A new look at Pattani Malay Initial Geminates: a statistical and machine learning approach

Francesco Burroni
Department of Linguistics and
Cognitive Science Program
Cornell University
203 Morrill Hall,
Ithaca, NY, USA 14850
fb279@cornell.edu

Sireemas Maspong
Department of Linguistics and
Southeast Asia Program
Cornell University
203 Morrill Hall,
Ithaca, NY, USA 14850
sm2627@cornell.edu

Pittayawat Pittayaporn
Department of Linguistics and
Southeast Asian Linguistics Research Unit
Faculty of Arts, Chulalongkorn University
Phayathai Road, Pathumwan,
Bangkok, Thailand 10330
pittayawat.p@chula.ac.th

Pimthip Kochaiyaphum
Department of Linguistics and
Southeast Asian Linguistics Research Unit
Faculty of Arts, Chulalongkorn University
Phayathai Road, Pathumwan,
Bangkok, Thailand 10330
pimthip.ko@student.ac.th

Abstract

In this paper, we present a statistical and machine learning approach to the acoustic discrimination of a cross-linguistically unusual phonological contrast, initial geminates vs. singletons in Pattani Malay. We show that the only statistically significant difference between geminates and singletons is the duration of the consonant itself. No differences in F0 and intensity were observed on the following vowel, *contra* earlier reports. We further investigated the robustness of this contrast using linear discriminant analysis. Results show that discrimination is above chance, but poor (~62%). The large overlap between the two categories may be partly due to the naturalistic nature of our speech samples. However, we also found that the contrast is neutralized in some minimal pairs. This merger is surprising since initial geminates are often the sole realization of lexical and morphosyntactic contrasts. We suggest that the singleton/initial geminate contrast is now best characterized as a marginal contrast. We hypothesize that this marginally contrastive status may be the result of an on-going sound change, perhaps connected with the more modest role that initial geminates play in Pattani Malay morphophonological alternations.

1 Introduction

Pattani Malay (PM), an Austronesian language spoken in Southern Thailand (Uthai 2011), exhibits a cross-linguistically unusual phonological 'length' contrast for all word-initial consonants, e.g., [matɔ] 'eye' *vs* [mːatɔ] 'jewelry'. The long forms of initial consonants, usually termed initial geminates (IGs), have been reported to differ from singletons along multiple acoustic dimensions. With regards to duration, PM IGs have been reported to be, on average, three times longer than their singleton counterparts (Abramson 1987). Durational differences are hardly a surprising finding since closure duration is usually considered the most reliable acoustic correlate of phonological length cross-linguistically (Ladefoged and Maddieson

1996). If previous work is representative, however, the IG/singleton duration ratio of 3:1 in PM would be on the extreme side of the spectrum (Ladefoged and Maddieson 1996).

Interestingly, duration is not the only cue that distinguishes IGs from singletons in PM. IGs have been reported to produce acoustic effects on the following vowel as well. In particular, previous research has reported that vowels following IGs display longer duration, higher fundamental frequency (F0), and higher intensity (Abramson 1987; Abramson 1998; Phuengnoi 2010). These F0 and intensity cues alone have been shown to be reliable enough for native speakers to correctly identify IGs vs singleton onsets; even in environments where durational cues are ambiguous, such as in absolute utterance-initial position where closure duration cannot be distinguished from preceding silence (Abramson 2003). Similar acoustic features in production and perceptual results have been reported for another closely related variety, Kelantan Malay (Hamzah et al. 2019; Hamzah et al. 2020).

The concomitant manifestation of IGs in the form of local durational differences and of effects on intensity and F0 of the following vowel has led scholars to hypothesize that PM speakers may be in the process of reanalyzing consonantal length as a prosodic contrast based on stress/pitch accent, or that the language may even be on its way to tonogenesis (Abramson 2004).

The possibility that IGs may be the target of ongoing sound change warrants by itself a fresh look at the realization of this unusual phonological contrast. However, we should be cautious in considering previous work the last word on PM IGs. For one thing, previous studies were based on a limited number of speakers (4 for Abramson, 7 for Phuengnoi). Moreover, the difference between IGs and singletons was studied only in words produced in isolation or in words that appeared in a carrier sentence. Finally, in previous studies, speakers were explicitly instructed about the production of the contrast in question. All these factors combined may have led to an exaggeration of the differences between IGs and singletons.

Given such limitations in previous studies, we investigate again the acoustic correlates of IGs in PM by comparing words with and without IGs, but we do so in more ecologically valid speech, which was elicited outside the lab using natural sounding sentences. To characterize the differences between IGs and singletons we make use of both statistical and machine learning techniques.

Statistical analyses showed that IGs are longer than their singleton counterparts, but the difference is much smaller than reported by previous studies. We also found no difference in F0 and intensity on the vowel following IGs vs singletons, *contra* the reports of previous studies.

Additionally, to quantify the robustness of the IG/singleton contrast and to find out which dimensions best discriminate the two categories, we performed classification using linear discriminant analysis (LDA) with a variety of models that employ different combinations of acoustic features. We found that the model performances are above chance, but still poor, peaking at only about 62% accuracy for the best feature combinations.

We speculate that the limited statistical differences and low accuracy of the LDA may be partly due to the naturalistic nature of the speech materials we collected and to ongoing neutralization of the contrast in some minimal pairs. We conclude by discussing several hypotheses concerning the mechanisms that may be at the heart of the observed neutralization.

2 Acoustic Analyses

2.1 Methodology

14 native speakers of PM (6M; 8F) were asked to pronounce 13 disyllabic minimal pairs differing only for their word-initial onsets, which were either geminate or singleton, as shown in Table 1. Stimuli were presented orally with natural-sounding Thai sentences containing the target words. Participants were asked to translate the sentence into PM. Each sentence was repeated six times.

singleton (CVCV)	**gloss**	**geminate (Cː VCV)**	**gloss**
pagi	‘morning’	*pːagi*	‘early morning’
paka	‘to use/wear’	*pːaka*	‘usable’
tanɔh	‘land’	*tːanɔh*	‘outside’
dapo	‘kitchen’	*dːapo*	‘at the kitchen’

singleton (CVCV)	**gloss**	**geminate (Cː VCV)**	**gloss**
katoʔ	‘hammer’	*kːatoʔ*	‘frog’
kabo	‘Java kapok’	*kːabo*	‘beetle’
gaɟɨ	‘wage’	*gːaɟɨ*	‘saw’
ɟalɛ	‘path’	*ɟːalɛ*	‘to walk’
ɟuɣi	‘to steal’	*ɟːuɣi*	‘thief’
misa	‘mustache’	*mːisa*	‘to grow a moustache’
labɔ	‘profit’	*lːabɔ*	‘spider’
bulɛ	‘moon’	*bːulɛ*	‘month’
buŋɔ	‘flower’	*bːuŋɔ*	‘to bloom’

Table 1. Stimuli

Audio was collected at 44.1 kHz in Praat (Boersma and Weenink 2020). All recordings were made in quiet rooms at the Prince of Songkla University Pattani Campus.

Segmental boundaries were obtained in Praat TextGrids by forced alignment using the Montreal Forced Aligner (McAuliffe et al. 2017). The TextGrids were inspected and manually corrected when necessary. The corrected TextGrids containing segmental boundaries and the audio signals of each word were read back in MATLAB® for analysis.

Eight acoustic measurements were collected:

(1) Duration of initial segments (ms)
(2) Duration of initial syllables (ms)
(3) F0 mean of initial syllables (semitone)
(4) Intensity peak of initial syllables (dB)
(5) F0 mean over initial 10% of vowel following target consonants (semitone)
(6) Intensity mean over initial 10% of vowel following target consonants (dB)
(7) Difference between semitone transformed mean F0 of initial and final syllable
(8) Ratio of mean RMS amplitude of initial to final syllable

F0 was calculated using a MATLAB® implementation of Talkin’s robust algorithm for pitch tracking (Talkin 1995) contained in the Voicebox toolbox for MATLAB®. (Brookes 1997). F0 was further processed within all trials and separately by participant by removing all data points with standard deviation scores greater than 2 from the mean; datapoints deviating ±10 Hz from neighboring samples were also excluded. When the F0 vector of a word contained less than 5 datapoints per each syllable, the contour was no longer processed, as interpolation over the entire word would not be reliable. In the other cases, F0 was subsequently interpolated using spline interpolation and smoothed using a median filter. F0 was transformed by converting from Hz to semitones according to the equation $\frac{12}{log_{10}2} \times log_{10}\left(\frac{Hz}{\mu Hz}\right)$ in Zhang (2018).

Sound Pressure Level (SPL) normalized intensity was calculated by transforming the root mean squared intensity of the signal to dB and normalizing to human auditory threshold using the formula $20 \times log_{10}\frac{P}{I0}$. In this formula P represents the power of the signal and $P0$ represents the normalizing term for the auditory threshold of a 1000 Hz sine wave, equal to 2×10^{-5} (Huang et al. 2001).

Statistical analyses were conducted by fitting linear mixed effect regressions. We compared a model where the fixed effect was the presence/absence of IGs to an intercept-only model. Random effects were subject, word, and position of the word in the phrase (medial or final). Random intercepts were present in the model for each random effect. Random slopes were added when they resulted in a better fit as determined *via* a loglikelihood ratio test. Loglikelihood ratio tests were, thus, used to assess statistical significance and to determine the random effect structure.

2.2 Results

Consonant Duration: Comparing the initial segment in the IG and no IG condition, we found that IGs are significantly longer than singletons ($\chi^2(1) = 4.03$, p = .04) with an effect size estimated at 17 ms, as illustrated in Figure 1.

Syllable Duration: The presence of IGs does not significantly affect the duration of the initial syllable ($\chi^2(1) = 1.34$, p= .24), as illustrated in Figure 2.

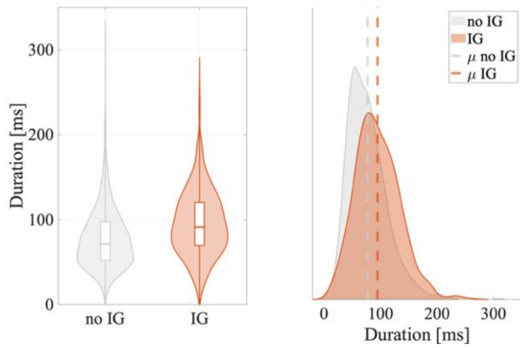

Figure 1. Comparison of initial segment duration (ms)

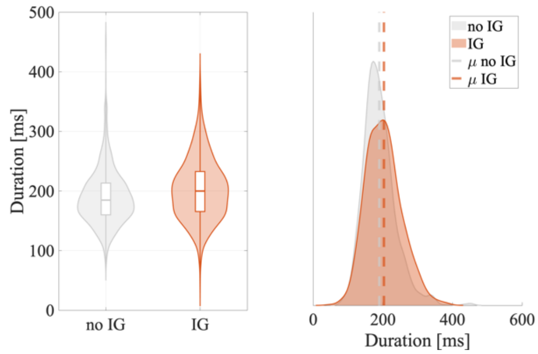

Figure 2. Comparison of syllable duration of initial syllables (ms)

F0: The presence of IGs does not significantly affect the mean F0 of the initial syllable ($\chi^2(1)$ = 0.16, p= .69), as illustrated in Figure 3.

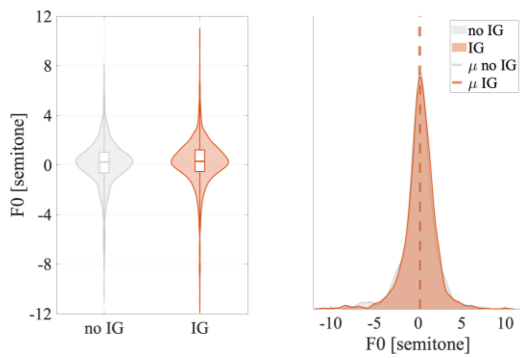

Figure 3. Comparison of mean F0 of initial syllables (semitones)

Intensity: IGs do not significantly affect the maximum SPL normalized intensity of the initial syllable ($\chi^2(1)$ = 0.49, p= .48), as illustrated in Figure 4.

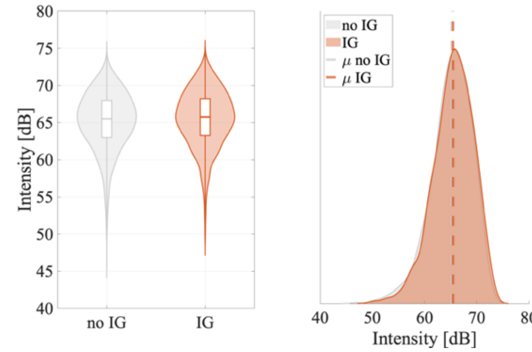

Figure 4. Comparison of maximum SPL normalized intensity of initial syllables (dB)

To further investigate whether the effects of IGs on the following vowel may be limited to the region immediately following the release, we also examined mean F0 and intensity over the first 10% of the vowel, following previous work on Kelantan Malay (Hamzah et al. 2020).

We found no significant differences between mean F0 over the initial 10% of the vowel following IGs vs. singletons ($\chi^2(1)$ = 0.06, p= .79). F0 contours over the vowel are presented in Figure 5.

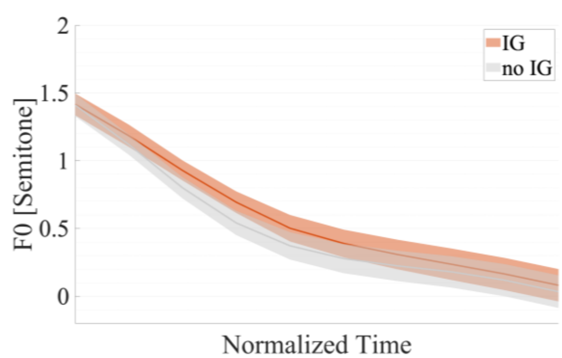

Figure 5. Comparison of time normalized F0 trajectory of initial vowel in semitone. Shaded areas represent ±2 Standard Errors

We also found no significant difference between mean SPL normalized intensity over the initial 10% of a vowel following IGs vs. singletons ($\chi^2(1)$ = 0.95, p= .33). The intensity contours of the following vowel are presented in Figure 6.

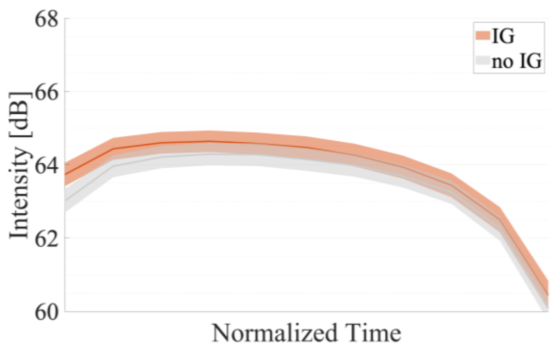

Figure 6. Comparison of time normalized SPL normalized intensity trajectory in dB. Shaded areas represent ±2 Standard Errors

Finally, also following previous work (Abramson 1998; Hamzah et al. 2020), we examined whether differences between IGs and singletons may be manifested more globally in the F0 difference and RMS amplitude ratios of the two syllables. We found no differences for both F0 ($\chi^2(1) = 0.007$, p= .93) and RMS amplitude ($\chi^2(1) = 0.07$, p= .79), as illustrated in Figure 7.

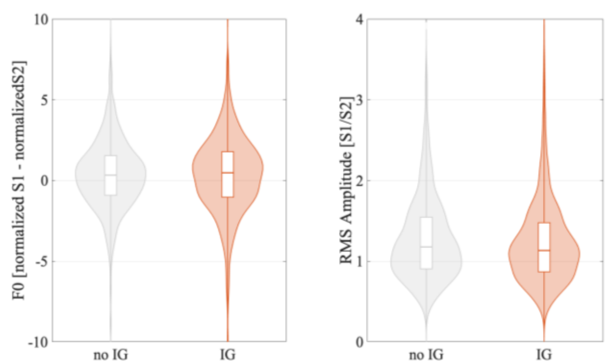

Figure 7. F0 difference and RMS amplitude ratio

2.3 Summary

We found that the durations of IGs and singletons are significantly different, but, unlike in previous studies, the duration of IGs is not three times longer than singletons. The durational differences are estimated at about 17 ms. Furthermore, there is a significant overlap between the two distributions. Contrary to previous descriptions, the presence or absence of IGs does not have a significant effect on syllable duration, mean F0, or peak intensity of the following vowel; no effect is observed even if only 10% of the vowel is examined. We also observed no significant differences in the F0 difference and amplitude ratios of the two syllables.

In sum, we found only very small durational differences between IGs and singletons and the other acoustic measurements do not display significant differences.

3 Linear Discriminant Analysis

To further address the question of whether the singleton/IG contrast in PM is comparable in terms of its magnitude to the singleton/geminate contrast of other languages, we performed classification of IGs vs. singletons using linear discriminant analysis (LDA). In a nutshell, LDA is a classification technique (and also a dimensionality reduction technique) that uses linear combinations of features to maximize the separation between two or more categories. LDA is of interest here because it has been successfully applied to the study of various phonetic contrasts, including geminate vs non-geminate contrasts in both word-medial, in Japanese (Idemaru and Guion-Anderson 2010) and Lebanese Arabic (Khattab and Al-Tamimi 2014), and word-initial position, in Salentino (Burroni and Maspong to appear). We tried to extend this methodology to characterize the word-initial geminate contrast of PM.

3.1 Methodology

We fitted LDA models using cross-validation to evaluate the accuracy of our models. We randomly assigned 80% of the data to a training set and the remaining 20% to a test set. 10,000 such LDA models were fitted for each combination of predictors. The mean accuracy and standard deviations reported here were taken over these 10,000 iterations.

To determine which acoustic dimensions were more apt to discriminate the singleton/IG contrast, we considered that duration of the first segment (CDur) and ratio of the duration of the first segment to the entire word (CDur / WordDur) are the only two statistically significant differences present in our data. We then tested whether adding information concerning the duration (σ_i Dur), mean F0 (σ_i MeanF0), and maximum intensity (σ_i MaxInt) of the target syllable would improve LDA classification. All features were z-scored by participants before performing LDA, as this procedure is known to improve LDA classification.

3.2 Results

We found that the model performance is above chance (that is, above 50%), but still quite poor, as summarized in Table 2, peaking at only about 62% accuracy for the best linear combination of features: the duration of the first segment (CDur) alone or in combination with the duration ratio of the first segment to the entire word (CDur / WordDur).

Model Structure	Mean Accuracy	Standard Deviation
CDur + CDur/WordDur + σ_i*Dur +* σ_i*MaxInt +* σ_i*MeanF0*	58.84%	2.18%
CDur + CDur/WordDur + σ_i*Dur +* σ_i*MeanF0*	58.20%	2.07%
CDur + CDur/WordDur + σ_i*Dur +* σ_i*MaxInt*	58.88%	2.20%
CDur + CDur/WordDur + σ_i*Dur*	58.19%	2.10%
CDur + CDur/WordDur	**61.40%**	2.06%
CDur/WordDur	59.84%	2.14%
CDur	**62.36%**	2.11%

Table 2. Accuracy of LDA models for different combinations of features

Optimizing the hyperparameters of the model does not greatly improve performance in the identification of IGs as is clear from the confusion matrix of the optimized model presented in Figure 8.

If we inspect the predicted boundary between the two classes, as shown in Figure 9, the reason for the low performance of the model becomes clear: IGs and singletons are not linearly separable in the investigated acoustic dimensions, thus, they cannot be captured by an LDA classifier.

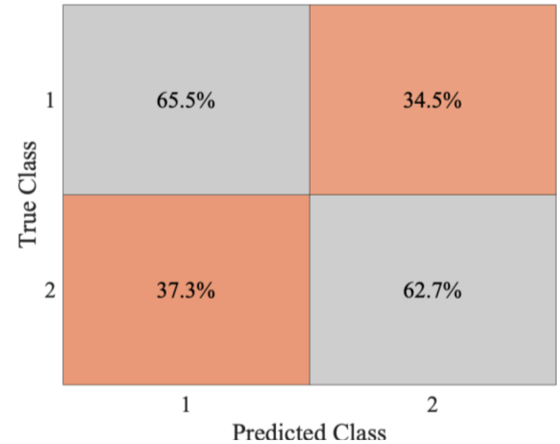

Figure 8. Confusion matrix showing the number of IGs (class 1, top) and singletons (class 2, bottom) classified correctly (gray diagonal) and incorrectly (orange diagonal).

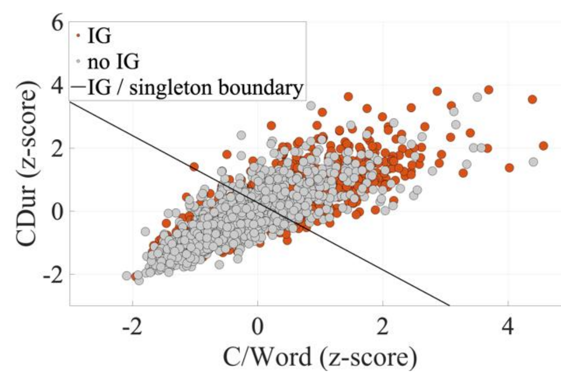

Figure 9. Output of LDA showing large overlap between categories

The low LDA accuracy for geminates contrasts sharply with high accuracy reported for other languages. For instance, for medial geminates in Japanese, accuracy is at ~85-95% (Idemaru and Guion-Anderson 2010) and, for IGs in Salentino, accuracy is at ~80% (Burroni and Maspong to appear).

3.3 Summary

In sum, the discrimination above chance shows that there is indeed a contrast between words with and those without IGs that can be picked up by a simple model, such as an LDA classifier. This is in line with previous phonetic and phonological research on PM and justifies looking for contrasts between words with and without IGs. On the other hand, the low classification accuracy suggests that the contrast is subtle.

We now discuss what factors may be responsible for the observed overlap between IGs and singletons.

4 Discussion

We have three non-mutually exclusive hypotheses to explain why the contrast between IGs and singletons looks much less robust than previously reported.

The first possibility that comes to our mind is that the differences between the result of our study and previous work is due to different methods of data collection. Previous work (Abramson 1987; Abramson 1998) examined IGs only in isolation and in a carrier sentence. Our data, on the other hand, presented IGs and their singleton counterparts in naturalistic sentences. Accordingly, the difference could be due to less carefully articulated speech.

A second possibility is that the contrast may be neutralized for some speakers. The size of our dataset does not allow for a full quantitative assessment of this claim; however, our impression is that almost all speakers produce IGs that are longer than singletons on average, as illustrated in Figure 10.

A third possibility is that the contrast only exists for a subset of minimal pairs. This means that, for many lexical items, the contrast between singletons and IGs is not realized.

Indeed, our data suggests that closure duration of the initial consonants is distinct only for a subset of minimal pairs, as illustrated in Figure 11.

Given this observation, we ask what generalizations may explain the observed neutralizations, as well as the non-neutralizations.

In the framework of Evolutionary Phonology (EP), IGs have been hypothesized to be diachronically unstable (Blevins 2004). Furthermore, EP holds that the stability of phonetic cues to IGs may be related to their wider role in the grammar. IGs survive only in languages where they represent the only cue to lexical contrasts and produce "sentential minimal pairs". In other words, IGs survive only when they compete lexically with singletons and cannot be disambiguated by context (Blevins and Wedel 2009; Burroni and Maspong to appear).

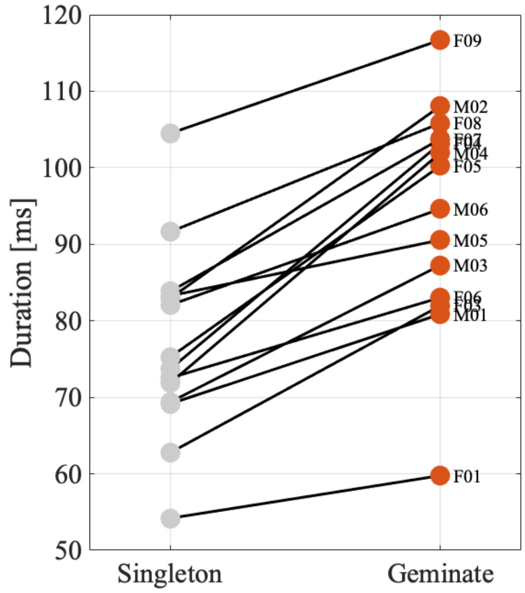

Figure 10. Mean duration of singletons (left) and IGs (right) by speaker (ms)

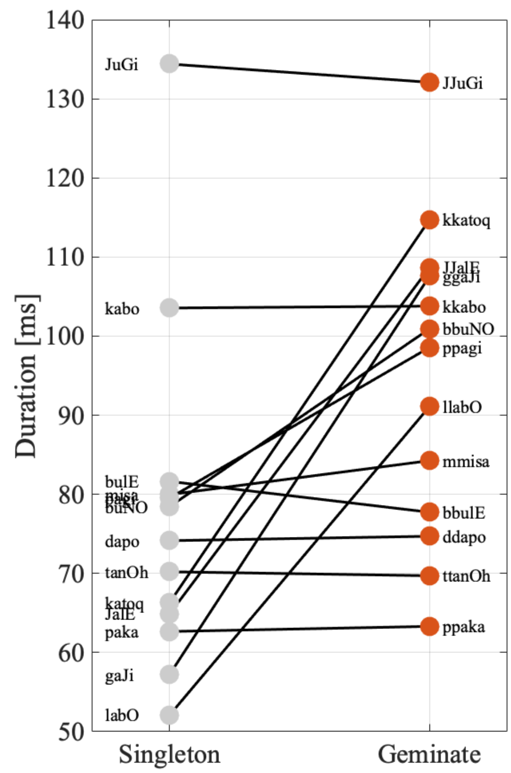

Figure 11. Mean duration of singletons (left) vs IGs (right) by word (ms)

Interestingly, PM seems a counterexample to this generalization, as IGs are being lost in this language, even though they are the unique realization of morphosyntactic contrasts. For instance, under an EP approach, the observed merger of [dapo] 'kitchen' and [dːapo] 'at the kitchen' is expected, since these forms appear in different positions and can be disambiguated by context. Similarly, the non-merger of [katoʔ] 'hammer' and [kːatoʔ] 'frog' is expected since these forms appear in the same context and the IG or lack thereof is the only cue distinguishing them. However, other mergers, such as [kabo] 'Java kapok (type of plant)' and [kːabo] 'beetle (type of bug)', are not expected, because context does not allow for disambiguation, thus, the neutralizing IG would be one that is a unique cue to the contrast, just like the non-merging one in [katoʔ]/[kːatoʔ]. However, the merger of [kabo]/[kːabo] may suggest some role for word frequency effects. Phillips (2006) explained that retrieving low-frequency word is a challenge for the learner. These difficulties, in turn, may lead to alterations of the phonetic forms of low frequency words on the model of unmarked patterns, that IGs may be altered to singletons. At any rate, for another counterexample to the EP claim that cues to IGs are dependent on lexical competition, we refer the reader to Burroni and Maspong (to appear). Since lexical competition alone does not explain the paradox of IGs merging with singletons in PM, other factors need to be considered.

It has been reported that PM speakers no longer make use of IGs for the purpose of morphological derivation due to contact with Thai (Uthai 1993), accordingly, it is possible that the contrastive phonological status of IGs is being eroded in connection with their reduced 'functional' role in the grammar. If IGs and singletons will be merging at evolutionary timescales, the loss of PM IG contrasts would be a striking example of sound change via lexical diffusion connected with a reduced functional load, an information theoretic measurement that has been argued to correlate with geminate to singleton ratio (Tang and Harris 2014) and resistance to merger (Wedel et al. 2013). Further research is necessary to test the merits of these hypotheses on the basis of a larger PM dataset. Corpus frequencies also need to be obtained in order to calculate information theoretic measurements, such as functional load (Surendran and Niyogi 2006).

At any rate, since the contrast between IGs and singletons is only observed for some minimal pairs, it may be best interpreted as a quasi-phonemic or marginal contrast (Hall 2013; Renwick and Ladd 2016). If this interpretation is correct, our acoustic results would align with recent work demonstrating that marginal phonological contrasts may display large overlaps when data is collected outside the lab, in more naturalistic contexts (Cohn and Renwick 2019).

5 Conclusion

In this paper, we have shown that the only significant difference between PM IGs and singletons in naturalistic speech is the duration of the consonants themselves. We have further shown that an LDA model is able to discriminate between syllables with and without IGs slightly above chance level (~62%). This is much below usual LDA performances for geminates in other languages.

The striking difference between our findings and earlier reports regarding the robustness of cues to IGs in PM calls for an explanation. One possibility is that previous experimental work may have exacerbated the difference between IGs and singletons. After all, highly controlled lab speech is very different from less carefully articulated naturalistic speech. IGs in PM could then be an example showing that a more nuanced characterization of phonological contrasts requires an integrated analysis of both laboratory and more naturalistic phonetic data, as advocated by Cohn and Renwick (2019).

However, we have also shown that, although speakers on average produce longer IGs than singletons, they produce the contrast only for a subset of minimal pairs. We have speculated that an appropriate characterization of the subsets that undergo and resist merger will require further collection of information theoretic measurements, such as functional load. One thing is relatively clearer: IGs are moving towards a more marginally contrastive role in the grammar of PM, a fact that may be reflected in their phonetic realization.

Acknowledgements

We would like to thank Santhawat Thanyawong for his help with data collection, Sam Tilsen, and members of the Cornell Phonetics Lab for their feedback. We are grateful to the Southeast Asian Linguistics Research Unit, Faculty of Arts, Chulalongkorn University for financial support.

References

Abramson, Arthur S. 1987. Word-initial consonant length in Pattani Malay. *International Congress of Phonetic Science* (ICPhS) 6, 68–70.

Abramson, Arthur S. 1998. The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length. *Annual meeting of the Southeast Asian Linguistics Society* (SEALS) 4, 1–20.

Abramson, Arthur S. 2003. Acoustic cues to word-initial stop length in Pattani Malay. *International Congress of Phonetic Science* (ICPhS) 15, 387–390.

Abramson, Arthur S. 2004. Toward prosodic contrast: Suai and Pattani Malay. In *International Symposium on Tonal Aspects of Languages: Emphasis on Tone Languages*, 1–4.

Blevins, Juliette. 2004. *Evolutionary Phonology: the emergence of sound patterns*. Cambridge; New York: Cambridge University Press.

Blevins, Juliette & Andrew B. Wedel. 2009. Inhibited sound change: An evolutionary approach to lexical competition. *Diachronica* 26(2), 143–183.

Boersma, Paul & David Weenink. 2020. *Praat: doing phonetics by computer*. [Computer software].

Brookes, Mike. 1997. *Voicebox: Speech processing toolbox for* MATLAB. [Computer Software].

Burroni, Francesco & Sireemas Maspong. to appear. Re-examining Initial Geminates: Typology, Evolutionary Phonology, and Phonetics. In *Historical Linguistics 2019*.

Cohn, Abigail C. & Margaret E. L. Renwick. 2019. Doing Phonology in the age of big data. In *Cornell Working Papers in Phonetics and Phonology 2019*, 1–36.

Hall, Kathleen Currie. 2013. A typology of intermediate phonological relationships. *The Linguistic Review* 30(2), 215 – 275.

Hamzah, Mohd Hilmi, J Fletcher & J Hajek. 2019. The secondary roles of amplitude and F0 in the perception of word-initial geminates in Kelantan Malay. *International Congress of Phonetic Science* (ICPhS) 19, 2735-2757.

Hamzah, Mohd Hilmi, John Hajek & Janet Fletcher. 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*. Cambridge University Press 50(1), 23–60.

Huang, Xuedong, Alex Acero, Hsiao-Wuen Hon & Raj Foreword By-Reddy. 2001. *Spoken language processing: A guide to theory, algorithm, and system development*. Prentice hall PTR.

Idemaru, Kaori & Susan Guion-Anderson. 2010. Relational timing in the production and perception of Japanese singleton and geminate stops. *Phonetica* 67(1–2), 25–46.

Khattab, Ghada & Jalal Al-Tamimi. 2014. Geminate timing in Lebanese Arabic: The relationship between phonetic timing and phonological structure. *Laboratory Phonology* 5(2), 231–269.

Kraehenmann, Astrid. 2011. Initial geminates. *The Blackwell companion to phonology*. Wiley Online Library.

Ladefoged, Peter & Ian Maddieson. 1996. *The sounds of the world's languages*. Oxford; Cambridge, MA: Blackwell Oxford.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner & Morgan Sonderegger. 2017. *Montreal Forced Aligner*. [Computer software].

Phillips, Betty S. 2006. *Word frequency and lexical diffusion*. New York: Palgrave Macmillan.

Phuengnoi, Nattaphon. 2010. *An acoustic study of stressed and unstressed syllables in Pattani Malay and Urak Lawoi'*. Bangkok: Chulalongkorn University thesis.

Renwick, Margaret E. L. & D. Robert Ladd. 2016. Phonetic distinctiveness vs. lexical contrastiveness in non-robust phonemic contrasts. *Laboratory Phonology* 7(1). 1–29.

Surendran, Dinoj & Partha Niyogi. 2006. Quantifying the functional load of phonemic oppositions, distinctive features, and suprasegmentals. In Ole Nedergaard Thomsen (ed.), *Competing Models of Linguistics Change: Evolution and Beyond. In commemoration of Eugenio Coseriu (1921-2002)*, 43–58. Amsterdam & Philadelphia: Benjamins.

Talkin, David. 1995. A robust algorithm for pitch tracking (RAPT). In *Speech Coding and Synthesis*. 495-518. New York, NY: Elsevier Science.

Tang, Kevin & John Harris. 2014. A functional-load account of geminate contrastiveness: a meta-study. In *Linguistics Association of Great Britain 2014*.

Uthai, Ruslan. 1993. *A comparison of word formation in Standard Malay and Pattani Malay*. Bangkok: Chulalongkorn University thesis.

Uthai, Ruslan. 2011. *Keistimewaan dialek Malayu Patani*. Bangi: Penerbit Universiti Kebangsaan Malaysia.

Wedel, Andrew, Abby Kaplan & Scott Jackson. 2013. High functional load inhibits phonological contrast loss: A corpus study. *Cognition* 128(2). 179–186.

Zhang, Jingwei. 2018. A Comparison of Tone Normalization Methods for Language Variation Research. *Pacific Asia Conference on Language, Information and Computation* (PACLIC) 32, 823-831.