
Formant Centralization Ratio: A Proposal for a New Acoustic Measure of Dysarthric Speech

RESEARCH NOTE

Shimon Sapir
University of Haifa, Haifa, Israel

Lorraine O. Ramig
Jennifer L. Spielman
University of Colorado at Boulder and
National Center for Voice and Speech,
Denver, CO

Cynthia Fox
National Center for Voice and Speech,
Denver, CO

Purpose: The vowel space area (VSA) has been used as an acoustic metric of dysarthric speech, but with varying degrees of success. In this study, the authors aimed to test an alternative metric to the VSA—the *formant centralization ratio* (FCR), which is hypothesized to more effectively differentiate dysarthric from healthy speech and register treatment effects.

Method: Speech recordings of 38 individuals with idiopathic Parkinson's disease and dysarthria (19 of whom received 1 month of intensive speech therapy [Lee Silverman Voice Treatment; LSVT LOUD]) and 14 healthy control participants were acoustically analyzed. Vowels were extracted from short phrases. The same vowel-formant elements were used to construct the FCR, expressed as $(F2u + F2a + F1i + F1u) / (F2i + F1a)$, the VSA, expressed as $ABS[(F1i \times (F2a - F2u) + F1a \times (F2u - F2i) + F1u \times (F2i - F2a)] / 2$, a logarithmically scaled version of the VSA (LnVSA), and the $F2i/F2u$ ratio.

Results: Unlike the VSA and the LnVSA, the FCR and $F2i/F2u$ ratio robustly differentiated dysarthric from healthy speech and were not gender sensitive. All metrics effectively registered treatment effects and were strongly correlated with each other.

Conclusion: Albeit preliminary, the present findings indicate that the FCR is a sensitive, valid, and reliable acoustic metric for distinguishing dysarthric from unimpaired speech and for monitoring treatment effects, probably because of reduced sensitivity to interspeaker variability and enhanced sensitivity to vowel centralization.

KEY WORDS: Parkinson's disease, speech acoustics, dysarthria, formant centralization ratio, vowel space area, $F2i/F2u$ ratio

Acoustic analysis has the potential of providing quantitative, objective, and precise means to help depict the presence, severity, and characteristics of motor speech disorders and to help monitor deterioration or improvement in speech with disease progression, recovery, or treatment effects (e.g., Kent, Weismer, Kent, Vorperian, & Duffy, 1999). The rationale for using acoustic analysis to assess motor speech function is straightforward: The speech signal contains measurable acoustic parameters that are lawfully related to some aspects of speech production and perception (Fant, 1960; Honda & Kusakawa, 1997). Thus, by studying speech acoustics one can make reasonable inferences about

Disclaimer

Lorraine O. Ramig and Cynthia Fox have ownership interest in LSVT Global LLC (a for-profit organization that runs training courses and sells products related to LSVT treatment). All members of this research team have fully disclosed any conflict of interest, and their conflict-of-interest management plan was approved by the Office of Conflict of Interest and Commitment at the University of Colorado at Boulder.

motor speech functions, normal and abnormal. Yet, as Kent and Kim (2003) commented, "Acoustic analysis, like any method, carries its own interpretative challenges and limitations, all the more so when it is applied to disordered speech with varying degrees of severity" (p. 428).

The present study deals primarily with issues and potential alternatives related to acoustic methods of measuring vowel articulation impairment in individuals with dysarthria secondary to idiopathic Parkinson's disease (IPD); however, the information gathered from this study may have implications for other types of dysarthria, such as those associated with amyotrophic lateral sclerosis (ALS), multiple sclerosis, traumatic brain injury, and cerebral palsy. The most relevant acoustic parameters for the perception and production of vowels are the frequencies of the first two formants, F1 and F2 (Hillenbrand, Getty, Clark, & Wheeler, 1995). These formant frequencies change in a fairly predictable way as a function of the movements of the articulators and as a function of changes in the three-dimensional configuration of the vocal tract that result from these articulatory movements. In general, the frequency of F2 increases, and that of F1 decreases, as the tongue moves forward (e.g., to form the vowel /i/), and the frequency of F2 decreases as the tongue moves backward (e.g., to form the vowels /u/ and /a/). Also, the frequency of F1 decreases when the tongue is elevated (e.g., to form the vowels /i/ and /u/) and increases when the tongue is lowered, alone or in concert with a downward movement of the jaw (e.g., to form the vowel /a/). Furthermore, the frequencies of both F1 and F2 decrease when the lips are rounded (e.g., to form the vowel /u/) and increase when the lips are retracted or become unrounded (e.g., to form the vowels /i/ and /a/; Kent et al., 1999).

Most types of dysarthria are characterized by *articulatory undershoot*, that is, a reduced range of articulatory movements, to the extent that the intended place and degree of vocal tract constriction are not fully achieved (Kent & Kim, 2003). This undershoot is likely to result in *vowel formant centralization*—that is, formants that normally have high frequencies tend to have lower frequencies, and formants that normally have low frequencies tend to have higher frequencies (Sapir, Spielman, Ramig, Story, & Fox, 2007). One common way to represent this centralization is with the *vowel space area* (VSA; Kent & Kim, 2003). Because of articulatory undershoot and consequent centralization of vowels, the VSA in the speech of individuals with dysarthria is expected to be compressed relative to that of normal speech (Kent & Kim, 2003). Improvement in speech that is due to natural recovery or treatment effects should be reflected in the expansion of the VSA toward normalcy (Sapir et al., 2003). Also, whereas conversational speech is likely to be characterized by some amount of articulatory undershoot, formant

centralization, and reduced VSA (cf. Fourakis, 1991), clear speech and hyperarticulated speech are likely to be characterized by increased articulatory precision, VSA expansion, and improvement in speech intelligibility (Ferguson & Kewley-Port, 2007).

In English, the VSA is usually constructed by the Euclidean distances between the F1 and F2 coordinates of the corner vowels /i/, /u/, and /a/ (triangular VSA), or the corner vowels /i/, /u/, /a/, and /æ/ (quadrilateral VSA) in the F1–F2 plane (Kent & Kim, 2003). In this study, we used the triangular VSA with the vowels /i/, /u/, and /a/. We also used a logarithmic version of this VSA (henceforth *LnVSA*), which means that the formant frequencies of the three vowels are logarithmically scaled (with a natural logarithm, or *Ln*) before the VSA is constructed. Such logarithmic scaling is important for reducing interspeaker variability, something we discuss later in this article. The mathematical expressions of the triangular VSA and LnVSA and the explanations for the logarithmic scaling as a means to transform differences in formant frequencies to ratios of formant frequencies are provided in detail in the Appendix.

Several studies have documented centralization of formants and/or compression of VSA in speakers with dysarthria (e.g., Liu, Tsao, & Kuhl, 2005; Weismer, Jeng, Lares, Kent, & Kent, 2001; Ziegler & von Cramon, 1983). Some of these studies have also demonstrated statistically significant positive correlations between VSA and speech intelligibility scores (e.g., Liu et al., 2005; Weismer et al., 2001, in individuals with ALS). Expansion of vowels and VSA following natural recovery or effective treatment also has been documented (e.g., Sapir et al., 2003; Ziegler & von Cramon, 1983). However, some studies have failed to find statistically significant differences between dysarthric and normal speech on some vowel acoustic measures, including VSA, although an overall trend toward centralization of vowels in the dysarthric speech was evident (e.g., Bunton & Weismer, 2001; Sapir et al., 2007; Weismer et al., 2001, in individuals with IPD). Moreover, in some studies the VSA accounted for only 6%–13% of the variance in measures of speech intelligibility (McRae, Tjaden, & Schoonings, 2002; Tjaden & Wilding, 2004).

The reasons for the inconsistent performance of the VSA are not clear. One likely explanation is that the VSA is highly sensitive to interspeaker variability, and this variability might mask, statistically speaking, true differences between dysarthric and normal speech. Interspeaker variability in vowel formant frequencies and VSA is assumed to be due to anatomical and physiological differences, such as those associated with gender and age (e.g., size and shape of the vocal tract; Hashi, Westbury, & Honda, 1998; Yang, 1996); idiosyncratic strategies of posturing the articulators (e.g., habitually speaking with a relatively fronted or retracted tongue posture for all

vowels; habitually coupling or decoupling lip rounding with tongue backing in the formation of /u/; e.g., de Jong, 1997; Hashi et al., 1998); idiosyncratic differences in interarticulatory coordination or coarticulation (de Jong, 1997); and idiosyncratic differences in vowel perception (e.g., discrimination or prototypic preference), this last having been shown to affect the vowel production map unique to the individual (Perkell et al., 2004). Other factors that might affect interspeaker variability include severity and/or pathophysiology of the dysarthria, idiosyncratic compensatory adjustments to the dysarthria, the nature of the speech task, the phonetic environment in which the vowels in the VSA are measured, and the specific methods of measuring the vowels (Rosen, Goozée, & Murdoch, 2008; Yunusova, Weismer, Westbury, Lindstrom, 2008). Given these facts, it is clear that, to improve differentiation of dysarthric from normal speech, the acoustic metric must be minimally affected by speaker-related variability and maximally affected by the articulatory impairment, as reflected by vowel formant centralization or other acoustic indexes that closely represent the impairment.

In this study, we wanted to test an acoustic metric we have developed that has been designed to maximize sensitivity to vowel centralization and minimize sensitivity to interspeaker variability. We call this metric the *formant centralization ratio* (FCR). The FCR is expressed as $(F2u + F2a + F1i + F1u) / (F2i + F1a)$, where F2u is the frequency of the second formant of the vowel /u/, F1i is the frequency of the first formant of the vowel /i/, and so on. The FCR is designed so that the formant frequencies in the numerator are likely to increase, and the formant frequencies in the denominator are likely to decrease, with vowel centralization. This arrangement should maximize sensitivity to vowel centralization (i.e., the FCR should increase with centralization and decrease with vowel expansion).

Note also that the FCR is expressed as a ratio. The expression of vowel formants as a ratio is one of the normalization procedures that have been used to reduce speaker-related variability in vowel perception studies

(Adank, Smits, & van Hout, 2004). Here we use it not with respect to vowel perception but as a simple way to reduce interspeaker variability in formant frequencies (Yang, 1996). When vowel formants are expressed as a ratio, the value of this ratio is likely to be similar across speakers, even though the formant frequencies of the same vowel across speakers are different. Thus, for example, if a man's F1i = 300 Hz and F2i = 2400 Hz, a woman's F1i = 350 Hz and F2i = 2800 Hz, and a child's F1i = 400 Hz and F2i = 3200 Hz, the ratio F2i/F1i will be the same for all speakers ($2400 / 300 = 8$, $2800 / 350 = 8$, $3200 / 400 = 8$), in spite of relatively large differences in F2i and F1i across the speakers. In fact, the coefficient of variation (CV) in this specific example is 0% ($CV = SD / M = 0 / 8 = 0\%$). Note also that if we replace the division operator ($F2i / F1i$) with a subtraction operator ($F2i - F1i$), the variance across the three speakers is much larger ($2400 - 300 = 2100$, $2800 - 350 = 2450$, $3200 - 400 = 2800$, mean difference = 2450, $SD = 350$, $CV = 350 / 2450 = 14\%$). Thus, although $F2i - F1i$ can also reflect vowel centralization, the interspeaker variability associated with it is much larger than that with the F2i/F1i ratio.

The effects of the FCR on interspeaker variability are considered in Table 1. These are average vowel formant frequencies of men, women, and children obtained from Hillenbrand et al.'s (1995) study (Table V, p. 3103). To the right of the formant data are the results of the VSA and FCR metrics applied to these data. Note that the formant frequencies in the children are higher than those of women, and those of women are higher than those of men, as would be expected from the anatomical differences in the vocal tract dimensions associated with gender and age. Note also that when the VSA is applied to the respective formant frequencies in each of the groups, it is highest in children, lowest in men, and in between for women. Thus, the VSA is highly sensitive to these group differences. This sensitivity is also reflected by the relatively large CV (26%) shown in Table 1. In contrast, when the FCR is applied to these formant data, there is little difference among the three groups, and the CV value is 1%. Thus, the FCR dramatically reduces interspeaker

Table 1. Mean formant data of men, women, and children from Hillenbrand et al.'s (1995) study.

Group	F1i (Hz)	F2i (Hz)	F1a (Hz)	F2a (Hz)	F1u (Hz)	F2u (Hz)	VSA (Hz ²)	FCR
Men	342	2322	768	1333	378	997	264423	0.99
Women	437	2761	936	1551	459	1105	399862	0.96
Children	452	3081	1002	1688	494	1345	448147	0.97
M	410	2721	902	1524	444	1149	370811	0.97
SD	60	381	121	179	60	178	95245	0.01
CV (%)	15	14	13	12	13	16	26	1

Note. The results of applying the triangular vowel space area (VSA) and formant centralization ratio (FCR) to these data are shown in the two rightmost columns. CV = coefficient of variation.

variability. Note also that the FCR values across men, women, and children in Table 1 are near 1.0. The fact that the FCR has values near 1.0 across these different groups of speakers suggests that it is insensitive, or only minimally sensitive, to gender and age effects. It also suggests that, across speakers (at least of the American English language), the sum of the frequencies of the formants in the numerator is very similar to that of the denominator. The FCR also has asymptotic meaning; specifically, in the extreme case of vowel centralization, the formants of the vowels /i/, /u/, and /a/ should collapse onto one location in the F1–F2 plane, whereby $F1i = F1u = F1a$ and $F2i = F2u = F2a$. In terms of the FCR formula, this means that the maximum FCR value should be 2, as indicated here:

$$\text{FCR} = (F2 + F2 + F1 + F1)/(F2 + F1) \quad (1)$$

$$= (2F2 + 2F1)/(F2 + F1) \quad (2)$$

$$= 2(F2 + F1)/(F2 + F1) \quad (3)$$

$$= 2.0$$

The FCR value of ~ 1.0 calculated with the mean data of Hillenbrand et al. (1995) suggests that this value may closely approximate the normal FCR value, at least for speakers of American English. What the asymptote should be at the other end of the FCR scale is not clear; theoretically, it should get infinitesimally close to 0 as the vowel space expands. Empirically, though, it is likely to be close to a value that is associated with clear speech and hyperarticulation of vowels. Ferguson and Kewley-Port (2007) noted that in clear speech the quadrilateral vowel space area increases by up to 10% compared with conversational speech. If this increase is typical of clear or hyperarticulated speech, then one should expect the FCR value associated with clear speech to be around .90.

To test the sensitivity of the FCR to formant centralization and its ability to differentiate dysarthric from normal speech, we elected to study vowel articulation in individuals with IPD and dysarthria and compare it with that of healthy control (HC) participants. The dysarthria associated with IPD has been characterized by various voice and speech abnormalities, including articulatory undershoot (Sapir, Ramig, & Fox, 2008). Thus, we expected the FCR to reflect such undershoot, by showing vowel centralization in the IPD speakers relative to the HC speakers. We also expected the FCR to show a decrease in vowel centralization following successful treatment of the dysarthria, such as the Lee Silverman Voice Treatment (LSVT LOUD; Sapir et al., 2007). We therefore elected to test the ability of the FCR to register treatment effects in individuals who have been treated with LSVT, by measuring changes from

pre- to posttreatment. The LSVT is an intensive regimen that trains individuals to speak in a healthy louder voice and with greater effort than they use in their hypophonic and hypokinetic speech (Ramig, Fox, & Sapir, 2008). The treatment is based on principles of motor learning and neural plasticity and has been proven highly effective in the reduction of speech problems, including hypokinetic vowel articulation in individuals with IPD (Fox et al., 2006; Sapir et al., 2007).

As argued earlier, the VSA is limited in its ability to differentiate dysarthric from healthy speech, most likely because of its high sensitivity to interspeaker variability. Alternatively, the VSA should be minimally affected by interspeaker variability when it is used to assess treatment effects, because the comparison is largely within rather than across speakers. Therefore, we expected the VSA to be sensitive to changes associated with treatment. We also expected the FCR to correlate well with the VSA when the correlated variable is the change induced by treatment.

To test the hypothesis that the VSA performs less effectively than the FCR, presumably because of high sensitivity of the VSA to interspeaker variability, we compared the VSA with a logarithmically scaled version of the VSA (LnVSA). As shown in the Appendix, logarithmic scaling of formant frequencies maps differences between frequencies into a ratio of these frequencies; once these frequencies are in a ratio form their interspeaker variability is likely to be reduced considerably, as discussed earlier (see also Yang's [1996] discussion of gender normalization procedures). Thus, we would expect the LnVSA to be less affected by interspeaker variability and to perform better than the VSA in the differentiation of dysarthric from normal speech. The sensitivity of a metric to interspeaker variability is indexed here by gender effects and the magnitude of the CV. High sensitivity to interspeaker variability should be reflected in significant gender effects and relatively large CV values. Low sensitivity to interspeaker variability should be reflected in the lack of gender effects and relatively small CV values. We also expected that the FCR and LnVSA would show stronger correlations with each other than with the correlation between the FCR and VSA, given that both the FCR and LnVSA are designed to reduce interspeaker variability, whereas the VSA is not.

We also elected to compare the FCR with the F2i/F2u ratio. The F2i/F2u ratio has been shown to effectively differentiate dysarthric speech of individuals with IPD from normal speech of healthy age- and gender-matched control participants and to effectively register treatment effects (Sapir et al., 2007). This metric has also been proven highly effective in differentiating abnormal articulation in children with Down syndrome from normal speech in typical children (Moura et al., 2008).

The F2 frequency range formed by the English vowels /i/ and /u/ is relatively large (~1500 Hz, from about 1000 Hz for the vowel /u/ to about 2500 Hz for the vowel /i/; Hillenbrand et al., 1995), and, as such, it might serve to index changes in the extent of articulatory movements. This ratio should be especially sensitive to anterior–posterior movements of the tongue, and rounding and unrounding of the lips, because these movements are most likely to affect F2i and F2u. Thus, the F2i/F2u ratio should decrease with articulatory undershoot and increase with improved articulatory movements. Other researchers have successfully used F2 parameters (e.g., F2u, F2i – F2u, F2 extent, F2 slope) to quantify and measure speech articulation impairment in dysarthric speakers (e.g., Rosen et al., 2008; Yunusova, Weismer, Kent, & Rusche, 2005). Thus, the F2i/F2u ratio seems a reasonable metric against which the convergent validity of the FCR might be tested. One might argue that the FCR is superfluous, given that the F2i/F2u ratio in a previous study (Sapir et al., 2007) reliably differentiated dysarthric and normal vowel articulation and effectively registered treatment effects. However, the F2i/F2u ratio is inclusive of only one formant and two vowels, whereas the FCR is inclusive of two formants and three vowels. Thus, the FCR has the advantage of being more effective than the F2i/F2u ratio in the detection of articulatory abnormalities when these abnormalities involve more than just F2i and F2u.

Method

Participants

The study participants included 38 individuals with IPD and dysarthria, of whom 19 received intensive voice/speech therapy (LSVT LOUD; henceforth the *PD-T group*; 10 men and 9 women) and 19 received no treatment (henceforth the *PD-NT group*; 9 men and 10 women). These groups were compared with another group of 14 neurologically healthy controls with normal voice and speech (the *HC group*, 7 men and 7 women), age- and gender-matched to the IPD groups. The acoustic data (*Ms* and *SDs* of the F1 and F2 frequencies of the vowels /i/, /u/, and /a/) and the biomedical data of the majority of these individuals (29 of the IPD individuals and all HC individuals) were already reported in a previous study (Sapir et al., 2007).

All participants were speakers of American English as their first language. The majority of these individuals were recruited from either Tucson, Arizona, or Denver, Colorado. The mean age of the PD-T group was 68.79 years (*SD* = 9.85), the mean stage of disease (based on Hoehn and Yahr's [1967] disability scale of 0–5, where 5 represents the most severe disability) in this group was 2.92 (*SD* = 1.08),

and the mean number of years since diagnosis was 6.97 (*SD* = 6.12). The mean age of the PD-NT group was 68.11 years (*SD* = 10.83), the mean stage of disease was 2.12 (*SD* = 0.65), and the mean number of years since diagnosis was 7.00 (*SD* = 5.08). The mean age of the HC group was 69.79 years (*SD* = 7.51).

In the majority of the participants with IPD, the dysarthria was rated as mild or moderate and characterized by reduced loudness, hoarseness, and monotone speech. Some individuals had other speech problems, mostly imprecise articulation. The participants with IPD were taking anti-Parkinson's medications at the time of data collection. They were all optimally medicated and stable at the time of the study.

Data Collection

For the Tucson participants, data collection for all three groups took place on 3 different days just before the time of treatment (T1) and on 2 different days just after the end of treatment (T2). For the Denver participants, data collection for all three groups took place on a single day before the beginning of treatment (T1) and on a single day just after the end of treatment (T2). Those recordings took place within 2–3 days before or after treatment. Only the PD-T group received treatment, and the specific dates of recordings were different across participants, yet the overall time schedule, as described previously, was the same for all participants.

The data in this study were based on multiple repetitions of three phrases obtained in the Tucson recordings (“The blue spot is on the key,” “The potato stew is in the pot,” and “Buy Bobby a puppy”), and on multiple repetitions of one phrase obtained in the Denver recordings (“The stew pot is packed with peas”). Each phrase in the Tucson recordings was repeated by each participant 3 times on each day of recording (i.e., 9 times total before and 6 times total after the time of treatment), and the single phrase in the Denver recordings was repeated by each participant 10 times on each day of recording (i.e., 10 times before and 10 times after time of treatment). The Tucson recordings were obtained in a sound-treated booth using a head-mounted condenser microphone (AKG C410) positioned 6 cm from the lips and a digital audiotape two-channel recorder (Sony PC-208AUC). The data were digitized from the digital audiotape to a computer at a sampling rate of 22 kHz using Goldwave software. Similar recording methods were used in the Denver study, but the acoustic signals were collected directly to a computer using an AKG C420 head-mounted microphone and sampled at 44.1 kHz with Kay Elemetrics CSL model 4300B hardware and software. All files were downsampled to 22 kHz for formant analysis.

Acoustic Analyses and Measurements

The vowels /i/, /u/, and /a/ were extracted from the words *key*, *stew*, and *Bobby* (Tucson recordings), or from the words *peas*, *stew*, and *pot* from the single phrase (Denver recordings), respectively. Regardless of the type and number of phrases uttered and the number of samples used, all vowels were extracted, and all F1 and F2 values were measured in the same manner. Formant frequency analysis was done using TF32, a Windows-based version of CSpeech software (Milenkovic, 2001). Forty percent of the analyzed data were also analyzed using MATLAB (Version 5.3) to assess reliability of the measures. F1 and F2 frequency values for /i/ and /a/ were measured for a 30-ms segment at the temporal midpoint of each vowel. For the vowel /u/, F1 and F2 were measured from a 30-ms segment at the end of the vowel. This segment was chosen to avoid the intrusion of the formant transition immediately preceding the /u/ in *stew*. Tests of the validity and reliability of the acoustic measures were described in a previous publication (Sapir et al., 2007). These tests indicate high intra- and inter-judge reliability for the F2 measures (Pearson product-moment correlations = .96–.99) and moderate to high reliability for the F1 measurements ($r_s = .83$ –.95) across the different vowels (/i/, /u/, /a/) and groups (IPD, HC). Standard errors of measurement were relatively small for both F2 (range: 20–26 Hz) and F1 measurements (range: 19–42 Hz).

Statistical Analyses

The vowel-formant data (e.g., frequency, in Hz, of F1a, F2i, F1u, etc.) were separately averaged for each individual for T1 and T2. We then constructed the VSA, F2i/F2u ratio, and FCR from these averages. In the case of the LnVSA, the formant frequencies were first transformed to a logarithmic scale and then used to construct the LnVSA. These data were then subjected to statistical analyses as detailed here. We used the Kolmogorov–Smirnov test to determine normality of the distribution of the data. We evaluated separately the differences among the three groups (PD-T, PD-NT, HC) for each of the dependent variables (VSA, F2i/F2u ratio, FCR) for the T1 (pre-treatment) data using a one-way analysis of variance (ANOVA). We conducted a repeated measures multivariate ANOVA to assess T2 (post-treatment) differences while accounting for T1 variation. We used Duncan's Multiple Range test for planned comparison analyses of significance, with α set at .05. Gender effects for the T1–T2 differences in the three groups were tested with a two-way ANOVA with interaction. The magnitude of difference between means was assessed with an effect size measure, using a pooled variance method (Cohen, 1988). By this method, an effect size of .80 is

considered large, .50 is considered medium, and .20 is considered small. Interspeaker variability was measured in terms of CV.

To assess the strength of the relationship among VSA, LnVSA, F2i/F2u ratio, and FCR, we correlated the T1–T2 differences of these metrics in the PD-T group using Pearson product-moment correlation analysis. We anticipated that if the four acoustic metrics measure similar phenomena, this should be reflected by a strong correlation. Poor correlation between two metrics might imply that these metrics measure different aspects of vowel articulation.

Results

Tests of Normality

The Kolmogorov–Smirnov test indicated that, with one exception, the VSA, LnVSA, F2i/F2u ratio, and FCR data were normally distributed in each of the three groups (PD-T, PD-NT, and HC) and at T1 and T2. The exception was the VSA at T2 in the PD-T group, which deviated from normality (kurtosis = -1.1138). Nevertheless, given that the majority of the data showed normal distribution, given that the majority of the analyses were done with the T1 data, and given the highly significant differences in the VSA from T1 to T2 in the PD-T group (discussed later), we elected to use parametric statistics for the data.

Differences Between Groups at T1

The means and standard deviations of the vowel-formant elements at T1 and T2 are shown in Tables 2 and 3. The means and standard deviations of the VSA, F2i/F2u ratio, LnVSA, and FCR data at T1 and T2 are shown in Table 4. The means and standard deviations (error bars) of the VSA, LnVSA, F2i/F2u ratio, and FCR data at T1 and T2 for the three groups (PD-T, PD-NT, and HC) are shown graphically in Figure 1.

As can be seen in Table 4 and Figure 1, at T1 the means of the VSA, LnVSA, and F2i/F2u ratio are smaller, and the mean of the FCR is larger, in the PD-T and PD-NT groups relative to the corresponding means in the HC group. These findings are consistent with vowel centralization in the PD groups. A one-way ANOVA of the T1 data indicated significant differences among the three groups for the FCR, $F(2, 49) = 8.01, p = .001$; F2i/F2u ratio, $F(2, 49) = 10.36, p = .0002$; and LnVSA, $F(2, 49) = 3.80, p = .0292$, but not for the VSA, $F(2, 49) = 2.12, p = .1303$. For the FCR, Duncan's paired comparison tests indicated a significant difference between the PD-T and HC groups, and between the PD-NT and HC groups, and no significant difference between the PD-T and PD-NT groups. The significant differences are associated

Table 2. Means (in Hz) and standard deviations of the F1 of the vowels /i/, /u/, and /a/ at Time 1 (T1; pretreatment) and Time 2 (T2; posttreatment) in the three groups.

Group	F1i		F1u		F1a	
	T1	T2	T1	T2	T1	T2
PD-T						
M (Hz)	330	328	361	369	756	803
SD (Hz)	67	62	60	61	114	105
CV (%)	20.3	18.9	16.6	16.7	15.1	13.1
PD-NT						
M (Hz)	338	331	363	370	786	781
SD (Hz)	30	26	45	47	103	104
CV (%)	8.8	7.8	12.4	12.6	13.0	13.3
HC						
M (Hz)	318	320	384	380	788	775
SD (Hz)	41	46	35	43	89	94
CV (%)	13.0	14.4	9.2	11.4	11.3	12.1

Note. PD-T = individuals with dysarthria secondary to Parkinson's disease who received treatment with Lee Silverman Voice Treatment; PD-NT = individuals with Parkinson's disease who did not receive treatment for their dysarthria; HC = healthy controls.

with large effect sizes (1.47 and 0.97, respectively). For the F2i/F2u ratio, the Duncan's tests indicate a significant difference between the PD-T and HC groups and between the PD-NT and HC groups but not between the PD-T and PD-NT groups. The significant differences are associated with large effect sizes (1.54 and 1.11, respectively). For the LnVSA, Duncan's test indicated a significant difference between the PD-T and HC groups but not between the PD-NT and HC groups or between the PD-T and PD-NT groups. The significant difference is

Table 3. Means (in Hz) and standard deviations of the F2 of the vowels /i/, /u/, and /a/ at T1 (pretreatment) and T2 (posttreatment) in the three groups.

Group	F2i		F2u		F2a	
	T1	T2	T1	T2	T1	T2
PD-T						
M (Hz)	2417	2490	1364	1252	1326	1315
SD (Hz)	303	328	193	213	182	137
CV (%)	12.5	13.2	14.1	17.0	13.7	10.4
PD-NT						
M (Hz)	2480	2481	1323	1330	1335	1330
SD (Hz)	335	329	216	217	135	134
CV (%)	13.5	13.3	16.3	16.3	10.1	10.1
HC						
M (Hz)	2565	2563	1189	1212	1307	1315
SD (Hz)	222	232	163	145	102	125
CV (%)	8.7	9.1	13.7	11.9	7.8	9.5

associated with a large effect size (0.96). Thus, at T1, the FCR and F2i/F2u ratio significantly and robustly differentiated dysarthric from normal groups, the LnVSA differentiated only partially between the groups, and the VSA failed to differentiate between dysarthric and non-dysarthric groups.

Gender Effects at T1

For the FCR data at T1, a two-way ANOVA with interaction indicated a significant main effect of group, $F(2, 46) = 8.66, p = .0006$, but not of gender, $F(1, 46) = 1.11, p = .2975$, and no Gender \times Group interaction, $F(2, 46) = 3.010, p = .0590$. For the F2i/F2u ratio, there was a main effect of group, $F(1, 46) = 10.23, p = .0002$, but not of gender, $F(2, 46) = 2.42, p = .1266$, and no Gender \times Group interaction, $F(2, 46) = 0.48, p = .6246$. For the VSA, there was a main effect of gender, $F(1, 46) = 8.43, p = .0056$, but not of group, $F(2, 46) = 2.47, p = .0956$, and no Gender \times Group interaction, $F(2, 46) = 0.070, p = .9325$. For the LnVSA, there were main effects of gender, $F(1, 46) = 7.50, p = .0087$, and group, $F(2, 46) = 3.98, p = .0254$, and no Gender \times Group interaction, $F(2, 46) = 0.24, p = .7849$. Thus, only the VSA and LnVSA were gender sensitive at T1.

Detecting Treatment Effects (Changes From T1 to T2)

A repeated measures multivariate ANOVA for between-subjects tests indicated significant between-group differences for the FCR, $F(2, 49) = 4.79, p = .0126$, and F2i/F2u ratio, $F(2, 49) = 7.83, p = .0011$, but not for the VSA, $F(2, 49) = 0.53, p = .5943$, or LnVSA, $F(2, 49) = 1.30, p = .2821$. A repeated measures ANOVA with univariate within-subjects tests indicated no main effect of time for the FCR, $F(1, 49) = 3.34, p = .0738$, F2i/F2u ratio, $F(1, 49) = 2.00, p = .1635$, VSA, $F(1, 49) = 3.92, p = .0534$, and LnVSA, $F(1, 49) = 3.25, p = .0765$, but a significant Time \times Group effect of the FCR, $F(2, 49) = 6.44, p = .0033$, F2i/F2u ratio, $F(2, 49) = 5.02, p = .0104$, VSA, $F(2, 49) = 10.64, p = .0001$, and LnVSA, $F(2, 49) = 7.77, p = .0012$. Effect size measures of the T1–T2 difference in the PD-T group were large for the FCR (0.84) and the F2i/F2u ratio (0.88), medium to large for the LnVSA (0.74), and medium for the VSA (0.58). Effect size measures of the T1–T2 differences in the PD-NT and HC groups were small in all metrics (absolute effect size < 0.25). Thus, all metrics register significant treatment effects in the PD-T group, but the FCR and F2i/F2u ratio registered a more robust effect than the VSA and LnVSA, as reflected by the effect size measures.

Gender Effects for the T1–T2 Difference

For the FCR, a two-way ANOVA with interactions indicated a main effect of group, $F(2, 46) = 6.43, p = .0034$,

Table 4. Means, standard deviations, and coefficients of variation of the FCR, VSA, logarithmically scaled VSA (LnVSA), and F2i/F2u ratio data at T1 and T2 and in the three groups.

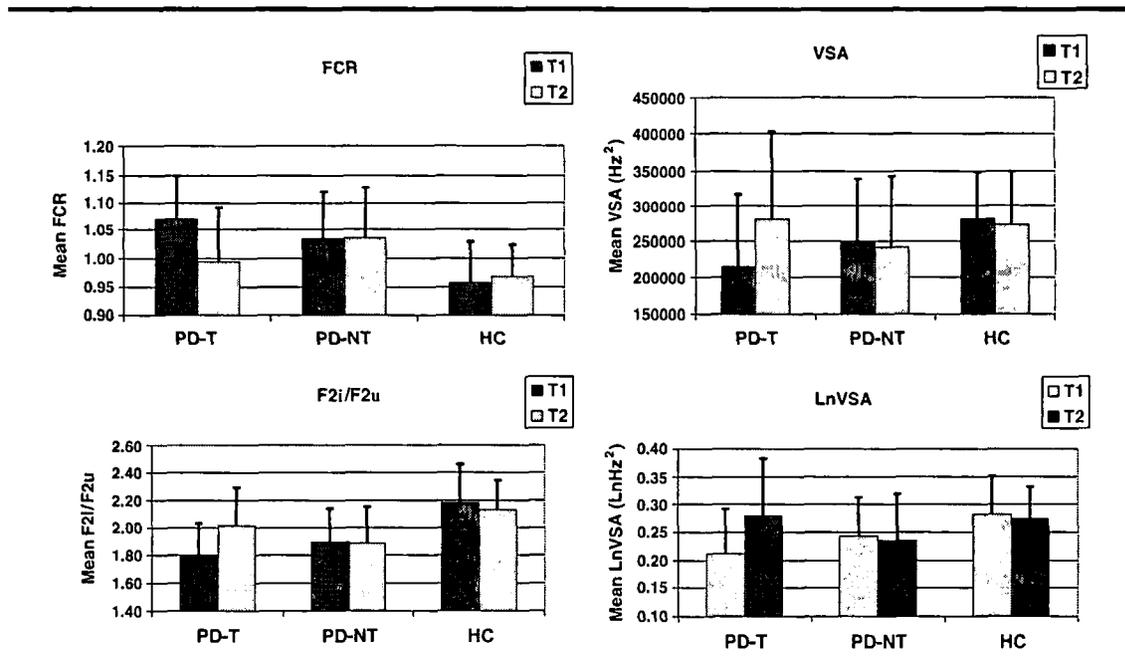
Group	FCR		VSA ^a		LnVSA ^b		F2i/F2u	
	T1	T2	T1	T2	T1	T2	T1	T2
PD-T								
M	1.07	1.00	217551	281724	0.21	0.28	1.79	2.02
SD	0.08	0.10	99982	121441	0.08	0.10	0.24	0.27
CV (%)	7.5	10.0	46.0	43.1	37.7	36.4	13.5	13.6
PD-NT								
M	1.03	1.04	233508	234683	0.24	0.24	1.90	1.89
SD	0.09	0.09	83369	92646	0.07	0.08	0.24	0.26
CV (%)	8.3	9.0	35.7	39.5	28.7	35.2	12.9	13.8
HC								
M	0.96	0.97	280420	272430	0.28	0.27	2.18	2.13
SD	0.07	0.07	77579	77184	0.07	0.06	0.27	0.22
CV (%)	7.6	6.0	27.7	28.3	24.0	21.5	12.5	10.1

^aIn Hz², rounded to 1-Hz accuracy. ^bIn logarithmically scaled Hz².

but not of gender, $F(1, 46) = 1.15, p = .2289$, and no Gender \times Group interaction, $F(2, 46) = 1.21, p = .3068$. For the F2i/F2u ratio, there was a main effect of group, $F(2, 46) = 4.75, p < .0134$, but not of gender, $F(1, 46) = 0.26, p = .611$, and no Gender \times Group interaction,

$F(2, 46) = 0.22, p = .8029$. For the VSA, there was a main effect of group, $F(2, 46) = 10.723, p < .0001$, but not of gender, $F(1, 46) = 0.13, p = .7180$, and no Gender \times Group interaction, $F(2, 46) = 1.05, p = .3598$. For the LnVSA, there was a main effect of group, $F(2, 46) = 7.53, p = .0015$,

Figure 1. Shown from top left, in a clockwise direction, the mean formant centralization ratio (FCR), vowel space area (VSA), logarithmically scaled VSA (LnVSA), and F2i/F2u ratio (error bars represent 1 SD) at Time 1 (T1, before treatment) and Time 2 (T2, after treatment) for individuals with dysarthria secondary to Parkinson's disease who received treatment with Lee Silverman Voice Training (PD-T group), individuals with Parkinson's disease who did not receive treatment for their dysarthria (PD-NT group), and healthy control (HC) participants.



but not of gender, $F(1, 46) = 1.37, p = .2483$, and no Gender \times Group interaction, $F(2, 46) = 0.56, p = .5756$. Thus, none of the metrics showed a gender effect when measuring treatment changes.

Pearson Correlations Among the FCR, F2i/F2u Ratio, LnVSA, and VSA for the T1–T2 Difference in the PD-T Group

We used the T2–T1 difference in the PD-T group to correlate pairs of metrics. There were high correlations among the metrics: FCR versus F2i/F2u ratio, $r = -.90, p < .0001$; FCR versus VSA, $r = -.85, p < .0001$; FCR versus LnVSA, $r = -.81, p < .0001$; F2i/F2u ratio versus VSA, $r = .85, p < .0001$; F2i/F2u ratio versus LnVSA, $r = .81, p < .0001$; and VSA versus LnVSA, $r = .89$.

Finally, as can be seen in Table 4, the CV values are in general largest in the VSA, smaller in the LnVSA, still smaller in the F2i/F2u ratio, and smallest in the FCR. Thus, if one considers the CV an index of interspeaker variability, this variability is largest in the VSA and smallest in the FCR. Note also that the CV values are in general larger in the PD groups than in the HC group.

Discussion

In this study, the FCR, like the F2i/F2u ratio, and unlike the VSA and LnVSA, effectively and robustly differentiated the groups with dysarthria (PD-T and PD-NT) from the HC group. Like the F2i/F2u ratio, VSA, and LnVSA, the FCR effectively registered treatment effects. Unlike the VSA and LnVSA, these treatment effects were registered by the FCR and F2i/F2u ratio with large effect sizes. Also, unlike the VSA and LnVSA, the FCR and F2i/F2u ratio were insensitive to gender effects and were associated with relatively small CV values. The LnVSA was more effective than the VSA in differentiating dysarthric from normal vowel articulation, but this was true for only the PD-T group. Finally, the FCR correlated highly with the other three metrics when the correlated variable was the change induced by treatment. Taken collectively, these findings, albeit preliminary, suggest that the FCR is a valid and highly sensitive metric of vowel articulation, normal and abnormal, and that its

performance is superior to that of the VSA and the LnVSA in differentiating dysarthric from healthy speech. The presence of gender effects in only the VSA and LnVSA, and the larger CV associated with their measurements, suggests that these two metrics were much more sensitive to interspeaker variability than the FCR and F2i/F2u ratio. This difference in sensitivity may account for the inability of the VSA, and the partial ability of the LnVSA, to differentiate between the dysarthric and nondysarthric groups. The fact that the LnVSA was associated with smaller CV values than the VSA and was more successful than the VSA in differentiating between the dysarthric and nondysarthric groups also supports the idea that the failure of the VSA to differentiate between dysarthric and nondysarthric speakers has to do, at least in part, with the high sensitivity of the VSA to interspeaker variability. Moreover, the fact that the VSA and LnVSA performed well in registering changes induced by treatment (a within-subject comparison) also speaks to the role of interspeaker variability in the performance of these acoustic metrics.

More evidence for the validity of the FCR comes from Hillenbrand et al.'s (1995) and Higgins and Hodge's (2002) studies. As can be seen in Table 1, the calculated values of the FCR from the formant frequencies in the men, women, and children in Hillenbrand et al.'s study (0.99, 0.96, and 0.97, respectively) are close to the mean FCR values of the HC group in the present study (0.96 at T1, 0.97 at T2). The formant data from Higgins and Hodge's study are shown in Table 5, along with the FCR, VSA, LnVSA, and F2i/F2u ratio values calculated from these data. The data are from young children with dysarthria secondary to cerebral palsy and from healthy controls. Note that the formant frequencies of the different vowels are centralized in the children with dysarthria relative to the controls, and this centralization is also reflected in the FCR, VSA, LnVSA, and F2i/F2u ratio. Note also that the FCR value for the typically developing children is 0.91, and for the children with dysarthria it is 1.14. These values are fairly similar to those in the present study for the normal and dysarthric speakers, respectively. The somewhat smaller FCR value for the typical children and larger FCR value for the dysarthric children relative to the data in the present study might be related to the fact that the children in Higgins and

Table 5. Mean formant data of children with dysarthric speech associated with cerebral palsy, and neurologically healthy children from Higgins and Hodge's (2002) study, with results of applying the VSA, LnVSA, FCR, and F2i/F2u ratio to these data (shown in the four rightmost columns).

Speech group	F1i (Hz)	F2i (Hz)	F1a (Hz)	F2a (Hz)	F1u (Hz)	F2u (Hz)	VSA (Hz ²)	LnVSA (Ln Hz ²)	FCR	F2i/F2u ratio
Normal	532	3528	1232	1556	520	1710	648132	0.31	0.91	2.06
Dysarthric	576	3406	930	1765	592	2029	230601	0.12	1.14	1.68

Hodge's study were Canadian and very young (5–6 years old) and the possibility that the dysarthria in the Higgins and Hodge study may have been more severe than the dysarthria in the present study.

In this study, the F2i/F2u ratio and the FCR were highly correlated ($r = -.90$, when the correlated variable was the treatment-induced change) and equally effective in differentiating the dysarthric and nondysarthric speakers and in registering treatment effects. The close performance of these two metrics suggests that, to a large extent, they reflected the same articulatory abnormalities, namely, restricted movements of the tongue in the anterior–posterior direction and restricted movements of the lips (rounding for /u/ and retraction for /i/). Given that the FCR and F2i/F2u ratio performed so much the same and showed a high correlation, one might argue that the FCR is superfluous, as the F2i/F2u ratio seems sufficient to capture the nature of articulatory abnormality. However, there might be individuals with dysarthria whose speech impairment may involve more vowels than the /i/ and /u/ and more than one formant. Thus, to capture such impairment, the FCR is likely to offer additional information and thus be more appropriate than the F2i/F2u ratio.

The FCR, VSA, and LnVSA are all based on the construct of vowel centralization. By this construct, one would expect that all the vowel formants will show centralization. However, such symmetry is unlikely to occur, as most studies of dysarthric vowel articulation indicate (see, e.g., Sapir et al., 2007; Weismer et al., 2001; Yunusova et al., 2008). Also, in some cases the acoustic measures of abnormal vowel articulation might be in the direction opposite of that expected from formant centralization. For example, in the present study, and in Weismer et al.'s (2001) study, the frequency of F1u in the speech of dysarthric individuals with IPD tended to be lower than normal, yet by the centralization construct the frequency of F1u should have been higher than normal. Yunusova et al. (2008) noted in their kinematic study that whereas the majority of individuals with dysarthria secondary to ALS had smaller than normal jaw or tongue movements, some individuals with dysarthria and ALS had jaw movements that were considerably larger than normal. Thus, the FCR, VSA, and LnVSA may not fully or faithfully capture the nature of the articulatory impairment in all speakers and all types of dysarthria. In future studies, it would be important to make accommodations for such asymmetry, especially if the asymmetry is very characteristic of a particular dysarthria.

These findings are a first effort at evaluating the FCR. We wish to stress that the FCR is not necessarily the best metric to differentiate dysarthric from healthy vowel articulation, and it is not necessarily the preferred

metric for all types of dysarthria and patient populations. Thus, without comparing the FCR with other metrics (other than those tested here), it is not possible to tell which of these metrics, or a combination of them, are most effective and reliable in measuring dysarthric vowel articulation. Also, there are numerous factors (e.g., speech task, phonetic environment, type and severity of dysarthria) that can affect vowel production and its acoustic manifestations. Thus, future studies should examine how the FCR performs under these conditions. Furthermore, in this study, we tested the VSA and FCR with the vowels /i/, /u/, and /a/; therefore, it would be important to assess how well these metrics might perform with more or different vowels. Finally, at this point it is not clear what specific articulatory abnormalities are represented by the FCR and how those might be related to perceived speech abnormality. It would therefore be important to correlate the FCR with physiological and perceptual measurements of vowel articulation, normal and abnormal.

Acknowledgment

This research was funded by National Institute on Deafness and Other Communication Disorders Grant R01 DC1150.

References

- Adank, P., Smits, R., & van Hout, R. (2004). A comparison of vowel normalization procedures for language variation research. *The Journal of the Acoustical Society of America*, 116, 3099–3107.
- Bunton, K., & Weismer, G. (2001). The relationship between perception and acoustics for a high–low vowel contrast produced by speakers with dysarthria. *Journal of Speech, Language, and Hearing Research*, 44, 1215–1228.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- de Jong, K. (1997). Labiovelar compensation in back vowels. *The Journal of the Acoustical Society of America*, 101, 2221–2233.
- Dunham, W. (1990). *Journey through genius: The great theorems of mathematics*. New York: Wiley.
- Fant, G. (1960). *Acoustic theory of speech production*. The Hague, The Netherlands: Mouton.
- Ferguson, S., & Kewley-Port, D. (2007). Talker differences in clear and conversational speech: Acoustic characteristics of vowels. *Journal of Speech, Language, and Hearing Research*, 50, 1241–1255.
- Fourakis, M. (1991). Tempo, stress, and vowel reduction in American English. *The Journal of the Acoustical Society of America*, 90, 1816–1827.
- Fox, C., Ramig, L., Ciucci, M., Sapir, S., McFarland, D., & Farley, B. (2006). The science and practice of LSVT/LOUD: Neural plasticity-principled approach to treating individuals with Parkinson disease and other neurological disorders. *Seminars in Speech and Language*, 27, 283–299.

- Hashi, M., Westbury, J., & Honda, K. (1998). Vowel posture normalization. *The Journal of the Acoustical Society of America*, 104, 2426–2437.
- Higgins, C., & Hodge, M. (2002). Vowel area and intelligibility in children with and without dysarthria. *Journal of Medical Speech-Language Pathology*, 10, 271–277.
- Hillenbrand, J., Getty, L., Clark, M., & Wheeler, K. (1995). Acoustic characteristics of American English vowels. *The Journal of the Acoustical Society of America*, 97, 3099–3111.
- Hoehn, M., & Yahr, M. (1967). Parkinsonism: Onset, progression, and mortality. *Neurology*, 17, 427–442.
- Honda, K., & Kusakawa, N. (1997). Compatibility between auditory and articulatory representations of vowels. *Acta Otolaryngologica*, 532(Suppl.), 103–105.
- Kent, R., & Kim, Y. (2003). Toward an acoustic typology of motor speech disorders. *Clinical Linguistics & Phonetics*, 17, 427–445.
- Kent, R., Weismer, G., Kent, J., Vorperian, H., & Duffy, J. (1999). Acoustic studies of dysarthric speech: Methods, progress, and potential. *Journal of Communication Disorders*, 32, 141–186.
- Liu, H., Tsao, F., & Kuhl, P. (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *The Journal of the Acoustical Society of America*, 117, 3879–3889.
- McRae, P., Tjaden, K., & Schoonings, B. (2002). Acoustic and perceptual consequences of articulatory rate change in Parkinson disease. *Journal of Speech, Language, and Hearing Research*, 45, 35–50.
- Milenkovic, P. (2001). *TF32*. Madison: University of Wisconsin Press.
- Moura, C., Cunha, L., Vilarinho, H., Cunha, M., Freitas, D., Palha, M., et al. (2008). Voice parameters in children with Down syndrome. *Journal of Voice*, 22, 34–42.
- Perkell, J., Guenther, F., Lane, H., Matthies, M., Stockmann, E., Tiede, M., et al. (2004). The distinctness of speakers' productions of vowel contrasts is related to their discrimination of the contrasts. *The Journal of the Acoustical Society of America*, 116, 2338–2344.
- Ramig, L., Fox, C., & Sapir, S. (2008). Speech treatment for Parkinson's disease. *Expert Reviews in Neurotherapeutics*, 8, 297–309.
- Rosen, K., Goozée, J., & Murdoch, B. (2008). Examining the effects of multiple sclerosis on speech production: Does phonetic structure matter? *Journal of Communication Disorders*, 41, 49–69.
- Sapir, S., Ramig, L., & Fox, C. (2008). Speech and swallowing disorders in Parkinson disease. *Current Opinions in Otolaryngology—Head and Neck Surgery*, 16, 205–210.
- Sapir, S., Spielman, J., Ramig, L., Hinds, S., Countryman, S., Fox, C., & Story, B. (2003). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on ataxic dysarthria: A case study. *American Journal of Speech-Language Pathology*, 12, 387–399.
- Sapir, S., Spielman, J., Ramig, L., Story, B., & Fox, C. (2007). Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research*, 50, 899–912.
- Tjaden, K., & Wilding, G. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, and Hearing Research*, 47, 766–783.
- Weismer, G., Jeng, J.-Y., Laures, J., Kent, R., & Kent, J. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica*, 53, 1–18.
- Yang, B. (1996). A comparative study of American English and Korean vowels produced by male and female speakers. *Journal of Phonetics*, 24, 245–261.
- Yunusova, Y., Weismer, G., Kent, R., & Rusche, N. (2005). Breath-group intelligibility in dysarthria: Characteristics and underlying correlates. *Journal of Speech, Language, and Hearing Research*, 48, 1294–1310.
- Yunusova, Y., Weismer, G., Westbury, J., & Lindstrom, M. (2008). Articulatory movements during vowels in speakers with dysarthria and healthy controls. *Journal of Speech, Language, and Hearing Research*, 51, 596–611.
- Ziegler, W., & von Cramon, D. (1983). Vowel distortion in traumatic dysarthria: A formant study. *Phonetica*, 40, 63–78.

Received August 31, 2008

Revision received January 25, 2009

Accepted July 4, 2009

DOI: 10.1044/1092-4388(2009/08-0184)

Contact author: Shimon Sapir, Department of Communication Sciences and Disorders, Faculty of Social Welfare and Health Sciences, University of Haifa, Haifa, Mount Carmel, 39105 Israel. E-mail: sapir@research.haifa.ac.il

Appendix. Mathematical expressions of the triangular VSA and LnVSA and the explanations for the logarithmic scaling as a means to transform differences in formant frequencies to ratios of formant frequencies.

The triangular VSA, constructed with the corner vowels /i/, /u/, and /a/, may be expressed mathematically as follows:

$$VSA = ABS\{[F1i \times (F2a - F2u) + F1a \times (F2u - F2i) + F1u \times (F2i - F2a)]/2\}, \quad (1)$$

where ABS = absolute value. This VSA can also be expressed as follows:

$$VSA = \text{sqrt}[S \times (S - EDiu)(S - EDia)(S - EDau)], \quad (2)$$

where sqrt = square root, and

$$EDiu = \text{sqrt}\{(F1i - F1u)^2 + (F2i - F2u)^2\} \quad (3)$$

$$EDia = \text{sqrt}\{(F1i - F1a)^2 + (F2i - F2a)^2\} \quad (4)$$

$$EDau = \text{sqrt}\{(F1a - F1u)^2 + (F2a - F2u)^2\} \quad (5)$$

$$S = (EDiu + EDia + EDau)/2, \quad (6)$$

where EDiu = the Euclidean distance between the vowels /i/ and /u/, EDia is the Euclidean distance between the vowels /i/ and /a/, and EDau is the Euclidean distance between the vowels /a/ and /u/ in the F1-F2 plane. The VSA formulas are in accordance with Heron's formula for a triangular area (Dunham, 1990).

The logarithmic version of the triangular VSA (LnVSA) is expressed mathematically as follows:

$$\text{LnVSA} = \text{sqrt}[\text{Ln}S \times (\text{Ln}S - \text{Ln}EDiu)(\text{Ln}S - \text{Ln}EDia)(\text{Ln}S - \text{Ln}EDau)], \quad (7)$$

where

$$\text{Ln}EDiu = \text{sqrt}\{(\text{Ln}F1i - \text{Ln}F1u)^2 + (\text{Ln}F2i - \text{Ln}F2u)^2\} \quad (8)$$

$$\text{Ln}EDia = \text{sqrt}\{(\text{Ln}F1i - \text{Ln}F1a)^2 + (\text{Ln}F2i - \text{Ln}F2a)^2\} \quad (9)$$

$$\text{Ln}EDau = \text{sqrt}\{(\text{Ln}F1a - \text{Ln}F1u)^2 + (\text{Ln}F2a - \text{Ln}F2u)^2\} \quad (10)$$

$$\text{Ln}S = (\text{Ln}EDiu + \text{Ln}EDia + \text{Ln}EDau)/2. \quad (11)$$

Note that by the mathematical equivalence $\text{Ln}(A/B) = \text{Ln}(A) - \text{Ln}(B)$, we can express the Euclidean distances in Equations 8, 9, and 10 as formant ratios:

$$\text{Ln}EDiu = \text{sqrt}[\text{Ln}(F1i/F1u)^2 + \text{Ln}(F2i/F2u)^2] \quad (12)$$

$$\text{Ln}EDia = \text{sqrt}[\text{Ln}(F1i/F1a)^2 + \text{Ln}(F2i/F2a)^2] \quad (13)$$

$$\text{Ln}EDau = \text{sqrt}[\text{Ln}(F1a/F1u)^2 + \text{Ln}(F2a/F2u)^2] \quad (14)$$

COPYRIGHT INFORMATION



Title: Formant Centralization Ratio: A Proposal for a New Acoustic Measure of Dysarthric Speech

Source: J Speech Lang Hear Res 53 no1 F 2010 p. 114-25

ISSN: 1092-4388

Publisher: American Speech Hearing Language Association
10801 Rockville Pike, Rockville, MD 20852

(C) The American-Speech-Language-Hearing Association is the publisher of this article and holder of the copyright. Further reproduction of this article in violation of copyright is prohibited without the consent of the publisher. To contact the publisher: <http://www.asha.org/>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden. The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.