Unifying Initial Geminates and Fortis Consonants via Laryngeal Specification: Three case studies from Dunan, Pattani Malay, and Salentino

Francesco Burroni,[1] Raksit Lau-Preechathammarach,[2] Sireemas Maspong[1]
[1]Cornell University, [2]University of California, Berkeley

1 Introduction

1.1 Initial geminates and fortis consonants

In the phonological literature, initial geminates and fortis consonants are treated as two separate categories. Initial geminates are phonologically represented by associating consonants to two timing units (See Muller (2001) for a more detailed discussion of representations of initial geminates). Fortis consonants, on the other hand, are often represented with a [+tense] feature or described as involving greater "articulatory force". Although initial geminates and fortis stops are implicitly analyzed with different phonological representations, they share many of the same acoustic cues (Jakobson et al., 1952; Trubetzkoy, 1969; Jaeger, 1983; Kohler & van Dommelen, 1987; Arvaniti & Tserdanelis, 2000; Cho et al., 2002; DiCanio, 2012; Al-Tamimi & Khattab, 2015; Ghosh, 2015). These are summarized in Table 1.

	Initial geminates	Fortis consonants
Phonological representations	two timing slots or a mora	[tense]/[constricted glottis]/[stiff]
Acoustic correlates	↑ duration, ↑ f0, ↑ intensity, ↓ spectral tilt, VOT (language dependent)	↑ "articulatory energy" / "resistance", ↑ duration, ↑ f0, ↑ intensity, ↓ spectral tilt, VOT (language dependent)

Table 1: Phonological representations and acoustic correlates of initial geminates and fortis consonants

The main acoustic correlate of initial geminates is assumed to be longer duration while other cues, such as higher fundamental frequency (f0), higher intensity, lower spectral tilt, and less positive VOT, are described as secondary, variable cues (Kraehenmann, 2011). On the other hand, the main correlate of fortis consonants is less clear, being generally described as high "articulatory energy" or high "resistance". The acoustic details of the descriptions, however, contain many of the same cues as those described for initial geminates: longer duration, higher f0, higher intensity, lower spectral tilt, and less positive VOT.

Based on the acoustic correlates of initial geminates and fortis consonants, an overarching question is whether initial geminates are distinct from fortis consonants. This question is not entirely new. One case study is Korean stops. Korean displays a three-way fortis, lenis, and aspirated contrast. There are two main proposals for the phonological representations of "fortis" stops in Korean (Kang et al., forthcoming). One of the earliest proposals by Kim (1965) introduced "tensity" or a [tense] feature to classify fortis stops. Similar proposals use [Constricted Glottis]

* We would like to thank our Dunan, Pattani Malay, and Salentino consultants for providing their time to work with us, the members of the Cornell Phonetics Lab and UC Berkeley's Phorum, Bob Ladd, Susan Lin, Pittayawat Pittayaporn, Larry Hyman, and three anonymous reviewers for comments on our work, Keith Johnson and Ron Sprouse for the automatic burst detection script, and Pimthip Kochaiyaphum for help in collecting data from Pattani Malay consultants.

and/or [Stiff Vocal Folds] features (Halle & Stevens, 1971; Iverson, 1983; Lombardi, 1994; Cho & Keating, 2001; Cho et al., 2002). Another line of analyses proposes that Korean fortis stops are underlyingly geminates (Avery & Idsardi, 2001; Ahn & Iverson, 2004). Similar analyses proposed a redundant representation with a [tense] feature with two timing slots (Han, 1996). In looking at Swiss German, Ladd & Schmid (2018) suggested that the boundary between initial geminates and fortis consonants is tenuous.

In this paper, we present converging evidence from three understudied languages, which have been reported to have initial geminates: Dunan (Japonic), Pattani Malay (Austronesian), and Salentino (Indo-European). We demonstrate that initial geminates may also be reliably cued by f0 and intensity. Reliance on these acoustic cues is strongly reminiscent of descriptions of the acoustic cues of fortis consonants in, for instance, Korean. Accordingly, we propose a rich, unified phonological representation of initial geminates and fortis consonants, where several cues are available to implement phonological contrast and the difference may lie in language-specific weights of these cues.

1.2 Language background Dunan is a Southern Ryukyuan language spoken on the island of Yonaguni in Okinawa Prefecture, Japan. Yamada et al. (2015) estimate approximately 400 speakers (about 25% of the population of the island). As the youngest speakers are in their sixties, there is no intergenerational transmission and the language is “severely endangered” (Catalogue of Endangered Languages, 2022). Dunan has a three-vowel inventory /a i u/[1] with no distinctive vowel length. The consonant inventory is laid out in Table 2. As geminate consonants are the object of study in this paper, they have been bolded. While Bentley (2008) describes these consonants as geminates, Yamada et al. (2015) describes them as fortis. /pp/ and /cc/ do not have singleton counterparts. Word-medial stops are neutralized to geminates. Dunan initial geminates historically arise from the syncope of high vowels between obstruents in word-initial syllables.

	Labial	Dental	Palatal	Velar	Glottal
Stop	pp, b	t, tt, d	cc	k, kk, g	
Nasal	m	n		ŋ	
Fricative		s			h
Tap		r			
Approximant	w		j		

Table 2: Dunan consonant inventory (adapted from Yamada et al. (2015:451))

Pattani Malay is one of the few Austronesian languages spoken in mainland Southeast Asia. Approximately 1.47 million people speak Pattani Malay at home according to the 2010 census. Pattani Malay, spoken in southern Thailand, is a dialect continuum with Kelantan Malay and other northeastern peninsular Malay varieties spoken in Malaysia (Collins, 1989). Pattani Malay is well-known as a language with geminates attested only in word-initial position. All possible onset consonants in Pattani Malay have a geminate counterpart, as shown in Table 3 adapted from Uthai (2011).

Salentino is a Romance variety belonging to the Extreme Southern Italian group; it is spoken in the cities of Lecce, Brindisi, and Taranto. As an Extreme Southern Italian variety, Salentino is most closely related to Southern Calabrian and Sicilian varieties. An official number of speakers is not available, however, estimates are in the range of 1-1.5 million speakers. A unique feature of Salentino in the Apulian area is the presence of contrastive initial geminates (Romano, 2003, 2015), which are described as robustly cued when compared with other Southern Italian languages that also have this constrast, e.g., such as Altamurano (Bertinetto & Loporcaro, 1999).

Initial geminates in Salentino have several possible lexical and non-lexical origins. They can arise from (i) expressive gemination,; (ii) consonantal assimilations at the beginning of a word, once initial vowel aphaeresis has taken place; (iii) Raddoppiamento Sintattico, an initial gemination process active after stressed vowels and a closed set of function words that do not have final stress

[1] /o/ occurs in a particle -do and a few interjections.

		Labial	Dental	Palatal	Velar	Glottal
Stop	voiceless	p, pp	t, tt	c, cc	k, kk	ʔ, **ʔʔ**
	voiceless aspirated	p^h, pp^h	t^h, tt^h	c^h, cc^h	k^h, kk^h	
	voiced	b, bb	d, dd	ɟ, **ɟɟ**	g, gg	
Fricative	voiceless		s, ss			h, hh
	voiced		zz		ɣ,**ɣɣ**	
Nasal		m, mm	n, nn	ɲ, **ɲɲ**	ŋ, **ŋŋ**	
Liquid			l, ll, rr			
Approximant		w, ww		j, jj		

Table 3: Pattani Malay consonant inventory (adapted from Uthai (2011))

(Loporcaro, 1997).

Unlike in Dunan and Pattani Malay, Salentino speakers distinguish singletons from geminates both word-initially and word-medially. Not all possible onset consonants in Salentino have an initial geminate counterpart, as shown in Table 4.

		Labial	Labiodental	Alveolar	Post-Alveolar	Retroflex	Velar
Stop	voiceless	p, pp		t, tt		ʈ,**ʈʈ**	k, kk
	voiced	bb		d, dd		ɖ,**ɖɖ**	g, gg
Fricative	voiceless		f, ff	s, ss			
	voiced		v, vv	z			
Affricate	voiceless			ts, tts	tʃ, **ttʃ**		
	voiced			dz, ddz	dʒ, **ddʒ**		
Nasal		m, mm		n, nn	ɲ **ɲɲ**		ŋ
Liquid				l, ll, r, rr			
Approximant		w			j, jj		

Table 4: Salentino consonant inventory

2 Methodology

2.1 Data collection Pattani Malay and Salentino data were collected in controlled, more laboratory-like settings, while Dunan data was gathered from fieldwork recordings.

2.1.1 Dunan The data for Dunan was culled from 19 fieldwork sessions with Yoneshiro Sueko, a speaker in her late seventies, carried out by Lau-Preechathammarach in the speaker's home. Sessions were recorded at a sampling rate of 44.1 kHz with a Zoom H4n Recorder and a Sony Electret Condenser Microphone (ECM-CR120). The content of the recorded sessions included elicitation of words and sentences, as well as narratives and stories from the speaker. As such, the nature of the data is starkly different from the data collected from Pattani Malay and Salentino. Sentences were aligned and transcribed in ELAN (Wittenburg et al., 2006) by Lau-Preechathammarach and Maspong. Following transcription, the Montreal Forced Aligner (MFA) (McAuliffe et al., 2017) was used to automatically align words and phones in order to create a miniature corpus. Target words containing initial /t tt k kk/ were identified and words with a glide (/w/ or /j/) following the initial consonant were omitted, leaving 2711 (666 t; 556 tt; 1346 k; 143 kk) target words. Because voiceless stops consist of silence, we removed tokens for which the beginning of the stop could not be identified. This process involved two steps: (1) Stops over 300 ms were excluded first. (2) With the remaining set, the mean and standard deviation for each stop was calculated and tokens over 3 standard deviations from the mean were removed. Because singletons are described as slightly aspirated (Yamada et al., 2015), we also applied a stop burst identification script, written by Keith Johnson (p.c.), to automatically detect the timing of stop bursts. The Voice Onset Time (VOT) was labeled as beginning from the timepoint of the burst and ending at the right edge of the consonant, as marked by the MFA. The burst script was unable to identify a burst for 306 tokens, so only 2405 (644 t; 510 tt; 1124 k; 127 kk) tokens were available to analyze VOT. The acoustic measures f0 and Intensity were measured over the span of the vowel following the initial stop for all 2711 tokens. 1225 (326 t; 313 tt; 530 k; 56 kk) tokens with both reliable closure onsets and a detectable burst remained for which stop duration could be calculated. Figure 1 illustrates the results of the automatic segmentation and burst detection for a word beginning with a singleton (left) and one beginning with a geminate (right). We recognize that geminate /kk/ is severely underrepresented and so the results must be interpreted with caution.

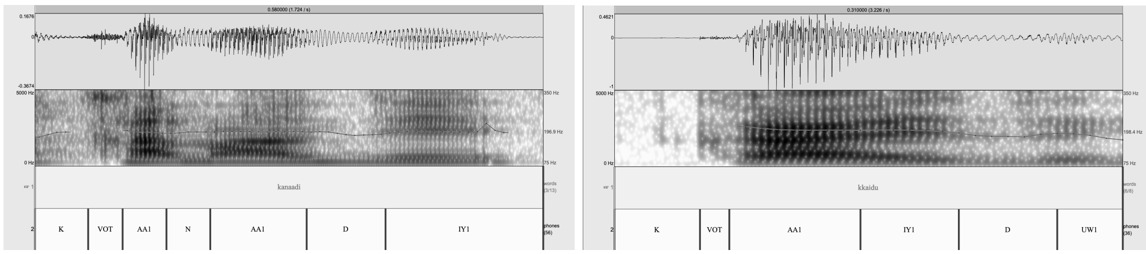


Figure 1: Automated VOT alignment for word beginning with singleton (left) and geminate (right)

2.1.2 Pattani Malay The data from Pattani Malay was collected from 14 native speakers of Pattani Malay (6M; 8F). They were asked to pronounce 13 disyllabic minimal pairs differing only in their word-initial onsets which were either geminates or singletons, as shown in Table 5. The stimuli were elicited in two phrasal positions: phrase-medial and phrase-final positions. Stimuli were presented orally with natural-sounding Thai sentences containing the target words. Participants were asked to translate the sentence to Pattani Malay. Each sentence was repeated six times.

Singleton (CVCV)		Geminate (CCVCV)	
pagi	‘morning’	ppagi	‘early morning’
paka	‘to use/wear’	ppaka	‘usable’
tanɔh	‘land’	ttanɔh	‘outside’
dapo	‘kitchen’	ddapo	‘at the kitchen’
katoʔ	‘hammer’	kkatoʔ	‘frog’
kabo	‘Java kapok’	kkabo	‘beetle’
gaɟi	‘wage’	ggaɟi	‘saw’
ɟalɛ	‘path’	ɟɟalɛ	‘to walk’
misa	‘mustache’	mmisa	‘to grow a mustache’
labɔ	‘profit’	llabɔ	‘spider’
bulɛ	‘moon’	bbulɛ	‘month’
buŋɔ	‘flower’	bbuŋɔ	‘to bloom’
ɟuɣi	‘to steal’	ɟɟuɣi	‘thief’

Table 5: Pattani Malay stimuli

Audio data were collected at 44.1 kHz in Praat (Boersma & Weenink, 2020) by using a Rode NT-USB, a cardioid condenser microphone. The recordings were done 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. There are 1,092 tokens in total.

2.1.3 Salentino Salentino data was collected from a native speaker, a university-educated male who was 26 at the time of the recording and hails from Specchia (Lecce) in Southern Salento. We collected 14 minimal pairs, representing all possible initial geminates and singleton pairs, Table 6. Each of the 28 unique items was embedded in a carrier sentence, [li'ddʒimu XXX bbonu] ‘we read

XXX well' and repeated 11 times. The order of presentation of each word was randomized.

Singleton (CVCV)		Geminate (CCVCV)	
punta	'point,tip'	ppunta	'I appoint'
mattu	'crazy M.SG.'	mmattu	'I place inside'
fjanku	'flank'	ffjanku	'next to'
vita	'life'	vvita	's/he screws'
takku	'heel'	ttakku	'I attack/stick'
dumu	'I tame'	ddumu	'I light up'
nata	'born F.SG.'	nnata	'year'
riva	'bed of a river'	rriva	's/he comes near'
lɛnta	'slow F.SG.'	llɛnta	's/he lets loose'
sulu	'alone'	ssulu	'keep a card'
ʧitu	'vinegar'	ʧʧitu	'I kill'
dʒustu	'right'	dʒdʒustu	'I repair'
kasu	'cheese'	kkasu	'I make dwell'

Table 6: Salentino stimuli

The recording was conducted in a quiet room at the participant's house. Audio was collected using a Sennheiser head-mounted microphone. The sampling rate was 44.1 kHz. The data was manually segmented using Praat TextGrids (Boersma & Weenink, 2020) by Burroni.

2.2 Acoustic analysis The TextGrids containing segmental boundaries and the audio signals were read back in MATLAB® for acoustic measurements and analysis. We obtained three measurements that were reported to be associated to initial geminates and fortis consonants: (i) consonant duration (except target words that appear in utterance-initial position), (ii) f0 contour of the vowel following singleton and geminate onsets, and (iii) intensity contour of the following vowels.

The effects of initial geminates on consonant duration were tested with linear mixed-effect regression models. We started with a baseline model that contained random effects only. For Salentino and Dunan, each with only one participant, we included random intercepts for the manner/place of the onset segments. For Pattani Malay with 14 participants, random intercepts for each participant were also included. We then tested using ANOVA to see whether adding a fixed effect for the presence of initial geminates improved the model fit. Random slopes for the fixed effect by speaker and onset manner/place were also included in this model.

f0 values were collected by using the MATLAB pitch() function, using the normalized correlation algorithm (Atal, 1969). Sound Pressure Level-normalized intensity values were calculated by transforming the root mean squared (RMS) amplitude of the signal, extracted using the MATLAB rms() function, to dB and normalizing to the human auditory threshold using the formula $20 \times log_{10}\frac{I}{I0}$, where I is the amplitude of the signal and $I0$ represents the normalizing term for the auditory threshold of a 1000 Hz sine wave, equal to 2×10^{-5} (Huang et al., 2001). For the Pattani Malay data, f0 and intensity values were normalized using z-scores. Finally, f0 and intensity contours were time-warped to obtain a fixed length required by the growth curve analysis.

f0 and intensity contours were analyzed using third-order orthogonal polynomial growth curve analysis. The effects of initial geminates were tested with a nested model, in the same fashion as the statistical analysis of consonant duration. The baseline model only contains random effects. We have random intercepts only for the vowel quality. We then tested using ANOVA whether adding a fixed effect for the presence of initial geminates improved the model fit. If the fixed effect improved the model fit, we added the interaction of the presence of initial geminates with orthogonal linear, quadratic, and cubic time (Mirman, 2014). The random effect of individual subjects is not included, because the f0 and intensity values had been previously z-scored by participant.

2.3 Trade-off analysis Given the multidimensionality of the phonetic cues involved in the geminate contrasts for all three languages, we also decided to explore whether these cues trade off with each other. If durational differences between singletons and geminates decrease, f0 and intensity differences are expected to increase in order to preserve the contrast. f0 and intensity, on the other hand, are expected to be positively correlated with each other since both because f0 contributes to the energy present in the signal (Albert & Nicenboim, 2020) and because speakers have been shown to manipulate f0 and intensity in tandem to cue prominence (Tilsen, 2016).

We calculated the following Pearson R correlation scores for each language: (1) f0 vs. Duration, (2) Intensity vs. Duration, (3) f0 vs. Intensity. The mean f0 and intensity values across the duration of the vowel were used to provide a single value to compare with duration. For Dunan, since VOT appears to be an important cue for the geminate contrast, we calculated f0 and intensity correlations with closure duration, VOT duration, and the combined closure+VOT duration.

2.4 Machine Learning We further supplemented statistical analyses with machine learning classification. We used Support Vector Machines (SVM) to assess: (i) the robustness of singleton-initial geminate contrasts and (ii) the contribution of each acoustic cue (duration, intensity and f0). Following previous work on the topic (Burroni et al., 2020, Burroni & Maspong, to appear), we investigated several combinations of acoustic cues to determine those that best predict singleton vs geminate. We choose SVMs over Linear Discriminant Analyses, employed in previous work, because SVMs, unlike LDAs, do not make assumptions regarding the underlying data distribution. The features that we used to predict the category labels are:

(1) Duration (2 features): consonant (C) raw duration, and ratio of C to the entire syllable

(2) Intensity (10 features): Intensity contours over the following vowel (V), downsampled to 10 datapoints.

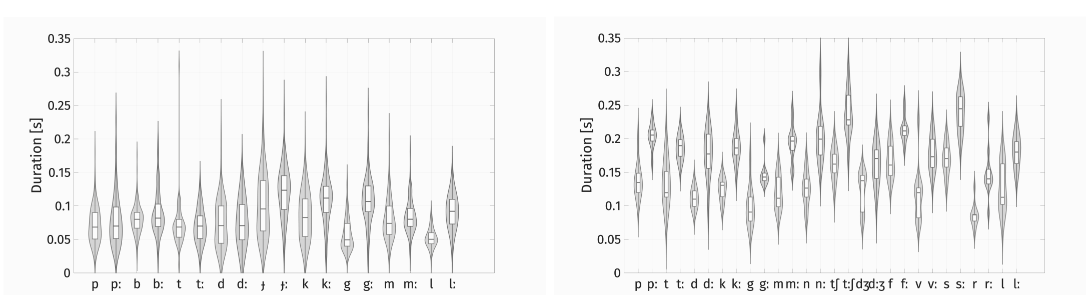
(3) f0 (10 features): f0 contours over the following vowel (V), downsampled to 10 datapoints.

For all languages, features were z-scored, as the procedure is known to improve classification. This was done separately by speaker for the Pattani Malay dataset. f0 and intensity contours were also downsampled as dimensionality reduction is recommended to improve SVM accuracy.

For each language, we trained 10 SVM models. In the results section, we report their mean accuracy and standard deviation. All SVM models were trained in MATLAB, using an L1 soft-margin minimization via quadratic programming optimization routine from the MATLAB optimization toolbox. Each of the 10 models by language was trained and tested using 10-fold cross validation. The mean and standard deviation accuracy we report are from cross-validation.

3 Results

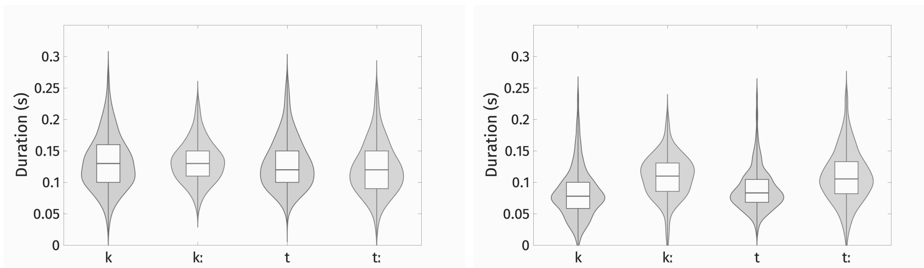
3.1 Acoustic analysis We observed a significant effect of the presence of initial geminates on duration of the onset consonant in Pattani Malay ($X^2(1) = 7.32, p = .004, CI = [.005, .028]$) and Salentino ($X^2(1) = 47.93, p < .001, CI = [.055, .067]$). In other words, initial geminates are longer than singletons in Pattani Malay, with estimated effect size of 17 ms, and in Salentino, with estimated effect size of 61 ms. Figure 2 shows the duration of initial geminates and singletons for each consonant. It is clear that the difference between initial geminates and singletons in Pattani Malay is not as robust as in Salentino. Note that the duration here includes both closure duration and burst duration for the case of stops.



(a) Pattani Malay (b) Salentino

Figure 2: Duration of initial geminates and singletons in Pattani Malay and Salentino

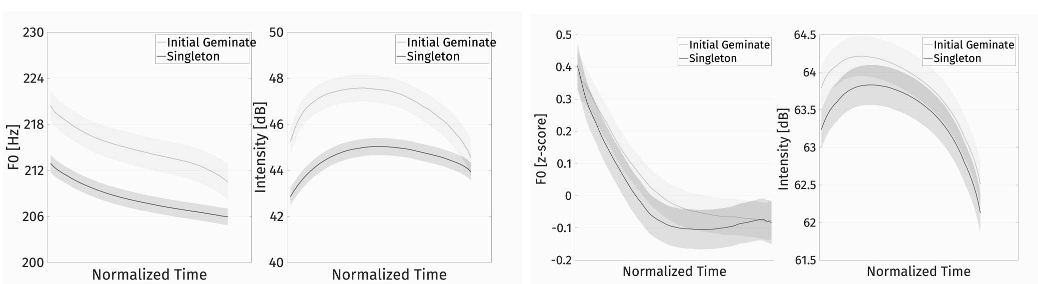
For Dunan, on the other hand, we observe significant effect of the presence of initial geminates on the duration of onset consonants (closure and burst duration combined) ($X^2(1) = 8.21, p = .004, CI = [-.012, -.002]$), but the effect is the opposite from the other two languages: singletons show longer duration than geminates (Figure 3 (left). A significant effect of the presence of initial geminates is also observed on the duration of closure duration (excluding the burst duration) ($X^2(1) = 79.27, p < .001, CI = [.020, .029]$). In this case, initial geminates display longer closure duration than singletons (Figure 3 (right). Recall that in the previous literature, Dunan initial geminates show shorter positive VOT. The different directions of the effect, then, suggests a trade-off effect between closure duration and burst duration in Dunan.



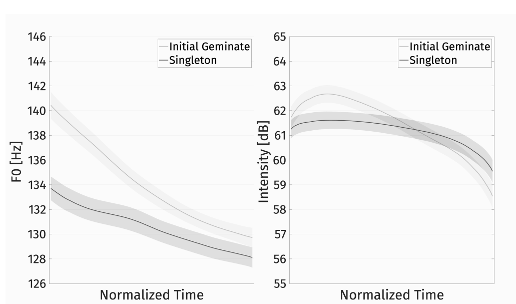
(a) closure+burst (b) closure only

Figure 3: (a) Closure+burst duration and (b) closure duration of initial geminates and singletons in Dunan

f0 and intensity show similar trends (Figure 4). We observed significant effects of initial geminates on f0 and intensity in all languages: f0 and intensity are higher on vowels following initial geminates in Dunan (f0: $t = 30.97, p < .001, Est. =$ 5 Hz; intensity: $t = 46.73, p < .001, Est. =$ 2.43 dB), Pattani Malay (f0: $t = 6.87, p < .001, Est. =$ $0.2 - 1.2$ Hz; intensity: $t = 18.77, p < .001, Est. =$ .39 dB), and Salentino (f0: $t = 68.72, p < .001, Est. =$ 3.5 Hz; intensity: $t = 11.47, p < .001, Est. =$ 2.5 dB).



(a) Dunan (b) Pattani Malay



(c) Salentino

Figure 4: f0 contours (left) and intensity contour (right) of (a) Dunan, (b) Pattani Malay, and (c) Salentino

For f0, the interactions initial geminate × orthogonal linear time and initial geminate × quadratic time are also significant for Pattani Malay (linear: $t = -4.03, p < .001$; quadratic: $t = -4.12, p < .001$) and Salentino (linear: $t = -32.67, p < .001$; quadratic: $t = 9.35, p < .001$), which shows that the slope and the curvature of the f0 of vowels following initial geminates are different from the slope and curvature of vowels following singletons. The interaction of initial geminate x cubic time is only significant for Pattani Malay ($t = 2.10, p = .04$).

As for intensity, the interaction of initial geminate × linear time and initial geminate × quadratic time is significant in all languages: Dunan (linear: $t = -11.99, p < .001$; quadratic: $t = -3.70, p < .001$), Pattani Malay (linear: $t = -3.33, p < .001$; quadratic: $t = 2.37, p = .02$), and Salentino (linear: $t = -28.73, p < .001$; quadratic: $t = -7.29, p < .001$), indicating that the slope and the curvature of intensity after initial geminates and singletons are also different. The interaction of initial geminate × cubic time is significant only in Salentino ($t = 5.59, p < .001$).

In sum, the statistical results show that duration is not the only cue in the production of initial geminates. f0 and intensity differ more robustly than generally assumed as a by-product of longer duration. Furthermore, speakers reliably sustain higher f0 and intensity long after release of initial geminates. These long lasting effects are unexpected if they are merely a phonetic effect from long closure duration.

3.2 Trade-off analysis Pearson R correlations with p-value significance levels for each language are presented in Table 7. In Dunan, the negative correlation between f0 and all three durational measures for geminates is clear. While there exists a significant negative correlation between f0 and VOT for singletons, it is weak at -.1. The significant correlations between intensity and duration for also weak (being at an absolute value of no more than .1). Pattani Malay shows weak significant positive correlations between duration and both f0 (only for singletons) and intensity, while Salentino shows no significant correlations. f0 and intensity are significantly positively correlated for all three languages (except for Salentino geminates), although the correlations for Pattani Malay are weak.

	Dunan		Pattani Malay		Salentino	
	Singletons	Geminates	Singletons	Geminates	Singletons	Geminates
Total duration ∼ f0	-.04	-.29***	.14***	0	.08	-.03
Closure duration ∼ f0	.03	-.19***	-	-	-	-
VOT ∼ f0	-.1***	-.2***	-	-	-	
Total duration ∼ Intensity	.07*	-.09	.1*	.13***	-.01	.09
Closure duration ∼ Intensity	.08*	-.07	-	-	-	-
VOT ∼ Intensity	.08	-.1*	-	-	-	-
f0 ∼ Intensity	.26***	.46***	.11***	.16***	.49***	-.04

Table 7: Pearson R correlations for trade-off analysis ($^* = p < .05, ^{**} = p < .01, ^{***} = p < .001$)

Scatterplots for the relationship between duration and f0 as well as between duration and intensity in Dunan (only total consonant duration is visualized to avoid redundancy and for consistency with the other two languages) are presented in Figure 5. The slanted distribution of the duration ∼ f0 relationship for Dunan geminates is particularly convincing evidence for a trade-off relationship. The relationship between f0 and Intensity is visualized in Figure 6. The slanted distributions for Dunan and Salentino (only singletons in the latter) are notable.

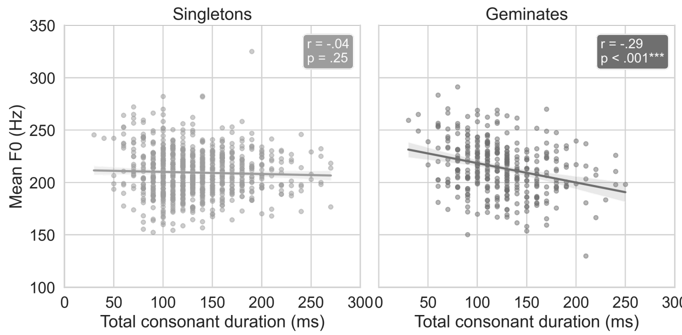
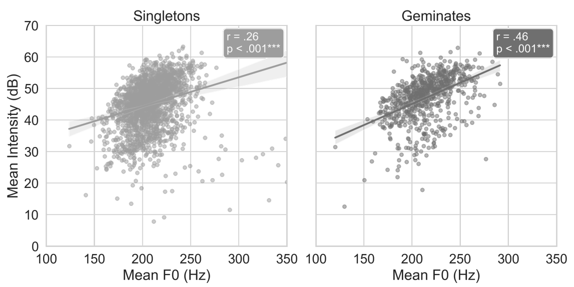
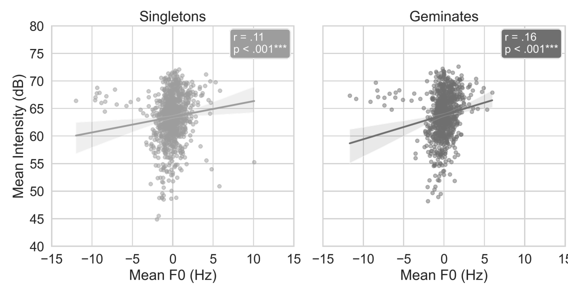


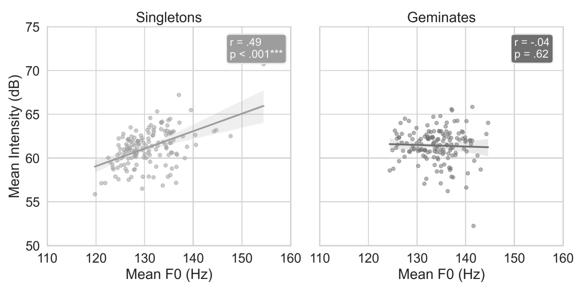
Figure 5: f0 correlation with duration in Dunan



(a) Dunan



(b) Pattani Malay



(c) Salentino

Figure 6: f0 correlation with Intensity

In sum, Dunan appears to be the only language of the three that shows a clear tradeoff between geminate duration and f0 values. As geminate durations approach those of singletons, f0 increases to compensate.

3.3 Machine Learning The SVM model for Dunan shows that the best accuracy (76.18%) is achieved by relying on all three cues: duration, intensity, and f0, Table 8. However, good accuracy can also be achieved on the basis of duration alone (71.96%).

Cues	Mean Accuracy	Std Accuracy
Duration, Intensity, f0	76.18	0.54
Duration	71.96	0.39
Intensity, f0	59.05	0.75
f0	52.70	0.48
Intensity	51.74	1.09

Table 8: SVM results for Dunan

The SVM model for Pattani also demonstrates that the best accuracy (63.7%) is achieved by relying on all three cues: duration, intensity, and f0, Table 9. However, similar accuracy can also be achieved on the basis of duration alone (61.68%).

Cues	Mean Accuracy	Std Accuracy
Duration, Intensity, f0	63.46	0.23
Duration	61.68	0.18
Intensity, f0	53.91	0.54
Intensity	53.80	0.55
f0	52.63	0.07

Table 9: SVM results for Pattani Malay

Finally, the SVM model for Salentino also demonstrates that the best accuracy (81.36%) is achieved by relying on all three cues: duration, intensity, and f0, Table 10. However, similar accuracy can also be achieved on the basis of duration alone (77.70%). Interestingly, however, an accuracy of almost 75% (74.81%) can be reached also on the basis of Intensity and f0 cues alone, without any durational information.

Cues	Mean Accuracy	Std Accuracy
Duration, Intensity, f0	81.36	0.58
Duration	77.70	0.45
Intensity, f0	74.81	0.96
f0	72.42	0.74
Intensity	71.47	0.85

Table 10: SVM results for Salentino

Taking stock of the three SVM models above, we can conclude that the ML analysis supports the statistical findings. In all three languages, the models that perform best rely on all three acoustic dimensions. This fact suggests that intensity and f0 add important complementary information for the realization of initial geminates contrasts beyond duration.

Additionally, note also that f0 and Intensity alone yield almost 75% accuracy in Salentino. The fact that 3 out of 4 geminates can be accurately labelled on the basis of f0 and intensity alone suggests that these are not merely secondary cues to duration; even in a language where initial geminates contrasts are described as fairly prototypical, duration-based ones (Romano, 2003).

4 Discussion

Our results accord with previous literature demonstrating “secondary” cues for geminate contrasts in various languages. Where we depart from other literature, however, is our argument that the term “secondary” underestimates the consistency of these cues. Our acoustic analysis demonstrates shared cues in three typologically diverse languages. While f0 and intensity differences could arguably derive from automatic, phonetic effects of laryngeal activity, if this were the case we would be expect them to be localized near the consonant release. The persistence of f0 and intensity differences over the course of the following vowel suggests that these “secondary” cues are incorporated in the phonology. Further evidence for the multidimensional usage of cues comes from our classification analysis. Even in Salentino, which appears to have “prototypical geminates”, a classifier trained on f0 and Intensity alone fares much better than chance. Furthermore, the notable improvement for all three languages of the Duration + Intensity + f0 model over one using just Duration suggests that f0 and intensity provide important information for categorization that is not redundant with duration.

The systematicity of the acoustic correlates of initial geminates in these languages suggests that a single feature, such as [+long] or two timing slots are insufficient to capture the crosslinguistic patterns. Rather, our evidence falls in favor of enriched phonological representations that include more than just duration. Such a representation places geminates in closer alignment with laryngeal contrasts such as register, which also involves a bundle of cues (Brunelle & Kirby, 2016). We echo previous work that has called attention to the multidimensionality of phonological contrasts (see, for example, Cho et al. (2019) for the insufficiency of VOT to capture crosslinguistic variation). How these co-varying cues come to characterize these contrasts may be better understood under a theory of emergent features (Mielke, 2008).

Alternatively, we may also consider phonetically richer phonological representation for initial geminates that present a unified framework of phonetics and phonology, like Articulatory Phonology (Browman & Goldstein, 1992). In this framework, the durational properties of initial geminates contrast could be captured in terms of geminate-specific closure/release/vowel couplings (Nam 2007, Burroni 2022). Additionally, f0 and intensity targets could be incorporated as dedicated "gestures" with aero-acoustic specifications. These gestures may be reminiscent of the "accentual gestures" recently proposed to model prominence (Tilsen, 2019). What remains to be determined in future articulatory work is whether the f0/intensity gestures are still coupled to initial geminates consonants or they have been fully "transphonologized" as a property of the following vowel, in a way that may lead to the emergence of new phonological contrast at the segmental or prosodic level (Abramson, 2004). This is a question that needs to be probed in future articulatory work.

Our results also tentatively provide evidence in favor of the point recently made by Ladd & Schmid (2018) that the distinction between geminates and fortis consonants may be a terminological, rather than a phonetic/phonological issue. The term fortis has been used to characterize contrasts that rely on a variety of cues, a number of which have been shown to be robust differentiators of the initial geminate contrasts we have investigated in this paper. The constellation of acoustic cues that appear to overlap between geminates and fortis consonants, as well as register and voicing distinctions crosslinguistically, suggests that the categories may be less discrete than previously assumed. If we understand fortis to be a general term encompassing various phonetic distinctions, we may understand geminates as special cases where durational cues outweigh other ones.

5 Conclusion

In this paper. we have shown that even for "prototypical initial geminates", f0 and intensity are more than secondary or "auxiliary". These cues are robust in all three languages and persist over the following vowel. The importance of f0 and intensity cues is observed in both the statistical and machine learning analyses. Furthermore, in one of the languages, Dunan, we also found evidence for compensatory trade-offs between duration and intensity/f0, suggesting that these dimensions are "synergistically" manipulated to cue initial geminate contrasts.

Our analyses suggest a greater overlap between acoustic correlates for initial geminates and fortis consonants than previously hypothesized. Accordingly, we have suggested two possible ways to conceptualize unified representations for initial geminates and fortis stops, a feature-based one and gesture-based one. In both cases, a constellation of possible phonological specifications is at play.

If we adopt this perspective, some typologically surprising facts find a straightforward explanation. Voiceless initial geminates are cross-linguistically common, even though they are hard to perceive, for obvious reasons. If, however, f0 and intensity tradeoffs with duration come into play as other important cues to initial geminates, the difficulty of perceiving duration is no longer as issue, because additional information is present in the acoustic signal throughout the following vowel. Moreover, if initial geminates and fortis stops are not distinct, it is also not surprising that, to our knowledge, no language has been reported has having both initial geminates and fortis stops.

References

Abramson, Arthur S (2004). Toward prosodic contrast: Suai and Pattani Malay. International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages.

Ahn, Sang-cheol & Gregory K. Iverson (2004). Dimensions in Korean Laryngeal Phonology. Journal of East Asian Linguistics 13, 345–379.

Al-Tamimi, Jalal & Ghada Khattab (2015). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: the case of fricative consonants. The Journal of the Acoustical Society of America 138:1, 344–360.

Albert, Aviad & Bruno Nicenboim (2020). The double life of language systems: Modelling sonority with complementary symbol and signal based models. PsyArXiv.

Arvaniti, Amalia & Georgios Tserdanelis (2000). On the phonetics of geminates: evidence from Cypriot Greek. Proceedings of the 6th International Conference on Spoken Language Processing, Beijing, China, 559–562.

Atal, B. S. (1969). Automatic Speaker Recognition Based on Pitch Contours. Journal of the Acoustical Society of America 52, 1687–1697.

Avery, Peter & William J. Idsardi (2001). Laryngeal dimensions, completion and enhancement. Hall, T.A. (ed.), Distinctive feature theory, De Gruyter Mouton, Boston, 41–70.

Bentley, John R. (2008). A Linguistic History of the Forgotten Islands: A Reconstruction of the Proto-Language of the Southern Ryukyus. Global Oriental, Folkestone.

Bertinetto, PierMarco & Michele Loporcaro (1999). Geminate distintive in posizione iniziale: Uno studio percettivo sul dialetto di Altamura (Bari. Annali delle Scuola Normale Superiore di Pisa 4, 304–322.

Boersma, Paul & David Weenink (2020). Praat: doing phonetics by computer. URL http://www.praat.org/.

Browman, Catherine P & Louis Goldstein (1992). Articulatory phonology: An overview. Phonetica 49:3-4, 155–180.

Brunelle, Marc & James Kirby (2016). Tone and phonation in Southeast Asian languages. Language and Linguistics Compass 10:4, 191–207.

Burroni, Francesco (2022). A split-gesture, competitive, coupled oscillator model of syllable structure predicts the emergence of edge gemination and degemination. Proceedings of the Society for Computation in Linguistics 5:1, 11–22.

Burroni, Francesco & Sireemas Maspong (to appear). Re-examining Initial Geminates: Typology, Evolutionary Phonology, and Phonetics. Historical Linguistics 2019, Canberra, Australia.

Burroni, Francesco, Sireemas Maspong, Pittayawat Pittayaporn & Pimthip Kochaiyaphum (2020). A new look at Pattani Malay Initial Geminates: a statistical and machine learning approach. Pacific Asia Conference on Language, Information and Computation 34.

Catalogue of Endangered Languages (2022). Dunan. University of Hawai'i at Mānoa. https://www.endangeredlanguages.com/lang/3170 (accessed: 01.31.2022).

Cho, Taehong & Patricia A. Keating (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. Journal of Phonetics 29, 155–190.

Cho, Taehong, Sun-Ah Jun & Peter Ladefoged (2002). Acoustic and aerodynamic correlates of Korean stops and fricatives. Journal of phonetics 30:2, 193–228.

Cho, Taehong, Douglas H Whalen & Gerard Docherty (2019). Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. Journal of Phonetics 72, 52–65.

Collins, James T. (1989). Malay dialect research in Malaysia: The issue of perspective. Bijdragen tot de Taal-, Land- en Volkenkunde [Journal of the Humanities and Social Sciences of Southeast Asia and Oceania] 145:2, 235–264, URL http://www.jstor.org/stable/27864031. Publisher: [Brill, KITLV, Royal Netherlands Institute of Southeast Asian and Caribbean Studies].

DiCanio, Christian T (2012). The phonetics of fortis and lenis consonants in Itunyoso Trique. International Journal of American Linguistics 78:2, 239–272.

Ghosh, Aanusha (2015). Acoustic correlates of voicing and gemination in Bangla. Proceedings of the 12th International Conference on Natural Language Processing, 413–418.

Halle, Morris & Kenneth N. Stevens (1971). A note on laryngeal features, QPR Research Laboratory of Electronics. Quarterly Progress Report of the Research Laboratory of Electronics, MIT, vol. 101, 198–213.

Han, Jeong-Im (1996). The phonetics and phonology of "tense" and "plain' consonants in Korean. Ph.D. dissertation, Cornell University.

Huang, Xuedong, Alex Acero & Hsiao-Wuen Hon (2001). Spoken language processing: a guide to theory, algorithm, and system development. Prentice Hall PTR, Upper Saddle River, NJ, URL https://hdl.handle.net/2027/mdp.39015051284142. Http://newcatalog.library.cornell.edu/catalog/4116648.

Iverson, Gregory K. (1983). On glottal width features. Lingua 60, 331–339.

Jaeger, Jeri J (1983). The fortis/lenis question: evidence from Zapotec and Jawoñ. Journal of Phonetics 11:2, 177–189.

Jakobson, Roman, C. Gunnar M. Fant & Morris Halle (1952). Preliminaries to speech analysis: The distinctive features and their correlates. Acoustics Laboratory, Massachusetts Institute of Technology.

Kang, Yoonjung, Jessamyn Schertz & Sungwoo Han (forthcoming). The phonology and phonetics of Korean stop laryngeal contrasts. The Cambridge Handbook of Korean Linguistics, Cambridge University Press.

Kim, Chin-Wu (1965). On the autonomy of the tensity feature in stop classification (with special reference to Korean stops). Word 21, 330–359.

Kohler, Klaus J. & W. A. van Dommelen (1987). The effects of voice quality on the perception of lenis/fortis stops. Journal of Phonetics 15:4, 365–381.

Kraehenmann, Astrid (2011). Initial geminates. The Blackwell Companion to Phonology, Wiley-Blackwell, Oxford, vol. 2, 1124–1146.

Ladd, D. Robert & Stephan Schmid (2018). Obstruent voicing effects on F0, but without voicing: Phonetic correlates of Swiss German lenis, fortis, and aspirated stops. Journal of Phonetics 71, 229–248.

Lombardi, Linda (1994). Laryngeal features and laryngeal neutralization. Garland, New York.

Loporcaro, Michele (1997). L'origine del raddoppiamento fonosintattico: saggio di fonologia diacronica romanza, vol. 115. Verlag.

McAuliffe, Michael, Michaela Socolof, Sarah Mihuc, Michael Wagner & Morgan Sonderegger (2017). Montreal Forced Aligner. URL http://montrealcorpustools.github.io/Montreal-Forced-Aligner/.

Mielke, Jeff (2008). The emergence of distinctive features. Oxford University Press.

Mirman, Daniel (2014). Growth Curve Analysis and Visualization Using R. Chapman & Hall/CRC the R series, Boca Raton.

Muller, Jennifer (2001). The phonology and phonetics of word-initial geminates. Thesis.

Nam, Hosung (2007). A gestural coupling model of syllable structure. Ph.D. thesis, Yale University.

Romano, Antonio (2003). Indici acustici di alcune geminate iniziali salentine. La coarticolazione, ETS, Pisa, 233–241.

Romano, Antonio (2015). Proprietà fonetiche segmentali e soprasegmentali delle lingue parlate nel Salento. Idomeneo 19, 157–185.

Tilsen, Sam (2016). A shared control parameter for f0 and intensity. Proceedings of Speech Prosody, vol. 2016, 1066–1070.

Tilsen, Sam (2019). Space and time in models of speech rhythm. Annals of the New York Academy of Sciences 1453:1, 47–66.

Trubetzkoy, Nikolai S. (1969). Principles of phonology. University of California Press, Berkeley.

Uthai, Ruslan (2011). Keistimewaan dialek Malayu Patani. Penerbit Universiti Kebangsaan Malaysia, Bangi.

Wittenburg, Peter, Hennie Brugman, Albert Russel, Alex Klassmann & Han Sloetjes (2006). ELAN: A professional framework for multimodality research. 5th international conference on language resources and evaluation (LREC 2006), 1556–1559.

Yamada, Masahiro, Thomas Pellard & Michinori Shimoji (2015). Dunan grammar (Yonaguni Ryukyuan). Heinrich, Patrick, Shinsho Miyara & Michinori Shimoji (eds.), Handbook of the Ryukyuan languages: History, structure, and use, De Gruyter Mouton, 449–478.