AN INVESTIGATION OF LOCUS EQUATIONS AS A SOURCE OF RELATIONAL INVARIANCE FOR THE STOP PLACE DIMENSION

H. Sussman

University of Texas at Austin, Texas, U.S.A.

Locus equations, straight line regression fits to data points formed by plotting onsets of F2 transitions along the Y-axis & corresponding midvowel target frequencies along the X-axis, were generated across 20 speakers using speech tokens /b(V)t/, /d(V)t/, & /g(V)t/ with 10 vowel contexts. Slopes & y-intercepts were significantly different across stop place & correctly predicted stop categorization. Locus equations provide a higher-order category-level metric capable of capturing relational invariance for place of articulation.

1. INTRODUCTION

The coarticulatory nature of speech has led to the theoretical impasse known as the "invariance problem"- i.e., perceptual constancy despite physical variation in the signal. Phonetic segments are realized in an overlapped, dynamic, & context-sensitive fashion, while conceptualizations in the abstract depict them as discrete, static, & context-independent. The elusive quest for invariance, the search for stable acoustic cues that isomorphically encode the phonetic segment has been ongoing since the early 1950's. The 'limus test' for invariance has been place of articulation for stop consonants. The purpose of this research was to offer a refocused conceptualization of a traditional candidate for acoustic invariance - the F2 transition as its onset & trajectory vary with the following coarticulated vowel. A formal metric that succinctly captures the relative changes ocurring in F2 transitions will be investigated as a potential, higher-order cue invariantly signalling place of articulation in voiced stops. This metric was initially formulated by Lindblom [3] & termed "locus equations." Despite a surface similarity to the "virtual locus" concept [1], Lindblom's metric was not intended to formalize a fixed & context-independent acoustic correlate of stop place, but rather to illustrate the context-dependence existing between onset frequency of F2 and its location in the vowel nucleus of the syllable. The regression plots showed extreme linearity and tight clustering of data points. Moreover, the slope & y-intercept differed as a function of stop place for Swedish /b/,/d/, & /g/ followed by 8 vowels. One purpose of the present study was to determine if American English stop + vowel syllables would also show the extreme linearity & orderliness exhibited by Lind-

blom's data. Another rationale was that a higher-order linguistic abstraction could be used to investigate the invariance issue. All previous studies have examined acoustic cues derived at & characterizing the single phonetic segment. The locus equation metric is derived over & characterizes an entire stop place category. A nontrivial aspect of the invariance dilemma might very well relate to the proper level of abstractness of the linguistic elements for which the invariant acoustic properties are sought

2. PROCEDURE

2.1 Subjects

Twenty subjects, 10 male & 10 female were used, ranging in age from 18 to 46. Varied dialects of American English were spoken.

2.2 Stimulus materials

Subjects were asked to produce CVC syllables in a carrier phrase format "Say CVC again." Words were typed on a list in five randomized orderings. Initial stops were /b/,/d/, and /g/ followed by 10 medial vowels contexts/i, I, e, , ae, a, o, , , u/. The final consonant was always /t/. Thus, there were 10 /b(v)t/ tokens, 10 /d(v)t/ tokens, and 10 /g(v)t/ tokens, each repeated five times yielding a total of 150 utterances per subject.

2.3 Instrumentation

Each speaker's productions were recorded in a soundproof booth using a high quality microphone & cassette tape recoder. The recorded signal was sampled at 10kHz & digitized using an Apple MacIntosh II computer with MacAdios II hardware. The MacSpeech Lab II package was used for all display, editing, measurement & playback routines. F2 measures were obtained from three sources: (1) direct on-screen wideband spectrograms; (2) LPC spectra; (3) wide & narrowband FFTs.

2.4 Data Sample Points

The two formant measurement points were F2 onset defined as the frequency value of F2 at the first glottal pulse following the release burst and F2 vowel defined as the frequency of F2 at the midvowel nucleus. F2 vowel measurement points were not fixed in time: if F2 was "steady-state" a midpoint on the formant was taken; if F2 was diagonally rising or falling a midpoint position was similarly used; if F2 was 'parabolic' a minima/maxima point was taken for F2 vowel.

3. RESULTS

Sixty locus equation scatterplots were generated. Extremely tight clustering of points about the regression line were found throughout all speakers, regardless of gender. Collapsing across repetitions & subjects, group mean locus equation plots are presented in Figure 1 for initial labial, alveolar, and velar stop place. It can be seen that male & female coordinates lie along the same linear function with female values lying further out each axis. Labial /b/ had the steepest slope (.91) followed by /g/ (slope =.79, and then alveolar /d/ (slope = .54). An ANOVA on both slope & y-intercept parameters revealed significant main effects for place of articulation ($p<.001$).

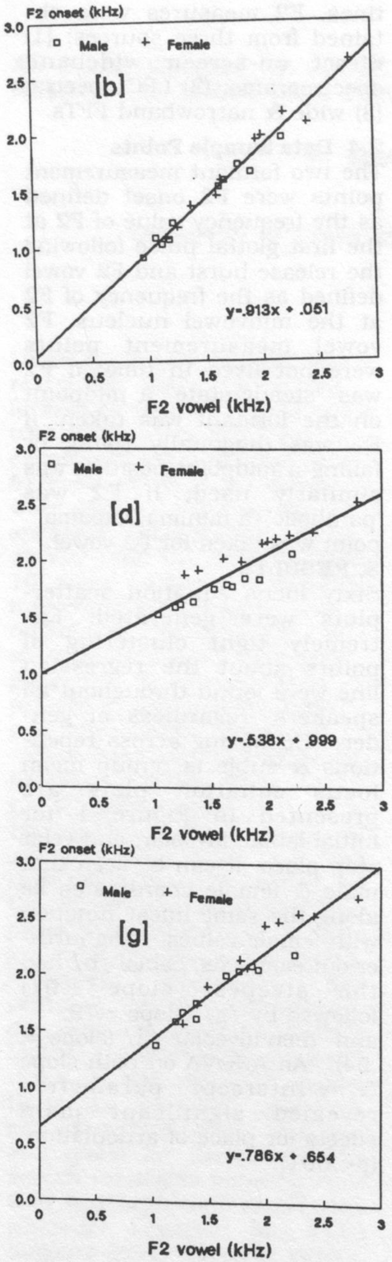

Fig.1: Group mean locus equation scatterplots for /b/, /d/, & /g/.

To summarize up to this point plotting F2 onsets obtained at the CV boundary in relation to midvowel 'target' frequencies (F2 vowel) yields linear relationships that systematically vary with stop place. Thus, despite the extreme context-dependency of the coarticulated CV gesture a relational form of invariance is captured that is independent of the following vowel. At the single CV level no absolute signal invariance is present; only when the higher-order stop + vowel phoneme category is represented acoustically does a relational-type of invariance first emerge.

3.1 Discriminant Analyses

To test the categorical classification success of token-level versus category-level predictor variables a set of linear discriminant function analyses were run. At the single token level F2 onset & F2 vowel frequencies were used as predictors for place of articulation (chance = 33.3%). Percent correct classification rates were 83.1, 79.4, and 67.9 for labials, alveolars, & velars respectively. When the 60 derived slopes & y-intercepts (3 stop place locus equation functions per subject X 20 speakers) were used as predictors 100 % corect classification was obtained.

3.2 CV "Prototypes"

Figure 2 shows canonical group mean regression lines for each stop place category. Velar /g/ is shown broken down into a more accurate subgrouping of allophonic variants of /g/ preceding front vowels (palatal -/g/p) & /g/ preceding back vowels (velar -/g/v). These mean regression functions obtained across 10

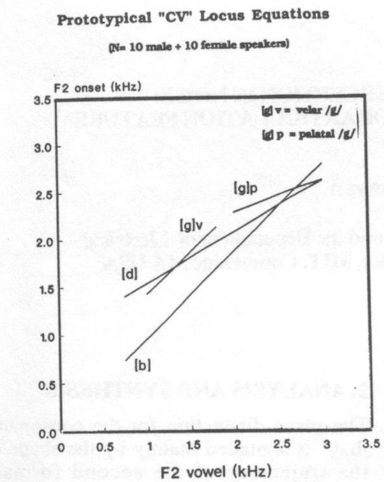

Fig. 2 Locus equation "prototype" functions for /b/, /d/, /g/-velar, & /g/-palatal place of articulation.

male & 10 female speakers are currently conceptualized as representing "prototypes" for CVs, & as such may contain the theoretical framework for understanding & studying the auditory representation of dynamically coarticulated CVs.

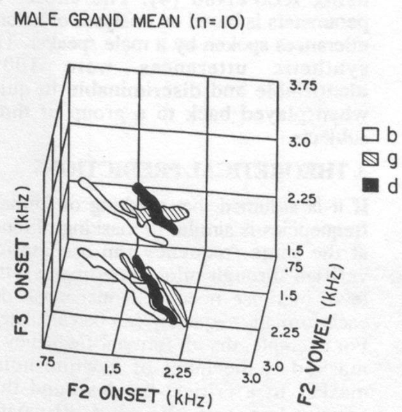

Fig. 3 3-D scatterplot of F2 onset, F2vowel, & F3onset frequencies for /b/, /d/, & /g/ followed by 10 medial vowels.

3.3 3-D Acoustic Space

Adding F3 onset (Hz) to the locus equation parameters provided a 3-D perspective of acoustic/phonetic space. Fig.3 illustrates a typical scatterplot averaged across the 10 male speakers. Distinct & minimally overlapping 'cloud' & respective 'shadow' representations were consistently found for both individual & group data as a function of the 3 stop place categories.

4. DISCUSSION

The data of this study demonstrate acoustic orderliness for stop place categories emerging from both locus equation & 3-D scatterplots of F2 & F3 data. A context- independent phonemic class descriptor, & hence a logical alternative to gestural-related invariance notions [2] has been demonstrated. The acoustic signal, despite coarticulatory complexity, contains a systematic set of correlational attributes of formant information capable of coding place of articulation without recourse to a gestural-level recoding of the signal. Stop place categories are sufficiently contrastive in their physical instantiation in the speech sound wave to permit direct auditory decoding.

5. REFERENCES

[1] DELATTRE. P., LIBERMAN, A.,& COOPER,F. (1955), "Acoustic loci & transitional cues for consonants," *JASA*,27,769-773.
[2] LIBERMAN,A.,& MATTINGLY,I. (1985), "The motor theory of speech perception revised," *Cognition*, 21, 1-36.
[3] LINDBLOM,B. (1963), "On vowel reduction," *STL-No.29*. Stockholm, Sweden.