WORD-INITIAL VOICELESS STOP GEMINATES IN KELANTAN MALAY: ACOUSTIC EVIDENCE FROM AMPLITUDE/F0 RATIOS

[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

The present study explores non-durational correlates of the word-initial voiceless stop singleton/geminate contrast in Kelantan Malay (KM) by focusing on the relative values of amplitude and F0 across two syllables of disyllabic words. We are particularly interested in identifying any inter-syllabic differences that could potentially enhance the acoustic salience of voiceless stop geminates in utterance-initial position, given the lack of acoustic closure duration in this specific context.

Results indicate that relative amplitude and relative F0 across syllables do contribute to the differentiation of word-initial voiceless stop singletons and geminates in KM, with amplitude and F0 ratios being usually significantly higher in the geminate context. The effect is further enhanced for utterance-initial tokens and it is also consistent across places of articulation. The conclusion is drawn that the word-initial voiceless stop length contrast in KM can also be characterized by differing amplitude/F0 ratios alongside the well-known primary acoustic correlate of closure duration.

Keywords: geminate, voiceless stops, Kelantan Malay, amplitude, F0

1. INTRODUCTION

Some studies have suggested that supplementary properties distinguishing singleton and geminate consonants can be extended across syllable boundaries via, among others, the relative amplitude and relative F0 relationship (e.g., [1]). In an instrumental study conducted for Pattani Malay, a closely related Malay variety with which KM shares many phonological features, Abramson [1] reports higher amplitude and F0 ratios for words beginning with voiceless stop geminates than for those beginning with their singleton counterparts.

The relative contribution of F0 and/or amplitude across syllables has also been suggested as a possible enhancing criterion in some languages with word-medial geminates. In Japanese, the relative values between syllables are interpreted in terms of a pitch contour shape; it was stated in [2] that F0 is higher in vowels preceding voiceless stop geminates than in those following them, suggesting that a pitch contour falls from the first syllable to the second in the geminate context. Likewise, the intensity of the pre-geminate syllable relative to that of the post-geminate syllable is also reported to be greater in this language. In a follow-up experiment [3], falling pitch is considered as an important auxiliary cue for Japanese voiceless stop geminates.

On the basis of these findings, it appears that relative amplitude and relative F0 may also help define and enhance the word-initial voiceless stop length contrast in KM and other languages. The analysis of the word-initial stop contrast, 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., [4]). That is, the effect of domain-initial strengthening on secondary cues such as amplitude and F0 is expected to be more apparent in the case of utterance-initial voiceless stop geminates since the closure onsets are acoustically missing for perceptual use in this utterance environment. Previous empirical studies on KM have explored some specific acoustic characteristics of word-initial voiceless stop geminates in terms of closure duration [5], VOT duration [6], post-consonantal vowel duration [7], burst amplitude [8] and post-consonantal amplitude and F0 [9]. In the current study, we aim to determine the extent to which the voiceless stop length contrast in KM is further characterized and enhanced by systematic differences in the amplitude/F0 ratios across two syllables.

2. METHOD

2.1. Materials

An acoustic phonetic experiment was conducted in order to explore the potential roles of amplitude/F0 ratios across syllables in further marking the word-initial voiceless stop singleton/geminate contrast in KM. As presented in Table 1, the KM data in this study contained twelve tokens consisting of six minimal pairs opposing singletons to geminates in

two vowel contexts: the high front vowel /i/ and the low central vowel /a/. The singletons and geminates represent three types of voiceless stops: /p, t, k/. All tokens were disyllabic words with either C(C)VCV or C(C)VCVC structures presented in two utterance conditions: utterance-initial and utterance-medial positions (see § 2.2 below).

Table 1: List of stimuli and their glosses.

	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

2.2. Speakers and Data Collection

Sixteen KM native speakers, eight males and eight females, participated in the experiment. All of them had spent their childhood in the state of Kelantan, Malaysia, and no speaker displayed any other regional accent or any speech disorder at the time of recording. They were all university students, with a mean age of 24 years, and were paid for their participation. The recording took place in a professional studio in Melbourne (6 speakers) and a quiet room on the university campus in Kelantan (10 speakers).

In all sessions, speakers were instructed to read the tokens six times each in isolation and a carrier sentence. Each item was randomly presented on a computer screen in Standard Malay orthography. The carrier sentence /diɔ katɔ (the target word) tigɔ kali/ (/he said (the target word) three times/) was written down on a piece of A4 paper. The total data collected consisted of 2,304 utterances (6 singletons + 6 geminates x 2 utterance positions x 16 speakers x 6 repetitions).

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 [10]. Following [1], amplitude/F0 ratios were obtained by dividing the amplitude/F0 values (measured in decibels and hertz, respectively) at vowel onset in the first syllable by those of the second syllable. A ratio value greater than 1 would imply that the first syllable was produced with greater amplitude/F0 values than the second syllable. On the other hand, a value below 1 would indicate that there was a lower level of energy or pitch in the initial syllable as compared to the following syllable.

The results were statistically analyzed using R. Two series of mixed-effects models [11] were used for statistical evaluation on the amplitude/F0 ratio data sets using lme4 package in the statistical package R. In each model, 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 Place of Articulation (/p, t, k/). Speaker was treated as a random factor.

3. RESULTS

3.1. Overall results

Results from mixed-effects models are shown in Table 2 below.

Table 2: Estimated coefficients, standard errors and associated *z*-statistics of fixed effects for two data sets: amplitude ratio (top); F0 ratio (bottom).

AMPLITUDE RATIO

Fixed effects				
	Estimate	Std. Error	*z*-value	Pr(>|*z*|)
(Intercept)	1.042	0.004	242.6	<.001
Length	-0.054	0.002	-21.83	<.001
Utterance Position	0.006	0.003	2.191	<.05
Place of Articulation	-0.061	0.003	-19.88	<.001
Length x Utterance Position	-0.077	0.003	-22.76	<.001
Length x Place of Articulation	-0.062	0.004	-16.19	<.01

F0 RATIO

Fixed effects				
	Estimate	Std. Error	*z*-value	Pr(>|*z*|)
(Intercept)	1.069	1.015	69.93	<.001
Length	-0.079	0.004	-19.89	<.001
Utterance Position	-0.055	0.004	-13.30	<.001
Place of Articulation	-0.022	0.005	-4.219	<.001
Length x Utterance Position	-0.104	0.005	-19.51	<.001
Length x Place of Articulation	-0.087	0.007	-12.74	<.05

It can be seen that Length is a significant main factor for both ratio types [amplitude ratio: $F(1,2301)=446.78$, $p<.001$; F0 ratio: $F(1,2238)=284.34$, $p<.001$], indicating that both amplitude and F0 ratios vary as a function of the voiceless stop length contrast. Other fixed factors also exhibit significant main effects: Utterance Position [amplitude ratio: $F(1,2301)=4.5347$, $p<.05$; F0 ratio: $F(1,2238)=124.53$, $p<.001$]; and Place of Articulation [amplitude ratio: $F(2,2300)=188.74$, $p<.001$; F0 ratio: $F(2,2237)=8.7552$, $p<.001$].

With regard to the interaction between factors, the results show that, for both ratio types, there is a significant interaction between Length and Utterance Position [amplitude ratio: $F(1,2299)=90.9403$, $p<.001$; F0 ratio: $F(1,2236)=31.705$, $p<.001$] and also between Length and Place of Articulation [amplitude ratio: $F(2,2297)=6.5533$, $p<.01$; F0 ratio: $F(2,2234)=0.6777$, $p<.05$], indicating that the degree of the voiceless stop length contrast in terms of relative values differs depending on utterance environment and phoneme identity.

In the following sections, the effects of Utterance Position (§3.2) and Place of Articulation (§3.3) are described in detail.

3.2. Utterance position effects

Figure 1 illustrates the effect of utterance position on amplitude ratio (top) and F0 ratio (bottom). Detailed measurements are summarized in Table 3.

Figure 1: Mean ratios of amplitude (top) and F0 (bottom) according to utterance position.

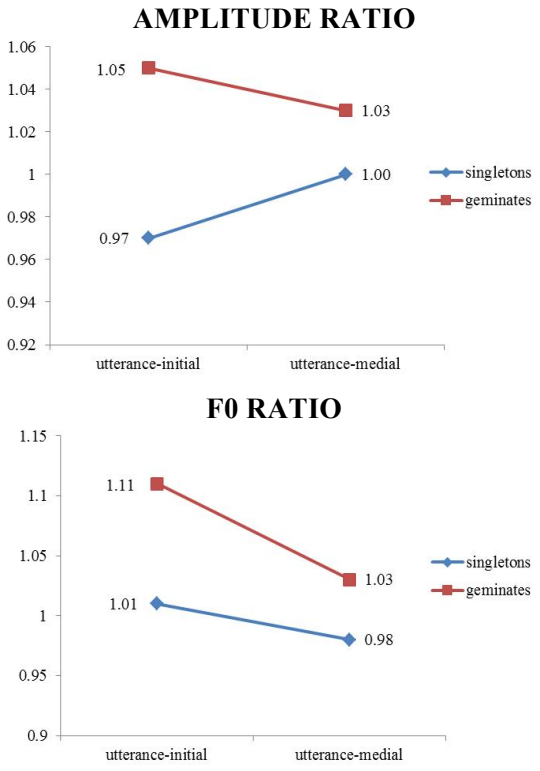

Table 3: Mean amplitude and F0 ratios according to utterance position. Standard deviations are indicated in parentheses.

Utterance-initial		Utterance-medial	
Singleton	**Geminate**	**Singleton**	**Geminate**
AMPLITUDE RATIO			
0.97 (0.06)	1.05 (0.06)	1.00 (0.06)	1.03 (0.06)
F0 RATIO			
1.01 (0.09)	1.11 (0.12)	0.98 (0.09)	1.03 (0.12)

It can be observed that the ratios are always significantly higher (all $p<.001$) in the geminate environment (red lines) than in the singleton environment (blue lines) across utterance positions. Moreover, there are always greater contrast enhancements between voiceless stop singletons and geminates in utterance-initial position across ratio types (mean differences: amplitude ratio=0.08; F0 ratio=0.10, both $p<.001$). The other important finding is that such enhancements in utterance-initial position are implemented differently between the two ratio types. On one hand, there is a lack of enhancement of the amplitude ratio for words beginning with voiceless stop singletons in which the mean amplitude ratio is instead decreased in utterance-initial position (i.e., 0.97). On the other hand, the utterance-initial enhancement is always present for F0 ratio, especially for words beginning with voiceless stop geminates in which there is a particularly large increase in the mean F0 ratio (i.e., 1.11); this specific heightening associated with voiceless stop geminates can be interpreted as the largest pitch fall across categories in the data set from the first to the second syllable of a disyllabic word, thus reinforcing the voiceless stop singleton/geminate contrast in this utterance context. In utterance-medial position, there are smaller degrees of separation between words beginning with singletons and geminates for both ratios of amplitude (0.03) and F0 (0.05).

3.3. Place of articulation effects

Figure 2 demonstrates the distribution of relative values of amplitude and F0 according to place of articulation. Table 4 summarizes the detailed measurements. It can be seen that the greater paradigmatic contrast enhancements between voiceless stop singletons and geminates in utterance-initial position, as observed in the previous section, are generally consistent across places of articulation (range of mean differences=0.06 to 0.11, all $p<.001$). The clearest differences can be observed in the cases of (1) bilabials for amplitude ratio (mean difference=0.10) and (2) bilabials and velars for F0 ratio (mean differences=both 0.11). The mean ratio differences between singletons and geminates for other places of articulation (amplitude ratio: alveolars=0.08, velars=0.06; F0 ratio: alveolars=0.09) are also still larger compared to the medial context.

In utterance-medial position, the mean differences between voiceless stop singletons and geminates are all reduced across places of articulation for both ratio types but still significant (range of mean differences=0.02 to 0.06, all $p<.001$).

The smallest differences in magnitude between singletons and geminates are more evident in the amplitude ratio data, particularly for velars in which the mean ratio difference between length categories is only 0.02. As for the F0 ratio data, the mean differences between utterance-medial singletons and geminates are not markedly different across places of articulation.

Figure 2: Distribution of ratios for amplitude (top) and F0 (bottom) according to place of articulation.

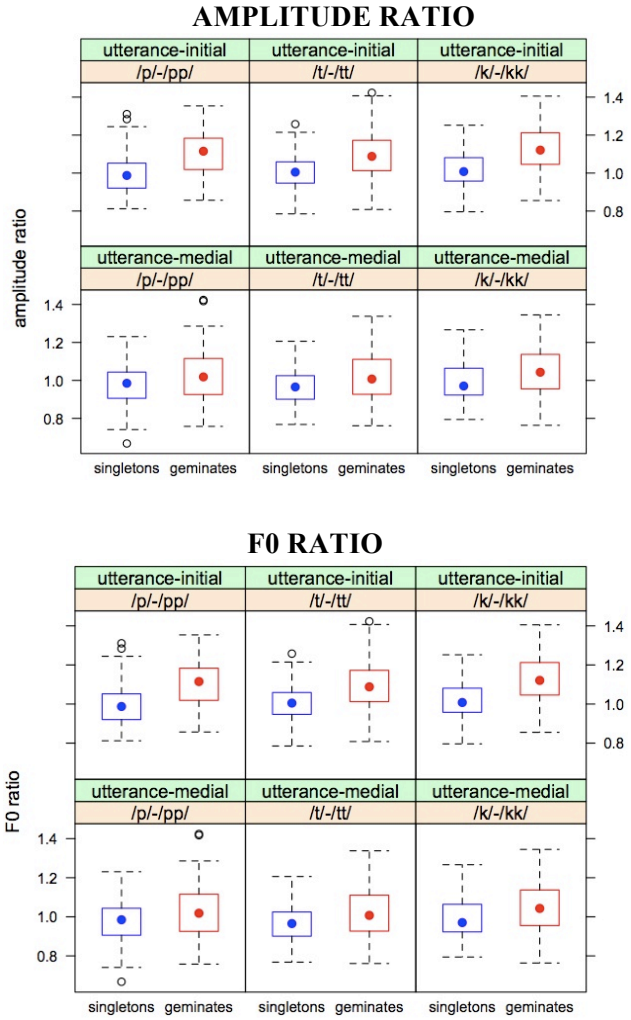

Table 4: Mean amplitude and F0 ratios according to place of articulation. Standard deviations are indicated in parentheses.

	Utterance-initial		Utterance-medial	
	Singleton	**Geminate**	**Singleton**	**Geminate**
AMPLITUDE RATIO				
/p/	0.99 (0.06)	1.09 (0.07)	1.04 (0.07)	1.07 (0.07)
/t/	0.95 (0.05)	1.03 (0.06)	0.97 (0.05)	1.00 (0.05)
/k/	0.98 (0.05)	1.04 (0.05)	1.00 (0.03)	1.02 (0.03)
F0 RATIO				
/p/	1.00 (0.10)	1.11 (0.12)	0.98 (0.10)	1.02 (0.12)
/t/	1.01 (0.09)	1.10 (0.12)	0.97 (0.09)	1.02 (0.11)
/k/	1.02 (0.09)	1.13 (0.11)	0.99 (0.09)	1.05 (0.12)

4. DISCUSSION AND CONCLUSIONS

In this paper, we have examined the degree to which inter-syllabic differences in amplitude and F0 are potentially associated with the word-initial voiceless stop singleton/geminate contrast in KM. The finding for KM is generally consistent with the interpretation of ratio results in Pattani Malay [1] in that the initial syllables of disyllabic words beginning with voiceless stop geminates are always associated with greater salience (i.e., higher amplitude and F0 ratios) as compared to those beginning with their singleton counterparts. In addition, the larger fall in amplitude and F0 values across syllables always accompanies KM disyllabic words beginning with voiceless stop geminates regardless of utterance position. The F0 pattern observed across syllables in KM is also generally similar to the observation of a falling pitch pattern in Japanese word-medial voiceless stop geminates reported in [2] and [3], although a finer analysis of F0 is required to establish the precise relationship between consonant gemination in KM and this particular pitch pattern.

In the present study, it was also found that there is a greater contrast enhancement between voiceless stop singletons and geminates in terms of amplitude/F0 ratios specifically in utterance-initial position. Further, this trend is also true across all places of articulation. This finding, which also accords well with the concept of domain-initial strengthening (e.g., [4]), can be related to the absence of "audible duration" information in the onset of word-initial stop closures in utterance-initial position. That is, greater amplitude/F0 ratios may reflect greater articulatory effort required to deal with perceptual difficulties in determining the presumably longer closure duration in utterance-initial voiceless stop geminates. Speakers therefore produce more forceful articulations that produce the secondary acoustic cues observed in this study. It can be speculated that, to certain extent, native listeners may also rely on these auxiliary cues to perceive the voiceless stop contrast; further investigation in geminate perception is warranted to verify this speculation. The reduced differentiation of amplitude/F0 ratios observed for utterance-medial tokens is not surprising in light of the clear perceptibility of durational differences in utterance-medial stop closures, as already shown in the earlier findings of exceptionally high singleton/geminate closure duration ratios in KM [5, 6] and also Pattani Malay [12]. In other words, amplitude/F0 ratios only serve as secondary or possibly concomitant cues alongside clearly perceptible closure duration differences in this particular utterance environment.

5. ACKNOWLEDGEMENTS

This research was conducted at the Phonetics Laboratory, the University of Melbourne. We would like to thank sixteen Kelantan Malay speakers for their participation.

6. REFERENCES

[1] 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*, 1-20.

[2] Idemaru, K., Guion, S. G. 2008. Acoustic covariants of length contrast in Japanese stops. *Journal of the IPA* 38(2), 167-186.

[3] Kubozono, H., Takeyasu, H., Giriko, M., Hirayama, M. 2011. Pitch cues to the perception of consonant length in Japanese. *Proc. 17th ICPhS* Hong Kong, 1150-1153.

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

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

[6] Hamzah, H., Fletcher, J., Hajek, J. 2011. Durational correlates of word-initial voiceless geminate stops: The case of Kelantan Malay. *Proc. 17th ICPhS* Hong Kong, 815-818.

[7] Hamzah, H., Hajek, J., Fletcher, J. 2012a. A taste of prosody: Possible effects of the word-initial singleton-geminate contrast on post-consonantal vowel duration in Kelantan Malay. *Proc. 6th International Conference on Speech Prosody* Shanghai, 490-493.

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

[9] Hamzah, M. H., Fletcher, J., Hajek, J. 2014. Amplitude and F0 as acoustic correlates of Kelantan Malay word-initial geminates. *Proc. 15th Australasian International Conference on SST* Christchurch, 63-66.

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

[11] 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.

[12] Abramson, A. S. 1987. Word-initial consonant length in Pattani Malay. *Proc. 11th ICPhS* Tallinn, 68-70.