Differentiating the Effect of Speech Tempo on CV Coarticulation

Augustine Agwuele[1], *Harvey Sussman*[2], *Björn Lindblom*[3] & *Amanda Miller*[4]

Texas State University, San Marcos[1]: aa21@txstate.edu; University of Texas at Austin[2]: sussman@mail.utexas.edu;
University of Texas at Austin[3]: lindblom@ling.su.se; Cornell University[4]: am332@cornell.edu

ABSTRACT

When suprasegmental and contextually induced variations interact, CV coarticulation undergoes more complex, and often subtle, resistance and assimilation. This study empirically documents the separate effect of increase in speech tempo on C-locus and V-nucleus F2 patterns using novel analysis techniques. Results showed systematic alterations to F2locus as a result of (a) rate-induced reduction of the vowel space and (b) rate-induced changes to coarticulation per se.

Keywords: Locus equations, coarticulation, speech tempo, modified locus equation metrics.

1. INTRODUCTION

One persistent problem undermining the phonetician's conceptualization of the production and perception of sequences of consonants and vowels is how to resolve the context-induced variability integral to the speech signal. The locus equation (LE) paradigm [1] has sought a solution to this non-invariance problem. Results from experiments using LE's metrics have provided some lawful orderliness to the production and perception of stop consonants across various vowel contexts.

LEs are linear regression fits to scatterplots of coordinates representing, separately for each consonantal category, all F2 transition onsets, plotted on the y-axis, in relation to mid-vowel frequencies plotted on the x-axis [2]. These linear equations are of the form F2onset = k*F2vowel + c (the constants k and c are slope of the regression line and its intercept, respectively). The slope parameter is generally interpreted as a phonetic index of the place of articulation [3], as well as quantifying the extent of CV coarticulation as it correlates the extent of F2 at CV with F2 at the vowel-nucleus, its 'putative' target [4].

In quantifying the degree of coarticulation, LE slopes values range from 0 to 1. Slopes nearer to 0 (minimum coarticulation) are said to indicate lesser dependence of CVonset on the vowel, while slopes closer to 1 (maximum coarticulation) specify the contrary. Thus, LE slopes statistically restate the extent of a consonant's dependence on the following vowel and also measures the effect of a vowel on the preceding consonant. Thus the locus equation allows inferences about the extent of variation in consonant production that's induced by the upcoming vowel context without claims about other separate, but interactive sources of signal variability such as suprasegmental factors, phonetic context, and speech style. [5, 6, 7]

Consider for example the influence of prosody on syllables and its consequence for coarticulation. Stressed syllables are known to be less susceptible to coarticulation [6, 7]; their midpoint F2 values compared to unstressed vowels more closely approximate their "putative" targets. In the F1*F2 plane, stressed segments cause expansion to the vowel space [8]. The ensuing articulatory distance between C & V not only affords them independent movements, it reduces the anticipatory influence of a stressed V on the locus of the preceding C. The resultant 'decrease in coarticulation,' as reflected by lower LE slopes, is now not mainly due to vowel context effects, but also to prosody, which also shapes the place of the vowel in the vowel space. Whereas contextual vowel effects persist, they are now adjusted by the prosodic overlay.

Recognizing this problem, existing LE accounts were amended [5] so as to determine how the stop consonant is affected by prosodic overlay apart from stress-induced vowel expansion. The modified metrics accounted for F2onset variations in [bV] tokens in terms of proportional scaling, whereby, bVonsets were borne along by the vowel. For [dV] & [gV] tokens, the metrics showed F2loci values higher than their predicted correlates, especially when back Vs follow the stops.

This study applies the amended metrics to CV-coarticulation under speech perturbations that reduce, rather than expand, vowel space. The aim is to dissociate rate-induced acoustic effect on F2onset from that affecting vowel nucleus. The research question is: Do increases in speech tempo affect F2loci differently or are they mainly dragged along by their tightly bonded vowel?

2. METHODOLOGY

2.1. Speech sample and elicitation of tempos

The test material consisted of V_1#CV_2 sequences that were embedded in English carrier sentences. V_1 varied 3 vowels (i,ɑ,ʊ), which were chosen to maximally exploit the acoustic space. The medial consonants were /bdg/. V_2 contained the following 10 vowels: (i, ɛ, æ, ʊ, o, ɔ, ʌ, eʲ, ɑʲ, aᵘ). An example of the carrier sentence is:

"I saw [*three beads*] in the mall again." The items in brackets were the focus of investigation. The subjects (P & J), not knowing this, were asked to read the sentences in three manners: normal, fast, and as fast as possible without losing clarity. Each subject performed at a self chosen pace. Each produced a total of 810 tokens [$3V_1$ * 3C * 3rates* $10V_2$ * 3 repetitions]. In the discussion, the rates are referenced as N, F, and Fst, for normal, fast, and fastest respectively.

2.2. Recording & measurement procedures

Two adult female speakers of American English were recorded in a sound attenuated room using a uni-directional high quality microphone (SHURE BG 3.1) and a digital master recorder (Fostex D-5). The recorded signals were sampled at 22kHz, digitized, and filtered using Praat [9], which was also used for all acoustic measurements.

Acoustic measurements were made from wide band spectrograms. F2 & F1 values were obtained for V_1 &, V_2mid, V_1off, and V_2ons following already established procedures [10]. The duration of segments V_1, C, & V_2 were measured.

3. Results

3.1. Duration

Data pooled across subjects and stops showed a consistent decrease for each segment as the rate of speech increased from N to F to Fst. V_1, C, and V_2 decreased respectively in the following order: /125 > 86 > 73/; /113 > 67 > 51/, and / 210 > 125 > 105/, all in (ms.). The results of series of paired t-tests, conducted to assess the significance of these decreases for each segment across speech tempo, showed that N differed significantly from F; N from Fst; and F from Fst, at $p < 0.0001$, for each subject. In all cases, the df. was 517.

3.2. Locus equations

Locus equation plots were obtained for each CV_2 sequence for each speaker and across the three speech rates. Table 3.1 is the summary of the LE coefficients and their means.

Table 3.1: Summary of LE slopes, intercepts, and R^2

Subj/	***[N]***		***[F]***		***[Fst]***	
cons	***k***	***c***	***k***	***c***	***k***	***c***
P /b/	0.71	*354*	0.78	*235*	0.85	*108*
J/b/	0.75	*339*	0.82	*218*	0.85	*132*
Mean	**0.73**	***347***	**0.8**	***227***	**0.85**	***120***
Mean R^2	*0.96*		*0.95*		*0.92*	
P /d/	0.53	*1002*	0.59	*829*	0.65	*681*
J/d/	0.49	*1136*	0.69	*669*	0.86	*284*
Mean	**0.51**	***1069***	**0.64**	***749***	**0.76**	***483***
Mean R^2	*0.83*		*0.83*		*0.86*	
P /g/	0.53	*1002*	0.97	*174*	1.01	*51*
J/g/	0.99	*260*	0.99	*168*	1.00	*102*
Mean	**0.71**	***631***	**0.98**	***171***	**1.01**	***77***
Mean R^2	*0.85*		*0.91*		*0.95*	

As shown in Table 3.1, LE slope for each CV_2 and speaker increased as speech rate increased from N to F and to Fst., with slight inter-speaker variation. In the traditional locus equation parlance, these increases reflect a rise in the degree of CV coarticulation. Statistical test of the slopes collapsed across subjects and stops showed no significance for the difference between N and F slopes. N and Fst slopes differed significantly [t (-2.76) = 5; p = 0.04]; so did F from Fst slopes [t (-2.75) = 5; p = 0.04].

Not clear however, is whether the observed difference between each speech tempo is a true reflection of coarticulatory processes or whether they were mere artifacts of rate-dependent vowel changes. The remainder of the paper focuses on determining whether the C-loci, as reflected by F2onset, were actively adjusting their values to changes in speech tempo or whether they were mainly responding to the coarticulated vowels.

For the sake of simplicity and brevity, the rest of the analyses distinguish mainly between two instead of three rates: (N)ormal and Upped-Tempo (UT). Upped tempo values were derived from the means of corresponding tokens of F and Fst rates. First, how did rate affect the vowel space?

3.3. Effect of increased rate on F1xF2 plane

In order to observe the effect of increased speech tempo on the vowels, changes in F1/F2 vowel space were examined by obtaining F1 & F2 values from V_2midpoints in N and UT rate conditions.

Figure 3.3: F1xF2 plane: Reduction of vowel space

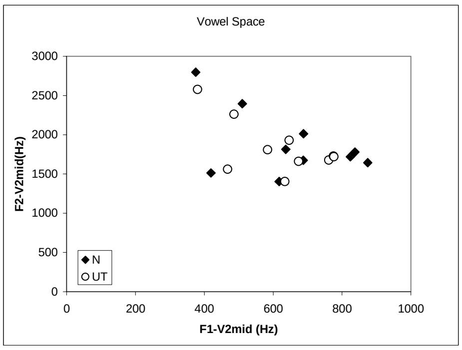


Figure 3.3 is a summary plot of the 10 V_2. The data was pooled across subjects and stops. Results show a consistent reduction to the space occupied by vowels under UT (unfilled diamonds), relative to those under N rate condition (filled diamond). The reduction is consistent with previous results [1].

3.4. Modified locus equation analysis

In order to dissociate the effect of increased speech tempo on F2-onsets, modified locus equations were used. Multiple regression analyses (MRA) with F2ons as the dependent variable and F2 of V_1 and V_2 nuclei as the predictor variables were run. The analyses produced the following baseline equation:

$$F2\text{-}onset(V_2) = a + b^* F2mid(V_2) + c^* F2mid(V_1).$$

Table 3.4: Results of the multiple regression analysis

	/b/		/d/		/g/	
	N	***UT***	***N***	***UT***	***N***	***UT***
a	295	105	918	411	141	-53
b	0.73	0.82	0.48	0.64	0.94	0.96
c	0.03	0.05	0.10	0.15	0.09	0.12

In Table 3.4 a = Y-intercept, b = anticipatory-, c = perseveratory influence. The table compares the coefficients for N and UT rate conditions. The results show a greater degree of V_2 anticipatory effects at the CV-onset. There is also evidence of the persistent effects of V_1 at the CV_2-boundary. However, V_2 exerted a greater influence on CV_2loci than did V_1 across /bdg/. For example, as rate increased from N to UT, the influence of V_2 on CV_2loci increased by 0.10Hz compared to 0.04Hz for V_1.

Using b & c from table 3.4; 'expected' onsets values were generated by substituting F2-V_2mid (UT) and F2-V_1mid (UT) frequency values into the baseline equation as shown below:

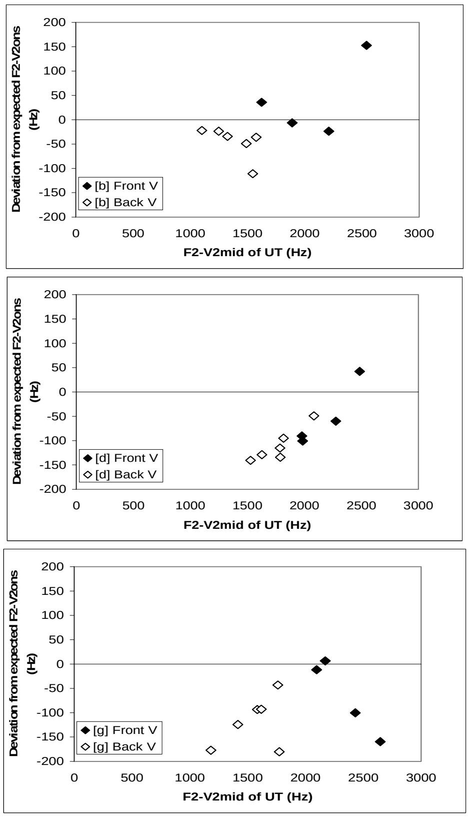
$$F\text{-}2onset(V_2) = a + b^* F2mid(\underset{\substack{\uparrow \\ F2mid(V_2)\ (UT)}}{V_2}) + c^* F2mid(\underset{\substack{\uparrow \\ F2mid(V_1)\ (UT)}}{V_1})$$

The substitution with reference to Lindblom [5] is predicated on the assumption that coarticulation, as traditionally defined by the LE slopes, does not change due to increases in speech tempo. As such, the substitution into the baseline equation is legitimate. Secondly, In the MRA, V_2-F2onset is a function of the preceding V_1-F2mid and of the upcoming reduced V_2-F2mid. Keeping the "degree of coarticulation" constant mandates leaving theV_2 and V_1 dependencies unchanged. The implemented substitution thus retains contextual effects, but is now conflated with rate effect. It is expected that the newly derived F2onsets will fall on the line representing the LE in two dimensions. If on the other hand, the 'predicted' F2onsets were plotted against 'observed' F2onsets from UT and there is a divergence; it would be fair to attribute such frequency differences solely to rate effects at CV_2 juncture, because the V_2 slope coefficient in the baseline equation was originally derived from F2midvowel frequencies without the rate increase.

The aim of this substitution strategy was to use the match, or mismatch, between 'expected' and 'observed' F2onset values to tease apart the vowel reduction effect by itself from the possible coarticulation changes that independently influence F2onsets. If there is agreement between 'observed' and 'expected' values, then variations found for F2onset would be solely due to rate-induced vowel space reduction and not to any systematic adjustment of the consonants to coarticulatory changes caused by speech tempo. On the other hand, if systematic discrepancies are found, then such may constitute evidence for rate-dependent changes that directly affect the CV_2onsets.

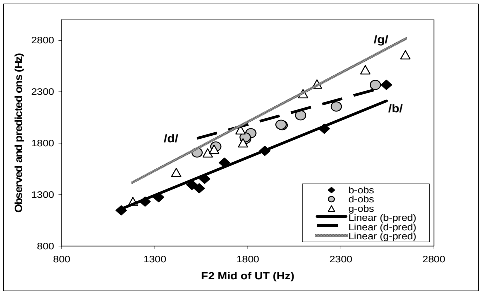
3.5. Isolating the effect of rate on CV$_2$ loci

Figure 3.5a, b, c: Comparisons of "expected" and "observed" F2-CV$_2$onsets



In Fig. 3.5, the mean deviations of 'observed' from 'expected' F2onset values were averaged across the two speakers and plotted against F2mid (Hz) of V$_2$-UT. Fig 3.5 shows, from top to bottom, scatterplots for /b,d,g/ indicating the extent of the deviation of observed F2onsets from the expected values. Except for /bi,bɛ,di, &gi/ all tokens fell below the line of equality (0). In sum observed F2ons exceeded expected values from the MRA.

Figure 3.6: Regression of "expected" and "observed



In Fig 3.6 the straight lines are the expected regression function as predicted by the MRA. The data points are the observed F2-onset values. This plot is a different representation of Fig. 3.5.

/b/, filled diamonds, shows minute deviation for all points, indication minor rate effect for F2onset. For /d/, grayed circle, all vowels except /i/ fall below the expected line. This indicates significant rate effect on F2onset. In /g/, unfilled triangle, almost all data points, except for two front vowels, were lower than their corresponding expected values, thus exhibiting significant rate effect on F2ons.

Were the values in Figures 3.5 regressed in the LE tradition, the resultant higher LE slopes [see Table 3.1] would be interpreted as reflecting increases in the degree of CV$_2$ coarticulation as a result of the vowels moving closer to & exerting greater influence on their C-onset linkages. Such a conclusion requires maintaining that rate only changes the vowel space, while C-loci remained stable. The presented results suggest otherwise.

4. Conclusion

CV$_2$ coarticulation was studied under increased speech tempo using modified LE metrics to isolate the effect of tempo on the CV$_2$onsets separate from V$_2$mid. Results show (i) a disproportional effect of tempo on V and CV$_2$ons, (ii) an adjustment of CV$_2$-ons to rate induced coarticulatory changes. These effects are systematically independent of contextual vowel effects.

5. REFERENCES

[1] Lindblom, B. 1963. "On vowel reduction." Report No. 29, Speech Transmission Laboratory, The Royal Institute of Technology, Sweden.
[2] Sussman, H. M., Fruchter, D., Hilbert, J., and Sirosh, J. 1998. "Linear correlates in the speech signal: The orderly output constraint," Behav. Brain Sci. 21, 241–299
[3] Neary, T. M. & Shammass, S.E. 1987. "Formant transitions as partly distinctive invariant properties in the identification of voiced stops," JASA. 15 (4) 17-24
[4] Krull, D. 1987. "Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech." PERILUS 5: 43-61.
[5] Lindblom, B, Agwuele, A., Sussman, H, & Eir Cortes, L 2007. "The effect of Stress on consonant vowel coarticulation. To appear in JASA 121 (6) 2007.
[6] De Jong K. 1995. "The supraglottal articulation of prominence in English: Linguistic stress as localized hyperarticuation" JASA, 97, 491-504.
[7] Agwuele, A. 2003. "The effect of Stress on consonantal loci" in, Proceedings of the 15th ICPhS. Vol 1, 787-790
[8] Tiffany, W. R. 1959. Non random sources of variation in vowel quality. J. Speech Hear Res. 2, 305-317
[9] Boersma, P. & Weenink, D. 2004. Praat [Version 4.2]. Doing phonetics by computer from http://www.praat.org.
[10] Sussman, H.M., McCaffrey, H.A., & Matthews, S.A. 1991. "An investigation of locus equations as a source of relational invariance for stop place categorization. JASA. 90, 1309 -1325