The effects of measurement error and vowel selection on the locus equation measure of coarticulation[a)]

Allen Montgomery
Department of Communication Sciences and Disorders, University of South Carolina, Columbia, South Carolina 29208

Paul E. Reed[b)]
Linguistics Program, University of South Carolina, Columbia, South Carolina 29208

Kimberlee A. Crass and H. Isabel Hubbard
Department of Communication Sciences and Disorders, University of South Carolina, Columbia, South Carolina 29208

Joanna Stith
Listen Foundation, Greenwood Village, Colorado 80111

(Received 10 December 2013; revised 18 July 2014; accepted 10 September 2014)

Effects on the slope of introducing error in the F2 Hz values in locus equations (LEs) and of using fewer than ten vowels were investigated. For each of the initial consonants /b, d, g/, 2000 simulated sets were generated using Monte Carlo techniques. The sets were altered with 50, 100, or 200 Hz error being randomly applied to each F2 value within a set. Selected vowels were then removed from the sets and the effects on the slopes were measured. Results suggest that the LE slopes are generally resistant to error and reduced number of vowels. Effects of adding 50 Hz of random error to the F2 values in sets using eight or ten vowels were minimal, yielding a mean absolute change in slope less than 0.07. © *2014 Acoustical Society of America*. [http://dx.doi.org/10.1121/1.4896460]

PACS number(s): 43.70.Bk, 43.70.Jt [CHS] Pages: 2747–2750

I. INTRODUCTION

A. Description of locus equation procedure to measure coarticulation

The locus equation (LE) is designed to measure the anticipatory effect of the vowel on the initial consonant in consonant-vowel-consonant (CVC) words. The initial consonants /b, d, g/ are widely used in combination with up to ten vowels in CVCs. Its value lies in its use of acoustic recordings of simple syllables and measurement of second formant (F2) frequencies. It is obviously an important contribution to the study of speech production, and is in active use (Iskarous *et al*. 2010; Iskarous *et al*. 2013; Lindblom and Sussman, 2012; Morrison, 2012; Rhone and Jongman, 2012; Noiray *et al*. 2013).

A typical LE is derived by analyzing recordings of a talker saying CVC monosyllables with a constant initial (and final) consonant in ten different vowel contexts. Values for the second formant (F2) at the onset of voicing and at mid vowel are then plotted against each other and the slope is calculated. To the extent that the vowel influences the articulation of the initial consonant, the linear slope is considered to be a measure of the amount of anticipatory coarticulation.

There are several ways to measure formant frequencies. Regardless of the source of the formant frequencies, however, the measurements contain errors usually of unknown magnitude and certainly varying with the method and supervisor (Alku *et al*., 2013; Duckworth *et al*., 2011; Zhang *et al*., 2013). Even with modern analytic software such as PRAAT (Boersma and Weenink, 2013), SPEECH ANALYZER (SIL International, 2012; Morrison and Nearey, 2011; Neary *et al*., 2002), or WAVESURFER (Sjolander and Beskow, 2013), errors may occur due to high fundamental frequency, poor recordings, and uncertainty concerning the initiation of voicing and the center of the vowel. These sources of error, whose statistical properties and occurrence are unknown, combine to produce the F2 error simulated in this study.

This study was designed to examine the stability of the LE slopes when random error of 50, 100, or 200 Hz is applied to each of the F2 values in LE sets (20 F2 values from onset and mid-vowel for ten vowels) in LE sets for initial /b/, /d/, and /g/ and when the number of vowels in the LE set is reduced systematically from 10 to 3.

Note that this study assesses only the effects of random error. Some forms of systematic error in the second formant frequency typically would not affect the slope of the locus equation, such as being approximately 70 Hz low on every measurement. However, more complex forms of systematic error, not modeled in this study, such as being low on high-frequency formants and high on low-frequency formants, would yield incorrect slope values.

II. METHODS

A. Generation of simulated locus equation sets of 20 F2 values

For each of three consonantal contexts the mean of the absolute values of changes in slope was obtained using a set

[a)]Portions of this work were presented at the American Speech and Hearing Association, Atlanta, November, 2012.
[b)]Author to whom correspondence should be addressed. Electronic mail: reedpe@email.sc.edu

of 2000 LEs, derived from existing data through Monte Carlo methods. Care was taken to ensure that the statistical descriptions of the resulting simulated distributions of 2000 LE sets for each consonantal context closely matched those of existing slopes obtained from the literature. A search yielded a total of 553 slopes, 161, 236, and 156 for /b/, /d/, and /g/, respectively. A complete listing of studies involving LEs is available at http://www.sph.sc.edu/comd/facultystaffdetails.php?ID = 369.

The generation of the collection of the simulated LE sets involved three steps: First, slopes derived from LEs using initial /b/, /d/, or /g/ monosyllables were collected from the studies in the literature. This allowed statistical distributions of slopes to be prepared for each consonant, hereafter called *existing* slopes. Second, 75 LE sets were generated through acoustic analysis of recordings for each consonant, forming *new* LE sets with 20 F2 values for each set. Finally Monte Carlo techniques were applied to each of the new LE sets to yield 2000 *simulated* LE sets for each consonant.

B. Generation of new LE sets

The new 75 LE sets based on 10 vowels was compiled through either recording normal speaking adults or through transcription of LE sets available in our laboratory. (Note that F2 frequency values in some new LE sets were adjusted to correct vowel outliers and to add missing data to achieve similarity to the distribution of existing slopes.) It was found that none of the three new mean slopes differed significantly from the corresponding existing mean slope (t-tests less than 1.2, probabilities > 0.24 meeting the criterion of a lenient α level of 0.20). In addition, all six distributions (the existing data and the new 75-set data for each consonant) satisfied the criteria for normality in that neither the skewness nor the kurtosis exceed its standard error (SE) by more than 2:1 (Myers *et al.*, 2010, p. 37) and all had Shapiro–Wilk probabilities greater than 0.20, again indicating the distributions were approximately normal. The new LE sets provide the basis for generating 2000 simulated LE sets for further study.

C. Generation of simulated LE sets

The 75 new LE sets based on each consonant were then subjected to Gaussian perturbation of F2 frequency. Gaussian random numbers were obtained from an EXCEL VBA program written for the study and had a mean of 0.003 and a SD of 0.9984 with normal skewness and kurtosis properties, as above—thus the values were normally distributed and ranged from approximately −3 to +3. Each F2 value in each LE set of 20 frequency values (F2-onset and F2-mid-vowel for 10 vowels) was altered through the addition (sometimes a negative number) of Gaussian random noise that averaged 5% of the F2 value. An F2 value (F) was altered with a distinct random number (R) by applying this formula to produce a perturbed F2 value (F_p),

$$F_p = F + 0.05RF. \quad (1)$$

For example, a value of 1750 Hz in an LE set in the 75 new sets and a random number of $R = -1.50602$ would yield: $F_p = 1750 + (0.05)(-1.50602)(1750) = 1618.2$ Hz. This process was repeated to produce 2000 LE sets of simulated 20 F2 values for each of the three consonants. A total of 40 000 unique random numbers was used in the generation of the simulated LEs for each consonant (20 F2 values for 2000 LE sets). The resulting distributions of the slopes were approximately normal, again reflecting the normal distributions of the existing slopes. The sets are now considered 2000 error-free samples of their respective populations based on ten vowels per LE, and formed the basis of all subsequent experimental manipulations. The manipulations involved adding known amounts of F2 error to all F2 values and reducing the number of vowels in the LE sets.

The magnitude of effects of error is estimated by calculating means of the absolute value of the change in individual slopes in the collection of equations. The mean of the absolute amount of change is reflective of the effect of the added error and vowel reduction and allows meaningful comparisons across conditions.

However, the question remains as to what is a *meaningful* amount of change in the mean slope. A search of the literature revealed comparisons of 24 mean differences due to experimental conditions that were not statistically significant (or were interpreted as not meaningful). Many of these studies used t-tests and analyses of variance with conservative *post hoc* tests to compare means of locus equation slopes (Baillargeon *et al.*, 2002; Sussman *et al.*, 2011; Sussman, *et al.*, 1998; Sussman, *et al.* 1999; Sussman *et al.*, 1995; Sussman *et al.*, 1997; Sussman *et al.*, 1998; Sussman *et al.*, 1991; Sussman and Shore, 1996). The differences in mean slope ranged from 0.01 to 0.11 with a mean of 0.06 slope units. On the other hand, 12 differences starting at 0.13 and higher were statistically significantly different and a conservative "cut-off" point of 0.10 slope units was selected to aid in interpretation of the confidence intervals generated with simulated data to assist in future studies.

However, the primary measure of the effect of the various experimental conditions below was the mean *absolute* difference of the change in slope in corresponding LEs in the sets of 2000 LEs. Accordingly, it was necessary to derive a similar cut-off value for the interpretation of a meaningful amount of influence on the absolute-value means. To do this the value of 0.1 as a cut-off for the interpretation of a meaningful difference between the signed means of two groups (above) was then used as the basis for interpreting the magnitude of the absolute differences obtained. The calculation is derived from the interpretation of confidence intervals described by Cumming (2012, pp. 178, 179).[1] The value for generating confidence intervals (CIs) that was obtained was 0.067. That is, if the result of adding error and reducing the vowel set to the F2 values in the LEs led to a mean based on the absolute values of greater than 0.067, it was considered a meaningful effect of experimental conditions.

III. INCLUSION OF ERROR WITH VOWEL REMOVAL

A. Procedure

Prior to the study two pilot studies were conducted to determine the effects of F2 error or vowel reduction independently. The results indicated that the LEs were relatively robust when either form of distortion was applied separately.

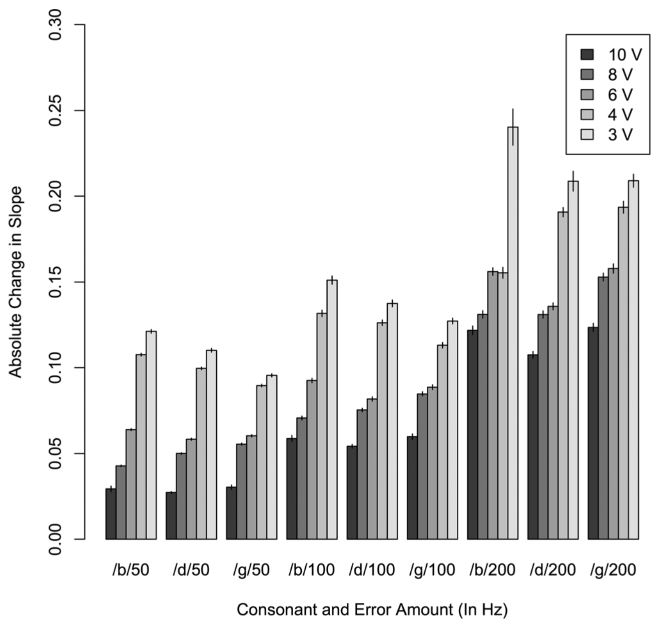

FIG. 1. The change in the mean absolute value of locus equations slopes affected by selective vowel removal and addition of 50, 100, or 200 Hz random F2 error. Each bar represents the application of only one of the three amounts of error. Change in slope is calculated by subtracting each slope in each condition from its corresponding slope based on ten vowels and no added error. Standard error bars are based on 2000 locus equations in each condition.

The values and procedures used in the pilot studies were then applied in the following study of the combined effects. Results of pilot studies can be found at http://www.sph.sc.edu/comd/facultystaffdetails.php?ID = 369.

The error-free sets of 2000 LEs for each consonant were perturbed with random positive or negative application of 50 Hz error to each of the F2 values. Then F2 values in the error-free sets were similarly perturbed with 100 Hz error and then with 200 Hz error, resulting in three sets representing constant amounts of randomly applied error. The values for random error were chosen to represent modest, normal, and extreme amounts of error. The values of 50, 100, and 200 Hz were based on Duckworth *et al.*, (2011).

Second, the typical number of vowels in the majority of LE sets in the literature was ten vowels (e.g., Sussman *et al.* 1991). To determine the effect of smaller vowel sets the number of vowels was systematically reduced to 3: (8 [i, e, ɛ, æ, ɑ, ɔ, o, u], 6 [i, e, æ, ɑ, o, u], 4 [i, æ, ɑ, u], and 3 [i, ɑ, u] vowels). Thus 12 sets of 2000 slope values were obtained for each consonant (four levels of vowel reduction with three amounts of random frequency error).

B. Results and discussion

The results are shown in Fig. 1. The LE measure is quite robust in the face of small amounts of error, and the eight-vowel sets reflect this stability. However, the introduction of moderate (100 Hz) and large (200 Hz) error yielded absolute amounts of change above the criterion of 0.067 slope units for all reduced vowel sets sizes (compared to the corresponding slopes based on all 10 vowels.) For example, in 8-vowel sets, the addition of 50 Hz error produced mean absolute differences of 0.043, 0.050, and 0.055 for /b/, /d/, and /g/, respectively; however, 100 Hz error caused 0.071, 0.075, 0.084 absolute change. Furthermore, 200 Hz perturbation caused 0.131, 0.131, and 0.152 change for /b/, /d/, and /g/, respectively. In addition, the use of six vowels also appears to be safe if F2 measurement error is 50 Hz or less, but all smaller vowel-set sizes and larger amounts of error are likely to produce unstable mean slope values. These conclusions are relevant only to LEs using the specific vowels selected in the simulations, which carefully retained the cardinal corner vowels /i/, /a/, and /u/ in the vowel sets.

C. Implications for experimental design

The simulations reported here may be helpful in the design of LE studies, if a laboratory is able to estimate or assume the amount of F2 error that is likely to be present in the measurements. Using the 2000 LEs slopes in each experimental condition, standard deviations were used to derive 95% confidence intervals (CIs) about the mean for realistic sample sizes. Table I shows the CIs for the /d/ context which is representative of the other consonantal contexts. Data for the /b/ and /g/ contexts can be found at the supplement website at http://www.sph.sc.edu/comd/facultystaffdetails.php?ID = 369. Each number in the lower section of Table I is the size of the one-half CI for a given condition and sample size. It is clear that the increasing reduction of the number of vowels used in the locus equations and the increasing addition of F2 error resulted in larger standard deviations and thus larger confidence intervals, with the effects of adding 200 Hz F2 error being especially obvious. These results may be used as

TABLE I. Means, standard deviations (absolute values), and one-half confidence intervals (CIs) in slope units for conditions of vowel reduction and added F2 Hz error in locus equations with initial consonant /d/.

No. of vowels		10 V	10 V	10 V	8 V	8 V	8 V	6 V	6 V	6 V	4 V	4 V	4 V	3 V	3 V	3 V
Random error (Hz)		50	100	200	50	100	200	50	100	200	50	100	200	50	100	200
Mean		0.408	0.403	0.377	0.403	0.370	0.373	0.414	0.406	0.383	0.424	0.418	0.401	0.441	0.444	0.430
StDev		0.151	0.161	0.189	0.169	0.172	0.208	0.173	0.183	0.216	0.198	0.217	0.281	0.220	0.243	0.344
Sample size																
$\frac{1}{2}$ CI	2000	0.006	0.006	0.007	0.006	0.006	0.008	0.006	0.007	0.008	0.007	0.008	0.010	0.008	0.009	0.013
	50	0.036	0.038	0.045	0.040	0.041	0.049	0.041	0.043	0.051	0.047	0.051	0.067	0.052	0.058	0.081
	25	0.052	0.055	0.065	0.058	0.059	0.071	0.060	0.063	0.074	0.068	0.075	0.097	0.076	0.084	0.118
	15	0.070	0.074	0.087	0.078	0.079	0.096	0.080	0.085	0.100	0.091	0.100	0.130	0.102	0.112	0.159
	10	0.090	0.096	0.113	0.101	0.102	0.124	0.103	0.109	0.129	0.118	0.129	0.168	0.131	0.145	0.205

follows: If an experimenter wants to use eight vowels and is willing to assume that no more than 50 Hz F2 error will be present, then using 15 participants in the study is estimated to result in a 95% confidence interval of +/−0.078 slope units. Finally, note that the addition of randomly applied 50, 100, or 200 Hz error to *every* F2 value in locus equations is a very thorough test of the derived locus equation slopes. Taking this assumption into account, these conclusions can be considered to be conservative estimates of the stability of the locus equation in general.

[1]Noting a strong relationship ($r = 0.95$) between the experimentally derived absolute means and the SDs of the 2000 slope values in each condition, it was appropriate to derive a cut-off for the absolute means based on our predictions of real-world CIs. To arrive at the cut-off of 0.067, we followed Cumming (2012, pp. 178,179). Cumming found that when 95% CIs about independent means overlap by one-half, the resulting *t*-tests of the means will be significant at approximately the $p = 0.05$ level. Since our values are one half of the full 95% CIs, the following equation was used: $CI/2 + CI = 0.10$. The 0.10 was the assumed difference deemed to be important and meaningful in previous studies. Solving for CI equals 0.067. Note that this number is independent of sample size and is dependent only on the assumed meaningful difference in independent means.

Alku, P., Pohjalainen, J., Vainio, M., Laukkanen A-M., and Story, B. (**2013**). "Formant frequency estimation of high-pitched vowels using weighted linear prediction," J. Acoust. Soc. Am. **134**, 1295–1313.

Baillargeon, M., McLeod, A., Metz, D. E., Schiavetti, N., and Whitehead, R. L. (**2002**). "Preservation of second formant transitions during simultaneous communication: A locus equation perspective," J. Commun. Disord. **35**, 51–62.

Boersma, P., and Weenink, D. (**2013**). "Praat: doing phonetics by computer (version 5.3.52) [computer program]," http://www.praat.org/ (Last viewed 6/24/2013).

Cumming, G. (**2012**). *Understanding New Statistics: Effect Sizes, Confidence Intervals and Meta-analysis* (Routledge, New York), pp. 158–159.

Duckworth, N. M., Mcdougall, K., de Jong, G., and Shockey, L. (**2011**). "Improving the consistency of formant measurement," Int. J. Speech Lang. Law. **18**, 35–51.

Iskarous, K., Fowler, C. A., and Whalen, D. H. (**2010**). "Locus equations are an acoustic expression of articulator synergy," J. Acoust. Soc. Am. **128**, 2021–2032.

Iskarous, K., Mooshammer, C., Hoole, P., Recasens, D., Shadle, C., Saltzman, E., and Whalen, D. H. (**2013**). "The coarticulation/invariance scale: Mutual information as a measure of coarticulation resistance, motor synergy, and articulatory invariance," J. Acoust. Soc. Am. **134**, 1271–1282.

Lindblom, B., and Sussman, H. M. (**2012**). "Dissecting coarticulation: How locus equations happen," J. Phonetics **40**, 1–19.

Morrison, H. M. (**2012**) "Coarticulation in early vocalizations by children with hearing loss: a locus perspective," Clin. Linguist. Phon. **26**, 288–309.

Morrison, H. M., and Nearey, T. M. (**2011**) "Formant measurer: Software for efficient human-supervised measurement of formant trajectories (release 2011-05-26) [computer program]," http://geoff-morrison.net/#FrmMes (Last viewed 6/15/2013).

Myers, J. L., Well, A. D., and Lorch, Jr. R. F. (**2010**). *Research Design and Analysis*, 3rd ed. (Routledge, New York), p. 37.

Nearey, T. M., Assman, P., and Hillenbrand, J. M. (**2002**). "Evaluation of a strategy for automatic formant tracking," J. Acoust. Soc. Am. **112**, 2323.

Noiray, A., Me′nard, L., and Iskarous K. (**2013**) "The development of motor synergies in children: Ultrasound and acoustic measurements," J. Acoust. Soc. Am. **133**, 444–452.

Rhone, A. E., and Jongman, A. (**2012**). "Modified locus equations categorize stop place in a perceptually realistic time frame," J. Acoust. Soc. Am. **131**, EL487–EL491.

SIL International (**2012**). "Speech Analyzer 3.1 [computer program]," SIL.org/computing (Last visited 1/28/2013).

Sjolander, K., and Beskow, J. (**2013**). "WaveSurfer 1.8.8 [computer program]," http://www.speech.kth.se/wavesurfer/ (Last viewed 12/9/2013).

Sussman, H. M., Bessell, N., Dalston, E., and Majors, T. (**1997**). "An investigation of stop place of articulation as a function of syllable position: A locus equation perspective," J. Acoust. Soc. Am. **101**, 2826–2838.

Sussman, H. M., Byrd, C. T., and Guitar, B. (**2011**). "The integrity of anticipatory coarticulation in fluent and non-fluent tokens of adults who stutter," Clin. Linguist. Phon. **25**, 169–186.

Sussman, H. M., Duder, C., Dalston, E., and Cacciatore, A. (**1999**). "An acoustic analysis of the development of CV coarticulation: A case study," J. Speech Lang. Hear. Res. **42**, 1080–1096.

Sussman, H. M., Fruchter, D., and Cable, A. (**1995**). "Locus equations derived from compensatory articulation," J. Acoust. Soc. Am. **97**, 3112–3124.

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.

Sussman, H. M., McCaffrey, H. A., and Matthews, S. A. (**1991**). "An investigation of locus equations as a source of relational invariance for stop place categorization," J. Acoust. Soc. Am. **90**, 1309–1325.

Sussman, H. M., and Shore, J. (**1996**). "Locus equations as phonetic descriptors of consonantal place of articulation," Atten. Percept. Psychophys. **58**, 936–946.

Zhang, C., Morrison, G. S., Ochoa, F., and Enzinger, E. (**2013**). "Reliability of human-supervised formant-trajectory measurement for forensic voice comparison," J. Acoust. Soc. Am. **133**, EL54–EL60.