

Statistical Models of F2 Slope in Relation to Severity of Dysarthria

Yunjung Kim^a Gary Weismer^b Raymond D. Kent^b Joseph R. Duffy^c^aDepartment of Communication Sciences and Disorders, Louisiana State University, Baton Rouge, La.,^bDepartment of Communicative Disorders, Waisman Center, University of Wisconsin-Madison, Madison, Wisc., and^cMayo Clinic, Rochester, Minn., USA

Key Words

Acoustic analysis · Dysarthria · Speech intelligibility

Abstract

Objective: This study investigated the distribution of second-formant (F2) slopes in a relatively large number of speakers with dysarthria associated with two different underlying diseases. **Patients and Methods:** Forty speakers with dysarthria (20 with Parkinson's disease, PD; 20 with stroke) and 5 control speakers without a history of neurological disease were asked to repeat six words (*coat, hail, sigh, shoot, row* and *wax*) 10 times. Acoustic analysis was performed to derive F2 slope, and speech intelligibility data were collected using a direct magnitude estimate technique to examine its relationship to F2 slope. **Results:** Statistical analysis revealed that both clinical groups showed significantly reduced F2 slopes compared to healthy speakers for all words but *row*. No group difference was found between speakers with PD and stroke; however, different words showed varying sensitivity to the speech motor control problems. The F2 slopes of only two words, *shoot* and *wax*, were significantly correlated with scaled speech intelligibility. **Conclusion:** The findings support the idea that distributional characteristics of acoustic variables, such as F2 slope, could be used to develop a quantitative metric of severity of speech motor control deficits in dysarthria, when the materials are appropriately selected and additional distributional characteristics are studied.

Copyright © 2009 S. Karger AG, Basel

Introduction

A number of studies of motor speech disorders have focused on acoustic measures that may be significant contributors to an acoustic model of speech intelligibility [1–4]. There are several potential benefits of such a model, but the most obvious one is the inference from the acoustic variables to the underlying articulatory events that are particularly important to the speech intelligibility deficit. Presumably, clinical attention to these particular articulatory behaviors, among the totality of articulatory behaviors that could be treated, would have the most dramatic impact on speech intelligibility.

One acoustic measure with a relatively straightforward articulatory interpretation is the slope of second formant (F2) transitions extracted from diphthongs or vocalic nuclei requiring relatively large changes in vocal tract configuration. Results from speakers with Parkinson's disease (PD) [5, 6], amyotrophic lateral sclerosis [7–9], multiple sclerosis [10, 11], and stroke and cerebellar disease [12] have shown reduced F2 slopes when compared to healthy controls. The articulatory explanation for reduced F2 slopes is relative slowness in changing the vocal tract configuration [13]. Articulatory kinematic data obtained from speakers with motor speech disorders [14] is consistent with this interpretation, showing slower lip, jaw, and tongue movements among a variety of speakers with dysarthria as compared to neurologically normal speakers.

F2 slopes also correlate rather highly with measures of speech intelligibility [6, 7]. This raises the possibility that F2 slope might serve as an index of speech motor control deficit – that is, severity of speech involvement – in persons with motor speech disorders. More importantly, F2 slope may be a simple *speech* measure of speech motor involvement, perhaps to be favored over the more common nonspeech oromotor measures (for example, lip or tongue strength, steadiness, sinusoidal tracking accuracy) often reported in the literature [15, 16].

Stable statistical estimates of F2 slope central tendency and variability are not readily available for speakers with motor speech disorders or healthy speakers. If F2 slope (or other measures) has potential as an estimate of speech mechanism involvement in persons with motor speech disorders, such statistical data are needed. These data could establish, for example, if the measure varies as a function of the word from which it is extracted, with different types of dysarthria, or with different types of diseases that cause dysarthria, among other factors.

The purpose of this study was to establish, for a group of normal speakers and for relatively large numbers of speakers with dysarthria associated with PD and stroke, some statistical properties of F2 slope. Specifically, we were interested in the distributional characteristics of F2 slope measures for different words requiring relatively large changes in vocal tract configuration throughout their vocalic nuclei. In addition, we examined the relationships between F2 slope and speech intelligibility in the two groups of speakers with dysarthria. In short, we wanted to explore the use of the F2 measure as a possible metric of speech mechanism involvement.

Method

Speakers/Listeners

A total of 40 male speakers with dysarthria secondary to PD ($n = 20$) or stroke ($n = 20$) participated in this study. These two neurological conditions were selected because their speech symptoms were expected to be quite different from each other [14]. The stroke group was heterogeneous with respect to location, size, and site of lesion. Five healthy male speakers in the age range of the dysarthric speakers were included as controls. The data from the relatively small number of healthy speakers provided an indication of the statistical properties of 'normal' F2 slopes, which are typically very stable both *within* and *across* healthy speakers [5, 17]. The speakers with dysarthria ranged in age from 46 to 79 years and were diagnosed by one of the authors as having several different types of dysarthria, including ataxic, hypokinetic, spastic, and mixed. Three graduate students participated in scaling speech intelligibility of the dysarthric speakers. They had completed a course in dysarthria, although none had been exposed on

a daily basis to speakers with dysarthria or had previous training in scaling speech intelligibility in dysarthria or other speech disorders.

Procedures

Subjects were asked to repeat each of the following single words 10 times in a row: *coat, hail, shoot, sigh, row, and wax*. These words were selected because their vocalic nuclei require relatively extensive changes in vocal tract configuration, either intrinsically or in relation to the surrounding consonant environment. These large articulatory changes result in large F2 transitions, which were the focus of the present study. The words were also chosen because previous analyses have reported them to be sensitive to the speech production deficit in dysarthria [7, 17]. The string of repeated words was used to generate a relatively large amount of data (F2 slopes) in an efficient way. The eight tokens in the middle of the repetition string were taken for analysis to eliminate possible utterance-initial or ending effects.

Subjects were also asked to recite three sentences (*The blue spot is on the key; The boiling tornado clouds moved swiftly; Combine all the ingredients in a large bowl*) from which intelligibility ratings were made. These sentences were chosen for intelligibility evaluation because they include phonetically complex sequences likely to be sensitive to dysarthria. Each listener scaled a total of 120 sentences (40 clinical subjects \times 3 sentences) in a soundproof booth after brief training on the scaling of speech intelligibility. Speech stimuli were played directly from a computer over a loudspeaker at a comfortable level of loudness.

The speech samples were collected in a quiet room with a high-quality microphone (Shure SM 58) and a digital audio tape recorder (Tascam DA-P1) at a sampling rate of 44.1 kHz and with 16-bit quantization. After the utterances had been collected on DAT, they were analyzed using the program TF32 [18].

Acoustic Analysis

Linear predictive code analysis, using the default value in TF32 of 26 coefficients, generated pitch-synchronous formant tracks for each word repetition. Tracks were hand-corrected by one of the authors (Y.K.) and a trained student. Acoustic measures included F2 starting frequency (the F2 value measured at the first glottal pulse of the corrected track), transition onset frequency (the F2 value measured at the onset of the operationally defined transition, see below), transition offset frequency (the F2 value measured at the offset of the operationally defined transition), transition extent (the frequency change between the transition onset and offset), and transition duration (the time between transition onset and offset). F2 slope was derived using transition extent and duration measures. Transition onsets and offsets were located according to the 20 Hz/20 ms rule [17].

Speech Intelligibility Scores

Modulus-free direct magnitude estimation was used to obtain scaled intelligibility scores for the three sentences identified above, produced by each speaker. Direct magnitude estimation is a method of perceptual ratio scaling in which an observer makes a numerical estimate of the sensory magnitudes associated with a set of stimuli [19]. Because each listener was free to choose his/her own modulus, the original scaled estimates for each judge were transformed to a common scale using a modulus-equaliza-

Table 1. Mean and standard deviation (in parentheses) of F2 slope (Hz/ms), transition duration, and transition extent for each word

	Slope, Hz/ms			Transition duration, ms			Transition extent, Hz		
	PD	stroke	HC	PD	stroke	HC	PD	stroke	HC
<i>coat</i>	2.42 (0.76)	2.31 (0.87)	3.51 (0.98)	92.64 (47.13)	97.32 (41.15)	96.48 (19.23)	220.8 (125.28)	218.07 (103.15)	336.45 (111.65)
<i>hail</i>	4.07 (0.90)	3.92 (0.97)	5.96 (1.05)	221.69 (45.72)	208.51 (39.48)	174.99 (35.38)	884.17 (201.07)	786.66 (212.69)	1,011.53 (98.78)
<i>row</i>	1.75 (0.48)	1.72 (0.48)	1.70 (0.33)	170.68 (65.32)	197.40 (82.30)	229.52 (32.98)	293.11 (104.71)	319.16 (113.64)	390.40 (89.34)
<i>sigh</i>	2.13 (0.57)	2.22 (0.73)	3.38 (0.62)	222.47 (62.22)	232.99 (76.76)	204.16 (32.63)	471.11 (177.72)	498.46 (178.29)	677.08 (99.33)
<i>shoot</i>	4.04 (1.75)	3.45 (1.14)	5.86 (1.82)	112.90 (41.93)	120.99 (49.72)	98.72 (22.28)	441.99 (194.74)	392.20 (185.63)	548.75 (107.37)
<i>wax</i>	3.73 (0.92)	3.47 (1.16)	6.08 (1.58)	213.32 (53.24)	234.57 (73.58)	173.71 (24.89)	770.76 (188.78)	782.39 (258.75)	1,025.53 (169.39)

Absolute values were taken for slope and transition extent. HC = Healthy controls.

tion procedure [20]. A t test showed that the scaled speech intelligibility measures were statistically equivalent for the two patient groups (PD vs. stroke, $t = 0.90$, $d.f. = 38$, $p = 0.37$).

Results

Means and standard deviations of F2 slope for each group and word are provided in table 1. F2 slopes are expressed in absolute values (Hz/ms) throughout this paper to eliminate the positive/negative sign caused by word-inherent F2-transition direction. Summary data for the components of F2 slope (TD = transition duration; TE = transition extent) are also provided.

Group Comparisons

For all words except *row*, absolute F2 slopes were greater for the HC group as compared to either of the patient groups; the magnitude of these differences varied across words. A statistical summary of group comparisons (t tests, alpha level per comparison set to 0.003 to yield a familywise error rate of roughly 0.05 for the set of 18 comparisons) is given for all six words in table 2. Figure 1 shows a box-and-whiskers plot of F2 slopes for selected words (*hail*, *wax*, and *row*). The absolute F2 slope differences between the healthy controls and patient groups are obvious for *hail* and *wax*, but not *row*. This graphical analysis is supported by the statistical analysis, showing no group difference for *row*. The HC-PD comparisons were also nonsignificant for *coat* and *shoot*, and for *coat* in the HC-stroke comparisons, even though the greater slopes for the HC productions of these words fit the pattern of the remaining, significant comparisons. An interesting finding was the absence of significant differences for all six PD-stroke comparisons. Across words, TD is

Table 2. t test results for F2 slope difference between speakers with PD and stroke, and healthy controls (HC)

	PD-stroke	PD-HC	Stroke-HC
<i>coat</i>	NS	NS	NS
<i>hail</i>	NS	**	**
<i>row</i>	NS	NS	NS
<i>sigh</i>	NS	**	**
<i>shoot</i>	NS	NS	**
<i>wax</i>	NS	**	**

Each comparison tested at $p < 0.003$ (critical t value for significance with $d.f. = 23$, $t = 3.104$). ** $p < 0.003$.

not consistently longer or shorter for the HC-PD or HC-stroke comparisons; TE, however, is always greatest for the HC group.

Figure 2 shows F2 slope data for *hail*, *sigh*, *wax* and *row* as cumulative probability distributions for all three groups. Cumulative probability functions rank F2 slope data as a series of normalized probabilities from the lowest to highest value in the distribution. The point where the function crosses a probability of 0.50 is the median, the steepness of the main part of the function provides an indication of the degree of data dispersion, and the shape of the function suggests the degree to which the distribution may be close to normal. The shape of a normal distribution is sigmoidal when plotted as a cumulative probability function.

In figure 2, cumulative probability functions for *hail*, *sigh* and *wax* are clearly displaced to the right for the HC group, relative to the PD and stroke groups, indicating the difference in their central tendencies. The PD and stroke functions for these words nearly overlap. The func-

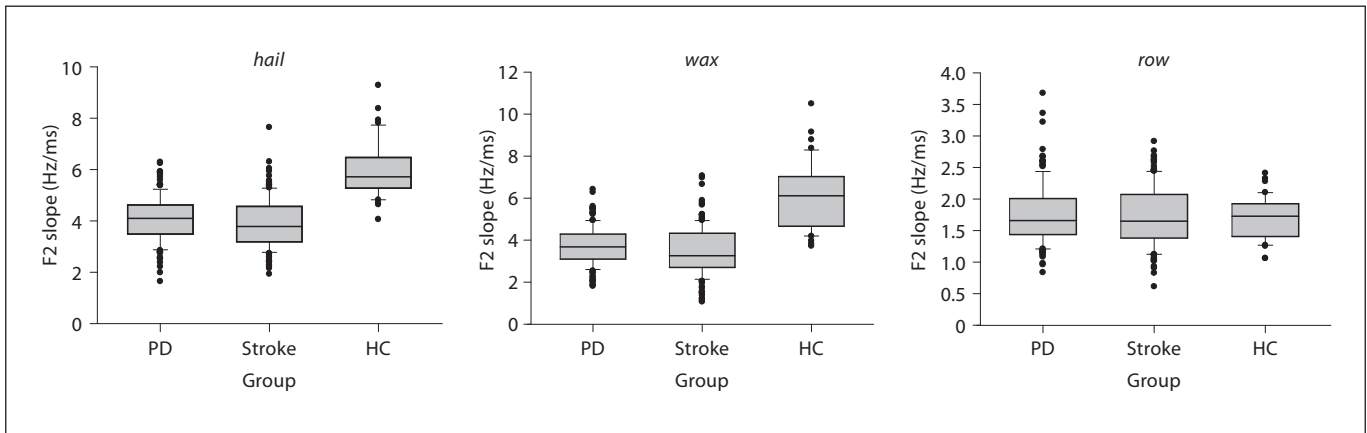


Fig. 1. Group comparison of F2 slope for words *hail*, *wax*, and *row* (median is indicated by solid line in the bar). HC = Healthy controls.

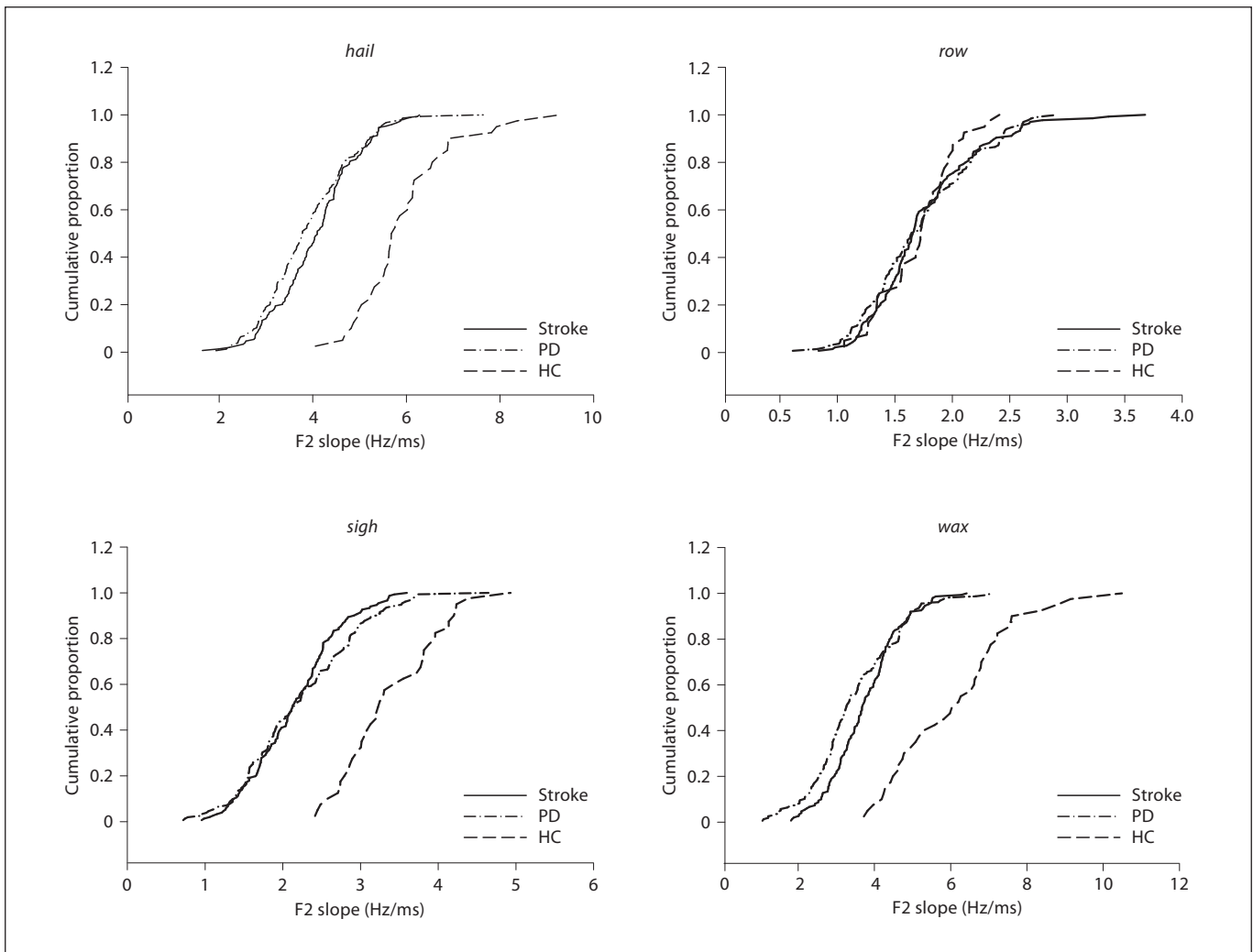


Fig. 2. Cumulative probability functions for F2 slope of words *hail*, *row*, *sigh*, and *wax*. HC = Healthy controls.

tions for *hail* and *wax* have a nearly sigmoidal shape for each group, suggesting roughly normal distributions for F2 slope values. Note also the similar steepness of the main part of the functions, suggesting similar variabilities across groups, at least for these words. In the case of *row*, the functions for all three groups are also roughly of sigmoidal shape and nearly completely overlapped. In this sense, *row* is exceptional in that there is no separation between the three groups. The results for these words show the range of effects for all six words.

The cumulative probability functions for *row* appear to be somewhat less steep for the PD and stroke groups, as compared to the HC group, indicating greater F2 slope variability for the patient groups.

Relationships between Intelligibility Scores and F2 Slope

Regression analysis revealed that F2 slope for only *shoot* ($r^2 = 14.3\%$) and *wax* ($r^2 = 13.9\%$) was significantly regressed against scaled speech intelligibility.

Discussion

A primary aim of this study was to explore a statistical approach to investigating acoustic characteristics of speech in persons with dysarthria. Such distributional data could be the basis for an index or indices of the speech motor control deficit in neuromotor speech disorders (see below).

Examination of F2 slopes in 40 speakers with dysarthria and 5 healthy control speakers revealed shallower slopes for speakers with dysarthria compared to healthy controls, a finding consistent with previous studies [5, 6, 12]. Somewhat surprisingly, however, there was no difference in F2 slope between speakers with PD and stroke, despite the expectation of different speaking rates between speakers in these two patient groups. Speaking rate in patients with stroke has typically been reported as slower than normal [see review in ref. 21, p. 156], whereas in PD, rate has been reported as slow, normal or even faster than normal [21, pp. 203–204], with the majority of reports suggesting normal or faster-than-normal rates [see ref. 6]. Although the relationship between speaking (or articulation) rate and formant transition slopes is not always straightforward [22], the effect of differential speaking rates on the current F2 slope measures cannot be ruled out.

A post hoc analysis of the relationship between F2 slope and the duration of vocalic nuclei in the current test

words, in which the duration of vocalic nuclei was regarded as an estimate of articulation rate [6], showed that the vocalic nucleus of each word spoken by the HC group was shorter than the corresponding durations for either patient group. Across words these differences ranged from 38 to 76 ms for HC-PD comparisons, and from 49 to 100 ms for HC-stroke comparisons. Significant HC-PD and HC-stroke differences, as revealed by t tests, were obtained for three of the words (*coat*, *hail*, *shoot*). The durations of the stroke group were longer than those of the PD group for all six words, the differences ranging from 11 to 53 ms; only the difference for *coat* was significant.

Correlation analysis of the token-to-token covariation of vocalic nucleus duration and F2 slope for each word produced by speakers in both patient groups revealed 4 significant relationships of the total of 12 evaluated (6 words \times 2 groups). The PD group had a single, significant correlation (*wax*), whereas the stroke group had three significant correlations (*hail*, *sigh*, *wax*). The strongest association among these 4 significant correlations accounted for only 44% of the variance between vocalic nucleus duration and F2 slope. In each of these significant correlations, longer durations of the vocalic nucleus were associated with shallower F2 slopes. A consistent relationship between TD and F2 slope did not emerge in the present data (table 1) [5].

Another conclusion from the current study is that it is critical to select appropriate speech materials and tasks for investigating acoustic characteristics of dysarthria, specifically in order to maximize the sensitivity of the material to speech motor control problems. The reduced values of F2 slope in speakers with dysarthria were more apparent in words such as *wax* and *hail* as compared to the other words (*sigh*, *coat*, and *shoot*). Although all words but *row* showed significant F2 slope differences between patient groups and healthy controls, F2 slopes were significantly correlated with scaled speech intelligibility only for *shoot* and *wax*. One possible reason for these word level effects on F2 slope and its relation to speech intelligibility is the magnitude of the intrinsic F2 slope of different vocalic nuclei. It may not be a coincidence that words for which speakers with dysarthria showed dramatically lower-than-normal F2 slopes (*wax*, *hail*) have the highest slope magnitudes among the six words in the data of healthy speakers. Presumably, these words require the most rapid and largest changes in vocal tract geometry for successful production (in the sense of producing the articulatory gestures that make the word intelligible). In contrast, the words not very sensitive to dysarthric speech (*coat*, *row*) showed the lowest F2 slopes in the

healthy control data (table 1). Rosen et al. [10] have also raised the issue of differential sensitivity of speech material to the dysarthria in people with multiple sclerosis.

The production task used in the current study was designed as a potential *speech* test for speech motor control integrity among adults with dysarthria. Speech intelligibility tests are a standard measure of severity of involvement in dysarthria, but, as noted in the literature [7, 23, 24], many factors extrinsic to the speech motor control deficit, such as linguistic content, listener experience, and so forth, affect intelligibility scores or scale values. On the other hand, arguments have been advanced that a 'pure' evaluation of speech motor control requires a separation of motor from linguistic components of speech production [25, 26]. According to this argument, this separation can only be achieved through the use of oromotor, non-verbal tasks. The relevance of such tasks has been challenged, however, as having little ecological validity for the motor control required when movements of the speech mechanism are made audible [16]. The approach explored in the present paper involves repeated words, each word requiring a relatively rapid change in vocal tract configuration throughout its vocalic nucleus. The multiple repetitions of each word by each patient and the use of a fairly large number of speakers permit some initial conclusions about the statistical properties of vocal tract gestures, as inferred from the slope of the second formant frequency [13]. First, the distributions typically seem to approximate normal, as indicated by the sigmoidal shape of the cumulative probability plots (fig. 2). Second, the shapes of the cumulative probability functions are roughly similar across the three speaker groups. When there are differences between groups, it is mainly in the dis-

placement of a cumulative probability function along the x axis (that is, the F2 slope axis), as exemplified by the group differences shown in figure 2 for *hail*, *sigh*, and *wax*. This finding, if replicated with additional speakers and for different acoustic measures, has an interesting statistical implication for a speech measure of speech motor control deficit [27]. Specifically, a distribution derived from normal speakers can be used to convert measures obtained from speakers with dysarthria to normalized (z) scores, which may serve as a distance metric of speech motor control deficit. The development of such a metric requires further exploration of the differential phonetic sensitivity to dysarthria, mentioned above, as well as the relation of the metric to other indices of dysarthria severity (such as speech intelligibility). One clear advantage of the metric is the use of *speech* to estimate the magnitude of a speech motor control deficit, as compared to an arbitrarily selected, oromotor nonspeech task that is almost certain to be novel for the typical patient [16]. A speech-based metric is not only ecologically valid, but does not require a patient to perform a task that is new and very different from the movement-control requirements of speech. Of course, it may still be of interest to compare a measure such as F2 slope with nonspeech motor tasks, as has been done with PD [28].

Acknowledgments

Authors thank Christina Kuo for assistance with data collection from listeners and manual correction of raw F2 trajectories. This study was funded by NIH DC00319 from the National Institute on Deafness and Other Communication Disorders.

References

- 1 Ansel BM, Kent RD: Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *J Speech Hear Res* 1992;35:296–308.
- 2 Liu HM, Tsao FM, Kuhl PK: The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *J Acoust Soc Am* 2005;117:3879–3889.
- 3 Weismer G, Martin RE: Acoustic and perceptual approaches to the study of intelligibility; in Kent RD (ed): *Intelligibility in Speech Disorders: Theory, Measurement, and Management*. Philadelphia, Benjamins, 1992, pp 67–118.
- 4 Whitehill TL, Ciocca V: Speech errors in Cantonese speaking adults with cerebral palsy. *Clin Linguist Phon* 2000;14:111–130.
- 5 Weismer G: Assessment of articulatory timing; in Cooper J (ed): *Assessment of Speech and Voice Production: Research and Clinical Applications*. NIDCD Monogr. Bethesda, National Institutes of Health, 1991, vol 1, pp 84–95.
- 6 Weismer G, Jeng J, Laures R, Kent RD, Kent JF: Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatr Logop* 2001; 53:1–18.
- 7 Kent RD, Weismer G, Kent JF, Martin RE, Sufit RL, Rosenbek JC, Brooks BR: Relationships between speech intelligibility and the slope of second-formant transitions in amyotrophic lateral sclerosis. *Clin Linguist Phon* 1989;3:347–358.
- 8 Kent JF, Kent RD, Rosenbek JC, Weismer G, Martin RE, Sufit RL, Brooks BR: Qualitative description of the dysarthria in women with amyotrophic lateral sclerosis. *J Speech Hear Res* 1992;35:723–733.
- 9 Weismer G, Martin R, Kent RD, Kent JF: Formant trajectory characteristics of males with amyotrophic lateral sclerosis. *J Acoust Soc Am* 1992;91:1085–1098.

- 10 Rosen KM, Goozée JV, Murdoch BE: Examining the effects of multiple sclerosis on speech production: does phonetic structure matter? *J Commun Disord* 2008;41: 49–69.
- 11 Tjaden K, Wilding GE: Rate and loudness manipulations in dysarthria: acoustic and perceptual findings. *J Speech Lang Hear Res* 2004;47:766–783.
- 12 Weismer G, Wildermuth J: Formant trajectory characteristics in persons with Parkinson, cerebellar, and upper motor neuron disease. 135th Meet Acoust Soc Am, Seattle, 1998.
- 13 Stevens KN: *Acoustic Phonetics*. Cambridge, MIT Press, 2000.
- 14 Weismer G: Motor speech disorders; in Hardcastle WJ, Laver J (eds): *The Handbook of Phonetic Sciences*. Cambridge, Blackwell, 1997, pp 191–219.
- 15 Ziegler W: Speech motor control is task-specific: evidence from dysarthria and apraxia of speech. *Aphasiology* 2003;17:3–36.
- 16 Weismer G: Philosophy of research in motor speech disorders. *Clin Linguist Phon* 2006; 20:315–349.
- 17 Weismer G, Kent RD, Hodge M, Martin R: The acoustic signature of intelligibility test words. *J Acoust Soc Am* 1988;84:1281–1291.
- 18 Milenkovic P: TF 32 [computer program]. University of Wisconsin-Madison, 2001.
- 19 Gescheider GA: *Psychophysics, Method and Theory*. Hillsdale, Erlbaum, 1976.
- 20 Engen T: Psychophysics. II. Scaling methods; in Kling JW, Riggs L (eds): *Woodworth and Schlossberg's Experimental Psychology*. New York, Holt, Rinehart, & Winston, 1971, pp 47–86.
- 21 Duffy J: *Motor Speech Disorders: Substrates, Differential Diagnosis, and Management*, ed 2. St Louis, Mosby, 2005.
- 22 Weismer G, Berry J: Effects of speaking rate on second formant trajectories of selected vocalic nuclei. *J Acoust Soc Am* 2003;113: 3362–3378.
- 23 Monsen RB: The oral speech intelligibility of hearing-impaired talkers. *J Speech Hear Disord* 1983;48:286–296.
- 24 Weismer G: Speech intelligibility; in Ball MJ, Perkins M, Müller N, Howard S (eds): *The Handbook of Clinical Linguistics*. Oxford, Blackwell, 2008, pp 568–582.
- 25 Ballard KJ, Robin DA, Folkins JW: An integrative model of speech motor control: a response to Ziegler. *Aphasiology* 2003;17:37–48.
- 26 Robin DA, Solomon NP, Moon JB, Folkins JW: Nonspeech assessment of the speech production mechanism; in McNeil MR (ed): *Clinical Management of Sensorimotor Speech Disorders*. New York, Thieme, 1997, pp 49–62.
- 27 Weismer G, Kim YJ: Classification and taxonomy of motor speech disorders: what are the issues?; in Maassen B, Van Lieshout, PHHM (eds): *Speech Motor Control: New Developments and Applied Research*. Oxford, Oxford University Press, in press.
- 28 Goberman AM: Correlation between acoustic speech characteristics and non-speech motor performance in Parkinson disease. *Med Sci Monit* 2005;11:CR109–CR116.