The impact of vowel length contrasts on locus equations and their implications for perceptual categorization

Mohammad Abuoudeh & Olivier Crouzet

Laboratoire de Linguistique de Nantes, LLING – EA3827, Université de Nantes

{mohammad.abuoudeh|olivier.crouzet}@univ-nantes.fr

Locus equations (LEs) represent the linear spectral relationship between the initial and the mid-point of the second formant in a CV sequence [Lindblöm, 1963; Sussman et al., 1991]. According to Sussman et al. [1991; 2012], LEs may constitute a source of relational invariance for the identification of stops' place of articulation and seem to be a good indicator of the degree of coarticulation in a CV sequence. As a matter of fact, it has been shown that LE coefficients (slope and y-intercept) may be broadly associated to place of articulation categories. In addition, when performing discriminant analyses with either, slopes and y-intercepts, or $F2_{ons}$ and $F2_{mid}$ as the predictors, good to perfect discrimination of stop place was achieved. LEs are also considered as a reliable measure of the degree of coarticulation between a consonant and a vowel. A slope value close to 1 indicates a high degree of coarticulation while a lower slope indicates a lower degree of coarticulation. LEs are used to describe CV interaction in many different speech production phenomena: stress and speaking style, speaking rate, emphatic stress and were usually associated with changes in low vs. high degrees of CV coarticulation.

As it appears that the *place of articulation* hypothesis may be related to changes in *degree of coarticulation*, we have started investigating the impact of temporal variation on LE parameters in speech production [Branzacio and Fowler, 1998; Agwuele et al., 2008;Weismer and Berry, 2003]. In a previous study where we tested the impact of vowel phonological length contrasts on LE parameters in Jordanian Arabic (JA) for 5 consonants /t, t', k, q, ʔ/ coarticulated with long vs. short vowels /i:, a:, u:, i, a, u/, it was observed that slopes and intercepts of consonants associated with long vowels differ significantly from slopes and intercepts of consonants associated with short vowels (respectively $F_{(1;2)} = 43:81$; $p < 0:05$, $F_{(1;3)} = 13:53$; $p < 0:05$). Our analysis showed that the LE slope of a consonant produced with long vowels is lower than the LE slope of the same consonant when it's produced with short vowels, which is in line with the *degree of coarticulation hypothesis*. Indeed, our prediction was that length opposition has duration and spectral changes which would impact the coarticulatory relationship between consonants and vowels. In addition, there is a main effect of consonant category on LE slopes ($F_{(4;8)} = 12:46$; $p < 0:01$), and that is also in line with hypotheses on the *relationship between LEs and place of articulation categories*. There is no interaction involving vowel length in our data. Therefore, the LE data of JA seem to be systematically influenced by alternations in vowel length and this may have strong implications on their relation to place of articulation categories.

Even though our study has provided insights into the relationship between spectral properties of Consonant Vowel sequences and the speech production time-course, it is not clear as to whether these differing slopes and intercepts may hinder place of articulation classification in terms of perceptual mechanisms. Indeed, although changes in mean slopes and intercepts depending of vowel length / duration confirm that LE parameters are related to degrees of coarticulation, this does not necessarily imply that such variation would lead to difficulties in classifying the consonantal place of articulation. The aim of the present work is to investigate this issue more closely.

Capitalizing on available acoustic data discussed in our previous study, a series of linear discriminant analyses (LDA) has been conducted in order to investigate the impact of LE changes induced by temporal variation on place of articulation classification. Two hypotheses have been investigated:

1. Using $F2_{ons}$ and $F2_{mid}$ as predictors for stop place of articulation has generally been shown to provide partially adequate classification cues (around 70% correct categorization). As vowel length impacts LE parameters, we investigated whether adding *vowel duration (in milliseconds)* into an LDA would increase classification accuracy compared to using spectral information only. If this were the case, it would provide arguments in favor of integrating continuous duration information into LE computation rather than F2 spectral measurements only;

2. As it is possible to train LDA on a subset of observations and, based on this training, to predict classification on an alternate subset, it is straightforward to train such a model for place of articulation categorization on either (respectively) short or long vowels and to evaluate this trained model by predicting classification of the complementary subset (respectively long vs. short vowels). If the impact of vowel length / duration on LE parameters is crucial to perceptual classification, we hypothesize that tests predicting *classification outside their training domain* should reach lower levels of accuracy than if they are performed *inside their training domain*;

Both LDA types were performed on a subset of our JA data (namely /t, k, q/ consonants only). It was decided to restrict this analysis to three categories in order to avoid a low classification accuracy that could be induced by a higher number of categories. This choice was based on contrast analyses showing that slopes and intercepts were maximally distinct for these 3 categories only (LE parameters did not significantly differ between, respectively /t/ and /t'/, /q/ and /ʔ/). Performing LDAs on 3 categories was also a means to facilitate comparisons between our work and previous studies involving this technique (e.g. Sussman et al. [1991]) as most of them involved 3 places of articulation.

Jackknife LDA tests for the 3 consonants show a rather low classification percentage compared with classical studies. When only $F2_{mid}$ and $F2_{onset}$ serve as predictors, classification accuracy reaches 48% only (compared to Sussman et al. [1991]'s 70%). Adding the duration as a third factor ($F2_{mid}$ + $F2_{onset}$ + duration) did not cause any increase in classification accuracy (45% correct categorization when these 3 predictors are entered into the analysis). Separating anterior and posterior vowels in two separate subsets broadly increased classification percentage (respectively 69% and 65% when predicting based on $F2_{mid}$ + $F2_{onset}$ alone). Though this manipulation did increase classification performance, adding duration as a third factor did not have any positive impact on classification accuracy. None of these results conforms to our first prediction.

In order to evaluate our second hypothesis, a series of 2-steps LDAs were performed with $F2_{mid}$ and $F2_{onset}$ as predictors with complementary training and test subsets (respectively long and / or short vowels). Test-predictions on the short vowels subset reached 48.9% accuracy when preceded by training on short vowels (in-domain) but only 40.7% accuracy when preceded by training on long vowels (out-domain). Though this observation seems to be in accordance with our second hypothesis, test-predictions for the long vowels subset does not seem to be impacted by the training subset (classification accuracies reach 47.9% and 47.5% for out-domain and in-domain training respectively).

According to our analyses, timing parameters mainly had a negligible effect on consonant classification.

However, the predicted behaviors may be partially hidden by the relatively low classification accuracies observed when compared with traditional studies. This issue may be related to the F2onset position choice, as calculating LE of $F2_{@burst}$ instead of $F2_{onset}$ results in a more accurate classification for voiceless consonants, Modarresi et al. [2005]. That could explain the relatively high slope values of our data and the low classification rate comparing to precedent studies. Further study is planned to calculate the LE of $F2_{@burst}$ instead of $F2_{onset}$, which would provide a better consonant classification.

References

A. Agwuele, H. Sussman, and B. Lindblöm. The effect of speaking rate on consonant vowel coarticulation. *Phonetica*, 65(4):194–209, 2008.

L. Branzacio and C. A. Fowler. On the relevance of locus equations for production and perception of stop consonants. *Perception & Psychophysics*, 60(1):24–50, 1998.

B. Lindblöm. Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35(11):1773–1781, November 1963.

B. Lindblöm and H. M. Sussman. Dissecting coarticulation: How locus equations happen. *Journal of Phonetics*, 40(1):1–19, January 2012.

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*, 90:1309–1325, November 1991.

G. Weismer and J. Berry. Effects of speaking rate on second formant trajectories of selected vocalic nuclei. *The Journal of Acoustical Society of America*, 113(6):3362–3378, June 2003.