A Taste of Prosody: Possible Effects of the Word-Initial Singleton-Geminate Contrast on Post-Consonantal Vowel Duration in Kelantan Malay

Hilmi Hamzah, John Hajek, Janet Fletcher

School of Languages & Linguistics, The University of Melbourne, Australia

mhhamzah@student.unimelb.edu.au, johnth@unimelb.edu.au, janetf@unimelb.edu.au

Abstract

The present study examines possible temporal effects of the word-initial geminate consonant contrast in Kelantan Malay (KM) beyond the duration of the target consonant itself. We are interested in identifying any duration interactions with following vowels that could potentially enhance the acoustic salience of geminate consonants, in particular word-initial voiceless geminate stops. Results indicate that post-consonantal vowel duration can vary as a function of preceding consonant length and according to prosodic context. This is especially the case for post-consonantal low central vowels in utterance-medial position in which vowel duration is significantly shorter following geminate stops. Although the effect was not consistent across all speakers, the results agree with findings from other languages in that durational effects of geminate timing extend to adjacent segments as well as to much larger stretches of utterance.

Index Terms: geminate, duration, vowel quantity, Malay

1. Introduction

The acoustic properties of geminate consonants have been a subject of considerable debate among researchers especially with regard to durational correlates of geminates that occur word-initially, and *potentially* utterance-initially. A range of factors influencing adjacent vowels have been identified as *potentially* important secondary cues in distinguishing singleton and geminate consonants. These include (a) prosodic context, and (b) vowel duration - at least in pre-consonantal position - known to be shorter before word-medial geminate consonants than before corresponding singletons in many languages, such as Italian and Cypriot Greek, e.g. [1].

With respect to word-initial geminate consonants, we are not aware, however, of any previous investigations of possible interactions between these consonants and the duration of following vowels - the focus of this study. In earlier work on Pattani Malay [2], spoken in Thailand and with which KM shares many linguistic features, the investigation of vowel duration was restricted to the possible interaction with the following short consonants in the next syllable, not with the preceding singleton and geminate consonants (e.g. /a/ + /g/ in /pagi/ vs. /ppagi/).

The current study was also partly motivated by informal feedback from other phoneticians who, while finding the perception of the singleton-geminate contrast in word-initial position in KM difficult, reported some sensitivity to possible differences in vowel duration following geminate stops when compared to singleton stops. Earlier work of Abramson (e.g. [2]) on Pattani Malay has confirmed some secondary role in the word-initial singleton-geminate contrast for acoustic parameters such as RMS amplitude and F0 on following vowels. It is possible that utterance position and vowel duration may also be additional interacting factors in the same #C(:)V context in KM.

Our previous findings [3, 4] have firmly established that closure duration is the most robust acoustic parameter for the singleton-geminate contrast in KM, while VOT also plays a highly significant role. By looking at vowel duration in more detail, it is hoped that the results of this study will further enhance our understanding of how singleton and geminate consonants are produced in KM (and in other languages).

2. Method

2.1. Materials

In order to test for possible effects of the singleton-geminate contrast on post-consonantal vowel duration in KM, we designed an acoustic phonetic experiment in which a list of twelve tokens was prepared consisting of six minimal pairs (See 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/	a beetle

Table 1. *The KM tokens and their glosses.*

All tokens were disyllabic words with either C(C)VCV or C(C)VCVC structures presented in two different utterance positions (see 2.2). Three voiceless stop pairs were chosen: /p/-/pp/, /t/-/tt/ and /k/-/kk/. Each pair was followed by two distinct vowels: high front vowel /i/ and low central vowel /a/. These vowels were selected to determine any possible interaction between vowel duration and height [5].

2.2. Speakers and Data Collection

Sixteen native speakers of KM (eight males, eight females), whose ages ranged from 21 to 28 years, were recruited for this study. Ten speakers were students at the Universiti Malaysia Kelantan in Kelantan, Malaysia, while six speakers were students at the University of Melbourne, Australia. In Kelantan, the data were recorded in a quiet room on the university campus, while in Melbourne, the recording took place in a professional studio. At the time of recording, no speakers reported any history of hearing or speech disorders.

In all sessions, speakers were asked to repeat each token in two different utterance conditions: utterance-initial position (the target word was preceded by a long silent pause); and utterance-medial position (the target word was preceded by a vowel). The carrier sentence used in utterance-medial position was /diɔ katɔ (the target word) tigɔ kali/ (he said (the target word) three times) (adapted from [6]). All tokens were randomly presented through powerpoint slides on a notebook using the Standard Malay orthography. The carrier sentence was written separately on a piece of A4 paper. Six repetitions were recorded for each speaker with a total of 2,304 utterances. The experiment took approximately thirty minutes for each speaker. All speakers were duly compensated for their time and efforts.

2.3. Data Analysis

The waveform files were digitized at 44.1 kHz, segmented into single utterances for each speaker and then coded accordingly. The annotation was conducted using PRAAT [7] following standard segmentation and labeling procedures (See Figure 1).

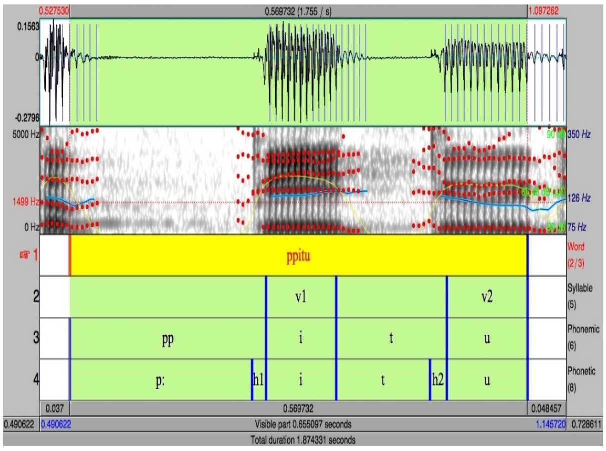

Figure 1. *An example of a sound file in a PRAAT window (a target word in a carrier sentence).*

The PRAAT template includes four tiers: the word tier, the syllable tier, the phonemic tier and the phonetic tier. The word tier (top tier) shows the segmentation and labelling of the target word. The syllable tier (second tier) highlights two vowels ('v1' and 'v2') in a disyllabic word. The underlying tier (third tier) represents the phonemic representation of the target word, and the surface tier (bottom tier) is the phonetic representation of the target word. VOT was marked as either 'h1' or 'h2' on the surface tier.

Post-consonantal vowel duration (marked as 'v1') was measured in all tokens produced in both utterance-initial and utterance-medial positions. It was calculated from the onset of voicing of the vowel (i.e. following the stop burst) until the offset of regular formant energy. All post-consonantal vowel duration values were extracted and analyzed statistically using the database analysis software EMU-R [8]. A mixed-effects model [9] was designed using the lme4 package in the statistical package R with Length (singleton or geminate), Utterance Condition (utterance-initial or utterance-medial) and Vowel Context (high front vowel or low central vowel) as fixed effects. Speaker was treated as a random factor. Where a significant result was found, Tukey's multiple comparisons of means were conducted. A repeated-measures ANOVA and *t*-tests were also employed for a more refined speaker-specific analysis.

3. Results

3.1. Utterance Condition and Vowel Context

The distribution of post-consonantal vowel duration in each utterance position and vowel context is illustrated in Figure 2. Tables 2 and 3 summarize the results of the statistical analysis of the corpus. Although consonant is not a significant main factor in determining vowel duration, the results show that it interacts highly significantly with Utterance Condition ($F(3,960)=65.17$, $p<.001$) and Vowel Context ($F(1,450)=301.91$, $p<.001$). A three-way interaction among these factors is also highly significant ($F(1,737)=15.30$, $p<.001$).

Random effects				
	Variance		Std. Deviation	
Residual	444.23		21.007	
Speaker	282.93		16.821	
Fixed effects				
	Estimate	Std. Error	*z*-value	Pr(>|*z*|)
(Intercept)	118.349	4.383	-7.440	0.001
Length	0.782	0.958	0.816	0.414
Length x Utterance Condition	-8.689	0.941	-9.232	0.000
Length x Vowel Context	-15.653	0.901	-17.38	0.000
Length x Utterance Condition x Vowel Context	-4.527	3.514	-6.619	0.000

Table 2. *Estimated coefficients, standard errors and associated z-statistics of fixed and random effects.*

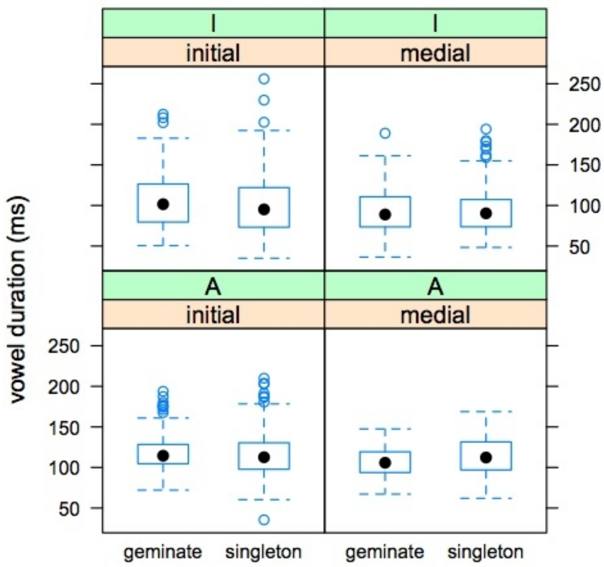

Figure 2. *Distribution of vowel duration values (ms) following singletons and geminates in each utterance condition and vowel context.*

Length	Position	Vowel	Mean (SD)
Singleton	Initial	/i/	101 (35)
Geminate			105 (32)
Singleton	Medial		94 (27)
Geminate			93 (26)
Singleton	Initial	/a/	117 (27)
Geminate			118 (22)
Singleton	Medial		114 (22)
Geminate			107 (17)

Table 3. *Mean vowel duration and standard deviation values (ms) following singletons and geminates in each utterance condition and vowel context.*

Tukey's multiple comparisons of means reveal that vowel duration values differ significantly following stop pairs in utterance-medial position and low central vowel context only (z=4.234, p<.001). However, the magnitude of durational difference is relatively small. Following these results and also due to space constraints, the next sub-sections will focus on the duration results for low central vowels in utterance-medial position.

3.2. Pooled Results

Figure 3 and Table 4 present the duration results for low central vowels in utterance-medial position following each singleton and geminate stop pair. The results are pooled across all speakers.

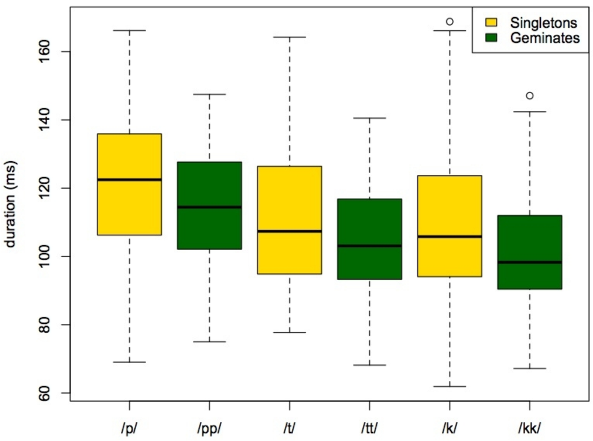

Figure 3. *Distribution of duration values (ms) of low central vowels following each stop pair in utterance-medial position for all speakers.*

Singleton	Mean (SD)	Geminate	Mean (SD)
/p/	121 (20)	/pp/	114 (16)
/t/	111 (21)	/tt/	104 (16)
/k/	110 (23)	/kk/	102 (18)

Table 4. *Mean duration and standard deviation values (ms) of low central vowels for all speakers*

It can be seen that the singleton-geminate contrast is reflected in the differences in post-consonantal vowel duration values in each stop pair. Overall duration values are *shorter* following geminate stops than singleton stops. A separate mixed-effects model was run on the data set with an additional fixed effect of Phoneme Type (/p/-/pp/, /t/-/tt/, /k/-/kk/). The results show that there is also a strong effect of Phoneme Type with vowel duration values shortest after velar stops (F(4,480)=42.15, p<.001). The durational differences are moderately significant in two pairs [/p/-/pp/ (t-ratio=2.88, p<.01) and /t/-/tt/ (t-ratio=2.71, p<.01)] and just significant in /k/-/kk/ (t-ratio=2.53, p<.05).

3.3. Speaker-specific Results

A speaker-specific analysis reveals that most speakers produce vowel duration differently at least some of the time, indeed following nearly half of the stop pairs. The adjusted p-values from t-test results for vowel duration following each stop pair produced by each speaker are displayed in Table 5 (shaded cells show significant pairs).

Speaker	/p/-/pp/	/t/-/tt/	/k/-/kk/
MS1	<.05	.065	.924
MS2	.909	<.004	.277
MS3	.301	<.005	.279
MS4	<.002	<.02	<.001
MS5	<.008	.168	.164
MS6	.173	.392	.186
MS7	.593	.168	<.03
MS8	.219	<.02	<.006
FS1	.711	.961	.222
FS2	<.04	.411	<.003
FS3	<.03	<.003	<.02
FS4	.378	.381	.820
FS5	.313	.076	.167
FS6	<.04	.762	.631
FS7	.407	.297	.157
FS8	<.001	<.006	<.007

Table 5. *The adjusted p-values for durational differences in low central vowels following each stop pair in utterance-medial position for each speaker.*

The results show that 19 out of 48 pairs are significantly different (40%): eleven of these are highly significant (appear in dark grey); and eight pairs are moderately significant (appear in light grey). As can be seen, the number of significant pairs is higher for bilabial stops (44%), while alveolar stops and velar stops share the same number of significant pairs (38%). It was also found that 89% of these significant pairs exhibit *shorter* vowel duration following geminate stops than singleton stops. Of sixteen speakers, eleven show significantly different vowel duration: six produce at least one significant pair (MS1, MS2, MS3, MS5, MS7, FS6); two produce two significant pairs (MS8, FS2); and three speakers produce all significant pairs (MS4, FS3, FS8).

A repeated-measures ANOVA was performed on all stop pairs for each speaker to test the overall effect of post-consonantal vowel duration. The results indicate that the effect is highly significant (p<.001) for six speakers (MS4, MS5, MS8, FS2, FS3, FS8), just significant (p<.05) for two speakers (MS1, FS5) and reaching significance (p=.07) for one speaker (MS7). The rest of the speakers do not produce any significant pairs. Gender differences are not observed since the number of significant pairs spreads almost equally among male and female speakers (male=53%; female=47%).

It is also worth noting that FS8 exhibits the strongest effect of post-consonantal vowel duration in each stop pair compared to other speakers (further details are given in Figure 4 and Table 6 on the next page). This particular speaker produces significantly different vowel duration values following all singleton and geminate stops (t-ratios=3.48 to 6.44, p<.01).

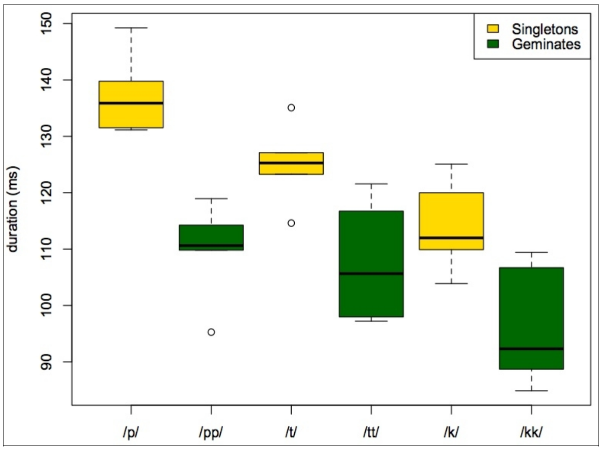

Figure 4. *Distribution of duration values (ms) of low central vowels following each stop pair in utterance-medial position for FS8.*

Singleton	Mean (SD)	Geminate	Mean (SD)
/p/	137 (7)	/pp/	110 (8)
/t/	125 (7)	/tt/	107 (10)
/k/	114 (8)	/kk/	96 (10)

Table 6. *Mean duration and standard deviation values (ms) of low central vowels for FS8*

Overall, despite the many significant differences across stop pairs (see Table 5), those reaching the suggested just noticeable difference (JND) value of 25 ms [10] were only found in three pairs (/k-/kk/ for MS8; /t/-tt/ for FS3; and /p/-/pp/ for FS8).

4. Conclusions

In this study, the first of its kind for this variety of Malay, our results suggest that for many speakers in our corpus, significant post-consonantal vowel duration differences can be associated with the short-long consonant contrast in word-initial position (at least after voiceless stops in KM) as tested here: vowels are shorter after word-initial geminates than after comparable singletons, just as they are often reported to be shorter before word-medial geminates, e.g. [1].

A number of conditioning factors on post-consonantal vowel duration were also identified and shown to have significant, albeit small, effects. In the first instance, utterance position is shown to be important with a stronger effect found in medial position. Such an outcome is surprising. Given the absence of closure duration information for voiceless stops in utterance-initial tokens, we would have anticipated instead a stronger effect in initial position - as speakers might be expected to enhance secondary cues such as vowel duration to reinforce perception of the short-long stop contrast in this specific context given the difficulty in determining consonant closure duration.

Vowel height effects are also apparent: duration differences are greater for /a/ than for /i/ - the result most likely of inherent duration properties associated with relative vowel height. With regard to stop place of articulation, there is a significant effect at all three places, but with the weakest effect after velars.

While speaker variability is notable, most subjects show at least some significant effects. However, the actual duration differences reported are not often at or above suggested JND, which indicates that vowel duration differences are not by themselves necessarily a reliable perceptual cue for the short-long consonant contrast. Our hypothesis, at least for the moment, given earlier work by Abramson on other prosodic cues in Pattani Malay (e.g. [2]), is that vowel lengthening/shortening may be part of a bundle of prosodic cues (e.g. F0 and amplitude) that are enhanced in effect when combined - with important secondary perceptual consequences for listeners. That the relatively small vowel duration differences we noted might function in this way in KM is anecdotally confirmed by the apparent sensitivity to them by non-native phoneticians when trying to distinguish long and short initial consonants, as reported in §1 above.

As part of a larger investigation into prosodic effects of word-initial gemination in KM, we also plan to investigate gemination-related duration effects over other prosodic units, i.e. syllable (e.g. type and position), and word (e.g. overall duration), to determine whether KM also exhibits similar kinds of compensation phenomena related to initial geminate consonants across a range of prosodic contexts.

Finally, given our hypothesis of a possible enhancing interaction between vowel duration and non-durational cues (F0 and RMS amplitude) on the same post-consonantal vowels, future work will also focus on determining the extent to which both F0 and RMS amplitude function as secondary acoustic and perceptual cues – both independently and possibly together with vowel duration.

5. Acknowledgements

We would like to thank the staff at the Universiti Malaysia Kelantan for their generous assistance during the recording in Kelantan, Malaysia. This work was also supported by the Arts Faculty and School of Languages and Linguistics at the University of Melbourne.

6. References

[1] Arvaniti, A. & Tserdanelis, G. 2000. On the phonetics of geminates: Evidence from Cypriot Greek. *Proc. 6th ICLP*, 559-562.

[2] Abramson, A.S. 1998. The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length, *Papers from the 4th Annual Meeting of the Southeast Asian Linguistics Society 1994*, 1-20.

[3] Hamzah, H. 2010. Durational properties of initial geminate consonants in Kelantan Malay. *Proc*. *SST 13*, 18-21.

[4] Hamzah, H., Fletcher, J. & Hajek, J. 2011. Durational correlates of word-initial voiceless geminate stops: The case of Kelantan Malay. *Proc. ICPhS XVII*, 89-92.

[5] Ladefoged, P. 2003. *Phonetic data analysis: An introduction to fieldwork and instrumental techniques*. Oxford: Blackwell.

[6] Abramson, A. S. 1986. The perception of word-initial consonant length: Pattani Malay. *JIPA* 16, 8-16.

[7] Boersma, P. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5(9-10), 341-345.

[8] Harrington, J. 2010. *The phonetic analysis of speech corpora*. Oxford: Blackwell.

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

[10] Klatt, D. H., & Cooper, W. E. 1975. Perception of segment duration in sentence contexts. In A. Cohen & S. Nooteboom (Eds.), *Structure and process in speech perception,* (pp. 69-89). Heidelberg: Springer.