo degemination, the observations made in Lapoliwa (1981) seem to be impressionistic. No acoustic data are available. No claim has been made regarding identical consonant sequences across word boundaries, as to whether degemination takes place. In this study, the durational patterns of one consonant vs. two-consonant sequence at the root-suffix and word boundaries are quantified.

There are several possibilities that we could anticipate finding in the data, with respect to degemination as a neutralizing phenomenon. The first possibility is that consonants at the morpheme boundaries undergo complete degemination. In this case, we

would expect to find no significant duration difference between two-consonant sequences and single consonants, as suggested by Lapoliwa (1981). So, for example, the durations of the bilabial nasal in *ragamu* and *ragammu* would not differ from each other, nor would those in *kala manis* and *kalam manis*. The second possibility is that the underlying consonantal sequences are maintained, resulting in fake geminates. Consequently, we might expect to see the duration differences between the two-consonant sequences and single consonants reflect those of a singleton-geminate contrast. Across languages, geminates are between one and a half time to three times longer than singletons (Ladefoged and Maddieson 1996). Thus we could expect that the duration of identical consonant sequences to be the same as that of non-identical consonant sequences. The third possibility is that a two-consonant sequence across a morpheme boundary is not completely neutralized to a single consonant. As a result, the duration of the underlying two-consonant sequence would tend to be greater than that of the underlying single consonant; however, these duration differences may not be of the magnitude of a singleton-geminate contrast.

At this point, let's consider three bodies of previous work relevant to the duration investigation in this study: a) acoustic studies of languages with a singleton-geminate contrast; b) phonetic studies of neutralization; and c) studies of the interaction of syllable structure and segment duration.

In most acoustic studies of geminate consonants, duration is the primary focus. The largest duration ratio that has been found between geminates and singletons is in Turkish, at 3:1 (Lahiri and Hankamer, 1988). In this language, geminate consonants may be underlying or derived. Underlying geminates are root-internal, while derived cases occur across the root-suffix boundary. As is the case with Turkish, geminate consonants in Bengali and Buginese may be underlying or derived. In Bengali, geminate consonants are found to be twice as long as their singleton counterparts (Lahiri and Hankamer, 1988). In Buginese, a Western Austronesian language spoken in South Sulawesi, the duration ratio of the geminate-singleton contrast root-internally is 1.9:1 for one speaker and 1.4:1 for the other (Podesva, 1998). The duration ratio of the geminate-singleton contrast at the prefix-root boundary is 1.7:1 for one speaker and 2:1 for the other; at the root-suffix boundary, it is 1.7:1 for both speakers. Another Western Austronesian language that has a singleton-geminate contrast is Madurese, spoken on the island of Madura and in some parts of East Java. Root-internal geminate consonants may be voiced or voiceless. The duration ratio of the geminate-singleton contrast ranges by phonation type, but it is approximately 1.7:1 for one speaker and 1.5:1 for the other (Cohn and Ham, this volume). In Sinhala, an Indo-Aryan language spoken in Sri Lanka, the duration ratio of geminate-

singleton contrast root-internally is about 1.7:1 (Letterman, 1994). Delattre (1971b) examines the geminate-singleton contrast of /r/ in German, which occurs word-medially and word-finally. The duration ratios range from 1.8:1 to 1.4:1.

Delattre (1971a) also examines the duration of identical consonants across word boundaries in Spanish, French, German and English. In the English data, the duration of the alveolar nasal, among others, in *I've seen Nelly* is contrasted with that in *I've seen Elly* and *We see Nelly*. Combined with the other consonantal cases, the average duration ratio in English for such cases is 1.4:1. This ratio is the smallest compared to the other three languages, which is 1.9:1 in French, 1.8:1 in Spanish, and 1.5:1 in German, for similar cases. The duration of the vowel preceding the two-consonant sequence was not found to be different from that preceding the single consonant.

In the present study, the issue of neutralization is central in considering the duration differences between one consonant and two identical consonants at the root-suffix and word boundaries. If the duration differences are non-existent, then we are faced with a case of complete neutralization. On the other hand, if the duration differences of these consonants cannot be ignored, we have a case of partial neutralization or a case of singleton-geminate contrast, depending on the magnitude of the duration ratios. Most studies of neutralization focus on the durations of word-final voiced vs. voiceless obstruents. In languages like German, Catalan, Russian, and Polish, among others, word-final devoicing has been analyzed as a case of phonological neutralization. However, several experimental studies (Dinnsen 1985; Slowiaczek and Dinnsen 1985; and the works cited therein) have shown that the neutralization in these languages is incomplete, in that the underlying voicing distinction is realized in the duration differences of the contrasted obstruents. Further, the duration of the vowel immediately preceding an obstruent may indicate the underlying voicing distinction of the obstruent, as is found to be the case in German (Port and O'Dell 1985) and in English (Peterson and Lehiste 1960; Chen 1970; Klatt 1973, to name a few). In Korean, underlying coronal obstruents /t, t^h, s/ are neutralized to [t] word-finally; the durations of these obstruents and of the preceding vowels show that this neutralization is complete (Kim and Jongman, 1996).

Vowel duration may also be affected by syllable structure. In many languages, vowels tend to be shorter in a closed syllable than an open syllable (Maddieson, 1985), though several languages contradict this tendency, e.g. Hungarian (Ham 1998) and Sinhala (Letterman 1994). Intervocalic geminates are usually syllabified such that the syllable boundary falls in the middle of the geminates. In Indonesian, the durations of the penult vowel in *ragammu* and *ragamu* may indicate whether the penult syllable of these words is open or closed. This may be combined with the durational facts from the target

consonants. If complete neutralization does not occur, the two-consonant sequence would be (slightly) longer than the single consonant; if complete degemination takes place, the two sequences would be practically the same.

The object of this study is to investigate whether or not identical-consonant sequences at the root-suffix boundary undergo phonological degemination in Indonesian, based on the durational measurements of the target consonants and the vowel preceding the target consonants. The structure of the paper is the following: the design and methodology of the experiment is laid out in section 2, the results of the experiment are discussed in section 3, and concluding remarks are presented in section 4.

2 Acoustic study

2.1 Speakers

Six speakers of Indonesian, three male and three female, were recorded. Five of the speakers have been in the US for 2-5 years, one speaker for more than 10 years. None had any known speech or hearing disorders. Some Indonesian speakers are monolingual; however, most are bilingual. In this study, three speakers (PH, WN, and BT) were monolingual, two speakers (NN and AM) were bilingual Indonesian and Javanese, and one speaker (YT) was bilingual Indonesian and Buginese.

2.2 Methodology

A set of bisyllabic roots and monosyllabic suffixes, and a set of paired bisyllabic words were selected. The combinations of the roots and suffixes and the sequences of bisyllabic words were chosen to create (near) minimal pairs, and the contrast of a sequence of two identical consonants and its single consonant counterpart could be observed. Six repetitions of the stimuli were randomized and embedded in the carrier phrase *Dibaca _____ saja* ‘Just read _____’. The total number of target words was 576 (16 words x 6 repetitions x 6 speakers) for the root-suffix cases, and 720 (20 words x 6 repetitions x 6 speakers) for the word-word cases.

This study, being part of a larger study of the acoustics of Indonesian consonants, reports on a subset of the recorded materials. Here the bilabial nasal and velar stop cases, when occurring in both environments, as shown in (5), are presented.

(5) a. Root-suffix cases

<i>Obstruents</i>		<i>Sonorants</i>	
kata <u>k</u> u	‘my saying/word’	ragam <u>m</u>	‘your body’
katak <u>kk</u>	‘my frog’	ragam <u>mm</u>	‘your style’

matak <u>u</u>	‘my eyes’	jalamu	‘your net’
watak <u>ku</u>	‘my personality’	salam <u>mu</u>	‘your greeting’

b. Word-word cases

<i>Obstruents</i>		<i>Sonorants</i>	
bata <u>ka</u> sar	‘rough brick’	kala <u>ma</u> nis	‘sweet time’
watak <u>ka</u> sar	‘rude personality’	kalam <u>ma</u> nis	‘sweet word’

The speakers were recorded in a sound proof booth using a cardioid microphone (Electrovoice, Model RE 20) and a Carver, model TD-1700 cassette recorder, reading the material which was presented in standard Indonesian orthography. Before being recorded, speakers were asked to make sure that all words in the list were known to them. They were also asked to read a few words to familiarize themselves with the material. The list was read at a comfortable reading rate throughout the recording.

The data were digitized at 11,025 Hz, and stored as files to be analyzed using WAVES+/ESPS on a SUN-SPARC station. Waveforms and spectrograms were obtained for each target word. For the purpose of this study, the durations of target consonants and the vowel preceding them are reported. The duration of VOT in the stop cases is also discussed where relevant. Labels were assigned at the onset and offset of each segment. These labels were used to generate duration measurements. A sample of the spectrogram with labels of *ragamu* ‘your body’ by speaker WN is shown in Figure 1.

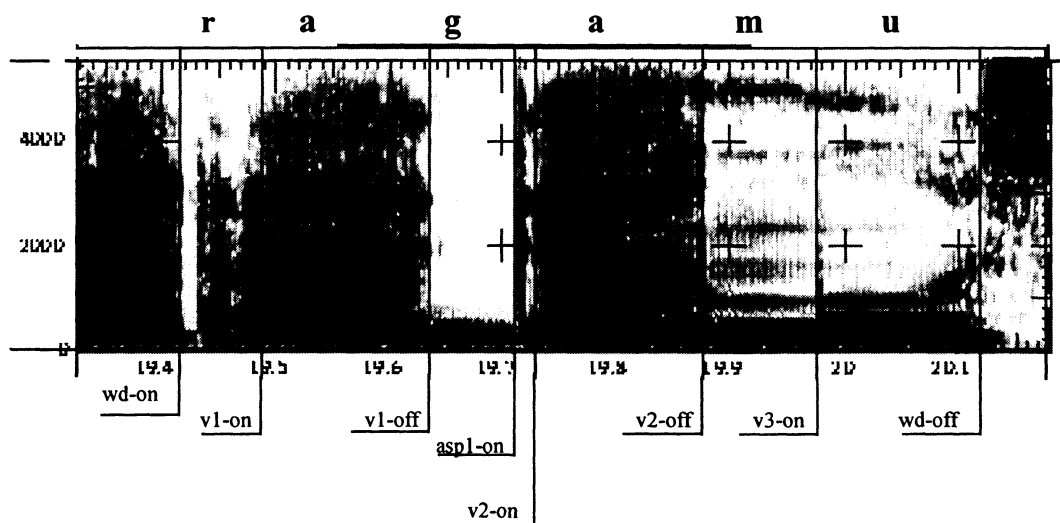


Figure 1. Spectrogram of *ragamu* by speaker WN. The penult vowel is delimited by the labels ‘v2-on’ and ‘v2-off’. The onset of the target consonant is labeled ‘v2-off’, and the offset ‘v3-on’.

The onset of a vowel was taken to be the onset of the second formant, and the offset of the vowel was taken to be the end of the second formant. The onset of a stop closure began where the preceding vowel ended, and the offset point was the beginning of the release burst. For nasals, the onset was taken to be the point where the energy of the second and third formants (F2 and F3) of the preceding vowel began to dissipate, and the offset was the point where the energy of F2 and F3 of the following vowel started to intensify. Based on both the waveform and spectrogram, the segmentation of stops is straightforward, given the clear spectral break from a vowel to the following stop and the clear energy burst following the release of the stop closure. The segmentation of bilabial nasals is also simple; in most cases, the first formant (F1) in the spectrogram shows clear movement from high F1 for the low vowel [a], to low F1 for the bilabial nasal, and to slightly higher F1 for the high vowel [u]. In many cases, the offset of a bilabial nasal is accompanied with a burst of release, which makes segmentation of this nasal less complicated. The results of the measurements are reported in the following section.

3 Results

3.1 Root-suffix cases

3.1.1 Consonant durations

The set of *root-suffix* cases consists of bisyllabic roots followed by a monosyllabic suffix. The occurring consonants in this set are the bilabial nasal and the voiceless velar stop. Table 1 shows the average durations of the consonants being contrasted. For the /mm/ case, the values are the closure duration average of the target consonant in *ragammu* and *salammu*; for the /m/ case, in *ragamu* and *jalamu*; for the /kk/ case, in *katakku* and *watakku*; for the /k/ case, in *kataku* and *mataku*. (For the average values of target consonants in individual words, see Appendix 1.)

	PH	WN	BT	NN	AM	YT	Mean
/mm/	90	156	101	107	98	199	125
/m/	78	110	84	91	72	101	89
Difference	12	46	17	16	26	98	36
/kk/	100	179	72	113			116
/k/	77	131	61	84			88
Difference	23	48	11	29			28

Table 1. The averaged closure durations (in ms) of target consonants and the difference of the averaged values between /mm/ vs. /m/, and between /kk/ vs. /k/, for each speaker.

For each speaker, the average duration of /mm/ is greater than that of /m/. However, the magnitude of difference varies among the speakers. Speaker YT shows the greatest durational difference, of 98 ms. Speaker PH, on the other hand, shows the least difference, of 12 ms. Statistically, the differences for each speaker are highly significant throughout ($P < 0.01$ by 2-sample T-test).

Similarly, the average duration of /kk/ is also greater than /k/. As is the case with the nasals, there is variation in the magnitude of difference among the four speakers. The greatest duration difference for /k/ is 48 ms for speaker WN, while the smallest is 11 ms for speaker BT. For each of the speakers, the differences are statistically significant ($P < 0.01$).

Note that in comparing the difference of the closure duration of the velar stop case, speakers AM and YT are excluded. For speaker AM, a bilingual speaker of Indonesian and Javanese, the velar stop in stem-final position is realized as a glottal stop: [kataʔku].¹ The same is the case with speaker YT, a bilingual speaker of Indonesian and Buginese. In the phonology of Javanese, stem-final velar stops are realized as a glottal stop. In Buginese, the consonants occurring in stem-final position are glottal stop and the velar nasal. In the speech of YT, stem-final /k/ is realized as a glottal stop.

There are systematic differences among the speakers with respect to the duration ratio of the two vs. one consonant, for both the nasal and stop cases, as shown in (6).

(6) *Duration ratios for each speaker*

a. nasal cases	1.2:1	PH, BT, NN
	1.4:1	WN, AM
	2:1	YT
b. stop cases	1.2:1	BT
	1.3:1	PH
	1.4:1	WN, NN

The duration ratio in the nasal case for speaker YT, a bilingual Buginese and Indonesian, is as large as that found in the Bengali geminate consonants (Lahiri and Hankamer, 1988). This may be explained by the fact that there are geminates in Buginese, and the influence of this language manifests itself in Indonesian. For speaker BT, the duration

¹ Interestingly, speaker NN, who is also a bilingual speaker of Indonesian and Javanese, did not pattern together with speaker AM. The difference between NN and AM is that NN was born in Jakarta, grew up speaking Indonesian, learned Javanese later and moved to Central Java; speaker AM, on the other hand, was born in Central Java, grew up speaking Javanese, and learned Indonesian later. In fact, speaker NN patterns together with the other monolingual speakers with respect to stem-final /k/.

ratio for both nasal and stop cases is 1.2:1. For speaker WN, it is 1.4:1 for both cases. For speaker PH and NN, on the other hand, the duration ratio for the nasal cases is relatively smaller than for the stop cases. These ratios suggest that two-consonant sequences tend to be longer than single consonants, even though for some consonants and for some speakers, these ratios are smaller than the duration ratio of a typical singleton-geminate contrast.

While the averages of two-consonant sequences tend to be longer than single consonants, is the duration difference sufficient for hearers to distinguish between these contrasting sequences? Klatt (1976) argues that the just noticeable difference (JND) for perception of duration is about 25 ms for a segment having a duration of about 100 ms or longer. Further, he suggests that the JND is about 20% of the duration of the segment, following Weber ratios (see also Lehiste 1970). Based on Table 1, if the 25 ms JND is used as the differential threshold, the following consonantal contrasts in the root-suffix boundary would be distinct: bilabial nasal for speakers WN, AM, and YT, and velar stop for speaker WN and NN. If Klatt's ratio criterion is used instead, the velar stop case for speaker PH would also be distinct. It would be worthwhile pursuing a study to determine the differential point for perception of one consonant vs. two-consonant sequences at the root-suffix boundary in Indonesian. It would also be interesting to compare the duration of identical consonant sequences with that of non-identical consonant sequences at the root-suffix boundary, but that is beyond the scope of the present study.

Duration averages alone tell only part of the story. To obtain a full picture of the difference between one consonant and two consonants, we need to see whether the ranges of values for these consonants are distinct. In Fig. 2, the ranges of duration for the words with /mm/ at the morpheme boundary: *ragammu* and *salammu*, and those with /m/: *ragamu* and *jalamu*, are compared. For each speaker, this figure shows whether there is an overlap between the longest /m/ tokens and the shortest /m+m/'s.

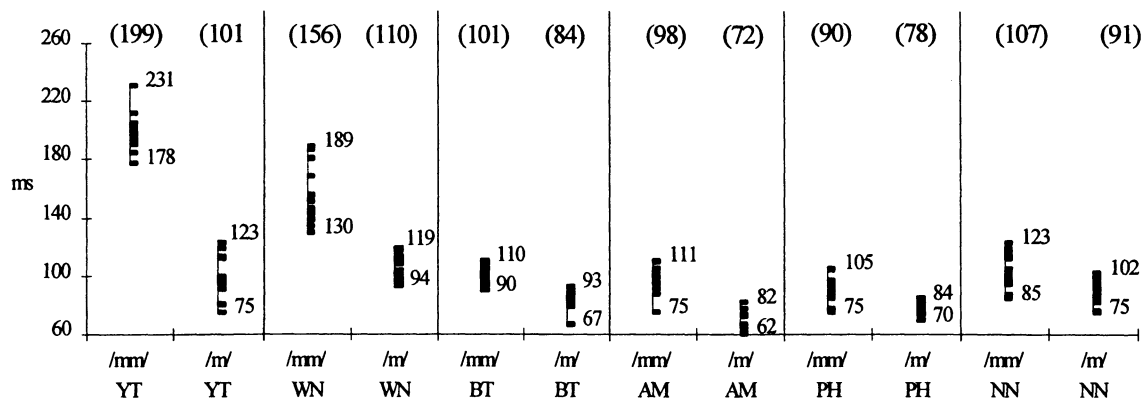


Figure 2. Ranges of duration (in ms) of target consonants in the words *ragammu*, *salammu*, *ragamu* and *jalamu*, pooled for one- vs. two-nasal case, for each speaker. There are twelve points for each nasal case (2 words x 6 repetitions); identical values are represented by one point. Values in parentheses along the top of the figure are average values.

The speakers in Fig. 2 are ordered by existence and extent of an overlap in duration ranges. In general, the distribution of values within the duration ranges is evenly spread. The most noticeable outlier is in speaker YT's /mm/ case, in which the greatest value is quite far from the other values within the same group. This outlier raises the average duration, though not by much. Upon closer examination, the word duration of this particular outlier is also the longest, compared to the duration of the other words. This may suggest that there is a little variation in speaking rate during recording, with a slower rate or more careful pronunciation resulting in longer duration.

As we can see, only speakers YT and WN show no overlap in the duration ranges for both the /mm/ and /m/ cases. The ranges for the other speakers overlap from 3 ms (speaker BT) to almost 20 ms (speaker NN). Speaker YT has the greatest distance between the contrasted duration ranges, and the difference of the averaged values between the /mm/ and the /m/ cases (shown in Table 1) is also the greatest, 98 ms. Speaker WN shows less distance between the two duration ranges relative to speaker YT. The difference of the average values for speaker WN, shown earlier in Table 1, is 46 ms. For the other speakers, whose duration ranges overlap, the difference of the average values is about half or less, compared to the difference for speaker WN.

The ranges of closure duration for /k/ and /kk/ cases for four speakers are shown in Fig. 3.

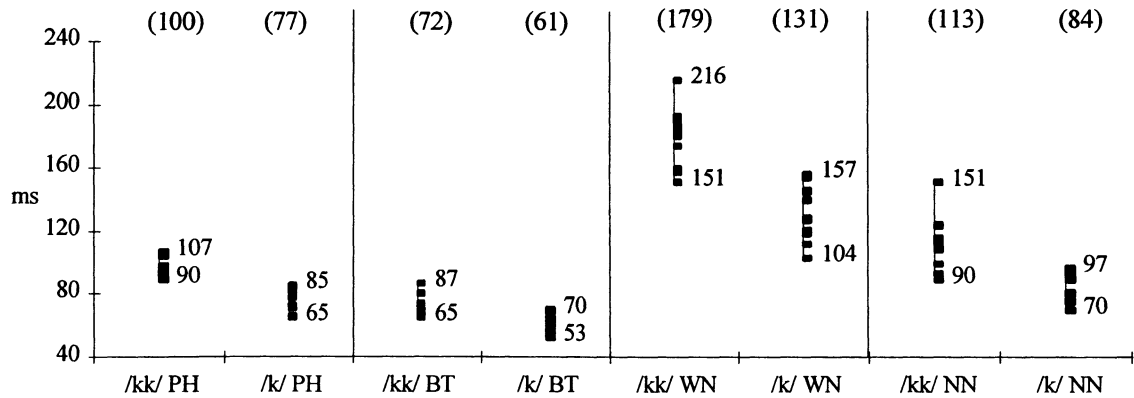


Figure 3. Ranges of closure duration (in ms) of target consonants in the words *katakku*, *watakku*, *kataku* and *matakku*, pooled for one- vs. two-stop case, for each speaker. There are twelve points for each nasal case (2 words x 6 repetitions); identical values are represented by one point. Values in parentheses along the top of the figure are average values.

For speaker WN, the longest duration of /kk/ is 216 ms, which is far apart from the other values within the same range for this speaker. The same is true in the duration range of /kk/ for speaker NN, where the highest value, 151 ms, is far above the rest. This means that the average values of /kk/ for speakers NN and WN are slightly higher than if these outliers are excluded. The duration ranges for the other speakers are evenly spread. Only speaker PH shows no overlap. The distance between the ranges is 5 ms. The other three speakers show overlap, and the degree of overlap is about the same, i.e. between 5-7 ms. These results suggest that the contrast between one consonant vs. two-consonant sequences is marginal for some speakers. In the following section, the duration of vowels preceding the target consonants are considered.

3.1.2 Vowel durations

While the contrast between one consonant and two-consonant sequences is marginal, the differences might be observed in the duration of the previous vowel, since vowels tend to be shorter in closed syllables (Maddieson, 1985). Figs. 4 and 5 show the averaged values of the vowel preceding the sequences of one vs. two consonants for both the nasal and stop cases. A longer vowel would indicate an open syllable, and therefore that a single C, rather than a geminate, follows. (For the average values of target consonants in individual word, see Appendix 2.)

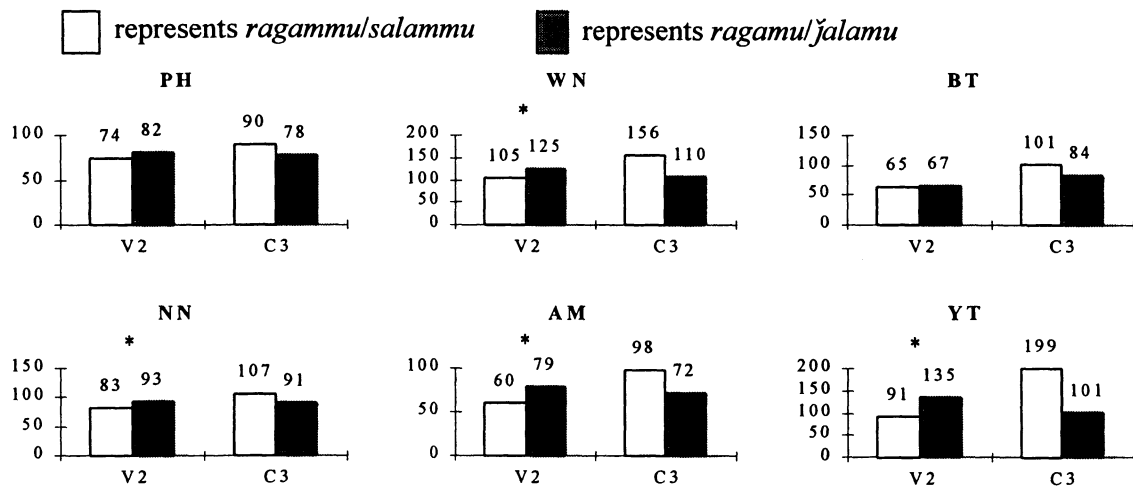


Figure 4. Average durations (in ms) of the target consonants (C3) and the vowel preceding them (V2), for the nasal cases. (*) over the V2 bars indicates $P < 0.01$.

For speaker BT, there is practically no difference in vowel duration, whether the vowel is followed by a sequence of one or two nasals at the root-suffix boundary. For speakers PH, WN, NN and AM, the vowel preceding the two-consonant sequence tends to be shorter than preceding the single consonant, though the difference is not statistically significant for speaker PH. These results suggest that for speaker BT, vowel duration does not play a role in distinguishing one consonant vs. a two-consonant sequence. For speaker PH, this may also be the case, since the difference is not statistically significant. For speakers WN, NN and AM, the duration difference may contribute to the distinction of the contrast, even though based on Klatt's ratio criterion, the vowel duration is distinct only for speaker AM. For speaker YT, for whom the difference in consonant duration is large, the vowel preceding the two-consonant sequence is about 1.5 times shorter than that preceding the single consonant.

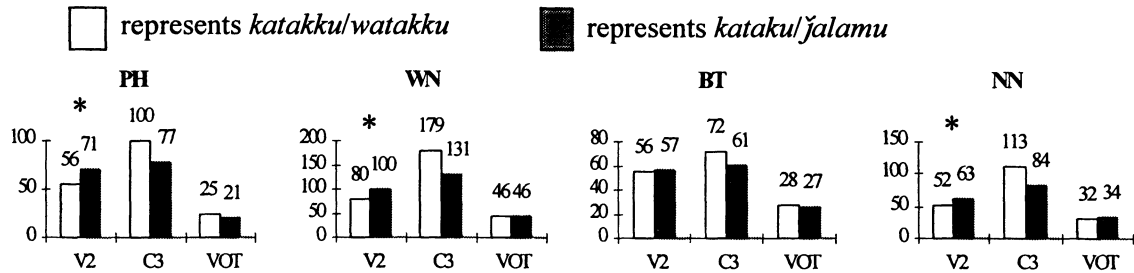


Figure 5. Average durations (in ms) of the vowel (V2) preceding the target consonants, the target consonants (C3) and VOT, for the stop cases. (*) over the V2 bars indicate $P < 0.01$.

For speaker BT, the vowels preceding one or two stops are not distinct, as is the case with this speaker's nasal. Consequently, the duration difference between vowels is much smaller than the JND. These facts suggest that the duration of vowel preceding single consonant vs. two-consonant sequences is not a salient indicator for speaker BT in distinguishing whether one or two consonants follow. For the other three speakers, the duration difference in the vowels is statistically significant. For speaker PH and WN, the difference of these vowels is larger than the differential threshold. For speaker NN, it is just a little bit smaller. The duration of VOT following the stop release varies among the speakers. For speakers PH and BT, the VOT following the two-stop sequence is slightly longer than when following the single stop. However, there is no difference in VOT duration in either sequence for speaker WN. For speaker NN, the VOT following the two-stop sequence is slightly shorter than when following its single counterpart. This result suggests that the duration of VOT does not play a role in determining whether the underlying sequence consists of one or two stops.

Fig. 6 shows the ranges of durations of vowels preceding one consonant or two consonants. This figure compares the duration ranges.

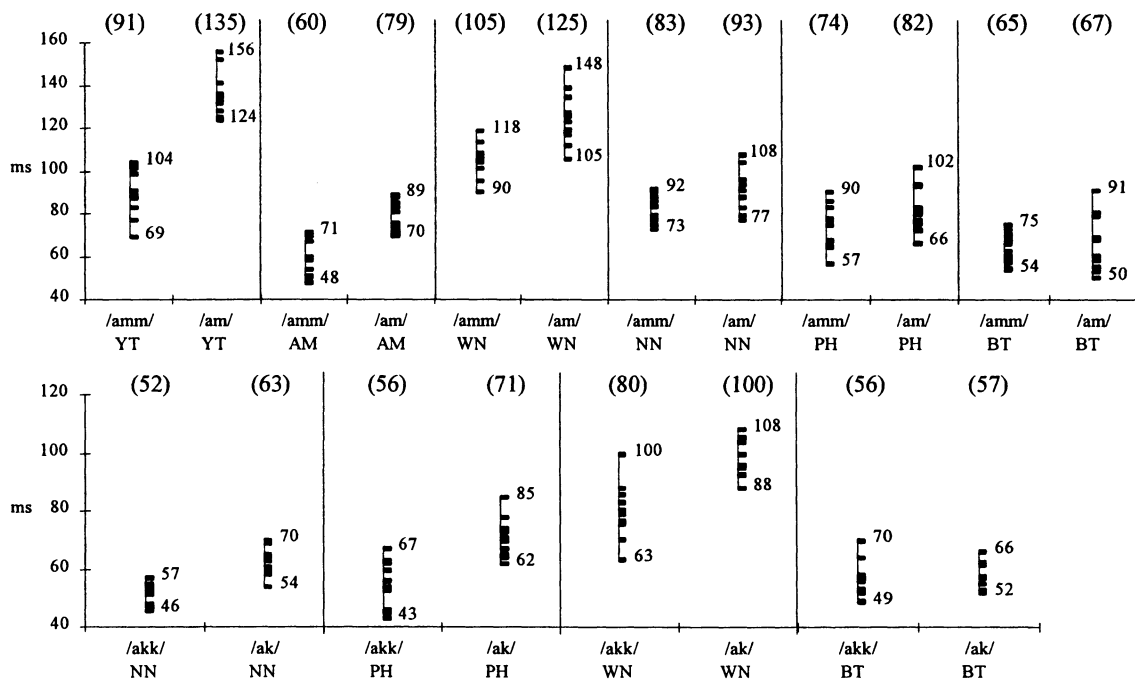


Figure 6. Ranges of duration (in ms) of the vowel /a/ preceding the sequence of one vs. two nasals or stops, for each speaker. There are twelve points for each consonantal case (2 words x 6 repetitions); however, identical values are represented by one point. The values within the parentheses are the average values from Figs. 4 and 5.

Speaker YT is the only one whose duration ranges for the vowel preceding the nasals do not overlap. The other speakers show overlap of the duration ranges from just 1 ms for speaker WN to 24 ms for speaker PH. For speaker BT, the duration range of the vowel preceding /mm/ falls completely within the range of the vowel preceding /m/. For the stop cases, the duration ranges overlap for the vowel preceding single consonants vs. two-consonant sequences, among the four speakers compared. The degree of overlap also differs among the speakers. For speakers NN, PH, and WN, the overlap is partial. For speaker BT, however, the overlap is complete.

These results show that there is a greater overlap in the duration ranges for vowels than for consonants. For the nasal, there seems to be a correlation between the magnitude of average duration difference and whether the values of the duration ranges are distinct, provided that outliers are excluded. When the duration difference is large, the duration ranges of single consonant vs. two-consonant sequences tend to be distinct; for example, speakers YT and WN (and perhaps speaker AM). In the stop cases, the duration differences for speakers WN and NN are greater than for speakers PH and BT; yet for

speakers WN and NN, the ranges of values are not distinct and for speakers PH, they are. This may be due to the different manners of articulation, or due to the fact that the duration ranges are not evenly spread for speakers WN and NN. These duration differences are statistically significant. In the vowel case, the duration ranges (almost) completely overlap, e.g. for speaker BT, and the duration difference is very small and not statistically significant. This suggests that when the duration difference is non-existent or small and not statistically significant, we would expect to see a great degree of overlap in the duration ranges. However, the reverse is not necessarily the case.

3.1.3 V+C durations

In sections 3.1.1 and 3.1.2, we looked at the duration differences and duration ranges of consonants and the vowels preceding them separately. We have also seen how variable they are among the speakers. One could also look at the durational relationship between a consonant and the preceding vowel syntagmatically by comparing two duration ratios (Maddieson, p.c.), i.e. the duration ratios of vowels preceding one vs. two consonants are compared with the duration ratios of vowels including the following one vs. two consonants, for each speaker. This comparison, presented in Fig. 7, may show uniformity among the speakers, which in turn would help to determine whether neutralization takes place.

	Ratios of <u>VC</u> : <u>VCC</u>	Speaker	Ratios of <u>VC</u> : <u>VCC</u>
a. nasal cases:	1.2:1	WN	1.1:1
	1.1:1	PH	1:1
	1:1	BT	*1.1:1
	1.1:1	NN	1.1:1
	1.3:1	AM	1.1:1
	1.5:1	YT	*1.2:1
b. stop cases:	1.3:1	WN	*1.1:1
	1.3:1	PH	1:1
	1:1	BT	*1.1:1
	1.2:1	NN	*1.1:1

Figure 7. Comparison between the duration ratios of the vowel preceding one vs. two consonants and the duration ratios of the vowel + the following one vs. two consonants across a root-suffix boundary. (*) next to ratio indicates P-value < 0.01.

The ratios of VC:VCC are those in which only the vowels are considered. The ratios of VC:VCC are those in which the durations of both the vowel and the following

consonant/s are taken into account. Relative to the duration ratios of the vowel alone, the duration ratios of vowel and the following consonant are less variable across speakers for both the nasal and stop cases. For most speakers, the duration ratio of $\underline{VC}:VCC$ is 1.1:1. This confirms the earlier observation that the difference between C and CC at the root-suffix boundary is not completely neutralized.

As we have seen, for speaker YT, the values in the ranges of duration of /m/ vs. /mm/ and of the vowel preceding these consonants are distinct. At the other end of the spectrum, for speaker BT, the two-consonant sequences tend to be longer, though marginally. There is practically no difference in vowel duration preceding these sequences, and the values in the duration ranges of both the target consonants and the vowel preceding them are not distinct (they are partially or completely overlapping).

To summarize, for speaker YT, fake geminates across the root-suffix boundary are maintained, with CC being twice as long as C, and the duration of the vowel (alone) preceding C being almost 1.5 times longer than that preceding CC. In contrast, for speaker BT, a sequence of identical consonants is simplified, with marginal durational differences between C and CC, and no difference between the vowels preceding C vs. CC. For the other speakers, the duration of a sequence of identical consonants tends to be longer than that of a single consonant, and the vowel preceding the former tends to be shorter than preceding the latter, though the differences are very small. Thus, for most speakers, an underlying two-consonant sequence is not neutralized to a single consonant, rendering the first of the two consonants in coda position. However, the durations of the preceding vowel do not seem to indicate strongly that the penultimate syllable of the target words listed in (5a) is CVC. Note that in several languages vowels do not undergo closed syllable shortening, e.g., Hungarian (Ham, 1998), Sinhala (Letterman, 1994).

In the following section, the discussion focuses on the question of whether identical consonant sequences at word-word boundary are maintained or whether they undergo neutralization.

3.2 Word-word cases

3.2.1 Consonant durations

In the previous section, the results of the root-suffix cases show that the duration differences between one vs. two consonants are statistically significant and that for most speakers, the duration of the vowel preceding these consonants is only slightly suggestive of the differences.

In this section, we will see if the word-word cases work the same way. Here, we have a minimal pair of *kalam manis* vs. *kala manis*, where we compare two vs. one bilabial

nasal across word boundary, and a near minimal pair of *watak kasar* vs. *bata kasar*, where the velar stops across word boundary are compared. For the velar case, speakers AM and YT are excluded due to glottalization of word-final /k/.

Consider first the durations of the target consonants and the vowels preceding the target consonants, presented in Figs.8 and 9.

□ represents *kalam manis* in Fig. 8, and *watak kasar* in Fig. 9
 ■ represents *kala manis* in Fig. 8, and *bata kasar* in Fig. 9

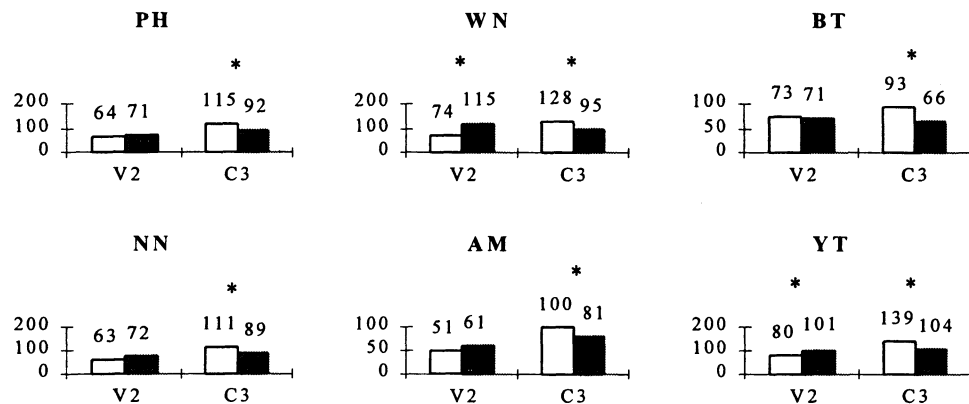


Figure 8. Average durations (in ms) of the target consonants (C3) and the vowel preceding them (V2), for the nasal cases. (*) over the bars indicates $P < 0.01$.

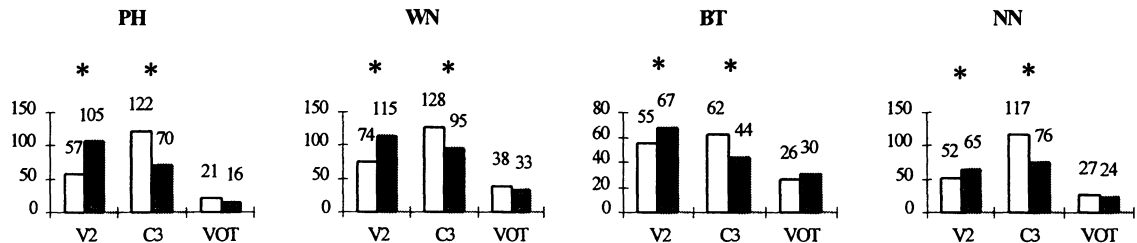


Figure 9. Average durations (in ms) of the target consonants (C3), the VOT, and the vowel (V2) preceding the target consonants, for the stop cases. (*) over the bars indicates $P < 0.01$.

Each speaker shows some difference in the averages between sequences of one and two consonants, in that the duration of a sequence of two consonants is greater than its one consonant counterpart. Statistically, the differences in both the nasal and stop cases are highly significant for each speaker. Based on Klatt's ratio criterion where the differential threshold is at least 20% of the segment duration, the average differences between single consonants and two-consonant sequences are above the JND, for most speakers. For

speaker AM, the difference of 19 ms is just 1 ms below the JND of 20 ms. The duration ratios of single consonants vs. two-consonant sequences vary for the speakers, as shown in (7).

(7) *Duration ratios for each speaker*

a. nasal case	1.2:1	NN, AM
	1.3:1	PH, YT
	1.4:1	BT
	1.5:1	WN
b. stop case	1.3:1	WN
	1.4:1	BT
	1.5:1	NN
	1.7:1	PH

For speaker WN, the duration ratio of the one vs. two-consonant contrast is greater in the nasal case than in the stop case. Speakers NN and PH, on the other hand, show greater duration ratios in the stop case, relative to the nasal one. For speaker BT, the duration ratios are the same for the nasal and the stop cases.

The average duration of VOT tends to be slightly longer following the two-stop sequences for speakers PH, WN, and NN. For speaker BT, on the other hand, the VOT following the two-stop sequence is shorter than that following one stop. Statistically, the differences in VOT following the two sequences are not significant. This suggests that VOT is not an indication of whether the preceding stop sequence consists of one or two segments. This is the case whether the stop sequence is at the word-word boundary or at the root-suffix boundary, as we have seen earlier.

3.2.2 Vowel durations

Vowels preceding two-consonant sequences tend to be shorter than preceding single consonants for most speakers, as shown earlier in Figs. 8 and 9, though not all of the differences are statistically significant. In the nasal case, the differences are significant only for speakers WN and YT. In the stop case, these differences are significant for all speakers. Based on Klatt's ratio criterion, the duration difference of the vowel preceding the nasal for speakers WN and YT, and the vowel preceding the stop for speakers PH, WN and NN are greater than the JND; for speaker BT, it is only slightly smaller. For speakers PH, BT, NN, and AM, the duration difference in the nasal case is smaller than the JND.

We have seen earlier that the average duration differences in the consonants at the word-word boundary are statistically significant and most of them are above the JND. But as before, we need to consider the ranges. It is interesting that the ranges of duration of one consonant are distinct from those of two-consonant sequences, in both the nasal and stop cases, for all speakers, as shown in Fig. 10.

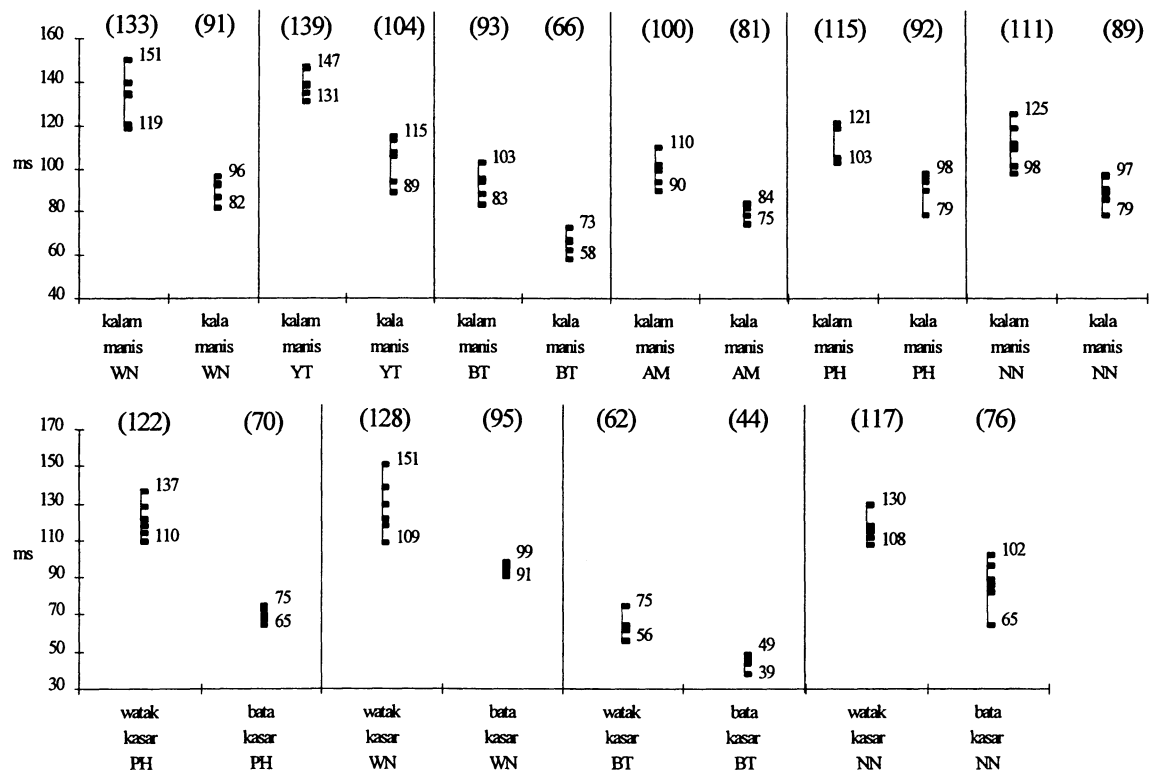


Figure 10. Ranges of duration (in ms) of single consonants vs. two-consonant sequences, for each speaker. Each point indicates the value of one of the six repetitions of the words; identical values are represented by one point.

Even though in some cases the differences of range are small, they are still distinct and the average differences are statistically significant.

In contrast to the consonantal cases, the ranges of duration for the preceding vowel are distinct only for some speakers, and not for the others. These ranges are shown in Fig. 11.

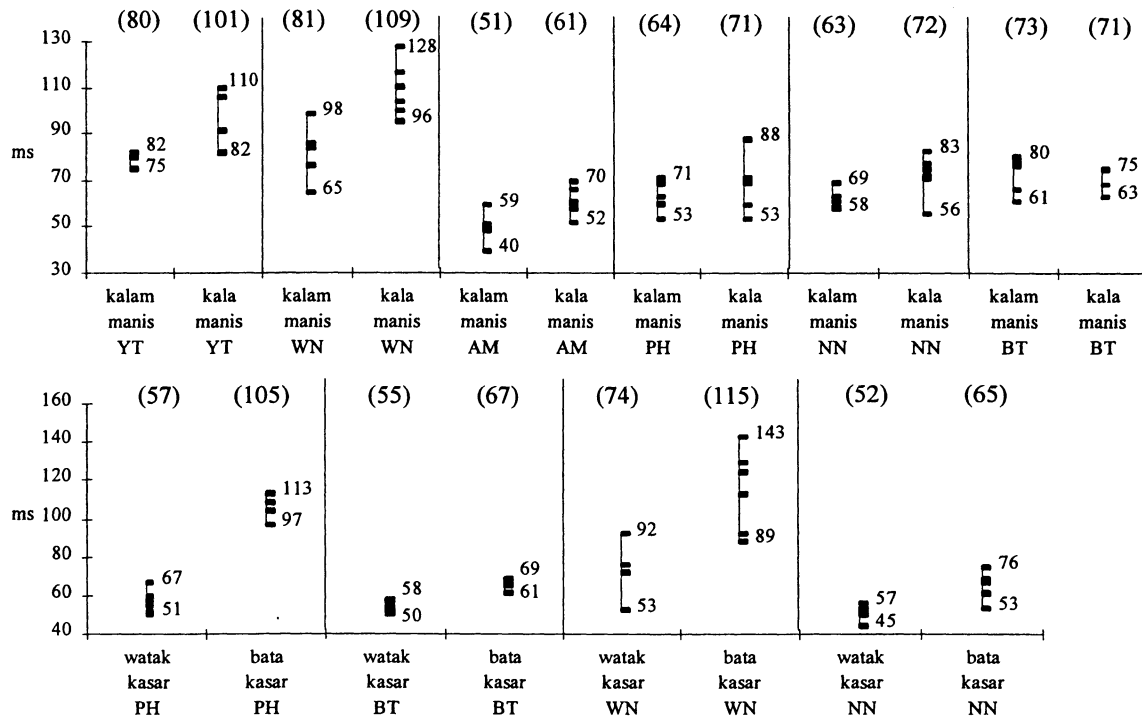


Figure 11. Ranges of duration (in ms) of vowels preceding single consonants vs. two-consonant sequences, for each speaker. Each point indicates the value of one of the six repetitions of the words; identical values are represented by one point.

In the nasal case, the values in the duration ranges for speaker YN show the least overlap. For speakers WN and AM, the duration ranges partially overlap; speakers PH, NN and BT show complete overlap. In the stop case, there is no overlap of the duration ranges for speakers PH and BT, and the distribution of the values is evenly spread. For speaker WN, the ranges overlap, which may be caused by the outlier, the highest value for this range, 92 ms, being so far apart from the rest. The ranges for speaker NN partially overlap.

Fig. 11 shows that the four speakers vary with respect to the magnitude of vowel shortening preceding a sequence of two nasals, compared to preceding a sequence of two stops. For speaker PH, the vowel preceding /kk/ is distinct from that preceding /k/; however, the vowel preceding /mm/ is not distinct from that preceding /m/, since the duration ranges completely overlap. This is also the case for speaker BT. This may suggest that for speakers PH and BT, the shortening effect for vowels preceding a two-consonant sequence is more evident when the consonant is a stop.

For speaker NN, the duration ranges of the vowel overlap preceding both the nasal and the stop, only partially in the stop case, and completely in the nasal case. For this speaker, vowels behave as if they are followed by a single consonant when the contrasted

consonants are nasals. However, when the consonants in these sequences are stops, the vowels preceding one consonant are distinct from those preceding the two-consonant sequence. For speaker WN, the duration ranges partially overlap in both the nasal and stop cases.

To summarize, vowels preceding two-consonant sequences tend to be shorter in duration relative to vowels preceding one consonant, for most speakers. In addition, when the nasal and the stop cases across a word boundary are compared, the difference in vowel duration preceding the nasals tends to be smaller than preceding stops. This suggests that vowels tend to be shortened more before a two-stop sequence than before a two-nasal sequence. Several factors may come into play, such as different manner of articulation (nasal vs. stop), and different place of articulation (bilabial vs. velar). The word boundary may also be a factor, since this phenomenon is not observed at the root-suffix boundary in most speakers (except speaker PH) (see Figs. 4 and 5). This is discussed in the following section.

3.3 Root-suffix vs. word-word cases

Comparing the duration differences at the two types of boundaries, the distinction between one consonant and two-consonant sequences is greater across the root-suffix boundary, markedly so for speaker YT and to a lesser extent for speakers WN and AM. In contrast, for speakers PH, BT, and NN, the duration differences for both the nasal and stop cases increase when they are across a word boundary, compared to when they are at the root-suffix boundary.

It is interesting to note that for speakers PH, BT, and NN, there seems to be a correlation between the differing magnitude of duration differences at morpheme boundaries and the overlapping of duration ranges. At the root-suffix boundary, duration differences are smaller and there is more overlapping of consonant duration ranges than at the word boundary. This suggests that the greater the magnitude of the duration difference between one consonant and two-consonant sequences, the less likely it is for the duration ranges of these consonants to overlap. Note, however, that this observation does not apply to speaker AM. For this speaker, the magnitude of duration difference across root-suffix boundary is greater than across word boundary, and yet the overlapping of duration ranges occurs in consonants at the root-suffix boundary, rather than at the word boundary. At this point, no explanation is offered for the behavior of this speaker.

4 Concluding remarks

This acoustic study examines whether identical consonants across the root-suffix boundary undergo phonological degemination. This question is extended to cases involving word boundaries. There are three possibilities previously mentioned, with respect to the phonetic realization of degemination. First, identical consonants at the morpheme boundaries undergo complete degemination, in which case we expect to find no duration difference between one consonant and two identical consonant sequences. Second, the underlying consonant sequence is fully preserved, resulting in fake geminates. Consequently, we might see the duration ratio of one consonant vs. two identical consonants to reflect that of a singleton-geminate contrast. Third, identical consonant sequences across root-suffix and word boundaries are only partially neutralized, in which case the duration of identical consonant sequences would tend to be greater than that of one consonant; however, the duration differences between the two may not be of the magnitude of a singleton-geminate contrast.

Based on the quantitative study presented here, we have seen that for most speakers, the duration of underlying two-consonant sequences tends to be longer than that of single consonants. However, as we have also seen, the duration differences between one and two consonants are smaller than those in a typical singleton-geminate contrast. This result leads us to argue that in Indonesian, which has no phonological geminate consonants, identical consonant sequences across morpheme boundaries undergo partial degemination.

Since a sequence of identical consonants across morpheme boundaries is not neutralized to one consonant, the first member of the sequence would be syllabified in coda position. The syllabification of the bilabial nasal in *ragammu* and *kalam manis* is as follows: *ra.gam.mu* and *ka.lam.ma.nis*. As Maddieson (1985) points out, vowels tend to be shorter in duration when they are in a closed syllable. In Indonesian, the duration of vowel preceding a two-consonant sequence tends to be shorter than preceding a single consonant, though not as dramatically as some of the cases discussed by Maddieson.

While the duration of two-consonant sequences is longer than single consonants, for some speakers, the duration differences fall below the differential threshold point, i.e. 20% of the segment duration, if one uses Klatt's ratio criterion. The question then is whether the duration of the consonant alone is sufficient as a perceptual cue to determine the distinction between one consonant and two identical consonants. Since the duration difference of the preceding vowel is not so large, it may or may not be the case that this vowel plays a role in the perception of one vs. two-identical consonant distinction. These

questions are currently being pursued in a perceptual experiment using tokens from natural speech, to find the differential threshold for one vs. two consonants in Indonesian. In this ongoing research, the duration of the preceding vowel is also taken into account.

In this study, I have considered only some cases that give rise to adjacent identical consonants. Of interest would be cases like *root + pun* and *root + lah*. These two suffixes convey emphatic meaning, which may or may not be realized phonetically as added duration. If they are, the duration of consonants in these cases may approach that in a singleton-geminate contrast. It is also of interest to add other clusters in the comparison.

We have seen that there is a great deal of variability among the speakers with respect to duration differences in the contrasted consonants and in the vowel preceding these consonants, as well as to the degree of overlap in the duration ranges. This variability may be caused by one factor or a combination of several factors. These factors include, but are not limited to, individual speaker variation, differing language backgrounds, the fact that the speakers are being recorded in a booth (thus the tendency for ‘correct pronunciation’), the fact that the words in the list are written with either one or two consonants. All of these are areas for further research, where some of these factors will be better controlled and others examined more systematically.

5 References

- Adisasmito-Smith, Niken. (1998) Constraints on the consonants in Indonesian. *Recent Papers in Austronesian Linguistics, UCLA Occasional Papers* 21, 1-18.
- Chen, Matthew. (1970) Vowel length variation as a function of the voicing of the consonant environment. *Phonetica* 22, 129-159.
- Cohn, Abigail and William Ham. (1998) *An acoustic study of Madurese consonants*. Presented at 8-ICAL, Taiwan, and in this volume.
- Delattre, Pierre. (1971a) Consonant gemination in four languages: An acoustic, perceptual, and radiographic study, Part 1. *IRAL* 9, 31-52.
- Delattre, Pierre. (1971b) Consonant gemination in four languages: An acoustic, perceptual, and radiographic study, Part 2. *IRAL* 9, 97-113.
- Dinnsen, Daniel. (1985) A re-examination of phonological neutralization. *Journal of Linguistics* 21, 265-279.
- Ham, William. (1998) *Phonetic and phonological aspects of geminate timing*. Doctoral dissertation, Cornell University.
- Kim, Hyunsoon and Allard Jongman. (1996) Acoustic and perceptual evidence for complete neutralization of manner of articulation in Korean. *Journal of Phonetics* 24, 295-312.

-
- Klatt, Dennis. (1973) Interaction between two factors that influence vowel duration. *Journal of the Acoustical Society of America* 54, 1102-1104.
- Klatt, Dennis. (1976) Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America* 59, 1208-1221.
- Ladefoged, Peter and Ian Maddieson. (1996) *The sounds of the world's languages*. Cambridge, MA: Blackwell Publishers.
- Lahiri, Aditi and Jorge Hankamer. (1988) The timing of geminate consonants. *Journal of Phonetics* 16, 327-338.
- Lapoliwa, Hans. (1981) *A generative approach to the phonology of Bahasa Indonesia*. Canberra: Australian National University. (Pacific Linguistics, Series D, No. 34)
- Lehiste, Ilse. (1970) *Suprasegmentals*. Cambridge, MA: MIT Press.
- Letterman, Rebecca S. (1994) A phonetic study of Sinhala syllable rhymes. *Working Papers of the Cornell Phonetics Laboratory* 9, 155-181.
- Maddieson, Ian. (1985) Phonetic cues to syllabification. In Fromkin, V. A. (ed.) *Phonetic Linguistics*. Orlando: Academic Press, 203-221.
- Peterson, Gordon, and Ilse Lehiste. (1960) Duration of syllabic nuclei in English. *Journal of the Acoustical Society of America* 32, 693-703.
- Podesva, Robert. (1998) *Buginese consonants: A phonetic and phonological study*. BA honors thesis, Cornell University.
- Port, Robert F., and Michael L. O'Dell. (1985) Neutralization of syllable-final voicing in German. *Journal of Phonetics* 13, 455-471.
- Slowiaczek, Louisa, and Daniel Dinnsen. (1985) On the neutralizing status of Polish word-final devoicing. *Journal of Phonetics* 13, 325-341.
-

Appendix 1

The averaged durations and the duration differences of /mm/ vs. /m/ and /kk/ vs. /k/, for each word.

	PH	WN	BT	NN	AM	YT	All speakers
ragam <u>mmu</u>	90	150	101	103	96	192	122
ragam <u>m</u>	78	109	83	94	74	97	89
Difference	12	41	18	9	22	95	33
salam <u>mmu</u>	89	161	100	110	99	206	128
jalam <u>m</u>	77	110	84	87	70	105	89
Difference	12	51	16	23	29	101	39
katak <u>kku</u>	96	184	70	121			118
katak <u>k</u>	73	131	60	86			88
Difference	23	53	10	35			30
watak <u>kku</u>	104	173	73	104			114
matak <u>k</u>	81	131	61	81			89
Difference	23	42	12	23			25

Appendix 2

The averaged durations and the duration differences of vowels preceding /mm/ vs. /m/, and preceding /kk/ vs. /k/, for each word.

	PH	WN	BT	NN	AM	YT	All speakers
ragam <u>mmu</u>	78	107	69	83	66	99	84
ragam <u>m</u>	89	129	67	88	79	137	98
Difference	11	22	-2	5	13	38	14
salam <u>mmu</u>	70	103	60	82	54	83	75
jalam <u>m</u>	74	120	66	98	78	133	95
Difference	4	17	6	16	24	50	20
katak <u>kku</u>	53	76	53	51			58
katak <u>k</u>	71	95	56	61			71
Difference	18	19	3	10			13
watak <u>kku</u>	58	84	58	52			63
matak <u>k</u>	71	104	58	64			74
Difference	13	20	0	12			11