Acoustic Correlates of Voicing and Gemination in Bangla

Aanusha Ghosh
Center for Linguistics
Jawaharlal Nehru University
New Delhi
aanusha.ghosh@gmail.com

Abstract

The goal of this paper is to conduct an acoustic phonetic investigation of both primary and secondary cues that aid in the distinction between voiced and voiceless geminates in Bangla. Results of the statistical analyses examining both duration and non-durational correlates show that besides closure duration secondary cues such as the amplitude of the stop release burst and the fundamental frequencies of the vowels immediately preceding and following the obstruent are acoustically significant in the distinction between voiced and voiceless geminates and singletons. Voiced stops have significantly greater burst amplitudes than voiceless ones, and geminates are flanked by vowels with significantly higher F0's than those flanking corresponding singletons.

We also briefly explore the effects of gemination on V-to-V coarticulation. The assumption is that longer consonantal duration will act as a more effective barrier to V-to-V coarticulation in the case of geminates as opposed to singletons. Particularly, in the case of dental geminates, we hypothesize that longer consonantal duration contributes to lingual fronting, which manifests itself as greater resistance to V-to-V coarticulation.

1 Introduction

Maintaining voicing in obstruents is articulatorily challenging. Sufficient transglottal air pressure drop is required to maintain voicing. This becomes harder for obstruents, which require rapid increase in intraoral air pressure. The difficulty increases in geminate obstruents as they have longer closure. (Ohala, 1983) Hence, crosslinguistically, voiced geminates are rarer than their voiceless counterparts (Hayes and Steriade, 2004). Bangla, an Eastern Indo-Aryan language spoken in Bangladesh and the Indian states of West Bengal, Tripura and Assam, and the fifth most spoken language in the world[1] with nearly 300 million speakers is one of the few languages which have both voiced and voiceless geminates in their inventory.

Work on gemination in Bangla has been sparse. The only existing study on Bangla geminates focuses on voiceless stops and does not take voiced geminates into account (Lahiri and Hankamer, 1988; Hankamer et al., 1989). This paper aims to fill that gap by taking voiced geminates and the effects they might have on acoustic cues into consideration. It also takes a look at how vowel to vowel coarticulation might be affected because of gemination, especially in the case of dental stops. The articulatory motivation is that there should be some lingual fronting in the case of dental geminates, in order to maintain air pressure for a longer duration, and this fronting should then pose higher resistance to V-to-V coarticulation.

The rest of the paper is organised as follows: In Section 2, we cover a brief overview of work done on geminates and outline the aims and motivations of the current study. Section 3 presents the materials and methods used for the study, while Section 4 discusses the parameters measured in the study. Section 5 elucidates the statistical analysis employed for drawing inferences from the data. Section 6 deals with the F2 locus equations fitted for investigating the variation of the degree of coarticulatory resistance. Section 7 collates the results, and Section 8 concludes the work with a discussion on the inferences drawn from analysis and explores issues that need further work.

2 Related work

Lahiri and Hankamer (1988) investigated various acoustic cues and their perceptual relevance on the distinction between geminates and non-geminates in Bangla and Turkish. Their study, which focused only on voiceless stops, confirmed that closure duration was a perceptually salient cue for distinguishing between geminates and singletons in Bangla and Turkish, but discounted the fact that V1 duration was a secondary cue that could be used to identify geminates. In a subsequent paper (Hankamer et al., 1989), they concurred that secondary cues do contribute information about the distinction between geminates and singletons when the primary cues are ambiguous. However, they never explicitly stated exactly what these

[1] https://en.wikipedia.org/wiki/Bengali_language

D S Sharma, R Sangal and E Sherly. Proc. of the 12th Intl. Conference on Natural Language Processing, pages 413–418,
Trivandrum, India. December 2015.

secondary acoustic cues might be, proposing instead that the set of secondary features that biased subjects when the primary cue of consonant duration was ambiguous, was possibly due to a combination of cues, each by itself too subtle for their measurements to detect.

The cues that have been found to be consistently significant in the geminate-singleton distinction include stop closure duration and the length of the preceding vowel (V1 duration). Geminates have been found to have longer stop closure duration and shorter V1 duration (Esposito and Di Benedetto, 1999; Stevens and Hajek, 2004). Shortening of V2 duration in the case of geminates has also been observed. However, it has not been significantly different from that of singletons (Esposito and Di Benedetto, 1999).

Voicing has also been observed to have an effect on stop closure duration in obstruents. Stevens and Hajek (2004) found voiceless geminates to be substantially longer than voiced ones in Sienese Italian.

Studies of word-initial voiceless geminates in languages like Pattani Malay (Abramson, 1986; Abramson, 1987; Abramson, 1992; Abramson, 1999) and Kelantan Malay (Hamzah, 2013; Hamzah et al., 2013; Hamzah et al., 2012) show that when listeners do not have the advantage of relying on consonant duration as an acoustic cue for gemination, they make use of secondary cues which help in disambiguation. In particular, amplitude and fundamental frequency (Abramson, 1992; Abramson, 1999) were found to be significant cues to word-initial consonant length. Stevens and Hajek (2004) found that long voiced stops are often partially devoiced. Moreover, he revealed that in the case of Sienese Italian, voiceless geminates were preaspirated and their voiced counterparts were devoiced, so that the phonetic contrast between voiced and voiceless stops no longer manifested in a difference in the presence or absence of closure voicing, but rather in the absence or presence of aspiration before the geminated consonant.

F2 locus equations, which are a source of relational invariance that determine place of articulation, are also signifiers of the degree of coarticulatory resistance for a given consonant. Fowler and Brancazio (2000) showed that the more highly resistant a consonant is, the lower its locus equation slope. High coarticulation resistance, measured as a low standard deviation of F2 at the consonant normalized by variability of F2 across vowels, directly leads to a low locus equation slope. Low coarticulation resistance would, by the same logic, directly lead to a locus equation slope close to 1, since the variability of F2 measured at consonant release would be nearly the variability measured at the midpoint of the vowel.(Iskarous et al., 2010)

Esposito and Di Benedetto (1999) observed that

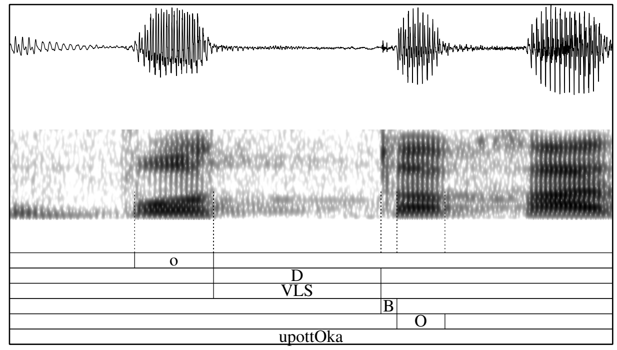

Figure 1: Waveform display of the word /upot̪ːɔka/ ('valley')

V1 formant frequencies showed no relationship with gemination, suggesting that this showed that no extra vocal effort was required for articulating geminates. They made no comment about formant transitions, however, or cast any light on the question of the variance of coarticulatory resistance in the presence of geminates.

In a study of the effect of consonant duration on V-to-V coarticulation in Japanese, Löfqvist (2009) found no significant effect of consonant closure duration on the degree of vowel-to-vowel coarticulation.

The aim of this paper is to answer the following questions:

- What are the major acoustic cues that set voiced geminates apart from voiceless ones?
- Are the secondary cues employed by word-medial geminates identical to those employed by word-initial geminates (in languages that do have them)?
- If voicing does indeed shorten the duration of stop closure, and gemination leads to partial devoicing (Stevens and Hajek, 2004), how are voiced geminates differentiated from voiceless singletons?
- Does gemination affect V-to-V coarticulation?

3 Materials and methods

The Shruti corpus, a Bangla corpus of read speech, built and maintained by Indian Institute Of Technology, Kharagpur (IITKGP) was used as the source of data for this study. The corpus consists of 7383 unique sentences. There are 34 speakers with ages varying from 20 to 40 years. 26 of the speakers are male, and 8 female. The speaker age in the corpus varies from 20 to 40 yrs. 700 words containing geminate and singleton stop consonant sequences from the

Shruti corpus were hand annotated for word, place and manner of articulation, consonant burst, and preceding and following vowels using Praat [2]. Each annotation file had six tiers of information: the vowel preceding the consonant, V1, the place of articulation of the stop, POA, the manner of articulation of the stop, MOA, the release burst, the following vowel, V2 and the word in which the geminate/singletons sequence appears.

The distribution of the data is given in Table 1.

	Singleton		Geminate	
Place of Articulation	Voiced	Voiceless	Voiced	Voiceless
Bilabial	32	20	44	10
Dental	44	149	36	165
Retroflex	—	48	18	25
Velar	11	59	17	20

Table 1: Distribution of singleton and geminate tokens across places of articulation

There are no voiced retroflex singleton tokens because the voiced retroflex ɖ does not appear as a singleton word-medially in Bangla.

4 Measurements

The following parameters were examined:

- Consonant duration
- Duration of the preceding vowel (V1)
- Duration of the following vowel (V2)
- Duration of the stop release burst
- RMS amplitude of the stop release burst
- Fundamental frequency of the preceding vowel
- Fundemental frequency of the following vowel

These measurements were taken for both voiced and voiceless geminates as well as voiced and voiceless singletons corresponding to four places of articulation: bilabial, dental, retroflex and velar.

5 Analysis

Three-way ANOVAs were done to test the significance of place of articulation, voicing and gemination on the duration of stop closure and length of the preceding and following vowels (V1 and V2). The interaction effects of all three factors on each of the parameters are given in Figures 2 and 3. In carrying out all of the following tests, the probability of type I error, α was fixed at 0.01 (1%).

[2]Boersma, Paul & Weenink, David (2015). Praat: doing phonetics by computer [Computer program]. Version 5.4.08, retrieved 24 March 2015 from http://www.praat.org/

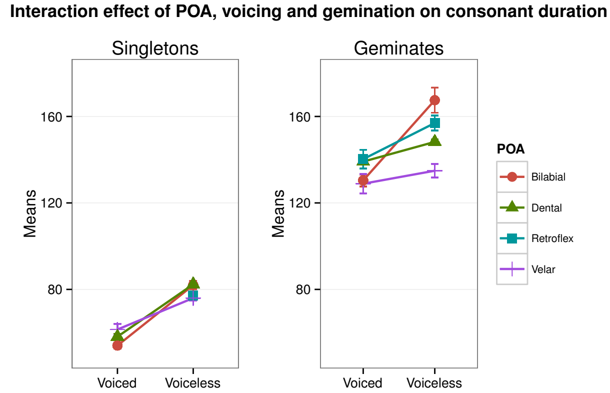

Figure 2: Effect of place of articulation, voicing and gemination on consonant duration

The ANOVA showed that gemination [F(1,806)=1146.749, p<0.0001], place of articulation [F(1,806)=6.242, p<0.001] and voicing [F(1,806)= 43.146, p<0.001] are all significant contributors towards variation of consonant duration.

In the case of V1 duration, gemination [F(1,810)=21.86, p<0.01], place of articulation [F(1,810)=24.109, p<0.00001], and voicing [F(1,810)=77.33, p<0.00001] are all highly significant factors.

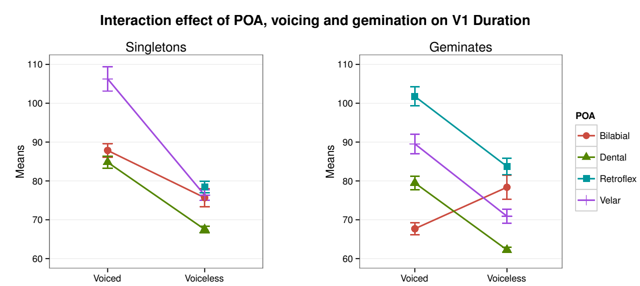

Figure 3: Effect of place of articulation, voicing and gemination on duration of preceding vowel

A three-way ANOVA carried out on V2 duration revealed that none of the factors appear to have any effect whatsoever of the duration of the V2.

Two-way ANOVAs were carried out to test the significance of voicing and gemination on duration and amplitude of the stop release burst, as well as the fundamental frequency (F0) of both the preceding and the following vowels. The effects of gemination [F(1,817)=19.102, p<0.01] and voicing [F(1,817)=71.13, p<0.0001] were found to be extremely significant contributors towards the variance of burst duration.

It was also found that voicing [F(1,817)=93.661, p< 0.0001] plays a very significant role in determining the burst amplitude.

The effect of gemination [F(1,817)=10.273, p<0.01] was found to be significant on the variation exhibited by F0 values of the preceding vowel.

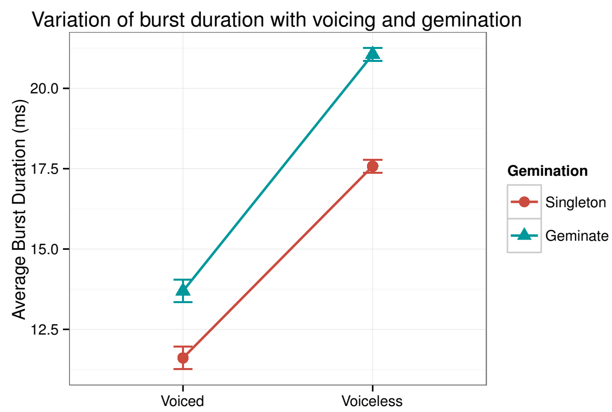

Figure 4: Effect of voicing and gemination on duration of stop release burst

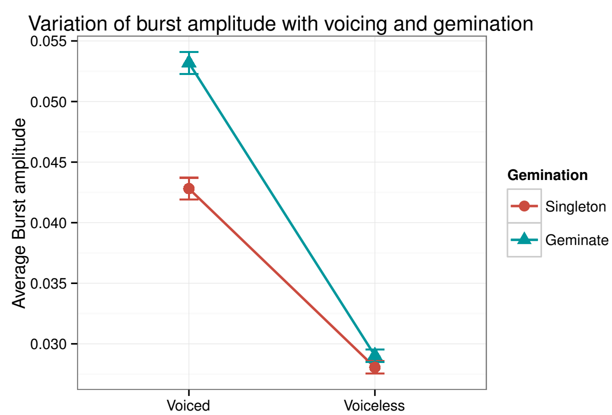

Figure 5: Effect of voicing and gemination on amplitude of stop release burst

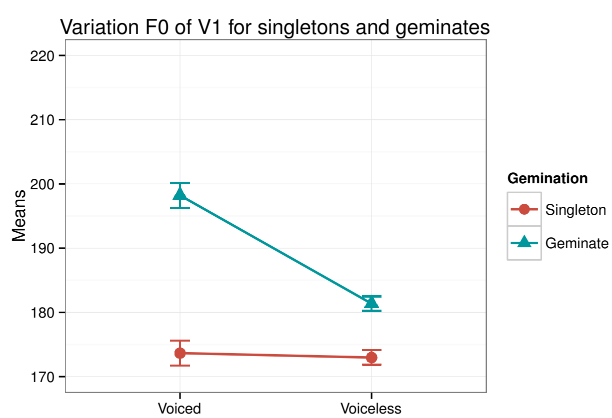

Figure 6: Effect of voicing and gemination on fundamental frequency of the preceding vowel

Voicing [$F(1,817)=4.613$, $p>0.01$] did not appear to have any effects on V1 duration.

Much like the preceding vowel, F0 of the following vowel is affected significantly by gemination [$F(1,817)=13.860$, $p<0.001$] but not by voicing [$F(1,817)=5.149$, $p>0.01$].

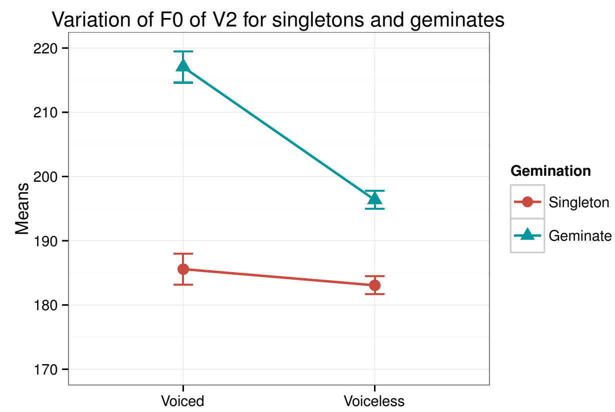

Figure 7: Effect of voicing and gemination on fundamental frequency of the following vowel

6 F2 Locus Equations

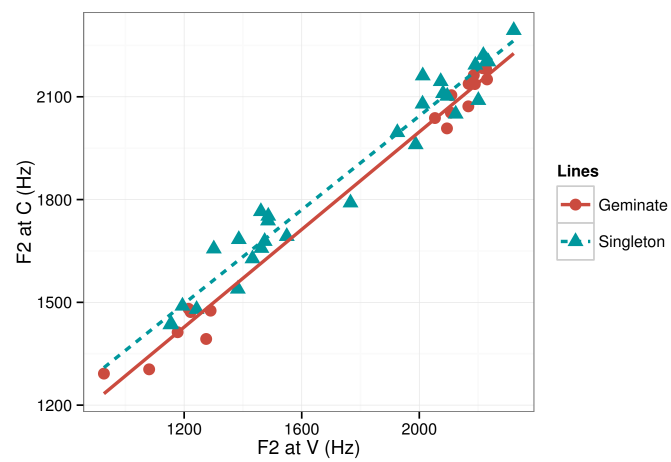

Figure 8: F2 Locus Equations for voiceless dental geminate and singleton

For modelling the variation of the degree of coarticulatory resistance of obstruents with gemination, locus equations of $F2_{onset}$ against $F2_{mid}$ of the vowel following the obstruent were plotted for both singleton and geminate tokens. Due to the sparsity of both inter and intra-speaker data, tokens from a single speaker, "Chan", was chosen for the study. The preceding vowel was fixed as the high front vowel /i/ and the obstruent was fixed as the voiceless dental /t̪/. 47 tokens were analysed, 18 of which were geminates and the others singletons. The unequal sample size is due to the fact that the data was taken from pre-existing corpus, and was thus subject to corpus-specific idiosyncrasies.

The second formant measures for the following vowel (V2) were taken as described in the previous chapter, with two measures for each formant - one at the 5% mark of the vowel for $F2_{onset}$ and another at the 50% mark of the vowel for $F2_{mid}$. The values were plotted and a regression line fitted through the data. The regression equation relating $F2_{\text{onset}}$ and $F2_{\text{mid}}$ is given by:

$$F2_{\text{onset}} = \beta + \alpha F2_{\text{mid}}$$

where α is the slope of the regression line, and β is the intercept.

The fitted regression lines for singletons and geminates (Figure 8) show no significant differences in their slopes. This implies that, at least for the given vocalic and consonantal environment, gemination does not affect the degree of coarticulation. However, more data from other speakers including different preceding vowels and obstruents is required to draw stronger conclusions in this regard.

7 Results

	F0 of surrounding vowels	
Burst Amplitude	High	Low
High	Voiced Geminate	Voiced Singleton
Low	Voiceless Geminate	Voiceless Singleton

Table 2: The roles of fundamental frequency and burst amplitude in the voiced-voiceless geminate-singleton distinction

Voicing has a significant effect on consonant duration, with both voiced singletons and voiced geminates having significantly shorter closure durations than their voiceless counterparts.

The duration of the preceding vowel is strongly affected by gemination. Geminates have shorter V1 duration. V1 duration is significantly affected by voicing as well — voiced obstruents are preceded by longer vowels. V1 duration is also contingent on the place of articulation, and displays interaction effects of place of articulation with both voicing and gemination.

Both gemination and voicing affect the duration of the stop release burst - burst duration is greater for geminates. Burst amplitude is significantly affected by voicing but not gemination — voiced stops have greater burst amplitude.

Fundamental frequencies of both vowels are affected by gemination. Vowels flanking geminates have higher fundamental frequency than those flanking singletons. Table 2 shows the variation of burst amplitude and fundamental frequency of surrounding vowels with voicing and gemination. Together, these two cues help in disambiguating voiced and voiceless singletons and geminates.

The F2 locus equations plotted for voiceless dental geminates and singletons with V1 fixed as /i/ show no appreciable difference in their slopes, indicating that geminates do not, after all, affect V-to-V coarticulation, at least as far as this particular vocalic and consonantal context is concerned.

8 Conclusion and further work

The perceptual experiments on voiceless geminate stops carried out by (Lahiri and Hankamer, 1988) and (Hankamer et al., 1989) focused exclusively on durational correlates. The presence of secondary acoustic cues, while acknowledged by them, were not credited with enough perceptual significance other than serving to bias listeners only when the primary cue of stop closure duration was ambiguous. However, the issue of voicing, the consequent perturbation of closure duration and its effects on the perception of geminates was left unaddressed in their work.

The results of this study clearly show that burst amplitude and fundamental frequency, both non-durational correlates, are significant contributors towards the distinction between voiced and voiceless geminates. Burst amplitude serves to disambiguate between voiced and voiceless stops, with the former having much greater burst amplitude than the latter. Fundamental frequency, on the other hand, is a key parameter in distinguishing geminates from non-geminates - F0 of vowels surrounding geminate consonants is consistently higher than for singletons. Burst duration is also a significant indicator for both gemination and voicing, with voiceless obstruents and geminates having longer burst duration than voiced stops and singletons respectively.

These results corroborate the findings of (Abramson, 1992; Abramson, 1999) and (Hamzah et al., 2012; Hamzah et al., 2013) regarding word-initial geminates in Pattani Malay and Kelantan Malay respectively. Thus, we can safely conclude that amplitude and F0 are powerful secondary cues that affect both word-initial and word-medial geminates.

We can also conclude that despite voicing shortening the duration of stop closure, geminates and singletons are still distinguishable by the fundamental frequencies of the surrounding vowels, which are significantly higher in the case of geminates.

From the preliminary study of the locus equations of the second formant of the following vowel, no difference was found between the slopes of the equations for geminates and singletons, indicating that geminates do no affect V-to-V coarticulation appreciably, at least as far as the voiceless dental /t̪/ is concerned. However,more data for other places of articulation and different vocalic contexts is required in this regard to draw a stronger conclusion.

One of the avenues of investigation that was left unexplored in this study, and which definitely merits further research, is a comparison of the degree of devoicing (a phenomenon noted in Sienese Italian and Japanese by Stevens and Hajek (2004) and Kawahara (2005) respectively) that occurs in voiced geminates and singletons in Bangla.

References

Arthur S Abramson. 1986. The perception of word-initial consonant length: Pattani malay. *Journal of the International Phonetic Association*, 16(01):8–16.

Arthur S Abramson. 1987. Word-initial consonant length in pattani malay. *Haskins Laboratories Status Report on Speech Research*, pages 143–147.

Arthur Abramson. 1992. Amplitude as a cue to word-initial consonant length: Pattani malay. *Haskins Laboratories Status Report on Speech Research*, pages 251–254.

Arthur S Abramson. 1999. Fundamental frequency as a cue to word-initial consonant length: Pattani malay. In *Proceedings of the 14th International Congress of Phonetic Sciences*, pages 591–594.

Anna Esposito and Maria Gabriella Di Benedetto. 1999. Acoustical and perceptual study of gemination in italian stops. *The Journal of the Acoustical Society of America*, 106(4):2051–2062.

Carol A Fowler and Lawrence Brancazio. 2000. Coarticulation resistance of american english consonants and its effects on transconsonantal vowel-to-vowel coarticulation. *Language and Speech*, 43(1):1–41.

Hilmi Hamzah, Janet Fletcher, and John Hajek. 2012. An acoustic analysis of release burst amplitude in the kelantan malay singleton/geminate stop contrast. In *Proceedings of the 14th Australasian International Conference on Speech Science and Technology*, pages 85–88.

Hilmi Hamzah, Janet Fletcher, and John Hajek. 2013. Amplitude and f0 as acoustic correlates of kelantan malay word-initial geminates.

Hilmi Hamzah. 2013. *The acoustics and perception of the word-initial singleton/geminate contrast in Kelantan Malay*. Ph.D. thesis.

Jorge Hankamer, Aditi Lahiri, and Jacques Koreman. 1989. Perception of consonant length: Voiceless stops in turkish and bengali. *Journal of Phonetics*, 17(4):283–298.

Bruce Hayes and Donca Steriade. 2004. The phonetic bases of phonological markedness. In B. Hayes, R. Kirchner, and D. Steriade, editors, *Phonetically Based Phonology*. Cambridge University Press.

Khalil Iskarous, Carol A Fowler, and Douglas H Whalen. 2010. Locus equations are an acoustic expression of articulator synergy. *The Journal of the Acoustical Society of America*, 128(4):2021–2032.

Shigeto Kawahara. 2005. Voicing and geminacy in Japanese: An acoustic and perceptual study. In K. Flack and S. Kawahara, editors, *University of Massachusetts Occasional Papers in Linguistics 31: Papers in Experimental Phonetics and Phonology*, pages 87–120.

A. Lahiri and J Hankamer. 1988. The timing of geminate consonants. *Journal of Phonetics*, 16:327–338.

Anders Löfqvist. 2009. Vowel-to-vowel coarticulation in japanese: the effect of consonant duration. *The Journal of the Acoustical Society of America*, 125(2):636–639.

John J Ohala. 1983. The origin of sound patterns in vocal tract constraints. In *The production of speech*, pages 189–216. Springer.

Mary Stevens and John Hajek. 2004. Comparing voiced and voiceless geminates in sienese italian: what role does preaspiration play? In *Proceedings of the 10th Australian International Conference on Speech Science and Technology. Sydney: S*, volume 340, pages 340–345.