Amplitude and F0 as acoustic correlates of Kelantan Malay word-initial geminates

[1]Mohd Hilmi Hamzah, [2]Janet Fletcher, [3]John Hajek

[1]Language Academy, Universiti Teknologi Malaysia, Malaysia
[2, 3] School of Languages & Linguistics, The University of Melbourne, Australia
hilmihamzah@utm.my, janetf@unimelb.edu.au, johnth@unimelb.edu.au

Abstract

This study examines the degree to which differences in amplitude and F0 can mark the word-initial singleton/geminate contrast in Kelantan Malay (KM). RMS amplitude and F0 values were measured for several consonant groups in utterance-initial and -medial positions. Results indicate that the contrast is reflected in systematic differences in the amplitude and F0 at the beginning of the following vowel. The findings support the view that, besides being primarily characterized by durational lengthening, word-initial geminates can be enhanced by non-durational correlates, such as amplitude and F0, particularly for voiceless stop tokens produced in utterance-initial position where domain-initial strengthening effects are also expected to be present.

Index Terms: amplitude, F0, geminate, Malay

1. Introduction

Many studies (e.g., [1, 2]) suggest that geminate consonants that occur in the word-edge environment, i.e., word-initial and word-final positions, are often "avoided", more so word-initially than word-finally. It was argued in [3] that geminates in word-initial position are considered unusual compared to those in other positions. The potential reason is that there will be fewer perceptual cues to closure duration in word-initial position, especially in the case of voiceless stops. In this latter context, the onset of closure is acoustically absent utterance-initially and therefore rendering closure duration perceptually indiscernible [2].

Previous instrumental studies analyzing languages with word-initial length contrasts (e.g., [4]) have made a number of claims that add to or even challenge the primary role of closure duration in defining the consonant length distinction. Additional acoustic correlates of word-initial geminates, such as the amplitude and F0 of the following vowel, have been found in a number of languages (e.g., [5]). In Pattani Malay [6], the Malay variety with which KM shares many phonological features, it was found that the RMS amplitude and F0 values are greater after word-initial geminates than their singleton counterparts. In other languages with word-medial geminate consonants, amplitude and F0 appear to be enhanced in one way or another surrounding geminates. In Japanese [7], for example, F0 is enhanced at vowel onset after geminates, which may be part of a strategy to compensate for the weakening of other cues, such as the voicing of the preceding closure phase.

Taking into account all these results, it appears that amplitude and F0 differences are potentially important acoustic properties that are worth examining in relation to a word-initial consonant length contrast. Further, the analysis of word-initial contrasts, such as in KM, is critical given earlier findings about an association between prosodic location (i.e., utterance-initial versus utterance-medial positions) and paradigmatic contrast enhancement (e.g., [8]). The current study, therefore, aims to determine whether the singleton/geminate consonant contrast in KM is accompanied by a larger set of acoustic parameters alongside closure duration, particularly in utterance-initial contexts. Results will add to the previous findings in KM that have shown some specific acoustic characteristics of word-initial geminates in terms of closure duration [9], VOT duration [10], post-consonantal vowel duration [11] and burst amplitude [12].

2. Method

2.1. Materials

An acoustic phonetic experiment was conducted in order to explore the potential roles of amplitude and F0 in marking the word-initial singleton/geminate contrast in KM. A list of thirty-eight tokens was prepared consisting of nineteen minimal pairs (presented in Table 1).

	Singleton		Geminate	
	Word	**Gloss**	**Word**	**Gloss**
/p/	/pitu/	door	/ppitu/	at the door
	/pagi/	morning	/ppagi/	early morning
/t/	/tido/	sleep	/ttido/	sleep by chance
	/tanɔh/	land	/ttanɔh/	outside
/k/	/kiɣi/	left	/kkiɣi/	to the left
	/kabo/	blurry	/kkabo/	beetle
/b/	/bini/	wife	/bbini/	married
	/bacɔ/	read	/bbacɔ/	is reading
/d/	/dike/	song	/ddike/	sing a song
	/dapo/	kitchen	/ddapo/	at the kitchen
/g/	/gigi/	teeth	/ggigi/	on the teeth
	/gaji/	salary	/ggaji/	sawing tool
/m/	/misa/	moustache	/mmisa/	moustached
	/maɣi/	come	/mmaɣi/	cupboard
/n/	/nikɔh/	marriage	/nnikɔh/	married
	/nanɔh/	pus	/nnanɔh/	getting pus
/l/	/lidɔh/	tongue	/llidɔh/	on the tongue
	/lapu/	lights	/llapu/	on the lights
/ŋ/	/ŋaŋɔ/	open the mouth	/ŋŋaŋɔ/	agape

Table 1. *List of stimuli and their glosses.*

All tokens were disyllabic words with either C(C)VCV or C(C)VCVC structures. Twenty phonemes were chosen and they were grouped according to voicing profile and manner of articulation: voiceless stops (/p/-/pp/, /t/-/tt/, /k/-/kk/); voiced stops (/b/-/bb/, /d/-/dd/, /g/-/gg/); and sonorants (/m/-/mm/, /n/-

/nn/, /ɲ/-/ɲɲ/, /l/-/ll/). Each phoneme was followed by two distinct vowels: the high front vowel /i/ and the low central vowel /a/, except /ɲ/-/ɲɲ/ (low central vowel /a/ only).

2.2. Speakers and Data Collection

The participants were sixteen native speakers of KM (8 males, 8 females) whose ages ranged from 20 to 28 (mean age: 22.4). Six of them were students from several universities in Melbourne, Australia, and ten were students from Universiti Malaysia Kelantan located in the state of Kelantan, Malaysia. All of them were born and raised in Kelantan, Malaysia.

For the speakers in Melbourne, the experimental materials were recorded individually in a professional recording studio in Melbourne. As for the speakers in Kelantan, they were recorded individually in a quiet room on the university campus in Kelantan. In all sessions, speakers were asked to repeat each token in two contexts: utterance-initial and utterance-medial positions. In the first context, the target word was preceded by a long silent pause, while in the second context, the target word was preceded by a vowel. The carrier sentence was: /diɔ katɔ (the target word) tigɔ kali/ "he said (the target word) three times."

All experimental tokens were presented in randomized order using a powerpoint presentation on a computer. The carrier sentence was written separately on a piece of A4 paper. Since there is no written counterpart of KM, Standard Malay orthography was used. The tokens were presented six times, each time in a different random order. The speakers were reminded to read them at a normal rate of speech. 228 utterances were recorded from each speaker in both contexts, yielding 7,296 utterances. The experiment took approximately one and a half hours for each speaker. They were compensated financially for their participation in the experiment.

2.3. Data Analysis

The waveform files were digitized at 44.1 kHz, segmented into single utterances for each participant and then coded accordingly. The segmentation and annotation were conducted using Praat version 5.1.11 [13]. RMS values, calculated in dB, were measured using R at vowel onset, i.e., at 0.1 (10%) into the vowel following the singleton/geminate consonant contrast. F0 values were calculated in hertz (Hz).

The results obtained from these measurements were statistically analyzed using R. A series of mixed-effects models [14] were used for statistical evaluation on various data sets using lme4 package in the statistical package R. The statistical analyses were conducted to test the significance of several main factors and their possible interactions. Due to lack of space, we present only the results for three main factors: Length (singletons and geminates), Utterance Position (utterance-initial and utterance-medial positions), and Manner of Articulation/Voicing (voiceless stops, voiced stops and sonorants). Speaker was treated as a random factor.

3. Results

3.1. Vowel onset amplitude

3.1.1. Utterance position effects

The mean RMS values at vowel onset after singletons and geminates according to utterance position are illustrated in Figure 1. Detailed measurements are given in Table 2.

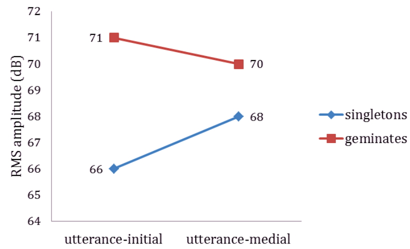

Figure 1: *Mean RMS values (dB) at vowel onset according to utterance position.*

Utterance-initial		Utterance-medial	
Singleton	**Geminate**	**Singleton**	**Geminate**
66 (6)	71 (6)	68 (6)	70 (6)

Table 2. *Mean RMS values (dB) at vowel onset according to utterance position. Standard deviations are indicated in parentheses.*

The results show that there is a significant main effect of Length on the RMS values (F(1,7293)=485.98, p<.001); mean values at vowel onset are always greater following geminates (red line) than following singletons (blue line) across utterance positions (both p<.001). There is a strong two-way interaction between Length and Utterance Position (F(1,7291)=64.404, p<.001); the magnitude of contrast between the two consonant length categories is larger in utterance-initial position (5 dB, p<.001) than in utterance-medial position (2 dB, p<.001). In utterance-initial position, the contrast magnitude reaches on average 5 dB, which is above the suggested just-noticeable-difference (JND) value of 3 dB for this parameter [15]. Note that the contrast enhancement shown in this position is due primarily to a slight fall in amplitude after singletons (blue line), i.e., there is a lowering of RMS values from 68 dB in utterance-medial position to 66 dB in utterance-initial position.

3.1.2. Manner of articulation/voicing effects

The RMS amplitude results for each manner of articulation/voicing type are shown in Figure 2 and summarized in Table 3. The mean RMS differences between singletons and geminates are all highly significant (p<.001) across consonant groups, with geminates being always associated with greater RMS values at the onset of the following vowel. There is a strong two-way interaction between Length and Manner of Articulation/Voicing (F(2,7289)=9.458, p<.001), which can be explained by a strong effect in the case of voiceless stops where the mean RMS differences between singletons and geminates are largest across utterance positions for this particular segment type (utterance-initial=5 dB; utterance-medial=3 dB, both p<.001). As for voiced stops and sonorants, the degrees of contrast between singletons and geminates are reduced somewhat, especially in utterance-medial position where the mean RMS differences do not reach the suggested JND value of 3 dB (voiced stops=2 dB; sonorants=1 dB, both p<.001).

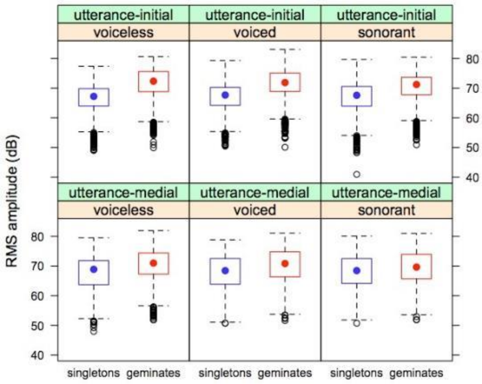

Figure 2: *Distribution of RMS values (dB) at vowel onset according to manner of articulation/voicing.*

Consonant group	Utterance-initial		Utterance-medial	
	Sin	Gem	Sin	Gem
Voiceless stops	66 (6)	71 (6)	67 (6)	70 (6)
Voiced stops	67 (6)	71 (6)	68 (6)	70 (6)
Sonorants	67 (6)	70 (6)	68 (6)	69 (1)

Table 3. *Mean RMS values (dB) at vowel onset according to manner of articulation/voicing. Standard deviations are indicated in parentheses.*

3.2. Vowel onset F0

3.2.1. Utterance position effects

The mean F0 values at vowel onset after singletons and geminates according to utterance position are illustrated in Figure 3. Detailed measurements are provided in Table 4.

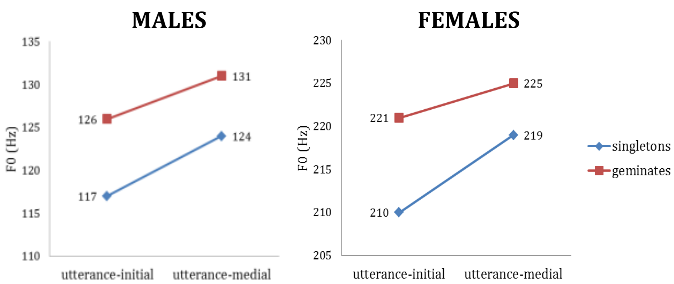

Figure 3: *Mean F0 values (Hz) at vowel onset according to utterance position: males (left); females (right).*

Utterance-initial		Utterance-medial	
Singleton	Geminate	Singleton	Geminate
MALES			
117 (11)	126 (10)	124 (11)	131 (9)
FEMALES			
210 (15)	221 (19)	219 (17)	225 (21)

Table 4. *Mean F0 values (Hz) at vowel onset according to utterance position. Standard deviations are indicated in parentheses.*

Length was found to be a significant main factor for both gender groups [males: $F(1,3594)=539.15$, $p<.001$; females: $F(1,3505)=199.80$, $p<.001$]; greater mean F0 values are always associated with geminates (red lines) across all contexts (all $p<.001$). The results are conditioned by Utterance Position [males: $F(1,3592)=16.182$, $p<.001$; females: $F(1,3503)=20.925$, $p<.001$]; greater F0 values are always found in utterance-medial position, which is most likely due to intonational effects associated with the full elicitation sentence (i.e., words were in focally prominent contexts). However, slightly greater enhancements of F0 contrast between singletons and geminates are always found in utterance-initial position (males=9 Hz; females=11 Hz, both $p<.001$). Note that, in utterance-initial position, the mean F0 differences are only reaching the suggested JND value for frequency, i.e., around 10 to 15 Hz depending on frequency regions (e.g., [16]). The degree of F0 contrast declines somewhat in the medial context (males=7 Hz; females=6 Hz, both $p<.001$).

3.2.2. Manner of articulation/voicing effects

The F0 results for each manner of articulation/voicing type are demonstrated in Figure 4 and summarized in Table 5.

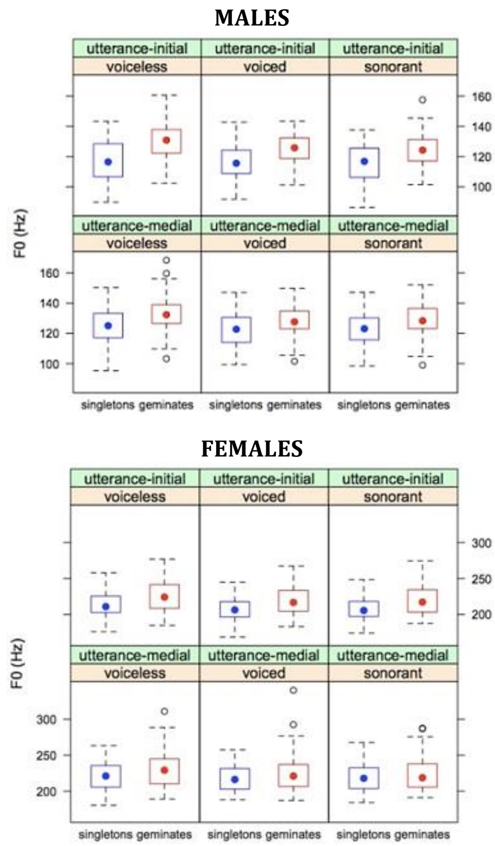

Figure 4: *Distribution of F0 values (Hz) at vowel onset according to manner of articulation/voicing: males (top); females (bottom).*

Consonant group	Utterance-initial		Utterance-medial	
	Sin	Gem	Sin	Gem
MALES				
Voiceless stops	118 (13)	130 (11)	125 (11)	133 (9)
Voiced stops	116 (10)	125 (9)	123 (10)	129 (8)
Sonorants	116 (11)	124 (9)	123 (10)	130 (9)
FEMALES				
Voiceless stops	213 (15)	226 (20)	222 (18)	230 (22)
Voiced stops	208 (15)	219 (18)	218 (16)	223 (20)
Sonorants	208 (15)	220 (19)	218 (17)	223 (20)

Table 5. *Mean F0 values (Hz) at vowel onset according to manner of articulation/voicing. Standard deviations are indicated in parentheses.*

It can be observed that the mean F0 values are always significantly greater ($p<.001$) after geminates across consonant groups, although the differences are relatively small. Length and Manner of Articulation/Voicing interact significantly [males: $F(1,3592)=14.234$, $p<.001$; females: $F(1,3503)=24.431$, $p<.001$], which can be accounted for by the F0 patterns after voiceless stops: the magnitudes of F0 contrast are reinforced more substantially after this particular stop type, which is consistent across utterance positions and gender groups (utterance-initial: males=12 Hz; females=13 Hz; utterance-medial: males=8 Hz; females=8 Hz, all $p<.001$). As for voiced stops and sonorants, the mean F0 differences are less in most contexts.

4. Discussion and Conclusions

In this study, we have examined the potential roles of two non-durational correlates (RMS amplitude and F0) in signaling the word-initial singleton/geminate contrast in KM in utterance-initial and -medial contexts. The overall results have shown that the consonant contrast is reflected through systematic variation in vowel onset amplitude and F0, although they are tempered by utterance position and manner of articulation/voicing types. It appears, therefore, that the consonant contrast in KM is characterized not only by durational parameters (e.g., [10]) but also by non-durational correlates in the following vowel. That is, KM geminates are signaled not only by longer closure duration, shorter VOT (for voiceless stops) and, to a certain degree, shorter post-consonantal vowel duration and greater burst amplitude (as presented in [9, 10, 11, 12]), but also by greater amplitude and F0 differences in the early part of the following vowel.

Further, because of these additional correlates, the saliency of the consonant contrast is preserved in utterance-initial position where the contrast is likely to be weakened. The fact that the singleton/geminate contrast in KM only occurs word-initially (and, by extension, utterance-initially) perhaps requires greater articulatory effort from native speakers to protect the distinction especially in utterance-initial position. This articulatory enhancement mirrors the additional non-durational reinforcement employed by the speakers of Pattani Malay [5] and also Cypriot Greek [17] to maintain the same word-initial consonant contrasts in these languages. This observation is also in line with the concept of domain-initial strengthening [8] and confirms the assertion in [18] that the phonological contrast is maximized in prosodically strong locations, i.e., utterance-initial position, although this requires further investigation in KM.

The current study also shows some important findings for voiceless stops in which there is a potentially different phonetic implementation strategy for this particular stop type at least, given the absence of acoustic closure duration in utterance-initial position. As for voiced stops and sonorants, the reduced effects in comparison to voiceless stops perhaps reflect the fact that there are already “internal cues” [7], i.e., voicing and resonance are long and audible enough during their initial closure phases to convey the contrast [9].

In a nutshell, these findings, together with those previously reported for KM (e.g., [10, 11, 12]), show that additional acoustic parameters may characterize the word-initial singleton/geminate contrast in KM alongside the major cue, i.e., closure duration. These additional cues contribute to acoustic enhancement of the consonant length contrast in KM, possibly functioning in combination with durational parameters to signal the consonant contrast, particularly in the case of voiceless stops. It is clear at this stage that there is a need for further articulatory examination in order to determine whether and how the difference in closure duration is maintained in utterance-initial voiceless stop tokens. Future work should also investigate spontaneous speech that may provide additional insights into word-initial consonant gemination in authentic communicative data where the influence of post-lexical prosody can also be examined.

5. Acknowledgements

The authors would like to thank Universiti Teknologi Malaysia for its financial support. This work was supported by the Potential Academic Staff (PAS) Grant no. Q.J130000.2741.00K78.

6. References

[1] Kraehenmann, A., “Swiss German stops: Geminates all over the word”, Phonology, 18(1), 109-145, 2001.

[2] Kraehenmann, A., and Lahiri, A., “Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops”, JASA, 123(6), 4446-4455, 2008.

[3] Ladefoged, P., and Maddieson, I., “Sounds of the world's languages”, Oxford: Blackwell Publishing, 1996.

[4] Ridouane, R., “Gemination in Tashlhiyt Berber: An acoustic and articulatory study”, JPhon, 37(2), 119-142, 2007.

[5] Abramson, A.S., “The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length”, in Papers from the 4th Annual Meeting of the Southeast Asian Linguistics Society, 1-20, 1998.

[6] Abramson, A. S., “Word-initial consonant length in Pattani Malay”, Proc. 11th ICPhS, 68-70, 1987.

[7] Kawahara, S., “Voicing and geminacy in Japanese: An acoustic and perceptual study”, University of Massachusetts Occasional Papers in Linguistics 31: Papers in Experimental Phonetics and Phonology, 87-120, 2005.

[8] Cho, T., and Keating, P., “Articulatory and acoustic studies on domain-initial strengthening in Korean”, JPhon, 29, 155-190, 2001.

[9] Hamzah, H., “Durational properties of initial geminate consonants in Kelantan Malay”, Proc. 13th Australasian International Conference on SST, 18-21, 2010.

[10] Hamzah, H., Fletcher, J., and Hajek, J., “Durational correlates of word-initial voiceless geminate stops: The case of Kelantan Malay”, Proc. 17th ICPhS, 89-92, 2011.

[11] Hamzah, H., Hajek, J., and Fletcher, J., “A taste of prosody: Possible effects of the word-initial singleton-geminate contrast on post-consonantal vowel duration in Kelantan Malay”, Proc. 6th Speech Prosody, 490-493, 2012.

[12] Hamzah, H., Fletcher, J., and Hajek, J., “An acoustic analysis of release burst amplitude in the Kelantan Malay singleton/geminate stop contrast”, Proc. 14th Australasian International Conference on SST, 85-88, 2012.

[13] Boersma, P., “Praat, a system for doing phonetics by computer”, Glot International, 5, 341-345, 2001.

[14] Baayen, R. H., Davidson, D. J., and Bates, D., “Mixed-effects modeling with crossed random effects for subjects and items”, Journal of Memory and Language, 390-412, 2008.

[15] Toole, F. E., and Olive, S., “The Modification of timbre by resonances: Perception and measurements”, JAES, 36 (3), 122-142, 1988.

[16] Klatt, D. H., “Discrimination of fundamental frequency contours in synthetic speech: Implications for models of speech perception”, JASA, 53, 8-16, 1973.

[17] Arvaniti, A., and Tserdanelis, G., “On the phonetics of geminates: Evidence from Cypriot Greek”, Proc. Interspeech 2000, 559-562, 2000.

[18] de Jong, K., “The supraglottal articulation of prominence in English: Linguistic stress as localized hyper-articulation”, JASA, 97(1), 491-504, 1995.