HAL
open science

Vowel length impact on locus equation parameters : An investigation on Jordanian Arabic

Mohammad Abuoudeh, Olivier Crouzet

To cite this version:

Mohammad Abuoudeh, Olivier Crouzet. Vowel length impact on locus equation parameters : An investigation on Jordanian Arabic. 15th Annual Conference of the International Speech Communication Association (ISCA; Interspeech 2014), International Speech Communication Association (ISCA), Sep 2014, Singapore, Singapore. 10.21437/Interspeech.2014-49 . hal-01078603

HAL Id: hal-01078603

https://hal.science/hal-01078603v1

Submitted on 29 Oct 2014

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

Vowel length impact on locus equation parameters : An investigation on Jordanian Arabic

Mohammad Abuoudeh, Olivier Crouzet

Laboratoire de Linguistique de Nantes − LLING / EA3827
Université de Nantes, Chemin de la Censive du Tertre
44312 Nantes Cedex − France

mohammad.abuoudeh@univ-nantes.fr, olivier.crouzet@univ-nantes.fr

Abstract

Investigations of changes in speech rate and speaking style have reported that locus equation parameters are sensible to temporal variation though they usually do not evidence any major impact on place of articulation classification. The aim of the present study was to investigate the role of vowel length contrasts on locus equation parameters in Jordanian Arabic (JA) speech production in order to study the influence of intrinsic temporal changes (vowel length contrasts) on locus equation parameters. Statistical analyses of slopes and intercepts show that the locus equation parameters of a consonant produced with long vowels are significantly different from those of the same consonant produced with short vowels. Though this obsevation confirms the impact of time variations on locus equations, further analyses will need to address issues related to the potential impact of duration-related spectral modifications and the role these changes may play on the perceptual categorization of consonantal place of articulation.

Index Terms: coarticulation, formant transitions, vowel length, locus equations, Jordanian Arabic.

1. Introduction

Formant transitions from a consonant to a vowel express crucial information on stop place of articulation [1]. Seminal investigations by [2] led to propose a fixed frequency virtual locus as an invariant descriptor of consonantal place of articulation. According to this hypothesis, second formant transitions of each of the tested stop consonants (/b, d, g/ in CV utterances) were predicted to exhibit a regular direction toward a fixed virtual frequency point, independently from vowel context. This account was later criticized by [3] who found that formant loci are not resistant to vowel context. He demonstrated that changing the first or the second vowel in VCV utterances led formant transitions to go through different frequencies depending on the vowels. In addition, [4] showed that even when analyzing simple CV utterances, natural speech stimuli were much more variable than what [2] hypothesized, confirming the intrisically dynamic nature of speech production.

Later capitalizing on a seminal proposal by [5], locus equations were proposed as a way of encoding the information contained at two different positions in the second formant [6] so as to express the fundamental coarticulatory relationship between the consonant and the vowel in a CV sequence, leading to an expression of the consonantal frequency locus as a function of the vowel formant frequency. Conventionally, locus equations are computed from the acoustic properties of one single consonant category produced in various vowel contexts. They express a linear relationship between $F2_{onset}$ (which represent the beginning of the transition of the second formant and can be measured at the first pulsation of the vowel that follows the consonant[1]) and $F2_{mid}$ frequency values (which represent the middle point –the steady-state frequencies– of the second formant). Addressing the linguistic significance of these equations, [6] found that locus equation coefficients (slope and y-intercept) vary in terms of place of articulation (for the /b, d, g/ stop consonants) but that they also tend to be stable over various modes of articulation [8]. Though this is still an ongoing debate [9], some accounts of locus equations state that they (or the acoustic properties they are derived from) may provide crucial information for the perceptual categorization of consonantal place of articulation, which has partly been confirmed by discriminant analyses that were based on either acoustic measurements (among which $F2_{onset}$ and $F2_{mid}$) or locus equation parameters (slopes and intercepts).

A less debated account of locus equations is that they offer a reliable method for measuring the degree of coarticulation [10, 9] of CV sequences. As a matter of fact, this coarticulatory influence would explain the apparent relationship between LE slopes and place of articulation. When the slope value is close to 1, it indicates a high degree of coarticulation; the consonant articulation is then massively associated with the vowel and, acoustically, $F2_{onset}$ frequencies are strongly correlated with $F2_{mid}$ ones. Such a case is found with e.g. bilabial consonants as the occlusion is strongly dependent on the planned vocal tract configuration that is being prepared for the vowel articulation. If the slope were close to 0, it would indicate a very low degree of coarticulation; the vowel and the consonant articulation would then be totally independent from one another and, acoustically, $F2_{onset}$ would be stable, no matter what the frequency of $F2_{mid}$ would be. Though coarticulation is never completely absent, lower slope values are found for consonants involving the tip / blade of the tongue (e.g. /d/) as the planification of the vowel articulation will have a smaller impact on reaching the closure target gesture, due to the possibility for the tongue to compensate for jaw movements.

Effects of different speaking styles on locus equation coefficients have been investigated by [11, 12, 13, 14] and they all found that locus equations are sensible to spontaneous vs. controlled speech variations. This is in line with statements that locus equations are good indicators for describing degree of coarticulation. Specifically, it was observed that the slopes of locus equations extracted from spontaneous speech are usually steeper than slopes derived from controlled speech. Ac-

[1] though see [7] and infra. for issues concerning unvoiced stops

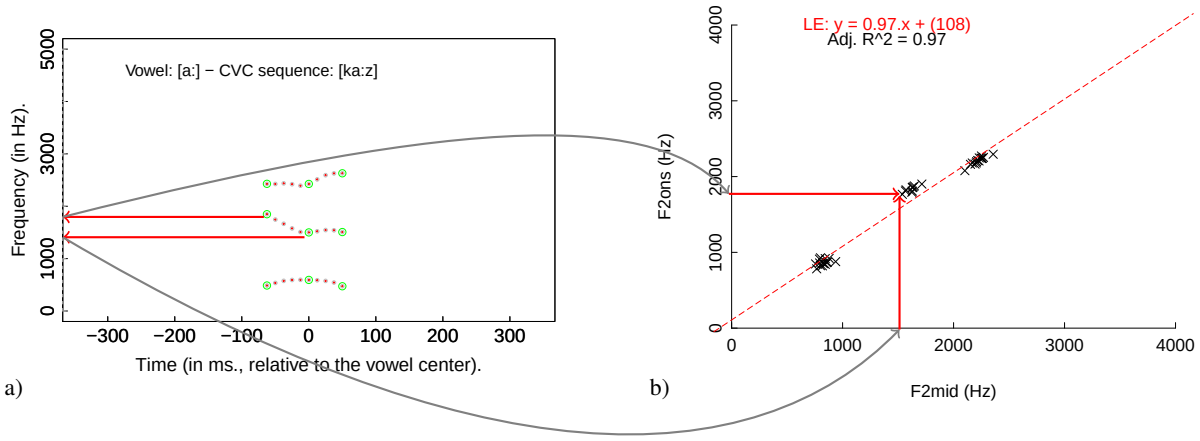

Figure 1: Illustration of the principles involved in the extraction of acoustic data to compute locus equations — (1a) A schematic spectrogram representing the first three LPC tracked formants and their transitions. $F2_{onset}$ and $F2_{mid}$ frequencies are measured. — (1b) The two extracted frequency values serve to draw one single point in a $F2_{onset}$ / $F2_{mid}$ space. The time dimension and the temporal distance between these two points are abstracted away.

cording to [12], such variations may come from either a higher degree of coarticulation in spontaneous speech as well as from vowel reduction or formant underhoot. Nonetheless, Sussman and colleagues [15] argued that the influence of speaking style variation on locus equations, though it is real, would be reduced in amplitude and would have almost no impact on place of articulation classification. It has also been shown that variations in speaking rate have a distinct influence on the slopes of locus equations [16] and, more recently, [17] observed that they even vary continuously with progressive variations in speaking rate.

While it has already been shown how temporal and / or articulatory variations may impact locus equations, research on the impact of temporal variations (stress and speaking style [11, 12, 13], speaking rate [16, 18, 19, 17] and emphatic stress [20]) on locus equations may usually be associated with a hypo– vs. hyper–articulation continuum. These contrasted speaking *states* are performed by speakers on the basis of extrinsic timing parameters[2]. In the case of vowel length contrasts, temporal variations are intrinsic to the characterization of the segment and, according to a recent study on the relationship between vowel length and V-to-V coarticulation [21], these time modifications are usually much more contrasted in terms of duration (short vowels are 60% the duration of long vowels on the average, while speaking rate variations only produce an average 20% ratio) while their impact on spectral phenomena is disputed (some authors have evidenced strong effets on formant frequencies, while others have found no impact).

Our aim was twofold: (1) investigating the influence of variations in the temporal properties of speech production when speakers maintain a relatively constant speaking style and rate by investigating the influence of vowel length oppositions and their corresponding durations on locus equations, and (2) to specifically address the role that such variations in locus equations parameters may play for the perceptual categorization of place of articulation. In order to address these issues, a speech production experiment was conducted on Jordanian Arabic speakers that was aimed at comparing locus equation parameters for stop consonants in the context of either short or long vowels.

[2]Though [17]'s data should be put appart as these were obtained from *continuous* variations in speaking rates.

2. Method

2.1. Subjects

Four native male speakers of Jordanian Arabic participated in this experiment. They were all students at the University of Nantes at the time of their enrollment, and were aged between 24 and 38 years old and had spent their youth in Jordany. All speakers reported no speech or hearing impairment and participated in the experiment on a voluntary basis. The first author is one of the participants.

2.2. Stimulus

Speakers were asked to produce C_0VC and CVC_0 syllables in a carrier phrase : /ħaka . . . marteːn/ (eng. *he said . . . twice*). C_0 was the target consonant while the alternate C was a random one. The sentences were sequentially displayed in arabic script (right-to-left) on a computer screen using a Python program. The order of presentation of the various items was randomized for each recorded speaker. The target (C_0) stops were /t, tˤ, k, q, ʔ/ (respectively alveolar, pharyngealized alveolar, velar, uvular and glottal place of articulation) followed or preceded by short or long vowels /i, a, u, iː, aː, uː, oː, eː/ while the other consonant (C) was varied randomly in order to select regular arabic words. The target consonant (C_0) appeared in either initial or final position of the CVC sequence. All these conditions (8 vowels, 5 consonants, 2 syllabic positions) were combined to select 80 ($8 \times 5 \times 2$) actual CVC Arabic words that were read by speakers. Each of these 80 sequences was repeated 15 times, providing a total number of 1200 recordings for each speaker.

2.3. Data analysis

Speech recordings were examined within the Praat [22] phonetic analysis software in order to extract the relevant transitional portion of each CVC utterance, locating the start and end of the second formant transition on a wide band spectrogram. Formant frequency tracks along with their corresponding

temporal positions were then extracted (LPC based formant extraction, window length = 0.025s, maximum formant frequency = 5000, maximum number of formants = 5) within the selected interval. Using the R programming language [23] , these were passed through a script that was designed to transform these raw acoustic data and locate $F2_{onset}$ and $F2_{mid}$ positions and frequencies for each of the 15 repetitions of a consonant in a particular context. Linear regressions of these data points were then produced in order to estimate the parameters of each of the 80 locus equations (5 consonants $\times$ 2 vowel lengths $\times$ 2 syllabic positions $\times$ 4 speakers). An example of such a linear regression is displayed in Figure 1, where the procedure for $F2_{onset}$ and $F2_{mid}$ extraction is illustrated. These two frequency values correspond to one single point on the $F2_{onset}$ / $F2_{mid}$ space.

In order to avoid the influence of extreme values on locus equation slopes, outliers were automatically identified based on sample residuals from the first linear regression. Then a second linear modelling was performed without the outliers. These outliers represent only 7% of the total data. As it may occur that our results were influenced by removing such outliers, we have consitently verified that the effects mentionned in this paper are still present when analyzing data with the outliers maintained. The two long vowels that have no short counterparts /eː, oː/ have been excluded from the statistical analyses that are dicussed in this paper so as to cancel out their impact in terms of formant frequency. The vowel length contrast is therefore orthogonal to the vowel timbre category as only the /i, a, u/ and /iː, aː, uː/ sets are taken into account.

3. Results

According to our expectations, the slopes of a consonant associated with long vowels tend to be lower than the slopes of the same consonant when it is associated with short vowels. Accordingly, the y-intercepts of consonants that are associated with long vowels are higher than when the same consonant is associated with short vowels (see Figure 2 and Table 1). One can notice the relatively high values of R^2 that are associated with these data. Looking at the potential relationship between place of articulation and vowel length effects, pharyngealized alveolar consonants (/tˤ/) seem to be the most influenced by vowel length alternations, while velar stops (k) seem to be much more stable across vowel lengths.

Looking at intercepts, similar tendencies seem to hold. Y-intercepts always seem to be higher for long vowels than for short ones. Again, velar stops seem to be less marked by this phenomenon as they only exhibit a small difference (around 30 Hz) while all the other places are associated with a 100–150 Hz change.

Two three-way ANOVAs (Vowel length $\times$ Stop place $\times$ Position) were conducted with "Speakers" as a random variable in order to compare the effects of vowel length on locus equation slopes and y-intercepts respectively. In order to satisfy the normality assumption, slopes and intercepts were transformed to respectively squared slopes and square-root intercepts but descriptive statistics are displayed using the original units. None of the interaction effects involving the "length" variable reach statistical significance. The 2nd order interactions (Length $\times$ Place $\times$ Position) are only marginally significant for slopes ($F_{(4,12)} = 2.93, p = 0.07$) as well as for intercepts ($F_{(4,12)} = 2.87, p = 0.07$). Looking at slopes specifically, none of the 1st order interactions are significant either (Length $\times$ Place: $F_{(4,12)} = 1.22, p > 0.1$; Length $\times$ Position: $F_{(1,3)} < 1$), while the Place $\times$ Position interaction is

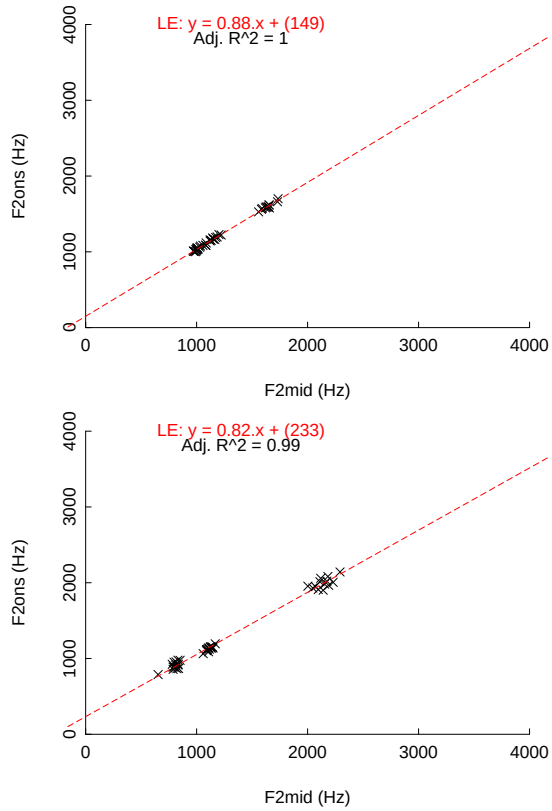

Figure 2: Two locus equations computed for the pharyngealized alveolar stop /tˤ/ produced with short vowels (up) and long vowels (down) in initial position by L2.

only marginally significant ($F_{(4,12)} = 3.03, p = 0.06$). As far as intercepts are concerned, only the Place $\times$ Position 1st order interaction reaches significance ($F_{(4,12)} = 8.04, p < 0.01$) but, in line with the slope data, none of the two others are significant (Length $\times$ Place: $F_{(4,12)} = 1.18, p > 0.1$; Length $\times$ Position: $F_{(1,3)} < 1$).

Looking at main effects, the slopes of long vowels significantly differ from those of short vowels across all consonants ($F_{(1,3)} = 65.68, p < 0.01$). Similarly, the y-intercepts of LEs associated with long vowels are significantly different from those of short vowels across all consonants, $F_{(1,3)} = 21.31, p < 0.05$. In addition, there is a main effect of consonant category on both locus equation slopes ($F_{(4,12)} = 25.56, p < 0.001$) and y-intercepts ($F_{(4,12)} = 20.26, p < 0.001$).

	Slope		Intercept		R^2	
C	Short V	Long V	Short V	Long V	Short V	Long V
t	0.78	0.67	375	549	0.94	0.95
tˤ	0.77	0.54	198	384	0.95	0.86
k	0.96	0.94	-33	3	0.95	0.94
q	0.84	0.78	43	158	0.93	0.92
ʔ	0.93	0.83	-62	119	0.98	0.94

Table 1: Mean values of slopes, intercepts, and R^2 of locus equations for each consonant coarticulated with short and long vowels.

4. Discussion

According to Sussman and his collaborators, locus equation slopes are a good indicator of stop place of articulation. Indeed, our statistical analyses indicate that slopes and intercepts of the investigated consonants are significantly different across all vowels. However, further work needs to address how these 5 individual consonant categories behave with respect to these parameters.

Though the present locus equation data are overall in line with this model, it appears that they are systematically influenced by alternations in vowel length. This influence is observed on both slopes and y-intercepts. In line with our predictions, locus equations are influenced by parameters that are associated with time whereas they represent linear relationships between spectral information. This may put strong constraints on the interpretation of locus equations for the identification of place of articulation in speech perception mechanisms.

However, we also observed that variations of vowel length are associated with modifications of spectral configuration measured at the vowel midpoint (remarkably for /i/): long vowels exhibit higher F2 frequencies than short vowels for back vowels /a, u/, and the opposite for front vowels /i/ in our sample, see Table 2. Though long vowels are (as expected) physically longer (mean value 127 ms) than short vowels (mean value = 68 ms; in a one-way ANOVA with repeated measures, this difference is significant ($F_{(1,3)} = 94.22, p < 0.01$). This vowel length difference also seems to be related to spectral properties of vowels. $F2_{mid}$ frequencies of long vowels are significantly different from those of short vowels ($F_{(2,6)} = 65.36, p < 0.001$). The impact of vowel length on locus equation slopes and intercepts may therefore be a consequence of this vowel modification, which will have to be investigated further.

	i	a	u
long	2157 (20)	1353 (27)	963 (26)
court	1711 (34)	1388 (28)	1090 (25)

Table 2: Mean (and Standard Error of Mean) for F2 values in Hz measured at the middle of the formants of long vowels compared to short vowels.

In parallel with this observation, a recent investigation [24] showed that modified locus equations that were obtained by extracting $F2_{mid}$ from different points closer in time to $F2_{onset}$ did nevertheless categorize place of articulation successfully. This account indicates (cf. also [20]) that modifications on $F2_{mid}$ have marginal effects on locus equation parameters, while modifications on the $F2_{onset}$ generate important effects on locus equations. Our data will need to be investigated further along these lines.

5. Conclusions

Locus equations are a good indicator of degree of coarticulation [16, 9, 25], that is in line with our results that show that slopes for consonants produced with short vowels are steeper than when the same consonant categories are coarticulated with long vowels. It would indicate that coarticulation overlap is more important in CVC utterances when V is short. These results on intrinsic time-related phenomena (vowel length contrasts) are similar to what have been found in the literature on extrinsic time-related phenomena (speaking rates and styles).

Further analyses are needed to investigate the potential impact of $F2_{mid}$ and / or $F2_{onset}$ respectively on the observed effects. Also, the impact of such variations on the perceptual classification of consonants will have to be addressed using discriminant analyses.

6. References

[1] F. S. Cooper, P. C. Delattre, A. M. Liberman, J. M. Borst, and L. J. Gerstman, "Some Experiments on the Perception of Synthetic Speech Sounds." *Journal of the Acoustical Society of America*, vol. 24, no. 6, pp. 597–606, 1952.

[2] P. Delattre, A. Liberman, and F. Cooper, "Acoustical loci and transitional cues for consonants," *The Journal of the Acoustic Society of America*, vol. 27, no. 4, pp. 769–773, July 1955.

[3] S. Öhman, "Coarticulation in VCV utterances: Spectrographic measurements," *The Journal of the Acoustical Society of America*, vol. 39, no. 1, pp. 151–168, January 1966.

[4] D. Kewley-Port, "Measurement of formant transitions in naturally produced stop consonant-vowel syllables." *Journal of the Acoustical Society of America*, vol. 72, no. 2, pp. 379–389, Aug. 1982.

[5] B. Lindblöm, "Spectrographic study of vowel reduction," *Journal of the Acoustical Society of America*, vol. 35, no. 11, pp. 1773–1781, November 1963.

[6] H. M. Sussman, H. A. McCaffrey, and S. A. Matthews, "An investigation of locus equations as a source of relational invariance for stop consonant place categorization," *Journal of the Acoustical Society of America*, vol. 90, pp. 1309–1325, November 1991.

[7] G. Modarresi, H. M. Sussman, B. Lindblöm, and E. Burlingame, "Locus equation encoding of stop place: revisiting the voicing/VOT issue," *Journal of Phonetics*, vol. 33, pp. 101–113, 2005.

[8] H. M. Sussman and J. Shore, "Locus equation as phonetic descriptors of consonantal place of articulation." *Perception & psychophysics*, vol. 58, no. 6, pp. 936–946, 1996.

[9] K. Iskarous, C. a. Fowler, and D. H. Whalen, "Locus equations are an acoustic expression of articulator synergy." *The Journal of the Acoustical Society of America*, vol. 128, no. 4, pp. 2021–32, Oct. 2010.

[10] C. Fowler, "Invariants, specifiers, cues: An investigation of locus equations as information for place of articulation." *Perception and Psychophysics*, vol. 55, pp. 597–610, 1994.

[11] D. Krull, "Second formant locus patterns as a measure of consonant-vowel coarticulation." *Phonetic experimental research at the institute of linguistics university of Stockholm-PERILUS*, vol. 5, pp. 57–75, 1987.

[12] ——, "Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech." *Phonetic experimental research at the institute of linguistics university of Stockholm-PERILUS*, vol. 10, pp. 87–108, December 1989.

[13] D. Duez, "Second formant locus-nucleus patterns: An investigation of spontaneous french speech." *Speech Communication*, vol. 11, no. 4-5, pp. 471–427, October 1992.

[14] H. M. Suusman, E. Dalston, and S. Gumbert, "The effect of speaking style on a locus equation caracterization of stop place of articualtion," *Phonetica*, vol. 55, pp. 204–225, 1998.

[15] H. M. Sussman, D. Fruchter, J. Hilbert, and J. Sirosh, "Linear correlates in the speech signal : The orderly output constraint," *Behavioral and Brain Siences*, vol. 21, pp. 241–299, 1998.

[16] L. Branzacio and C. A. Fowler, "On the relevance of locus equations for production and perception of stop consonants," *Perception & Psychophysics*, vol. 60, no. 1, pp. 24–50, 1998.

[17] J. Berry and G. Weismer, "Speaking rate effects on locus equation slope," *Journal of Phonetics*, vol. 41, pp. 468–478, 2013.

[18] A. Agwuele, H. Sussman, and B. Lindblöm, "The effect of speaking rate on consonant vowel coarticulation," *Phonetica*, vol. 65, no. 4, pp. 194–209, 2009.

[19] G. Weismer and J. Berry, "Effects of speaking rate on second formant trajectories of selected vocalic nuclei," *The Journal of Acoustical Society of America*, vol. 113, no. 6, pp. 3362–3378, June 2003.

[20] B. Lindblöm, A. Agwuele, and H. M. Sussman, "The effect of emphatic stress on consonant vowel coarticualtion," *The Journal of Acoustical Society of America*, vol. 121, no. 6, pp. 3802–3813, June 2007.

[21] P. Mok, "Effects of vowel duration and vowel quality on vowel-to-vowel coarticulation." *Language & Speech*, vol. 54, no. Pt 4, pp. 527–545, 2011.

[22] P. Boersma and D. Weening, "Praat: Doing phonetics by computer." Computer program, 2012. [Online]. Available: http://www.praat.org

[23] R Core Team, *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2012, ISBN 3-900051-07-0. [Online]. Available: http://www.r-project.org/

[24] A. Rhone and A. Jongman, "Modified locus equations categorize stop place in a perceptually realistic time frame," *The Journal of Acoustical Society of America*, vol. 131, no. 6, pp. EL487–EL491, May 2012.

[25] B. Lindblöm and H. M. Sussman, "Dissecting coarticulation: How locus equations happen." *Journal of Phonetics*, vol. 40, no. 1, pp. 1–19, January 2012.