Speech Prosody 2024
2-5 July 2024, Leiden, The Netherlands

C$F0$ effect and articulatory strength of geminate consonants

Sireemas Maspong, Francesco Burroni, James Kirby

Spoken Language Processing Group, Institute for Phonetics and Speech Processing, LMU München
{s.maspong|francesco.burroni|jkirby}@phonetik.uni-muenchen.de

Abstract

This study explores how fundamental frequency ($F0$) and articulatory strength are related in Italian geminate consonants. Consonant-intrinsic $F0$ (C$F0$) effects are examined with a focus on the hypothesis that geminates exhibit such effects as a consequence of their inherent "tense" articulation, manifested as higher $F0$ and a more constricted articulatory target. Simultaneous articulatory and acoustic data were collected from 10 native Central and Southern Italian speakers pronouncing six disyllabic nonce words ([ip(:)a, ib(:)a, im(:)a]) within a carrier sentence varying in speech rates. $F0$ values were extracted at 10 ms intervals before and after consonantal closure, while articulatory data, including Minimum Lip Aperture (LA) and Maximum Jaw Height (JH), were recorded using an AG501 Carsten EMA. Linear mixed-effect regressions were fit to the data. The findings reveal that geminates exhibit higher post-closure $F0$, lower LA, and higher JH compared to their singleton counterparts, supporting the hypothesis that geminate consonants possess "tense" properties. Additionally, weak positive correlations were observed between post-closure $F0$ and LA.

Index Terms: C$F0$, articulatory, geminates, Italian, microprosody

1. Introduction

1.1. C$F0$ and geminates

Numerous studies have demonstrated that fundamental frequency ($F0$) can be affected by intrinsic consonantal properties. Specifically, it has been observed that $F0$ following voiceless obstruents tends to be higher than $F0$ following voiced obstruents, as noted in [1, 2]. This effect of onset voicing on the $F0$ of subsequent vowels has been observed in both non-tonal [1, 3, 4, 5, 6, 7] and tonal languages [8, 3, 9]. For instance, in Italian, a "true voicing" language featuring a contrast between fully voiced consonants and voiceless unaspirated consonants, voiceless obstruent onsets /p, f/ exhibit higher $F0$ on the subsequent vowel compared to voiced obstruents /b, v/ and nasal /m/ [5].

The observed consonant-intrinsic $F0$ (hereafter C$F0$) of voiced and voiceless consonants is commonly attributed to articulatory mechanisms favoring or hindering voicing. These mechanisms include the adjustment of vocal fold tension associated with voiceless obstruents [10, 11], as well as larynx lowering and pharyngeal expansion in voiced consonants [12, 13, 14, 15].

$F0$ has also been identified as a non-durational acoustic cue of geminate consonants in various languages, both in word-medial position [16, 17, 18] and in word-initial position [19, 20, 21]. Specifically, $F0$ is higher on vowels adjacent to geminate consonants compared to those adjacent to singletons. It is important to note that $F0$ effects vary across languages. In some languages, the effect is observed on vowels preceding geminates (e.g., [16, 17, 18]), while in others, it is observed on vowels following geminates [19, 20, 21]. Conversely, some languages may not exhibit any $F0$ effect on the adjacent vowel in the presence of geminates [22].

Several hypotheses have been proposed to explain this phenomenon. One hypothesis posits that geminate production involves more than just a longer closure duration to encompass multiple articulators [21]. This line of thinking unifies the singleton-geminate contrast and the lenis-fortis contrast. In essence, the higher $F0$, along with other acoustic and articulatory correlates, observed in geminates can be attributed to their "tense" articulatory strength. Similarly, the elevated articulatory pressure and larger constriction area observed in geminates [23, 24] are also considered the results of "tense" production. A question that arises is whether the "tense" articulatory strength exhibited by geminates, characterized by multiple articulatory actions such as higher $F0$, increased articulatory pressure, and a larger constriction area, are interconnected. In essence, the question is whether the various articulatory actions observed are a consequence of the coordination of articulators working together to achieve a unified "task" in geminate production. This is similar to the production of bilabial consonants that involve articulatory actions of the jaw, the lower lip, and the upper lip [25].

An additional interesting aspect worth exploring is the potential interplay between gemination and voicing, given their potentially conflicting effects on the $F0$ of surrounding vowels. [22] explored this issue, suggesting that the production of voiced geminates involves an active strategy to sustain voicing throughout the prolonged closure of geminates. This prompts the question of whether the $F0$ perturbation effect differs in voiced singleton-geminate pairs compared to voiceless singleton-geminate pairs.

In the context of previous work on Kelantan Malay, it has been reported that $F0$ is consistently higher following geminates than following singletons across voiceless obstruents, voiced obstruents, and sonorants. However, the effect is more pronounced for the voiceless pairs [26]. A parallel observation extends to intraoral pressure during the production of singleton and geminates with different voicing categories in Italian: geminates exhibit higher intraoral pressure than singletons, with the effect being more pronounced for voiceless singleton-geminate pairs compared to their voiced counterparts [27].

1.2. Research Questions, hypotheses, and predictions

In this paper, we tackle the following three research questions.

1. How does gemination interact with C$F0$ in Italian?
2. Do Italian geminates exhibit tenser articulation than singletons?
3. What is the relationship, if any, between geminates effect on $F0$ and their articulation?

For 1), we can entertain the following three hypotheses and associated predictions. A) Gemination does not interact with

 10.21437/SpeechProsody.2024-68

C$F0$ and we will thus observe no effect of gemination on C$F0$. B) Gemination magnifies C$F0$ effects, increasing the already higher $F0$ after voiceless stops and decreasing the already lower $F0$ observed after voiced stops. C) Gemination raises $F0$ across the board regardless of voicing.

For 2), we can entertain the following two hypotheses and associated predictions. A) Italian geminates reflect a tenser articulation, hence they will display more constricted target. B) Italian geminates are simply longer consonants with articulation comparable to that of singletons, thus, articulatory targets will be comparable for the two categories.

For 3), we can entertain the following three hypotheses and associated predictions. A) There is no relationship between articulation of geminate and thier $F0$ effects, no correlations will therefore be observed. B) There is a "synergistic" relationship (or a trading relationship) between articulation and $F0$ effects. In the presence of weaker articulation, $F0$ will supply as a cue to gemination. Hence, a negative correlation should be observed between the two, indicating that they compensate for each other. C) There is a "non-synergistic" relationship between articulation and $F0$ effects. $F0$ effects are purely aerodynamically related to articulation; hence a positive correlation should be observed between the two, indicating that they proceed in tandem, as a consequence of "tense" articulation.

2. Methods

2.1. Data collection

We collected simultaneous audio and Electromagnetic Articulography (3D Carsten AG501 EMA) data from 10 native Italian speakers, speaking Central and Southern varieties (south of the Rimini-La Spezia line) where geminates are uncontroversially realized (e.g., [28]). Participants were instructed to produced six disyllabic nonce VCV words containing all singleton and geminate Italian bilabial consonants: [ipa, ipːa, iba, ibːa, ima, imːa]. We refer to /i/ as V1 and /a/ as V2. A high to low vowel transition was chosen to maximize tongue vertical movement and facilitate landmarking. Bilabial consonants were chosen to avoid competing demands on tongue movement from consonants and vowels. Additionally, bilabial consonants are contrastive for voicing with voiceless stops, voiced stops, and nasals. Target words were embedded in a carrier sentence [dika __ due vɔlte] "Please say __ twice". Trials were produced at 5 rates "very slow", "slow", "normal", "fast", "very fast", to introduce variability in duration. Each word was repeated 12 times at each rate in randomized order. We thus collected approximately 360 tokens per speaker. Following the exclusion of trials with speech errors, we arrived at a dataset comprising 3,593 tokens for subsequent analyses.

2.2. Data processing

Audio recordings were segmented using the MAUS forced aligner [29]. The segmentation was hand-corrected when necessary.

$F0$ was obtained using the Summation of Residual Harmonics (SRH) algorithm [30], which was implemented using the *pitch()* function in MATLAB. $F0$ tracking employed a window size of 42 ms with a 1 ms time step. The $F0$ extraction process consisted of a two-step procedure for each participant, following the approach proposed by Hirst [31]. Initially, a broad $F0$ range of 75-400 Hz was used for the first pass. The first ($Q1$) and the third quartiles ($Q3$) of $F0$ values for each participants were calculated. In the second round of $F0$ tracking, the $F0$ floor for each participant was set to $Q1 \times 0.75$, while the $F0$ ceiling was set to $Q3 \times 1.5$. Following z-scoring by participants, two $F0$ means were computed for subsequent statistical analysis: (i) the mean of $F0$ values during the last 10 windows (10 ms) of V1 preceding the consonantal closure, and (ii) the mean of $F0$ values during the first 10 windows of V2 following the consonantal closure. Note that the $F0$ values from windows spanning medial consonants were excluded from the analysis to avoid artifacts caused by tracking errors during silence or burst periods.

This paper examined two articulatory trajectories: Lip Aperture (LA) and Jaw Height (JH). LA was defined as the 3D Euclidean distance between the Lower Lip and Upper Lip sensors. Conversely, JH represents the height of the Jaw sensor. For the analyses, we extracted Minimum LA and Maximum JH during the bilabial closure of the target consonants ([p(ː), b(ː), m(ː)]). A small Minimum LA and an elevated JH indicate a more constricted closure of the bilabial consonants.

2.3. Data analyses

We initially conducted statistical analyses to investigate potential differences between geminates and singletons in terms of $F0$ values for V1 and V2, as well as constriction degrees represented by Minimum LA and Maximum JH. All dependent variables were entered in linear mixed effect regression models. The fixed effects were geminate status (categorical, with reference as singleton), consonant voicing (categorical, with reference as nasal consonant), and their interaction. A random intercept for subject was also included. Likelihood ratio tests were performed to compare all conceivable nested models and determine the best fitting models for each dependent variable.

Moreover, we explored whether $F0$ values of vowels preceding and following geminate consonants could be predicted based on constriction degrees. $F0$ values of V1 and V2 were employed as dependent variables in linear mixed-effect regression models. The fixed effects were constriction degree (Minimum LA or Maximum JH; continuous), consonant voicing, and their interaction. A random intercept for subject was included in this model as well. Likelihood ratio tests were performed to compare these models against their respective null models, where consonant voicing served as a fixed effect and subject as a random intercept. This analysis aimed to determine whether incorporating a fixed effect of constriction degree improves model fit.

3. Results

3.1. Fundamental frequency ($F0$)

3.1.1. Vowel preceding target consonant (V1)

The likelihood ratio tests identified the model with only consonant voicing as the best fitting model. This outcome indicates distinctions in V1 $F0$ values based on the voicing status of the following consonants. Specifically, V1 $F0$ values are higher preceding /p(ː)/ compared to those preceding /m(ː)/ ($t(3335) = 4.62, p < 0.001$), with an effect size of 0.12 z-scores. Similarly, V1 $F0$ values are also higher preceding /b(ː)/ compared to those preceding /m(ː)/ ($t(3335) = 2.63, p = 0.009$), with an effect size of 0.07 z-scores. Note that within the same model with /p(ː)/ as a reference, we also observed differences in V1 $F0$ values between /p(ː)/ and /b(ː)/ with /b(ː)/ showing lower V1 $F0$ ($t(3335) = -2.01, p = 0.045$), with an effect size of -0.06 z-scores. The absence of geminate status in the best fitting

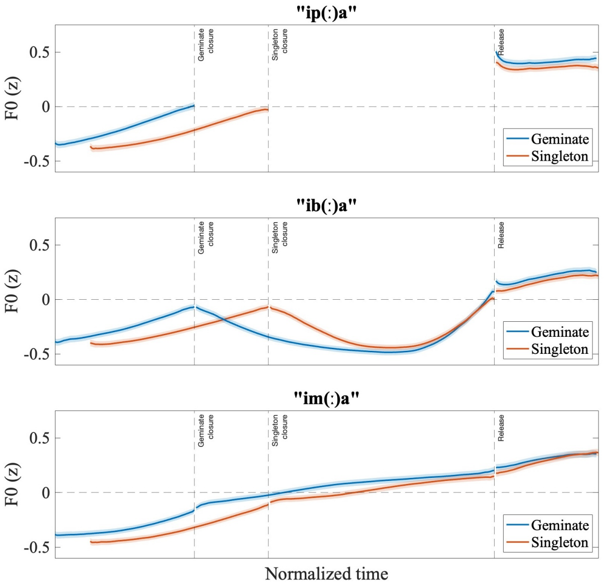

Figure 1: *F0 trajectories of target words with singleton and geminate consonants anchored at the release of the medial consonant. Normalized time based on the mean duration of each segment and stimuli type.*

model implies that V1 $F0$ values do not differ between words with singletons and words with geminates.

The $F0$ trajectories of V1 are illustrated in Figure 1 (pre-closure).

3.1.2. Vowel following target consonant (V2)

The likelihood ratio tests identified the model with geminate status and consonant voicing (excluding their interaction) as the best fitting model. This outcome suggests differences in V2 $F0$ values following consonantal closure between singletons and geminates. Specifically, words containing geminates exhibit higher $F0$ values during the initial 10 ms of V2 compared to words with singletons ($t(3286) = 2.79, p = 0.005$), with an effect size of approximately 0.07 z-scores. Additionally, we observed that V2 $F0$ values are higher following /p(:)/ compared to those following /m(:)/ ($t(3286) = 7.37, p < 0.001$), showing an effect size of 0.22 z-scores. Conversely, V2 $F0$ values are lower following /b(:)/ compared to those following /m(:)/ ($t(3286) = -2.71, p = 0.007$), with an effect size of -0.08 z-scores. Note that within the same model with /p(:)/ as a reference, we also observed that V2 $F0$ is lower following /b(:)/ compared to following /p(:)/ ($t(3286) = -10.04, p < 0.001$) with an effect size of -0.3 z-scores. The absence of interaction between geminate status and consonant voicing indicates that, while V2 $F0$ values differ across different voicings of the preceding consonant, the effect of geminate status is stable irrespective of voicing.

The $F0$ trajectories of V2 are illustrated in Figure 1 (post-closure).

3.2. Constriction degrees

3.2.1. Lip Aperture (LA)

The likelihood ratio tests identified the model with geminate status and consonant voicing (excluding their interaction) as the

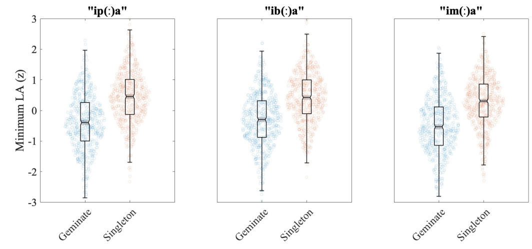

Figure 2: *Minimum Lip Aperture during consonantal closure.*

best fitting model. This result suggests differences in Minimum LA during consonantal closure between singletons and geminates. Specifically, geminates exhibit a lower Minimum LA (indicating more constriction) than singletons ($t(3589) = -25.14, p < 0.001$), with an effect size of approximately -0.8 millimeters. Furthermore, we observed that Minimum LA is higher (indicating less constriction) for /p(:)/ compared /m(:)/ ($t(3589) = 2.66, p = 0.008$), showing an effect size of 0.1 millimeter. Similarly, Minimum LA is higher for /b(:)/ compared to /m(:)/ ($t(3589) = 3.96, p < 0.001$), with an effect size of 0.15 millimeters. It is important to note that within the same model with /p(:)/ as a reference, we did not observe differences in Minimum LA between /p(:)/ and /b(:)/ ($t(3589) = 1.30, p = 0.19$). The absence of interaction between geminate status and consonant voicing indicates that, although Minimum LA varies across different voicings of the preceding consonant, the effect of geminate status remains consistent regardless of voicing.

Figure 2 illustrates the Minimum LA of geminates and singletons.

3.2.2. Jaw Height (JH)

The likelihood ratio tests identified the model with geminate status and consonant voicing (excluding their interaction) as the best fitting model. This result suggests differences in Maximum JH during consonantal closure between singletons and geminates. Specifically, geminates exhibit a higher Maximum JH (indicating more constriction) than singletons ($t(3589) = 4.05, p < 0.001$), with an effect size of approximately 0.13 millimeters. Furthermore, we observed that Maximum JH is lower (indicating less constriction) for /p(:)/ compared to /m(:)/ ($t(3589) = -8.12, p < 0.001$), showing an effect size of -0.32 millimeter. Similarly, Maximum JH is lower for /b(:)/ compared to /m(:)/ ($t(3589) = -2.34, p = 0.02$), with an effect size of -0.09 millimeters. It is important to note that within the same model with /p(:)/ as a reference, we observed that Maximum JH is higher (indicating more constriction) for /b(:)/ compared to /p(:)/ ($t(3589) = 5.78, p < 0.001$), showing an effect size of 0.23 millimeters. The absence of interaction between geminate status and consonant voicing indicates that, although Maximum JH varies across different voicings of the preceding consonants, the effect of geminate status remains consistent regardless of voicing.

Figure 3 illustrates the Minimum LA of geminates and singletons.

3.3. Relationship between $F0$ and constriction degrees

Given that only V2 $F0$ values exhibit significant differences between words with geminates and singletons, this section will exclusively address the relationship between V2 $F0$ and constriction degree measures of geminate consonants.

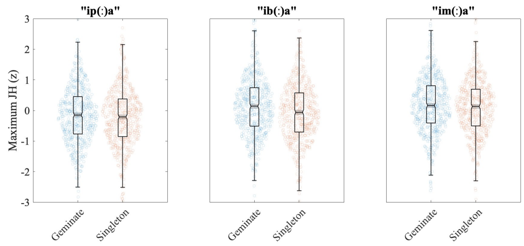

Figure 3: *Maximum Jaw Height during consonantal closure.*

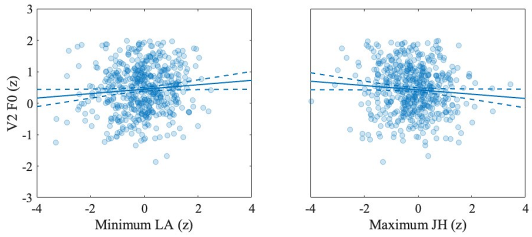

Figure 4: *Relationship between F0 and Minimum LA (left) and the relationship between F0 and Maximum JH (right) of /pː/. Solid lines represent regression lines and dashed lines represent 95% confidence intervals.*

3.3.1. V2 F0 and LA

The likelihood ratio tests identified the model with Minimum LA, consonant voicing, and their interaction as the best fitting model. This result reveals that V2 $F0$ values can be predicted from Minimum LA. In other words, there is a relationship between V2 $F0$ values and Minimum LA. However, the relationship is significant only for words with stop geminates /pː,bː/. This is suggested by the absence of significance for Minimum LA in the model with nasal geminate /mː/ as reference ($t(1625) = 0.43, p = 0.66$), while the interactions between Minimum LA and /pː/ ($t(1625) = 3.23, p = 0.001$) and /bː/ ($t(1625) = 2.21, p = 0.02$) are significant. In the same model with /p(ː)/ as a reference, the effect size is approximately 0.16 z-scores indicating that an increase of one z-scored Minimum LA correspond to an increase of 0.16 z-scored V2 $F0$. Similarly, in the same model with /b(ː)/ as a reference, the effect size is approximately 0.11 z-scores indicating that an increase of one z-scored Minimum LA correspond to an increase of 0.11 z-scored V2 $F0$. In sum, for /pː,bː/, an increase in Minimum LA (less constriction) corresponds to an increase in V2 $F0$, as shown in Figure 4 (left).

3.3.2. V2 F0 and JH

The likelihood ratio test reveals that adding the Maximum JH does not improve the model fit ($\chi^2(1) = 2.17, p = 0.14$). This result suggests that V2 $F0$ values cannot be predicted from Maximum JH. In other words, there is no significant relationship between V2 $F0$ values and Maximum JH, as shown in Figure 4 (right).

4. Discussions

Before delving into the answers to our research questions, it is noteworthy that our analyses reveal a consistent pattern: vowels following geminates in Italian exhibit higher $F0$ compared to those following singletons. This aligns with the behavior observed for geminates in various languages. Additionally, our analyses confirm the existence of C$F0$ effects of consonant voicing on vowels, both preceding and following the target consonants based solely on voicing.

In addressing the first research question, our findings are compatible with an absence of voicing-specific interaction with gemination concerning $F0$ effects. This outcome suggests that the $F0$ effect of gemination remains consistent across voicing categories. In simpler terms, gemination neither amplifies the voicing effect nor vice versa. Instead, geminate consonants elevate $F0$ uniformly across all voicing categories.

Beyond the elevated $F0$, our analyses reveals that geminate consonants are not characterized only by longer duration; they also exhibit a more constricted closure, as indicated by both LA and JH measures. The raised $F0$ and the more constricted articulation collectively suggests "tense" articulatory properties associated with geminate consonants.

Lastly, our analyses reveal a "synergistic" relationship between $F0$ and LA in geminate production. This pattern implies the coordination of multiple articulators working together, possibly not bio-mechanically linked, to achieve a singular "task" in geminate production. Specifically, when the closure is less constricted, $F0$ appears to compensate to accomplish geminate production. This compensatory behavior draws parallels to observations in the production of e.g., bilabial consonants, where restricted jaw movement prompts adjustments in the movements of the upper and lower lips to fulfill the task [32]. However, a note of caution is in place for this last analyses, as we do not observe a synergistic relationship between $F0$ and JH, which may be expected to be stronger than the one between $F0$ and LA. The absence of a relationship between $F0$ and JH in our data may be attributed to the limited range of observed values in Maximum JH. This limitation arises from the inherently more restricted vertical movement of the jaw compared to the upper and lower lips. Alternatively, it is also possible that the $F0$ is higher in the presence of less constricted targets (where lips are less compressed and reduce less the closure volume) simply as a consequence of an aerodynamic effect whereby less constricted target allow a higher transglottal airflow and thus interfere less with vocal fold vibration. This would also be more in line with the lack of effects on sonorants observed in the data.

5. Conclusions

In this paper, we have presented an investigation of the relationship between gemination and C$F0$ effects in Italian. We have found that bilabial geminates increase $F0$ on the following vowel. Additionally, we also found that this raising in $F0$ is accompanied by more constricted targets. Finally, we also found that correlation of $F0$ and LA seem to point to an articulatory trade-off relationship of the synergestic type; at least for some geminates and in some articulatory dimensions. But more work on this aspect is needed to rule out an aerodynamic effect. We have interpreted our findings in light of a "tenser" articulation of geminates and a synergy of various laryngeal and oral articulators underlying the production of geminate consonants. Future work probing more places of articulation, gemination in other languages, and laryngeal gestures is needed to assess the generality of our findings.

6. References

[1] A. S. House and G. Fairbanks, "The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels," *Journal of the Acoustical Society of America*, vol. 25, pp. 105–113, 1953.

[2] I. Lehiste and G. E. Peterson, "Some Basic Considerations in the Analysis of Intonation," *Journal of the Acoustical Society of America*, vol. 33, no. 4, pp. 419–425, 1961.

[3] J.-M. Hombert, "Towards a theory of tonogenesis: an empirical, physiologically and perceptually based account of the development of tonal contrasts in languages," Ph.D. dissertation, University of California, Berkeley, 1975.

[4] J. P. Kirby and D. R. Ladd, "Stop voicing and F0 perturbations: Evidence from French and Italian," *Proceedings of the Eighteenth International Congress of Phonetic Sciences*, 2015.

[5] ——, "Effects of obstruent voicing on vowel F0: Evidence from "true voicing" languages," *The Journal of the Acoustical Society of America*, vol. 140, pp. 2400–2411, 2016.

[6] S. Seyfarth and M. Garellek, "Plosive voicing acoustics and voice quality in Yerevan Armenian," *Journal of Phonetics*, vol. 71, pp. 425–450, 2018.

[7] J. Schertz and S. Khan, "Acoustic cues in production and perception of the four-way stop laryngeal contrast in Hindi and Urdu," *Journal of Phonetics*, vol. 81, p. 100979, Jul. 2020.

[8] J. Gandour, "Consonant types and tone in Siamese," *Journal of phonetics*, vol. 2, pp. 337–350, 1974.

[9] J. P. Kirby, "Onset pitch perturbations and the cross-linguistic implementation of voicing: Evidence from tonal and non-tonal languages," *Journal of Phonetics*, vol. 71, pp. 326–354, 2018.

[10] M. Halle and K. N. Stevens, "A Note on Laryngeal Features," in *MIT Research Laboratory of Electronics Quarterly Progress Report*, 1971, vol. 101, pp. 198–213.

[11] A. Löfqvist, T. Baer, and N. S. McGarr, "The cricothyroid muscle in voicing control," *Journal of Acoustical Society of America*, vol. 85, no. 3, pp. 1314–1321, 1989.

[12] F. Bell-Berti, "Contral of pharyngeal cavity size for English voiced and voiceless stop," *Journal of Acoustical Society of America*, vol. 57, no. 2, pp. 456–461, 1975.

[13] J. R. Westbury, "Enlargement of the supraglottal cavity and its relation to stop consonant voicing," *Journal of Acoustical Society of America*, vol. 73, no. 4, pp. 1322–1336, 1983.

[14] J. R. Westbury and P. A. Keating, "On the naturalness of stop consonant voicing," *Journal of Linguistics*, vol. 22, no. 1, pp. 145–166, 1986.

[15] S. Ahn, "The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazillian Portuguese," *Journal of Phonetics*, vol. 71, pp. 451–467, 2018.

[16] K. Idemaru and S. G. Guion, "Acoustic covariants of length contrast in Japanese stops," *Journal of the International Phonetic Association*, vol. 38, no. 2, pp. 167–186, Aug. 2008.

[17] M. G. Di Benedetto and L. De Nardis, "Consonant gemination in Italian: The affricate and fricative case," *Speech Communication*, vol. 134, pp. 86–108, Nov. 2021.

[18] ——, "Consonant gemination in Italian: The nasal and liquid case," *Speech Communication*, vol. 133, pp. 62–80, Oct. 2021.

[19] A. S. Abramson, "The complex acoustic output of a single articulatory gesture: Pattani Malay word-initial consonant length," in *Papers from the fourth annual meeting of the Southeast Asian Linguistics Society*. Tempe: Program for Southeast Asian Studies, Arizona State University, 1998, pp. 1–20.

[20] R. Ridouane, "Gemination in Tashlhiyt Berber: an acoustic and articulatory study," *Journal of the International Phonetic Association*, vol. 37, no. 2, pp. 119–142, 2007.

[21] F. Burroni, R. Lau-Preechathammarach, and S. Maspong, "Unifying Initial Geminates and Fortis Consonants via Laryngeal Specification: Three case studies from Dunan, Pattani Malay, and Salentino," in *Proceedings of the Annual Meetings on Phonology*, vol. 9, 2022.

[22] J. Al-Tamimi and G. Khattab, "Multiple cues for the singleton–geminate contrast in Lebanese Arabic: Acoustic investigation of stops and fricatives," *Hong Kong*, 2011.

[23] A. Kraehenmann and A. Lahiri, "Duration differences in the articulation and acoustics of Swiss German word-initial geminate and singleton stops," *Journal of the Acoustical Society of America*, vol. 123, pp. 4446–4455, 2008.

[24] W. Baik, "On tensity of Korean fricatives (Electropalatographic study)," *Korean Journal of Speech Sciences*, vol. 4, pp. 135–145, 1998.

[25] E. L. Saltzman and K. G. Munhall, "A dynamical approach to gestural patterning in speech production," *Ecological Psychology*, vol. 1, no. 4, pp. 333–382, 1989.

[26] M. H. Hamzah, J. Hajek, and J. Fletcher, "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*, vol. 50, no. 1, pp. 23–60, 2020.

[27] A. Butcher, "'Fortis/lenis' revisited one more time: the aerodynamics of some oral stop contrasts in three continents," *Clinical Linguistics & Phonetics*, vol. 18, no. 6-8, pp. 547–557, 2004.

[28] P. Mairano and V. D. Iacovo, "Gemination in Northern versus Central and Southern Varieties of Italian: A Corpus-based Investigation," *Language and Speech*, vol. 63, no. 3, pp. 608–634, 2020.

[29] F. Schiel, "Automatic Phonetic Transcription of Non-Prompted Speech," in *Proceedings of the ICPhS*, 1999, pp. 607–610.

[30] T. Drugman and A. Alwan, "Joint Robust Voicing Detection and Pitch Estimation Based on Residual Harmonics," in *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 2011, pp. 1973–1976.

[31] D. Hirst, "The analysis by synthesis of speech melody: from data to models," *Journal of Speech Sciences*, vol. 1, no. 1, pp. 55–83, 2011.

[32] J. A. S. Kelso, B. Tuller, E. Vatikiotis-Bateson, and C. A. Fowler, "Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 10, no. 6, pp. 812–832, 1984.