The Role of Closure Duration in the Perception of Word-Initial Geminates in Kelantan Malay

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

[1]School of Education and Modern Languages, Universiti Utara Malaysia, Malaysia
[2] School of Languages & Linguistics, The University of Melbourne, Australia
hilmihamzah@uum.edu.my, johnth@unimelb.edu.au, janetf@unimelb.edu.au

Abstract

This study examines the extent to which closure duration plays a role in perceptually cueing Kelantan Malay (KM) word-initial geminates. Three word-pairs (/k-kk/, /b-bb/, /l-ll/) embedded in a carrier sentence were chosen for manipulation. In each pair, the closure phase in word-initial consonants was manipulated. Results show that KM listeners shift their perception from singletons to geminates, or vice versa, in lengthened singletons and shortened geminates respectively. The findings support the view that closure duration is a robust acoustico-perceptual cue to word-initial consonant gemination.

Index Terms: closure duration, word-initial geminates, consonant contrast, geminate perception, Kelantan Malay

1. Introduction

It is well established across languages that the singleton/geminate contrast in consonants, either word-initially or word-medially, is distinguished primarily by a difference in closure duration (e.g., [1]). Perceptually, closure duration has also been shown to be a powerful cue to consonant gemination across languages (e.g., Cypriot Greek [2]; Pattani Malay [3]). In the case of word-initial geminates, the perceptual salience of this parameter has been demonstrated in a series of perception experiments in Pattani Malay (henceforth PM), the Malay variety with which KM shares many phonological features. These experiments (e.g., [3]) involved manipulating the closure phase of voiceless stop singletons and geminates embedded in utterance-medial position and determining the perceptual boundary between singleton and geminate consonants, i.e., the duration value that corresponded to a categorical shift in perception.

For instance, in a perception experiment using the manipulated word-pair /paka/ 'to use' versus /ppaka/ 'usable' [3], it was found that a difference in closure duration is sufficient to cue the PM short versus long distinction; lengthening the word-initial singleton /p/ leads to more responses for the geminate category, while shortening the word-initial geminate /pp/ results in more responses for the singleton category. In both cases, there are perceptual crossovers with complete perceptual shifts to the opposite consonant categories, in line with the acoustic findings for PM [4] and supported by other studies (e.g., [5]) dealing with the perceptual effect of closure duration on consonant gemination.

However, the PM study in [3] also shows that the crossover-points are different between the groups of stimuli; for the stimuli made from the original word with the singleton /p/, the 50% perceptual crossover point is at 120 ms, while for those created from the original word with the geminate /pp/, the perceptual shift begins at 104 ms, i.e., 16 ms earlier than that for the former group of stimuli. The difference between these two crossover points is statistically reliable (p<.001), suggesting that there may be additional acoustic cues at play that influence the placement of the category boundary.

In KM, the robust role of closure duration in characterizing the production of the singleton/geminate contrast has been demonstrated in earlier acoustic studies, e.g., [6]. In the current study, we aim to examine whether and how closure duration also serves as a perceptual cue to KM word-initial geminates among KM native listeners. We are especially interested in examining what contribution, if any, controlled changes in closure duration make to the perception of the word-initial singleton/geminate contrast in KM. In addition, we also aim to identify whether there are ambiguous zones between the stimuli created from original singletons and original geminates that may suggest the presence of other acoustic cues alongside closure duration.

2. Method

2.1. Materials

The word-pairs chosen for manipulation in this study are displayed in Table 1 below. The word-pairs consist of three consonant groups (i.e., voiceless stops /k-kk/, voiced stops /b-bb/ and sonorants /l-ll/). This choice of consonant group is based on Abramson's methodology in his investigation of word-initial geminates in PM [3]. Since the manipulation of acoustic closure duration for voiceless stop tokens is only possible for those recorded in utterance-medial position, all word-pairs were taken from this environment to ensure consistency.

Table 1. *Sources of stimuli*.

Consonant group	Singleton		Geminate	
	Word	**Gloss**	**Word**	**Gloss**
Voiceless stops	/kabo/	blurry	/kkabo/	a beetle
Voiced stops	/batʃɔ/	read	/bbatʃɔ/	is reading
Sonorants	/lapu/	lights	/llapu/	on the lights

The procedures for manipulation were achieved by using the manipulation editor in Praat version 5.1.11 [7]. Closure duration was varied along a set of continua in two ways. First, the closure duration values of each singleton were *lengthened* in a series of controlled steps. For the voiceless and voiced

stop singletons /k/ and /b/, their duration values were each lengthened from their original values in eleven 10-ms steps, resulting in twelve stimuli each including their original singletons. For the sonorant singleton /l/, its closure duration values were lengthened from the original values in ten 10-ms steps, resulting in eleven stimuli including the original singleton. The duration values of the longest variant for each singleton group were all similar to their original geminate counterparts, i.e., 173 ms for the longest /k/, 180 ms for the longest /b/, and 170 ms for the longest /l/.

Second, the closure duration values of each geminate were *shortened* in a series of controlled steps. For the voiceless and voiced stop geminates /kk/ and /bb/, their duration values were each *shortened* from their original values in eleven 10-ms steps, resulting in twelve stimuli each including their original geminates. For the sonorant geminate /ll/, its closure duration values were shortened from the original values in ten 10-ms steps, resulting in eleven stimuli including the original geminate. The duration values of the shortest variant for each geminate group were all similar to their original singleton counterparts, i.e., 63 ms for the shortest /kk/, 70 ms for the shortest /bb/, and 70 ms for the shortest /ll/. The total number of manipulated stimuli that consisted of all the singleton and geminate variants including their original words was 24 each for the voiceless and voiced stop pairs /k-kk/ and /b-bb/, and 22 for the sonorant pair /l-ll/. Altogether, the manipulation of closure duration for each word in all three pairs yielded 70 stimuli. They were presented three times to the listeners, creating a total of 210 manipulated trials.

2.2. Listeners and Data Collection

The participants for all perception experiments were 30 undergraduate students (15 males, 15 females), all native speakers of KM, at the Universiti Malaysia Kelantan, Kelantan, Malaysia. Their ages ranged between 20 to 25 years (mean age: 21.2). At the time of the experiments, they exhibited no symptoms of hearing disability. All of the listeners were born and raised in Kelantan, Malaysia.

The listeners participated individually in the perception experiment in a quiet room at the Universiti Malaysia Kelantan. They were seated at a desk and were fitted with a stereo headphone in order to listen to the experiment stimuli. All the stimuli were presented through a computer using Praat's Experiment Multiple Forced Choice listening experiment (version 5.1.11). Three experiment files were designed that fitted the three minimal pairs tested in this experiment. These files were played in Praat during the experiment.

The first author gave verbal instructions to the listeners in their native language, i.e., KM. The listeners were first trained with a few stimuli before the experiment took place so that they were comfortable with the experiment design. The listeners then read the instruction on a computer screen, informing them to listen to a sound and choose the word that most closely resembled to what they were going to listen. Since there is no written counterpart of KM, all the words were written in Standard Malay that corresponded to the stimuli in KM.

During the experiment, only one pair was tested at a time. Therefore, the same response categories appeared on the screen for each pair. For each stimulus, there was initial silence duration of 1.5 seconds. The listeners were allowed to replay each stimulus once. The experiment files were run separately but sequentially and controlled by the first author. There was one break after the 36th stimulus for each pair. The experiment lasted for approximately thirty minutes for each participant. All listeners were financially compensated for their participation.

2.3. Data Analysis

The results for each listener were saved in a separate response file. The responses for each listener and each stimulus were processed using Excel (version 14.1.0). Since the stimuli were presented three times in each experiment, the scores for the correct responses ranged from 0 to 3. These scores represented the correct responses for geminates. The total scores for each stimulus were then converted into percentages and plotted into response curves. Geminate responses to each series of stimuli made from original words with singletons or geminates were submitted to one-way ANOVA tests using SPSS (version 20.0.0) to determine their significance levels.

Following [3], the differences observed between the two series of stimuli in the 50% crossover points (i.e., the boundary on a response curve where the listeners' perception switches from singletons to geminates or vice versa) were calculated and compared statistically using ANOVA. Samples paired *t*-tests were also employed to test the level of significance of geminate responses between the two groups of stimuli at a specific step on a duration continuum.

3. Results

The response curves in Figure 1 illustrate the perception results for manipulated stimuli, showing mean percentages of geminate responses for (a) /kabo/-/kkabo/, (b) /batʃɔ/-/bbatʃɔ/ and (c) /lapu/-/llapu/. Detailed measurements underlying these response curves are provided in Table 2. The leftmost durations are the original durations for lengthened singletons, while the rightmost durations are the original durations for shortened geminates. Horizontal lines show the crossover zones between singletons and geminates at 50%, while vertical lines indicate the 50% crossover points between the two series of stimuli.

As shown in the response curves in Figure 1(a-c), there are in general similar patterns across all word-pairs: lengthening of original singletons brings about more geminate responses, while shortening of original geminates causes fewer geminate responses. The response curves are all categorical; in each case, there is a clear perceptual boundary with a single 50% crossover point for each series of stimuli. At the beginning of the duration continua where the original closure durations for words with singletons are located, geminate responses are almost virtually non-existent across word-pairs. That is, almost all the stimuli are identified as singletons. By contrast, at the far right end of the continua where the original closure durations for words with geminates are located, all the stimuli are identified almost 100% of the time as geminates.

For statistical treatment, geminate responses (i.e., the dependent variable) for each series of the stimuli across word-pairs were submitted to one-way ANOVA tests. Results show that the differences are highly significant ($p<.001$) across the board, indicating that closure duration plays a sufficient role in perceptually cueing the consonant length distinction across word-pairs tested in this experiment:

1. original /k/ ($F(11,682.7)=209.0$, $p<.001$); original /kk/ ($F(11,550.0)=206.6$, $p<.001$)
2. original /b/ ($F(11,688.0)=256.3$, $p<.001$); original /bb/ ($F(11,687.9)=177.9$, $p<.001$)

3. original /l/ (F(10,596.4)=112.1, p<.001); original /ll/ (F(10,602.5)=136.4, p<.001)

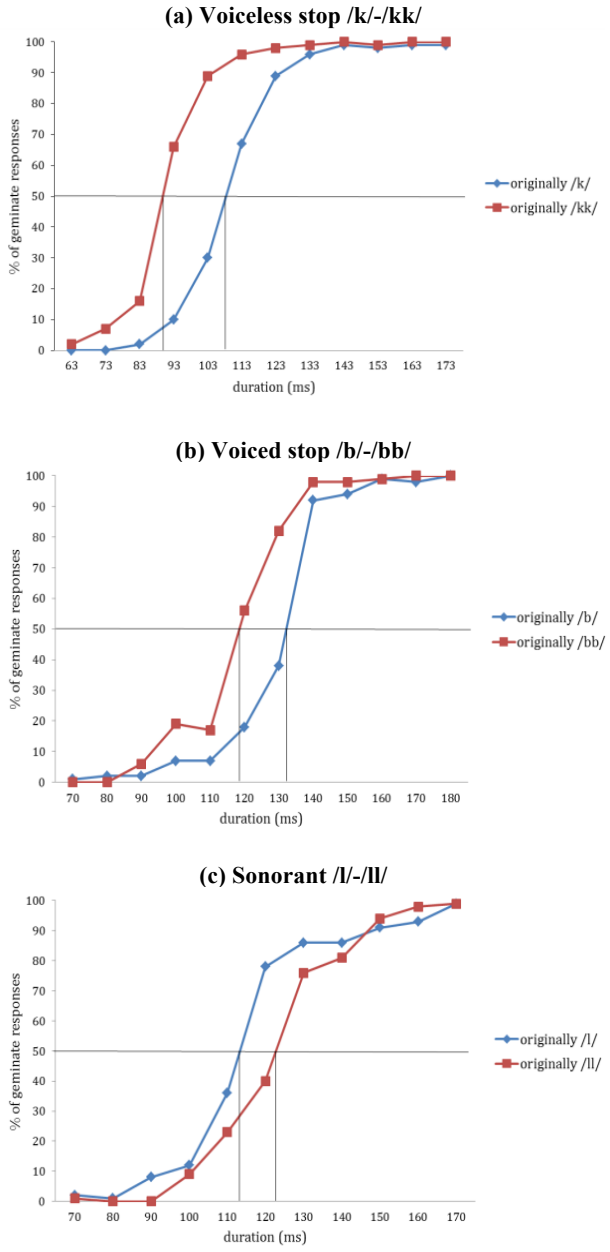

Figure 1: *Mean percentages of geminate responses to the manipulated stimuli.*

Table 2. *Number of tokens and mean percentages of geminate responses to the manipulated stimuli. Significant differences in the responses between the two types of stimuli are highlighted in grey ('***' highly significant; '**' moderately significant; '*' just significant; 'n.s.' not significant).*

(a) Voiceless stop /k/-/kk/

Duration (ms)	n	Originally /k/ (%)	n	Originally /kk/ (%)	Sig.
63	90	0	90	2	0.161 n.s.
73	90	0	90	7	0.056 n.s.
83	90	2	90	16	<0.01 **
93	90	10	90	66	<0.001 ***
103	90	30	90	89	<0.001 ***
113	90	67	90	96	<0.001 ***
123	90	89	90	98	<0.05 *
133	90	96	90	99	0.184 n.s.
143	90	99	90	100	0.326 n.s.
153	90	98	90	99	0.573 n.s.
163	90	99	90	100	0.326 n.s.
173	90	99	90	100	0.326 n.s.

(b) Voiced stop /b/-/bb/

Duration (ms)	n	Originally /b/ (%)	n	Originally /bb/ (%)	Sig.
70	90	1	90	0	0.326 n.s.
80	90	2	90	0	0.161 n.s.
90	90	2	90	6	0.184 n.s.
100	90	7	90	19	<0.05 *
110	90	7	90	17	0.083 n.s.
120	90	18	90	56	<0.001 ***
130	90	38	90	82	<0.001 ***
140	90	92	90	98	0.096 n.s.
150	90	94	90	98	0.264 n.s.
160	90	99	90	99	1.000 n.s.
170	90	98	90	100	0.161 n.s.
180	90	100	90	100	1.000 n.s.

(c) Sonorant /l/-/ll/

Duration (ms)	n	Originally /l/ (%)	n	Originally /ll/ (%)	Sig.
70	90	2	90	1	0.573 n.s.
80	90	1	90	0	0.326 n.s.
90	90	8	90	0	0.070 n.s.
100	90	12	90	9	0.448 n.s.
110	90	36	90	23	0.102 n.s.
120	90	78	90	40	<0.001 ***
130	90	86	90	76	0.054 n.s.
140	90	86	90	81	0.214 n.s.
150	90	91	90	94	0.522 n.s.
160	90	93	90	98	0.255 n.s.
170	90	99	90	99	1.000 n.s.

However, as shown in Figure 1 and Table 2, the identification curves for the manipulated stimuli with lengthened singletons and shortened geminates are not entirely identical. It can be observed across word-pairs that there are ambiguous zones between the two response curves, as indicated by two vertical lines in each figure. In the case of the voiceless stop pair /k/-/kk/ and the voiced stop pair /b/-/bb/, the perceptual crossovers at 50% from the singleton to the geminate category are earlier for the stimuli made from original geminates than for those made from original singletons: the duration differences between the two crossover points are 18 ms (for the voiceless stop pair) and 15 ms (for the voiced stop pair). Interestingly, for the sonorant pair /l/-/ll/, the crossover is unexpectedly earlier for the stimulus with the original singleton /l/ rather than for that with the original geminate /ll/: the duration difference is 10 ms between the two crossover points. Note that the ambiguous zone is largest for the voiceless stop pair (18 ms), followed by the voiced stop pair (15 ms) and then the sonorant pair (10 ms). ANOVA tests indicate that the differences between crossover points are significant across all word-pairs: /k/-/kk/ (F(1,540)=18.588, p<.001); /b/-/bb/ (F(1,540)=15.422, p<.001); /l/-/ll/ (F(1,540)=10.558, p<.01).

Additionally, geminate responses to different series of stimuli appear to diverge significantly at several duration points on the duration continua, as highlighted in grey in Table 2(a-c). For the voiceless stop pair /k/-/kk/, significant differences occur at five consecutive duration points: 83 ms (p<.01), 93 ms (p<.001), 103 ms (p<.001), 113 ms (p<.001) and 123 ms (p<.05). As for the voiced stop pair /b/-/bb/, the differences are significant at three points: 100 ms (p<.05), 120 ms (p<.001) and 130 ms (p<.001). Note the confusion at the duration point of 110 ms, as indicated by the zigzag line in Figure 1(b), in which listeners' geminate responses to the stimuli made from the original /bb/ drop slightly by 2%, although the differences are not statistically significant (p=0.083). Finally, in the case of the sonorant pair /l/-/ll/, the differences are only significant at one duration point, i.e., 120 ms (p<.001).

4. Discussion and Conclusions

In this study, we have looked at the effect of closure duration on the perception of the word-initial singleton/geminate contrast in KM. In doing so, we have examined carefully modified stimuli with incremented/decremented closure durations placed in utterance-medial contexts. The results show that closure duration is a highly significant acoustico-perceptual cue to KM word-initial geminate consonants. It appears that KM listeners respond reliably to opposite categories when the duration continuum shifts in a series of controlled steps, i.e., lengthened singletons bring about more geminate responses while shortened geminates lead to more singleton responses. The near-complete perceptual crossovers between singletons and geminates demonstrated across the manipulated stimuli made across word-pairs can be interpreted as strongly supporting the acoustic findings on closure duration for KM presented in [6], in which the durational contrast between singletons and geminates is evident in the consonant closures across all contexts in all phoneme categories in KM.

It appears that listeners are sensitive to closure duration information for all tested word-pairs so much so that, at the extreme ends of the duration continua, closure duration overrides any other possible cues, such as unmanipulated amplitude or F0, giving almost complete judgments of opposite categories, i.e., either singletons or geminates, depending on whether the manipulated stimuli are made from original singletons or geminates. This observation is comparable to Abramson's experiments with synthesized stimuli in PM [3]. More importantly, this observation is broadly in agreement with the universal claim in the literature on the robust role of closure duration in defining consonant gemination (e.g., [8]).

Our results have also shown that, despite the powerful effect of closure duration, there are displacements at the category boundary between the two series of stimuli. In KM, there is an 18-ms difference between the crossover points for the stimuli created from the original /k/ and /kk/, which is almost identical to that reported in PM [3] between the stimuli made from the original /p/ and /pp/ (the crossover-point difference is 16 ms, p<.001). Collectively, the displacements at crossover points shown in KM and PM are larger than in languages with word-medial geminates, such as in Bengali [5] in which the difference between crossover points between two series of stimuli (i.e., original singletons and original geminates) is reported to be 10 ms for /t/-/tt/. All other things being equal, this cross-linguistic difference is possibly due to the fact that consonant gemination in KM and PM only occurs in word-initial position which, in the case of voiceless stops, can be reliably accompanied by additional acoustic cues (e.g., VOT differences) even in utterance-medial position where closure duration cue is also present, as shown in [9,10,11] for KM.

Another critical point regarding the boundary displacement is that, in the case of voiceless stop stimuli, an earlier crossover in KM is exhibited in the stimuli made from the original word with a geminate (i.e., /kk/), which is again consistent with the data for PM [3]; the 50% crossover point is earlier for the stimuli synthesized from the original geminate /pp/ in PM. As for the case of the sonorant pair /l/-/ll/ in KM, the earlier crossover for the stimulus with the original singleton /l/ could be due to the presence of other acoustic correlates (e.g., increased amplitude in the following vowel) associated with its geminate counterpart, as shown in [10] for KM. All in all, the significant differences observed between the two crossover points across all word-pairs suggest that other durational or non-durational acoustic correlates (e.g., F0 differences in the following vowel) may also potentially cue the word-initial singleton/geminate contrast in KM. At this stage, it seems that KM listeners may be attuned to a range of acoustic cues associated with the word-initial consonant contrast in KM in addition to closure duration. Further perception experiments are needed, however, in order to verify this claim.

5. Acknowledgements

The authors would like to thank Universiti Utara Malaysia for its financial support. This work was supported by the university grant S/O code no. 13392.

6. References

[1] Lahiri, A., and Hankamer, J., "The timing of geminate consonants", JPhon, 16, 327-338, 1988.

[2] Muller, J.S., "The production and perception of word-initial geminates in Cypriot Greek", Proc. 15th ICPhS, 1867-1870, 2003.

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

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

[5] Hankamer, J., Lahiri, A., and Koreman, J., "Perception of consonant length: Voiceless stops in Turkish and Bengali", JPhon, 17, 283-298, 1989.

[6] Hamzah, M. H., Fletcher, J., and Hajek, J. "Closure duration as an acoustic correlate of the word-initial singleton/geminate consonant contrast in Kelantan Malay", JPhon, 58, 135–151, 2016.

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

[8] Ridouane, R., "Geminates at the junction of phonetics and phonology", in C. Fougeron, B. Kuhnert, M. D'Imperio, and N. Vallée (Eds.), Laboratory Phonology 10, 61–90, Mouton, 2010.

[9] Hamzah, M. 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.

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

[11] Hamzah, M. H., Fletcher, J., and Hajek, J., "Word-initial voiceless stop geminates in Kelantan Malay: Acoustic evidence from amplitude/F0 ratios", Proc. 18th ICPhS, 2015.