MALAYSIAN JOURNAL OF APPLIED LINGUISTICS
https://e-journal.uum.edu.my/index.php/mjoal

How to cite this article:
Hamzah, M. H. (2026). Native perception of word-initial consonant gemination in Kelantan Malay. *Malaysian Journal of Applied Linguistics, 4*, 98-114. https://doi.org/10.32890/mjoal2026.4.7

NATIVE PERCEPTION OF WORD-INITIAL CONSONANT GEMINATION IN KELANTAN MALAY

Mohd Hilmi Hamzah
Universiti Utara Malaysia, Malaysia

hilmihamzah@uum.edu.my

Received: 1/8/2026 Revised: 31/8/2026 Accepted: 27/9/2026 Published: 31/10/2026

ABSTRACT

The current study investigates the native perception of word-initial consonant gemination by Kelantan Malay speakers. Given the rare behaviour and the potential strengthening of the word-initial length contrast in utterance-initial contexts, this study aims to explore geminate perception of the word-initial consonant contrast from the perspective of native listeners. More specifically, it is of theoretical interest to test the perception of word-initial consonant gemination beginning with voiceless stops in which acoustic information is unavailable in this environment and therefore potentially perceptually indiscernible for listeners. Perception data were collected from thirty native speakers of Kelantan Malay in a forced-choice perceptual experiment involving original stimuli of isolated words in three consonant types (i.e., voiceless stops, voiced stops, and sonorants) and two vowel environments (i.e., /i/ and /a/). Results reveal that native Kelantan Malay speakers are able to identify accurately the original stimuli located in word-initial position across all consonant groups, phoneme types and vowel contexts, though geminates with voiceless stops do not fare as well as those with voiced stops and sonorants. The results provide considerable support for domain-initial strengthening that suggests the presence of primary and secondary acoustic parameters that may have important perceptual functions for the word-initial singleton/geminate consonant contrast.

Keywords: Consonant gemination, singleton, geminate, speech perception, Kelantan Malay.

INTRODUCTION

Word-initial consonant gemination is a cross-linguistically rare linguistic event as compared to consonant gemination that occurs in word-medial or word-final positions. In the acoustic phonetic literature, there is always a controversial debate whether word-initial consonant gemination can be defined by closure duration alone (i.e., primary cue) or by other secondary cues (e.g., pitch and intensity) that are usually linked to domain-initial strengthening (Cho & Keating, 2001). As closure duration is acoustically absent

in domain-initial contexts for a perceptual cue, a question is usually raised about the degree to which geminates beginning with voiceless stops are preserved or neutralised in this specific utterance environment (Abramson, 1998).

In Kelantan Malay (KM), the singleton/geminate consonant contrast occurs in word-initial position only. It is present across all consonant groups, as empirically established in Hamzah et al. (2016, 2020). It was reported that closure duration is a major acoustic correlate of word-initial consonant gemination in this language, with non-durational parameters such as root mean square (RMS) amplitude and fundamental frequency (F0) playing critical auxiliary roles in further strengthening the length contrast. The present study attempts to explore how KM native speakers perceive the rare presence of "strengthened" geminates in word-initial position. As such, the research questions of the current study are as follows:

1. What are the identification rates for the original stimuli beginning with singletons and geminates among native speakers of KM?
2. To what extent do the rates of identification differ according to consonant groups, phoneme types and vowel contexts?

LITERATURE REVIEW

It has been well established across languages (e.g., Kawahara, 2005; Tsukada et al., 2018) that a consonant contrast tends to occur in an environment that is more perceptually discernible (e.g., Pajak, 2010; Ridouane & Hallé, 2017). Word-medial consonant gemination is usually more favourable as it conforms to sonority hierarchy (i.e., the loudness of speech sounds) and therefore accommodates geminate perception. This is evident in the following minimal pair in Italian: /papa/ 'pope' vs. /pappa/ 'baby food' (Esposito & Di Benedetto, 1999). In this particular pair, it is relatively easier to distinguish the singleton /p/ and its geminate counterpart /pp/ as these target sounds are bordered by vowels.

Pertaining to word-initial consonant gemination, this linguistic phenomenon is often "avoided" and regarded as unusual (see, e.g., Frej et al., 2018; Kraehenmann, 2009). Many scholars (e.g., Ladefoged & Maddieson, 1996) claim that geminates occurring in word-initial position are phonologically and phonetically unpredictable. As mentioned earlier, word-initial consonant gemination is associated with lower perceptibility, particularly for voiceless stop geminates in which closure duration is harder to perceive. This is evident in the following minimal pair in Pattani Malay: /katoʔ/ 'to strike' vs. /kkatoʔ/ 'frog' (Abramson, 1987). In this pair, it is harder to perceive the singleton /k/ and its geminate counterpart /kk/ as these target sounds are not preceded by any vowels. It is more challenging if this pair is uttered in utterance-initial position as it is not possible to employ closure duration as a cue. However, as shown in Pattani Malay (Abramson, 1986) and Cypriot Greek (Muller, 2003), the word-initial length contrast can be perceived by native speakers including the contrast beginning with voiceless stops in utterance-initial environments, suggesting the presence of auxiliary cues that increase the perceptibility level of such a contrast.

Besides several papers associated with the present study (e.g., Hamzah et al., 2011, 2012, 2016, 2020), perception data on word-initial consonant gemination in KM is unfortunately limited (see Hamzah et al., 2016, for an illustration of waveform and spectrogram for KM word-initial consonant gemination). However, some perception data have been gathered in Pattani Malay (Abramson, 1986, 1987, 1998, 1999, 2003). Note that Pattani Malay is comparable to KM, though the former differs from the latter in terms of language contact (see, e.g., Sisamouth & Lah, 2015).

It was reported in Pattani Malay (Abramson, 1986) that original stimuli containing word-initial consonant gemination, including voiceless stops, are identified successfully by its native speakers (i.e., 80% of the time or better). A similar observation was reported by Muller (2003) for Cypriot Greek in which native speakers in this language can perceive word-initial singletons and geminates at high identification rates (i.e., 97% of the time), including those beginning with voiceless stops. Besides, Tsukada and Hajek (2023) have also demonstrated that Japanese native speakers can successfully discriminate the Japanese singleton/geminate contrast at high identification rates (i.e., 93% of the time).

Drawing on these past findings on geminate perception, it appears that word-initial consonant gemination is perceptible among native speakers, despite the lack of closure duration information as a cue in utterance-initial position for voiceless stops. In this context, there is a possibility that word-initial consonant gemination can be cued by a group of additional acoustic parameters alongside closure duration that assist native speakers in identifying the word-initial length contrast. The acoustic findings across languages (e.g., Abramson, 2003; Al-Tamimi & Khattab, 2018; Dmitrieva, 2018; Kraehenmann, 2001; Lahiri & Hankamer, 1988; Mitterer, 2018) suggest that there are indeed durational (e.g., closure duration) and non-durational cues (e.g., vowel onset amplitude) that are potentially strengthening the native perception of word-initial consonant gemination.

The present study is theoretically driven by the arguments related to domain-initial strengthening, as proposed by Cho and Keating (2001). Domain-initial strengthening refers to potential enhancement in domain-initial environments to assist speech perception (Yoo, 2021). When geminates start in utterance-initial position with voiceless stops, domain-initial strengthening is more evident. In this environment, the length contrast is claimed to be preserved and strengthened via a bundle of acoustic cues potentially associated with fortition (Cho & Keating, 2001). Cho (2005) described this process as "a prominence-enhancing strategy" (see also Gordon, 2009; Payne, 2005). That is, a target sound can be strengthened via additional acoustic attributes, such as higher intensity, which can make the sound more perceptible. Hence, it appears critical to examine how KM native speakers perceive consonant gemination in word-initial position that is also potentially linked to domain-initial strengthening (Cho & Keating, 2001). In domain-initial environments, the issue is more intriguing theoretically for the hypothetical contrast enhancement of voiceless stop geminates.

Drawing from Abramson's (1986) perceptual results in Pattani Malay, it is hypothesised that the identification rates for word-initial consonant gemination in KM are also high. However, these rates are expected to vary according to consonant groups, phoneme types and vowel heights. Motivated by Abramson's (1986) instrumental work, the present study aims to investigate the KM singleton/geminate contrast by examining its perceptual values among native speakers. The findings will offer an empirical foundation for the status of native perception of word-initial consonant gemination that is understudied in the phonetic literature.

METHODOLOGY

Stimuli

Following previous experimental works on geminate perception, an experiment was designed to examine a perceptual behaviour among native speakers in KM. The original stimulus materials consisted of thirty-eight natural tokens representing nineteen minimal pairs in KM (see Table 1). All tokens were produced in isolation.

Table 1

Stimuli Beginning with Singletons and Geminates (Bolded and Underlined)

Consonant group	Phoneme	Singleton		Geminate	
		Word	Gloss	Word	Gloss
Voiceless stops	/p/	/**p**itu/	door	/**pp**itu/	at the door
		/**p**aɡi/	morning	/**pp**aɡi/	early morning
	/t/	/**t**ido/	sleep	/**tt**ido/	sleep by chance
		/**t**anɔh/	land	/**tt**anɔh/	outside
	/k/	/**k**iɣi/	left	/**kk**iɣi/	to the left
		/**k**abo/	blurry	/**kk**abo/	beetle
Voiced stops	/b/	/**b**ini/	wife	/**bb**ini/	married
		/**b**aʧɔ/	read	/**bb**aʧɔ/	is reading
	/d/	/**d**ike/	song	/**dd**ike/	sing a song
		/**d**apo/	kitchen	/**dd**apo/	at the kitchen
	/g/	/**g**igi/	teeth	/**gg**igi/	on the teeth
		/**g**aʤi/	salary	/**gg**aʤi/	sawing tool
Sonorants	/m/	/**m**isa/	moustache	/**mm**isa/	moustached
		/**m**aɣi/	come	/**mm**aɣi/	cupboard
	/n/	/**n**ikɔh/	marriage	/**nn**ikɔh/	married
		/**n**anɔh/	pus	/**nn**anɔh/	getting pus
	/l/	/**l**idɔh/	tongue	/**ll**idɔh/	on the tongue
		/**l**apu/	lights	/**ll**apu/	on the lights
	/ŋ/	/**ŋ**aŋɔ/	open the mouth	/**ŋŋ**aŋɔ/	agape

The stimuli were all disyllabic in structure: C(C)VCV or C(C)VCVC. Twenty phonemes were selected and they were categorised according to consonant groups: voiceless stops (/p-pp/, /t-tt/, /k-kk/); voiced stops (/b-bb/, /d-dd/, /g-gg/); and sonorants (/m-mm/, /n-nn/, /ŋ-ŋŋ/, /l-ll/). Each target phoneme was followed by two vowels: the high front vowel /i/ and the low central vowel /a/, except /ŋ-ŋŋ/ (low back vowel /a/ only). All stimuli were frequently-used words and familiar to the listeners. The materials were previously recorded in a production experiment involving a male native speaker of KM (see Hamzah et al., 2016, 2020 for more information on production experiments).

Participants

Thirty native speakers of KM were recruited in a perception experiment (15 males, 15 females, mean age: 21.2), as seen in Table 2. They were born and grew up in Kelantan, Malaysia. During the experiment, the participants had no symptoms of hearing disorder.

Table 2

Demographic Details of Participants

MALES			FEMALES		
Listener	Age	District	Listener	Age	District
ML1	22	Bachok	FL1	20	Kota Bharu
ML2	21	Kuala Krai	FL2	21	Pasir Mas
ML3	21	Kuala Krai	FL3	23	Gua Musang
ML4	23	Jeli	FL4	21	Kuala Krai
ML5	21	Pasir Mas	FL5	21	Tumpat
ML6	22	Tanah Merah	FL6	21	Pasir Puteh
ML7	21	Jeli	FL7	21	Tumpat
ML8	23	Kota Bharu	FL8	25	Machang
ML9	20	Jeli	FL9	22	Bachok
ML10	21	Machang	FL10	21	Pasir Mas
ML11	21	Tanah Merah	FL11	21	Tanah Merah
ML12	20	Gua Musang	FL12	21	Kota Bharu
ML13	20	Tumpat	FL13	20	Pasir Puteh
ML14	22	Pasir Puteh	FL14	20	Kota Bharu
ML15	21	Machang	FL15	20	Gua Musang

Data Collection

The perception data were collected in an individual session for each participant in a public university in Kelantan, Malaysia. The participants listened to the stimuli with a stereo headphone via a notebook by running a Multiple Forced Choice listening experiment in Praat (version 6.3.17). Prior to data collection, the participants were trained with a few stimuli to familiarise them with the procedures. They were required to listen carefully and identify the written word from the two boxes that represented the singleton/geminate consonant contrast. KM has no orthographic system, so Standard Malay was used in the stimuli presentation. The instruction was also delivered in Standard Malay.

The stimuli were presented in random order, with each stimulus being presented three times, yielding 114 trials for each participant. There was a pause of 1.5 seconds between the stimuli. The participants could replay the stimuli, but they could only do so once for each stimulus. They were also given an option to rest for a short period. Each participant completed the task within approximately forty minutes. They received monetary compensation at the end of the experiment.

Data Analysis

The perception data were organised in Excel (version 14.7.7). The identification rates were represented through the number of correct responses for stimuli (either singletons or geminates). Since the number of stimuli presentation was three times, the scores for each stimulus were allocated between 0 to 3. The percentages of correct responses and their corresponding bar graphs were obtained from these scores. A mixed-effects model (Baayen et al., 2008) was then employed to analyse the identification rates statistically. This analysis was conducted via SPSS version 28.0.0. The fixed factors in the model were as follows:

1. Length (singleton and geminate)
2. Consonant Group (voiceless stop, voiced stop and sonorant)
3. Phoneme (/p-pp/, /t-tt/, /k-kk/, /b-bb/, /d-dd/, /ɡ-ɡɡ/, /m-mm/, /n-nn/, /ŋ-ŋŋ/, /l-ll/)
4. Vowel Height (/i/ and /a/)

The source of significant results (at least $p<.05$) was identified through Tukey's multiple comparisons of means. The listener was regarded as a random variable.

RESULTS

Overall Results

Figure 1 illustrates the overall results of identification rates (in percentages) for the KM singleton/geminate contrast, while Table 3 provides the measurements in detail. The percentages were calculated by averaging the correct responses from thirty participants. The data are pooled together across consonant groups.

Figure 1

Overall Results of Identification Rates (%) for Singleton and Geminate

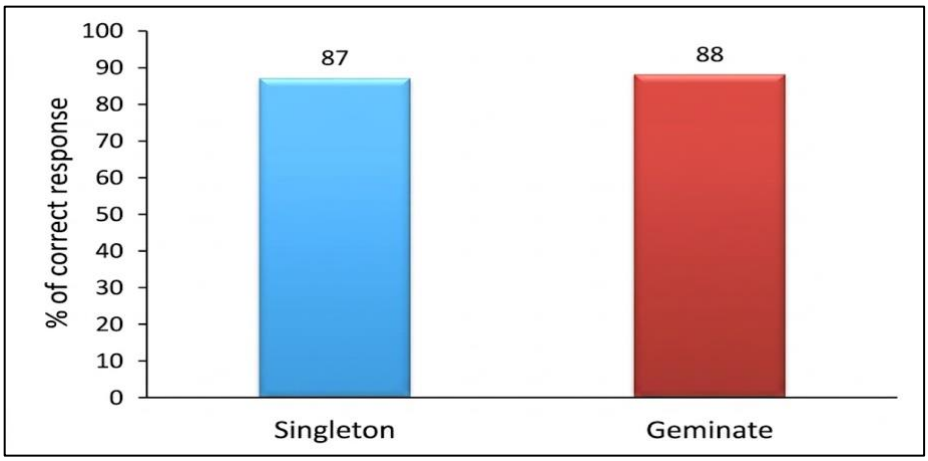


Table 3

Detailed Measurements of Identification Rates (%) for Singleton and Geminate ('n.s.' not significant)

n	Singleton (%)	*n*	Geminate (%)	Sig.
1710	87	1710	88	0.326 n.s.

As shown in Figure 1, the original stimuli were identified correctly for both singletons and geminates (87% and 88% of the time, respectively). The differences between two identification rates are not statistically significant ($p=0.326$), which implies a similar pattern of stimuli identification among the participants. A detailed scrutiny into the identification rates of each participant reveals that all of them identify the stimuli correctly and consistently (detailed results according to each participant can be found in Appendix 1). The identification rates were further analysed statistically via a mixed-effects model in order to examine the main effects of Consonant Group, Phoneme, Vowel Height, and how they interact with Length (singleton and geminate). Table 4 presents the statistical results.

Table 4

*Statistical Results from the Mixed-Effects Model ('***' highly significant; '*' just significant; 'n.s.' not significant)*

Random effects				
	Variance		Std. Deviation	
Residual	0.0018		0.0422	
Listener	0.0004		0.0213	
Fixed effects				
	Estimate	Std. Error	*z*-value	Pr(>|*z*|)
(Intercept)	0.035	0.005	83.11	<0.001 ***
Consonant Group	0.033	0.003	14.758	<0.001 ***
Phoneme	0.058	0.002	11.610	<0.001 ***
Vowel Height	-0.042	0.002	-8.32	<0.05 *
Length x Consonant Group	0.036	0.004	-18.339	0.645 n.s.
Length x Phoneme	0.045	0.005	-7.189	<0.05 *
Length x Vowel Height	0.005	0.001	-5.227	<0.05 *

As shown in Table 4, significant main effects are exhibited across the factors: Consonant Group ($F(2,554)=71.012$, $p<.001$); Phoneme ($F(8,540)=21.610$, $p<.001$); and Vowel Height ($F(1,1078)=15.145$, $p<.05$). Length interacts at just-significant levels with Phoneme ($F(8,540)=10.124$, $p<.05$) and Vowel Height ($F(1,1078)=11.225$, $p<.05$), suggesting that the identification rates for original stimuli beginning with singletons and geminates are somewhat affected by vowel context and phoneme type. Consonant Group, however, has no interaction with Length in spite of showing a significant main effect. This suggests that singletons and geminates are perceived in a similar way regardless of consonant types. The following sub-section describes this particular result in greater detail.

Effect of Consonant Group

Figures 2 presents geminate perception results according to consonant groups, while Table 5 shows the detailed measurements. It can be observed that the identification rates for voiceless stops were poorest across two length groups, i.e., singletons (75% of the time) and geminates (76% of the time), which explains the significant overall effect of Consonant Group ($p<.001$), as shown in the previous sub-section. These poor rates for voiceless stops are unsurprising given the lack of acoustic information for closure duration. The rates of identification, nevertheless, exceed chance level (50%). For voiced stops and sonorants, the identification rates reach 90% or better.

Figure 2

Identification Rates (%) for Singleton and Geminate According to Consonant Group

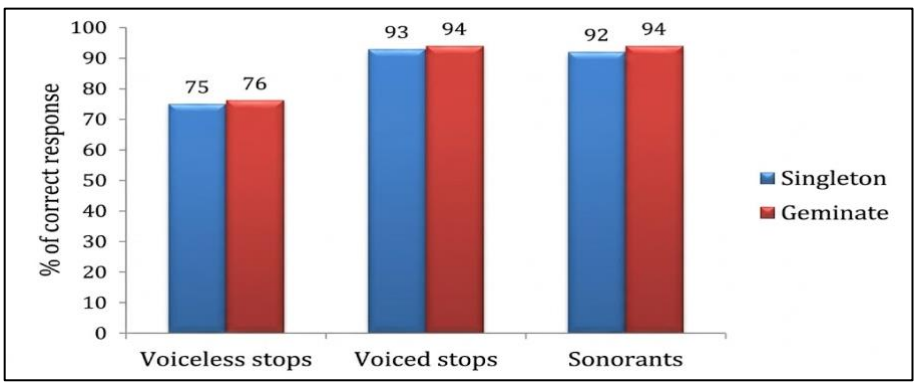


Table 5

Detailed Measurements of Identification Rates (%) for Singleton and Geminate According to Consonant Group ('n.s.' not significant)

Consonant Group	*n*	Singleton (%)	*n*	Geminate (%)	Sig.
Voiceless stops	540	75	540	76	0.846 n.s.
Voiced stops	540	93	540	94	0.385 n.s.
Sonorants	630	92	630	94	0.261 n.s.

A further check into Tukey's multiple comparisons of means reveals that voiceless stops stand out and are statistically different (p<.001) from other groups of consonants. Voiced stops and sonorants behave in a similar fashion and are not different statistically (p=0.930). It should be noted that both singletons and geminates are perceived correctly within each consonant group. Statistically speaking, the identification differences between two length groups are not significant across all consonant groups, as shown in Table 5 for voiceless stops (p=0.846), voiced stops (p=0.385), and sonorants (p=0.261). These results explain why Length does not interact with Consonant Group, as mentioned earlier.

Length, nonetheless, interacts significantly with both Phoneme (p<.05) and Vowel Height (p<.05), as observed in the results of the mixed-effects model earlier. On further scrutiny of detailed results for each phoneme and each vowel environment, it was discovered that, out of ten phoneme pairs (see stimuli in Table 1), four of them show unique patterns, which are further discussed in the following sub-section.

Effects of Phoneme and Vowel Height

The four selected phoneme pairs that will be highlighted in this section are /p-pp/, /t-tt/, /k-kk/, and /b-bb/, as shown in Table 6 below. Note that the first three pairs are voiceless stop pairs, while the final one is the voiced stop pair. The percentages shown in Table 6 were obtained by averaging the correct responses for all participants according to phonemes and vowel height (number of tokens each=90). The highlighted information (in grey) refers to percentages that are significantly higher than their respective contrasting phoneme. Appendix 2 and 3 provide the results in greater detail for each phoneme type and vowel height, respectively.

Table 6

Identification Rates (%) for Singleton and Geminate According to Phoneme and Vowel Height

Vowel height	/p-pp/		/t-tt/		/k-kk/		/b-bb/	
/i/	/p/ (97%)	/pp/ (56%)	/t/ (98%)	/tt/ (40%)	/k/ (58%)	/kk/ (83%)	/b/ (97%)	/bb/ (93%)
/a/	/p/ (80%)	/pp/ (88%)	/t/ (41%)	/tt/ (92%)	/k/ (76%)	/kk/ (96%)	/b/ (78%)	/bb/ (97%)

For the first phoneme pair /p/ versus /pp/, a significant identification difference is exhibited in the /i/ environment, in which the stimulus with /p/, i.e., /pitu/ 'door', receives a significantly higher identification rate (97%) as compared to its geminate counterpart /pp/, i.e., /ppitu/ 'at the door' which is just above chance level (56%). In the /a/ environment, the difference is not statistically significant for /p/ versus /pp/ (p=0.335).

For the second phoneme pair /t/ versus /tt/, the identification differences between two length categories are significant in both vowel environments. In the /i/ context, a significantly higher identification rate is found for the stimulus with /t/ (98%), i.e., /tido/ 'sleep', rather than that with /tt/ (40%), i.e., /ttido/ 'sleep by chance'. In the /a/ context, a significantly higher identification rate is found for the stimulus with /tt/ (92%), i.e., /ttanɔh/ 'outside', rather than that with /t/ (41%), i.e., /tanɔh/ 'land'.

For the third phoneme pair /k/ versus /kk/, the identification differences are present in both vowel environments in which the stimuli beginning with /kk/ are recognised with higher rates across the vowel contexts (/i/=83%, i.e., /kkiɣi/ 'to the left'; /a/=96%, i.e., /kkabo/ 'beetle'). Note the identification rate for /kiɣi/ 'left' (58%) that barely surpasses chance level.

For the final phoneme pair /b/ versus /bb/, a significant identification difference is observed in the /a/ environment, in which the stimulus beginning with /bb/, i.e., /bbaʧɔ/ 'is reading', receives a higher identification rate (97%) as compared to its singleton counterpart /b/ (78%), i.e., /baʧɔ/ 'read'. The identification difference is not statistically significant for /b/ versus /bb/ in the /i/ environment (p=0.326).

In sum, it can be observed that voiceless stop phoneme pairs are all affected by vowel height. These results confirm the speculation about the poor level of perceptibility for voiceless stops in word-initial position, given the lack of primary acoustic cue (closure duration), as highlighted earlier. The native perception for /t/ versus /tt/ needs further elaborations. The rates of identification for /ttido/ 'sleep by chance' (40%) and /tanɔh/ 'land' (41%) do not reach chance level. On further scrutiny via Praat, some noteworthy patterns were discovered.

A cross-examination between the stimulus /ttido/ 'sleep by chance' and its singleton counterpart /tido/ 'sleep' reveals that both stimuli are acoustically comparable, which perhaps leads to a perceptual crossover from /ttido/ to /tido/. The acoustic details are as follows:

1. Burst amplitude: /ttido/ (55 dB) versus /tido/ (54 dB)
2. Vowel onset amplitude: /ttido/ (71 dB) versus /tido/ (70 dB)
3. Vowel onset F0: /ttido/ (118 Hz) versus /tido/ (114 Hz)

Similarly, a further acoustic analysis between the stimulus /tanɔh/ 'land' and its geminate counterpart /ttanɔh/ 'outside' indicates that both stimuli liken each other with respect to some acoustic features (see below), which potentially switches native speakers' perception from /tanɔh/ to /ttanɔh/. The similarities between two stimuli can be observed in the following acoustic details:

1. Burst amplitude: /tanɔh/ (54 dB) versus /ttanɔh/ (55 dB)
2. Vowel onset amplitude: /tanɔh/ (69 dB) versus /ttanɔh/ (70 dB)
3. Vowel onset F0: /tanɔh/ (137 Hz) versus /ttanɔh/ (139 Hz)

Note that the above-mentioned acoustic details have been described in Hamzah et al. (2020) as significant secondary acoustic parameters for KM word-initial consonant gemination. That is, besides longer closure duration, word-initial consonant gemination in KM can be characterised by greater amplitude (stop burst and at vowel onset) and higher F0 (at vowel onset). To confirm the perceptual roles of these acoustic characteristics, more experimental data are needed.

With the exception of two phonemes (i.e., /ttido/ and /tanɔh/) discussed above, the rest of the phonemes in the voiceless stop group are identified accurately and surpass chance level (50%). That is, native speakers are able to recognise these phonemes in spite of the unavailability of a major perceptual attribute (i.e., closure duration). The other phonemes in other consonant categories (i.e., sonorants and voiced stops) are well perceived across the board (87% of the time or better; see Appendix 2).

DISCUSSION

The KM native speakers' rates of identification for the original stimuli were the subject of the current study's first research question. Based on the previous findings in Pattani Malay (Abramson, 1986), it was hypothesised that KM native speakers would accurately recognise original stimuli with high rates of identification. The clearest finding reported in the result section confirms this hypothesis: word-initial singletons and geminates are identified extremely well by KM native speakers across the board (around 90% of the time). The high rates of identification for the utterance-initial data are comparable to those observed in Pattani Malay (Abramson, 1986) in which word-initial consonant gemination is also accurately identified (virtually 100% of the time). It seems that native listeners of KM are accustomed to consonant gemination in word-initial position in spite of the previous arguments (e.g., Ladefoged & Maddieson, 1996; Pajak, 2010; Ridouane & Hallé, 2017) that highlight the poor perceptibility level of this particular length contrast.

With regard to the second research question (effects of consonant groups, phoneme type and vowel height), the results reveal that stimuli beginning with voiced stops and sonorants are correctly identified in KM (i.e., above 90%), which is probably due to longer voicing (for utterance-initial voiced stops) and clear sonority (for utterance-initial sonorants). These acoustic features are present throughout the initial consonant occlusions and therefore enable listeners to perceive the length contrast, as revealed in the acoustic findings for KM (Hamzah et al., 2016) as well as in Abramson's (1987) description for Pattani Malay (see also, e.g., Dmitrieva, 2018). Pertaining to stimuli with voiceless stops, native speakers of these related Malay varieties are able to correctly identify the stimuli despite the unavailability of closure duration in this environment. Although the identification rate for voiceless stop geminates in KM (76%) is lower than that in Pattani Malay (82%), the rate exceeds chance level at 50%.

The high identification rates found in KM also accord well with those reported in Cypriot Greek (Muller, 2003); native speakers of Cypriot Greek can also perceive successfully the utterance-initial voiceless stop geminates in this language (see also Tsukada & Hajek, 2023). Based on these findings, it seems that native listeners are sensitive to secondary acoustic parameters that likely strengthen and enhance the perception of voiceless stop geminates that occur in utterance-initial position, as described acoustically in Hamzah et al. (2020) for KM (see also Al-Tamimi & Khattab, 2018; Lahiri & Hankamer, 1988; Mitterer, 2018).

On phoneme/vowel effects on the identification rates for voiceless stops, especially the contradictory behaviour for stimuli beginning with /t/ and /tt/ in distinct vowel contexts (/i/ and /a/), this finding is unexpected. Based on the acoustic data of vowel duration related to KM word-initial consonant gemination (Hamzah et al., 2012), it was reported that the duration of vowels can be affected significantly by the preceding target consonant (i.e., either singleton or geminate). However, the effect is only observed when the length contrast occurs in utterance-medial position and for the vowel /a/ only. Besides, the duration of VOT (Voice Onset Time) is also a possible parameter related to vowel effects in the current study. As shown in Hamzah et al. (2011) for KM, VOT is longer before /i/ rather than before /a/. In the case of /t/ versus /tt/, longer VOT duration is associated with the singleton /t/, in which the effect is particularly greater in the utterance-initial context.

Given these findings, it can be speculated that KM listeners are probably affected by vowel context and VOT differences. In other words, the stimuli beginning with /t/ or /tt/ and followed by /i/ are probably perceived as singletons since VOT duration is longer for singletons in this context, while those followed by /a/ are mostly perceived as geminates since VOT duration is shorter for geminates in this environment. Visual inspection on the original stimuli used in the current study (i.e., /tido/ vs. /ttido/; /tanɔh/ vs. /ttanɔh/) shows that, in the /i/ context, the VOT duration is longer (/t/=27 ms; /tt/=20 ms), while in the /a/ context,

the VOT duration is shorter (/t/=16 ms; /tt/=10 ms). However, this speculation requires further perceptual tests on VOT duration and vowel height.

The inherent values of relative vowel height may also explain the effect of vowel height on the findings of the current study. It has been established that the low vowel /a/ and the high vowel /i/ differ in terms of duration and intensity, in which /a/ is longer and more intense than /i/ (see, e.g., Lehiste, 1970). It can be speculated that these inherent values may possibly influence how native speakers perceive the stimuli with /t/ and /tt/ in KM. Future experiments should examine vowel height in order to examine the degree to which it affects native listeners' perception of word-initial consonant gemination.

It is worth noting that the successful geminate perception for stimuli with voiced stops and sonorants in KM may counter the arguments made on the interaction between geminate perception and the hierarchy of sonority (e.g., Kawahara, 2005; Tsukada et al., 2018). It was argued that non-intervocalic geminates, including those beginning with voiced stops and sonorants, are relatively more challenging and difficult to perceive, which contradicts the data observed in KM in the current study.

With regard to domain-initial strengthening (Cho & Keating, 2001), it appears at this stage that the strengthening or enhancement of additional cues for utterance-initial voiceless stops can perhaps be "extended" to voiced stops and sonorants too. This speculation is linked to the inherent acoustic behaviour of word-initial consonant gemination, such as in KM, which is often disfavoured and more marked perceptually (e.g., Pajak, 2010). To this end, it can be further speculated that the KM singleton/geminate consonant contrast is experiencing a process of fortition, i.e., word-initial consonant gemination in KM may be characterised by not only durational changes in consonant closures but also "tenseness" through a bundle of acoustic cues like RMS amplitude, vowel onset F0, VOT, syllabic differences, among others. Future perception studies should take this "strengthening" factor into consideration when describing the perception of consonant gemination in word-initial position.

CONCLUSION

The perception of word-initial consonant gemination by native speakers of KM has been investigated in this study. The results show high rates of identification for the original stimuli beginning with singletons and geminates in word-initial position and, by extension, utterance-initial position, across all consonant groups including voiceless stops. These findings are in tandem with Pattani Malay (Abramson, 1986, 1987, 1998, 1999, 2003) and also with other studies related to native versus non-native perception of a phonological contrast (e.g., Kimiko & Hajek, 2023). More critically, the results have shed some light on the perceptibility of word-initial consonant gemination that is often "avoided" (e.g., Frej et al., 2018; Kraehenmann, 2009). There are of course some hints of domain-initial strengthening that may serve as "a prominence-enhancing strategy" in helping native speakers perceive the word-initial length contrast (e.g., Gordon, 2009; Payne, 2005).

Future research should address several important questions that the current study has raised. It appears warranted to examine non-native perception of the KM word-initial consonant contrast by speakers of other languages who are familiar or unfamiliar with word-initial consonant gemination. This cross-language speech perception will inform the extent to which acoustic cues associated with the word-initial length contrast are processed by speakers with different language backgrounds. It is clear also that there is a need for further perception studies involving manipulated stimuli in order to explore the potential perceptual roles of non-durational acoustic results reported for KM (e.g., Hamzah et al., 2020). Another suggestion for possible future research is to investigate the role of tempo.

According to the geminate perception results presented in this paper, the word-initial singleton/geminate consonant contrast in KM is characterised by a variety of additional acoustic cues in addition to closure duration, which are crucial to sustaining and enhancing such a contrast in word-initial position (see also Kraehenmann, 2001). As such, this study has affirmed the native speakers' ability to successfully perceive the much-debated word-initial consonant gemination that may be accompanied by other acoustic cues besides closure duration, especially for utterance-initial voiceless stops. More importantly, this study has made some contributions in the description of consonant length contrasts that take place in a word position that is considered rare and unusual.

ACKNOWLEDGMENT

This research was supported by Universiti Utara Malaysia (UUM) through University Grant (SO Code: 21899). The guidelines or codes of relevant ethics employing human subjects follow the Declaration of Helsinki whereby participants were informed of the nature of the research using a research protocol.

REFERENCES

Abramson, A. S. (1986). The perception of word-initial consonant length: Pattani Malay. *Journal of the International Phonetic Association, 16*, 8–16. https://doi.org/10.1017/S0025100300003054

Abramson, A. S. (1987). Word-initial consonant length in Pattani Malay. In T. Gamkrelidze (Ed.), *Proceedings of the 11th International Congress of Phonetic Sciences* (pp. 68–70). Academy of Sciences of the Estonian S.S.R.

Abramson, A. S. (1998). The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length. In U. Warotamasikkhadit & T. Panakul (Eds.), *Papers from the 4th Annual Meeting of the Southeast Asian Linguistics Society 1994* (pp. 1–20). Monograph Series Press. http://sealang.net/sala/archives/pdf4/abramson1998complex.pdf

Abramson, A. S. (1999). Fundamental frequency as a cue to word-initial consonant length: Pattani Malay. In J. J. Ohala (Ed.), *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 591–594). University of California Press.

Abramson, A. S. (2003). Acoustic cues to word-initial stop length in Pattani Malay. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (pp. 387–390). Universitat Autònoma de Barcelona.

Al-Tamimi, J., & Khattab, G. (2018). Acoustic correlates of the voicing contrast in Lebanese Arabic singleton and geminate stops. *Journal of Phonetics*, *71*, 306–325. https://doi.org/10.1016/j.wocn.2018.09.010

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*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Cho, T., & Keating, P. (2001). Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, *29*, 155–190. https://doi.org/10.1006/jpho.2001.0131

Cho, T. (2005). Prosodic strengthening and featural enhancement: Evidence from acoustic and articulatory realizations of /a,i/ in English. *Journal of the Acoustical Society of America*, *117*(6), 3867–3878. https://doi.org/10.1121/1.1861893

Dmitrieva, O. (2018). The role of perception in the typology of geminate consonants: Effects of manner of articulation, segmental environment, position, and stress. *Language and Speech*, *61*(1), 43–70. https://doi.org/10.1177/0023830917696113

Esposito, A., & Di Benedetto, M. G. (1999). Acoustical and perceptual study of gemination in Italian stops. *Journal of the Acoustical Society of America, 106*(4), 2051–2062. https://doi.org/10.1121/1.428056

Frej, M. Y., Caringnan, C., Proctor, M. I., & Best, C. T. (2018). Acoustic and ultrasound investigation of word-initial gemination in Moroccan Arabic. *The Journal of the Acoustical Society of America*, *144*(3), 1871–1871. https://doi.org/10.1121/1.5068233

Gordon, M. (2009). Prominence and gemination in Ingrian. *Linguistica Uralica, 45*(2), 81–100. https://doi.org/10.3176/lu.2009.2.01

Hamzah, M. H., Fletcher, J., & Hajek, J. (2011). Durational correlates of word-initial voiceless geminate stops: The case of Kelantan Malay. In W.-S. Lee, & E. Zee (Eds.), *Proceedings of the 17th International Congress of Phonetic Sciences* (pp. 815–818). City University of Hong Kong.

Hamzah, M. H., Hajek, J., & Fletcher, J. (2012). A taste of prosody: Possible effects of the word-initial singleton-geminate contrast on post-consonantal vowel duration in Kelantan Malay. In M. Qiuwu, H. Ding & D. Hirst (Eds.), *Proceedings of the 6th International Conference on Speech Prosody* (pp. 490–493). Tongji University Press.

Hamzah, M. H., Fletcher, J., & Hajek, J. (2016). Closure duration as an acoustic correlate of the word-initial singleton/geminate consonant contrast in Kelantan Malay. *Journal of Phonetics*, *58*, 135–151. https://doi.org/10.1016/j.wocn.2016.08.002

Hamzah, M. H., Fletcher, J., & Hajek, J. (2020). Non-durational acoustic correlates of word-initial consonant gemination in Kelantan Malay: The potential roles of amplitude and f0. *Journal of the International Phonetic Association*, *50*(1), 23–60. https://doi.org/10.1017/S0025100318000142

Kawahara, S. (2005). Voicing and geminacy in Japanese: An acoustic and perceptual study. In K. Flack & S. Kawahara (Eds.), *University of Massachusetts Occasional Papers in Linguistics 31: Papers in Experimental Phonetics and Phonology* (pp. 87–120). Graduate Linguistic Student Association Publications. http://user.keio.ac.jp/~kawahara/pdf/kawahara_UMOP31.pdf

Kraehenmann, A. (2001). Swiss German stops: Geminates all over the word. *Phonology*, *18*(1), 109–145. https://doi.org/10.1017/S0952675701004031

Kraehenmann, A. (2009). *The perception of the word-initial quantity contrast in voiceless Swiss German stops* [Paper presentation]. The 17th Manchester Phonology Meeting, University of Manchester, Manchester. http://www.lel.ed.ac.uk/mfm/17mfm-prog.pdf

Ladefoged, P., & Maddieson, I. (1996). *Sounds of the world's languages*. Blackwell Publishing.

Lahiri, A., & Hankamer, J. (1988). The timing of geminate consonants. *Journal of Phonetics, 16*, 327–338. https://doi.org/10.1016/S0095-4470(19)30506-6

Lehiste, I. (1970). *Suprasegmentals*. Massachusetts Institute of Technology Press.

Mitterer, H. (2018). Not all geminates are created equal: Evidence from Maltese glottal consonants. *Journal of Phonetics*, *66*, 28–44. https://doi.org/10.1016/j.wocn.2017.09.003

Muller, J. S. (2003). The production and perception of word-initial geminates in Cypriot Greek. In M. J. Solé, D. Recasens & J. Romero (Eds.), *Proceedings of the 15th International Congress of the Phonetic Sciences* (pp. 1867–1870). Universitat Autònoma de Barcelona.

Pajak, B. (2010). Contextual constraints on geminates: The case of Polish. In N. Baier & E. Donnelly (Eds.), *Proceedings of the 35th Annual Meeting of the Berkeley Linguistics Society (BLS)* (pp. 269–280). University of California, Berkeley. https://journals.linguisticsociety.org/proceedings/index.php/BLS/ article/download/3617/3322

Payne, E. (2005). Phonetic variation in Italian consonant gemination. *Journal of International Phonetic Association, 35*(2), 153–181. https://doi.org/10.1017/S0025100305002240

Ridouane, R., & Hallé, P. A. (2017). Word-initial geminates: From production to perception. In H. Kubozono (Ed.), *The phonetics and phonology of geminate consonants* (pp. 34–65). Oxford University Press. https://doi.org/10.1093/oso/9780198754930.003.0004

Sisamouth, W. K., & Lah, S. C. (2015). Attitudes towards Thai, Patani Malay, and English of Thai undergraduates: A case study at Prince of Songkla University Pattani Campus, Thailand. *Procedia - Social and Behavioral Sciences*, *208*, 240–252. https://doi.org/10.1016/j.sbspro.2015.11.200

Tsukada, K., Idemaru, K., & Hajek, J. (2018). The effects of foreign language learning on the perception of Japanese consonant length contrasts. In J. Epps, J. Wolfe, J. Smith & C. Jones (Eds.), *Proceedings of the 17th Australasian International Conference on Speech Science and Technology* (pp. 37–40). ASSTA. https://assta.org/proceedings/sst/SST2018/SST_2018_Proceedings_Rev_A_IDX.pdf

Tsukada, K., & Hajek, J. (2023). Cross-language perception of Japanese consonant length by speakers from Italian- and Mandarin-speaking backgrounds. *Second Language Research*, *39*(3), 925–938. https://doi.org/10.1177/02676583221108269

Yoo, K. (2021). *The production and perception of domain-initial strengthening in Seoul, Busan, and Ulsan Korean* [Unpublished doctoral dissertation]. University of Cambridge. https://doi.org/10.17863/CAM.66434

APPENDIX 1

Listener-Specific Results

MALES			FEMALES		
Listener	***n***	**Identification rates (%) averaged across singletons and geminates**	**Listener**	***n***	**Identification rates (%) averaged across singletons and geminates**
ML1	3420	92	FL1	3420	83
ML2	3420	85	FL2	3420	91
ML3	3420	93	FL3	3420	90
ML4	3420	80	FL4	3420	94
ML5	3420	86	FL5	3420	82
ML6	3420	92	FL6	3420	75
ML7	3420	89	FL7	3420	89
ML8	3420	81	FL8	3420	94
ML9	3420	85	FL9	3420	80
ML11	3420	82	FL11	3420	73
ML11	3420	94	FL11	3420	92
ML12	3420	90	FL12	3420	87
ML13	3420	89	FL13	3420	90
ML14	3420	95	FL14	3420	91
ML15	3420	93	FL15	3420	90

APPENDIX 2

Phoneme-Specific Results

Phoneme	*n*	Singleton (%)	*n*	Geminate (%)	Sig.
/p/	180	88	180	72	<0.05 *
/t/	180	69	180	66	<0.05 *
/k/	180	67	180	89	<0.01 **
/b/	180	87	180	95	<0.05 *
/d/	180	94	180	91	0.312 n.s.
/ɡ/	180	97	180	96	0.536 n.s.
/m/	180	95	180	94	0.813 n.s.
/n/	180	87	180	93	0.117 n.s.
/l/	180	96	180	98	0.293 n.s.
/ŋ/	90	90	90	90	1.000 n.s.

APPENDIX 3

Vowel-Specific Results

Vowel/phoneme	*n*	Singleton (%)	*n*	Geminate (%)	Sig.
/i/					
/p/	90	97	90	56	<0.001 ***
/t/	90	98	90	40	<0.001 ***
/k/	90	58	90	83	<0.01 **
/b/	90	97	90	93	0.326 n.s.
/d/	90	93	90	92	0.769 n.s.
/ɡ/	90	97	90	97	1.000 n.s.
/m/	90	97	90	92	0.211 n.s.
/n/	90	86	90	92	0.227 n.s.
/l/	90	97	90	98	0.573 n.s.
/a/					
/p/	90	80	90	88	0.335 n.s.
/t/	90	41	90	92	<0.001 ***
/k/	90	76	90	96	<0.01 **
/b/	90	78	90	97	<0.01 **
/d/	90	94	90	89	0.258 n.s.
/ɡ/	90	98	90	96	0.489 n.s.
/m/	90	93	90	97	0.326 n.s.
/n/	90	88	90	94	0.184 n.s.
/l/	90	96	90	99	0.375 n.s.
/ŋ/	90	90	90	90	1.000 n.s.