The Acoustics of Contrastive Prosody in Adults With Cerebral Palsy

Rupal Patel, Ph.D.

Northeastern University
Bouve College of Health Sciences
Department of Speech Language Pathology and Audiology
Boston, Massachusetts

Recent studies have documented that adults with cerebral palsy can control prosodic features such as pitch, loudness, and duration despite severely impaired segmental control. Although their productions are restricted in range, they are able to modulate prosody within that narrowed range. This control has been demonstrated for isolated vowel productions and in short phrases produced with interrogative versus imperative tone. The present study focused on finer grained manipulations of prosody such as those required for indicating contrastive stress within a phrase. The results reported in this article constitute preliminary findings of a larger study. Three speakers with severe dysarthria due to cerebral palsy and three nonimpaired healthy controls produced three, four-word phrases with contrastive stress placed on one of the four words. Acoustic analyses of their productions revealed that all speakers were able to place contrastive stress at all four phrase locations. Both groups used increased fundamental frequency (F0), intensity, and duration to mark the stressed word. In fact, speakers with dysarthria were able to achieve F0 values in the same range as control speakers. Additionally, speakers with dysarthria increased intensity to a greater extent than the control speakers, perhaps indicating a cue exchange or compensatory physiological modification. Implications for intervention are discussed.

Prosodic deviations in fundamental frequency (F0), intensity, and rate have been noted across various etiologies of dysarthria including cerebral palsy (cf. Darley, Aronson, & Brown, 1969; Hardy, 1983; Murry, 1983), across languages (see Le Dorze, Ouellet, & Ryalls, 1994 for French; Whitehill, Ciocca, & Lam, 2001 for Cantonese), and across severity levels of dysarthria (cf. Patel 2002a, 2002b, 2003; Vance, 1994; Yorkston, Beukelman, Minifie, & Sapir, 1984). Reduced prosodic range and modulation within that range, however, are separate issues. Speakers with compromised control of segmental units of speech may nonetheless be able to convey their intentions using prosodic contrasts. Prosodic cues vary more gradually and at slower time scales than segmental units such as phonemes and thus may be easier to produce by speakers with dysarthria, a motor speech impairment characterized by slow, weak, and imprecise movements.

In healthy speakers, contrastive stress is signaled by increased F0, increased intensity, and prolonged syllable duration on the word of focus (cf. Bolinger, 1961; Fry, 1955; Lehiste, 1970; Lieberman, 1960; Morton & Jassem, 1965). In contrast, stress patterning in dysarthria is poorly understood and largely overlooked. Previous studies have shown that increased physiological effort needed to signal syllabic prominence may lead to inaccurate, inconsistent, exaggerated, and bizarre stress patterning in dysarthria (Liss & Weismer, 1992, 1994; Netsell, 1973; Yorkston et al., 1984).

Journal of Medical Speech-Language Pathology
Volume 12, Number 4, pp. 189–193

This study aimed to determine whether speakers with severe dysarthria due to cerebral palsy can exploit their narrowed range of F0 and intensity, and slowed rate, in order to signal contrastive stress. Prosodic cues required for marking contrastive focus may also be highly informative cues for conveying communicative intent. Thus, identifying and harnessing the residual prosodic control abilities available to speakers with dysarthria may lead to novel intervention strategies for improved communication success and efficiency.

Specific experimental questions addressed in this work include: (1) Which acoustic cues do speakers with dysarthria use to mark contrastive stress? (2) Are some word positions easier to stress and if so, how are they marked? (3) How do speakers with dysarthria and healthy controls differ in the acoustic cues they use to mark the contrast?

METHOD

Participants

Speakers With Dysarthria

Three female monolingual speakers of American English with spastic dysarthria due to cerebral palsy (48–62 years; mean = 56 years) participated in this study. The author, a certified speech-language pathologist (SLP), screened all subjects using a battery of formal and informal evaluations to examine the nature and extent of speech impairment. An additional experienced SLP listened to samples of each speaker with dysarthria and verified the author's assessment. A modified version of the Assessment of Intelligibility of Dysarthric Speech (AIDS) (Yorkston & Beukelman, 1981) consisting of 25 single words was used to determine dysarthria severity based on ratings from three unfamiliar raters. Intelligibility ratings for all three speakers fell below 15%, indicating severe impairment. All participants passed an audiometric evaluation and demonstrated adequate receptive language and cognitive skills necessary for the experimental task.

Healthy Controls

Three female monolingual speakers of American English with normal hearing function (23–37 years; mean = 32 years) served as healthy controls.

Materials and Apparatus

Recordings of healthy controls were collected in an audiometric booth using a MiniDisc recorder (HHB 500 PortaDisc) and a unidirectional head-mounted cardiod microphone (Shure, SM10A) placed 2 cm from the corner of the speaker's mouth. Given geographical constraints and policies of the clinical site, recordings of speakers with dysarthria were collected in a quiet, isolated room at an outpatient facility in New Jersey using identical procedures and equipment as in the laboratory. Acoustic analyses on sample productions were used to verify adequate recording quality prior to data collection.

Procedure

All speakers produced three phrases, each four words in length, spoken with emphasis on the first, second, third, or fourth word. Contextual scenarios were used to elicit contrastive stress at each word position. Each phrase was also produced neutrally (i.e., "without emphasis on any particular word"; similar to Weismer & Ingrisano, 1979). Phrase length was minimized to facilitate dysarthric productions in light of poor breath support and coordination. To control for inherent differences in vowel F0, all phrases consisted of monosyllabic words with high front vowels ("He lives near me"). Five repetitions of each phrase and stress location were requested, resulting in a total of 75 recordings per speaker.

Acoustic Analyses

All utterances were sampled at 22,050 Hz. The Praat speech analysis software package (Boersma & Weenik, 2000) was used to manually mark the beginning and end of each word, using spectrographic as well as F0 contour and intensity cues. Interjudge reliability of boundary placement was calculated for 10% of the data; $r = 0.932$. Praat was also used to extract F0 and intensity contours for each phrase. A customized software program operated on the Praat output to calculate the following acoustic features per word: duration, peak F0 (F0peak), average F0 (F0ave), peak intensity (INTpeak), and average intensity (INTave).

RESULTS

Acoustic analyses were conducted on all 225 dysarthric (DYS) and all 225 healthy control (HC) productions. Separate repeated measures analyses of variance were conducted for each acoustic feature. Each analysis examined the effect of one between-subject factor (group; DYS vs. HC) and two within-subject factors. Factor 1 represented emphasis type, (i.e., whether emphasis was requested on the first [a], second [b], third [c], or fourth word [d], or if neutral production [n] was requested). Factor 2 represented stressed word position (word 1 [W1], word 2 [W2], word 3 [W3], and word 4 [W4]). The response variables were all continuous; F0 in Hz, intensity (INT) in dB, and duration in seconds. The F statistic was used to test the null hypothesis with $\alpha = 0.05$. For each acoustic feature, 36 pairwise contrasts (with Bonferroni corrected α levels) were conducted to examine differences within and across word, emphasis, and group. Figure 1 shows the results of the three most salient acoustic cues

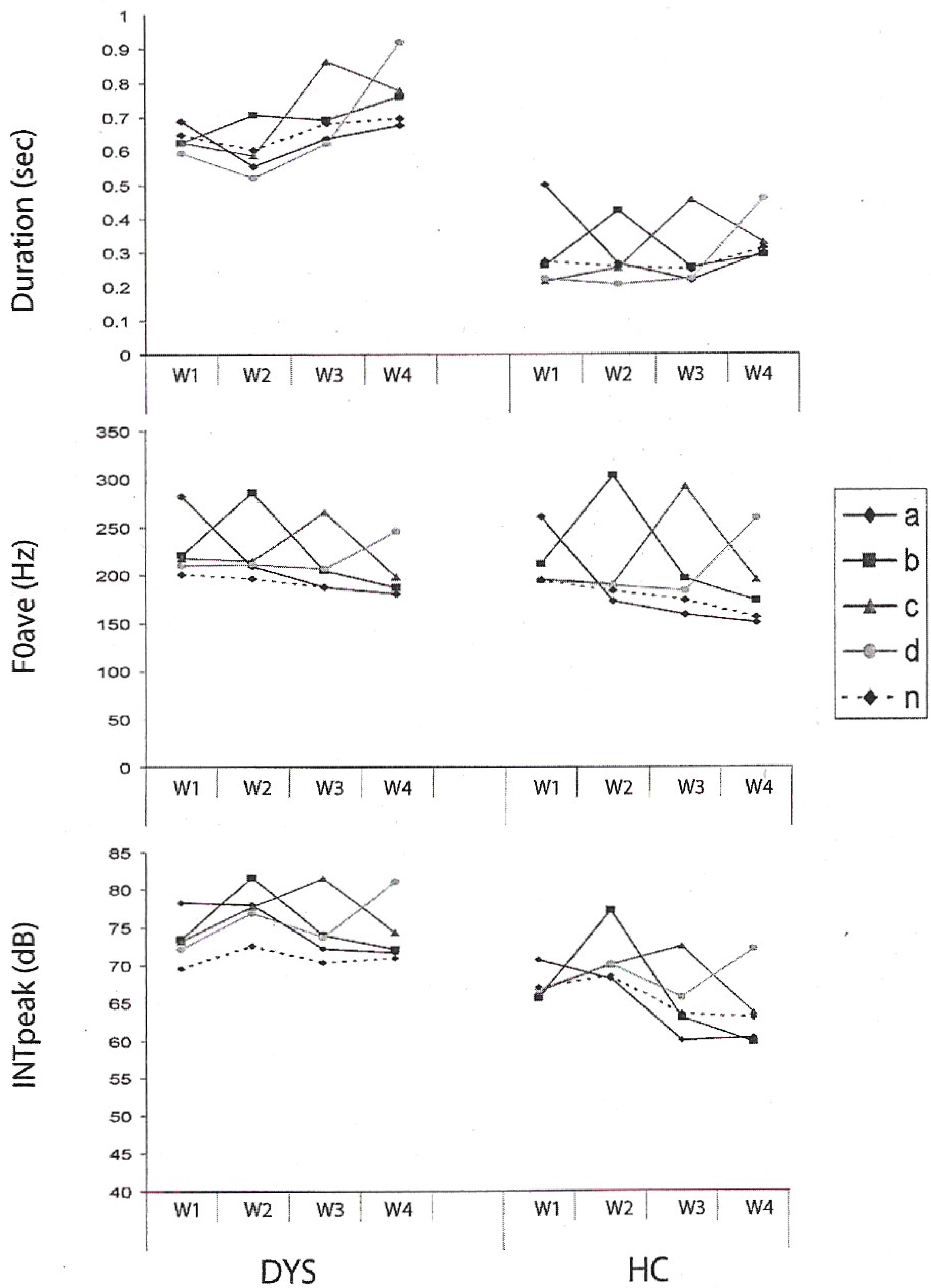

Figure 1. Top graph: Mean duration (seconds) at each word position for each emphasis type. Middle graph: Mean average fundamental frequency (Hz) at each word position for each emphasis type. Bottom graph: Mean peak intensity (dB) at each word position for each emphasis type. The results for speakers with dysarthria (DYS) are plotted on the left and for healthy control speakers (HC) on the right.

(Duration, F0ave, and INTpeak) used by DYS and HC to mark contrastive stress. Because F0peak and INTave results were essentially identical to the F0ave and INTpeak results, respectively, they have not been plotted. Statistical summaries for all five acoustic features are provided below.

Duration

Statistically significant main effects in duration were found for group ($F(1,4) = 117.32, p = 0.0004$) and emphasis ($F(4, 16) = 4.11; p = 0.0177$). The main effect for word was not significant; however, the two-way interactions between word and emphasis, and three-way interaction between word, emphasis, and group, were significant ($p < 0.05$). Both groups elongated word duration to indicate contrastive stress; however, DYS productions were approximately two times slower than HC productions for all emphasis locations.

F0

Statistically significant main effects in F0peak were found for word ($F(3,12) = 12.14; p = 0.0006$) and emphasis ($F(4,16) = 16.07; p < 0.0001$), but not for group ($F(1,4) = 2.01; p = 0.229$). Only the two-way interaction between word and emphasis was significant ($F(12,48) = 63.50; p < 0.0001$). Both groups used increased F0peak to signal emphasis at all word positions, and the degree of change in F0peak was similar across speakers in each group.

Statistically significant main effects in F0ave were found for word ($F(3,12) = 18.23; p < 0.0001$) and emphasis ($F(4,16) = 11.75; p = 0.0001$), but not for group ($F(1,4) = 1.02; p = 0.369$). The two-way interaction between word and emphasis was highly significant ($F(12,48) = 50.23; p < 0.0001$), and the three-way interaction between word, emphasis, and group was also significant ($F(12,48) = 2.10; p = 0.035$). Both groups use increased F0ave as an emphasis cue at all word positions, and the range and degree of F0ave change was similar across speaker groups.

Intensity

Statistically significant main effects in INTpeak were found for word ($F(3,12) = 11.45; p = 0.0008$) and emphasis ($F(4,16) = 26.76; p < 0.0001$), but not for group ($F(1,4) = 3.51; p = 0.134$). The two-way interactions between word and emphasis ($F(12,48) = 24.18; p < 0.0001$), and emphasis and group ($F(4,16) = 7.15; p = 0.0017$) were also significant. Although DYS and HC use increased INTpeak to indicate emphasis at all word positions, DYS productions tended to be higher in INTpeak overall.

Statistically significant main effects in INTave were found for word ($F(3,12) = 4.98; p = 0.018$) and emphasis ($F(4,16) = 10.56; p = 0.0002$), but not for group ($F(1,4) = 3.38; p = 0.139$). The two-way interactions between word and emphasis ($F(12,48) = 15.81; p < 0.0001$), and emphasis and group ($F(4,16) = 6.21; p = 0.0032$), as well as the three-way interaction between word, emphasis, and group was also significant ($F(12,48) = 2.24; p = 0.024$). Although DYS and HC used increased INTave to indicate emphasis, DYS consistently used higher INTave for all word positions. HC relied less on INTave and did not use it at all to mark stress at W2.

DISCUSSION

The present study examined whether speakers with dysarthria could modulate prosodic cues to mark contrastive stress within short phrases. Acoustic analyses revealed both groups were able to place contrastive stress at all four word positions. Both groups used increased F0, INT, and duration to mark the stressed word. In fact, speakers with dysarthria and healthy controls did not differ in the extent or range of F0 used to mark stress. Additionally, speakers with dysarthria increased intensity to a greater extent than healthy controls, perhaps indicating a cue exchange (Howell, 1993) or compensatory physiological modification. It is possible that dysarthric productions were higher in F0 because of concurrent changes in intensity. Perhaps they increased medial compression and subglottal pressure (through increased respiratory effort) to mark stress and thereby simultaneously achieved an increase in intensity and fundamental frequency. Speaker groups differed most in word duration, a finding consistent with previous reports of markedly slow rate in dysarthria. It is possible that this slowed rate in fact benefits speakers with dysarthria because it allows them sufficient time to accrue physiological resources such as vocal tension and subglottal pressure, which are required to modulate intensity and F0. Corroborating evidence from physiological measures will be necessary to help substantiate these hypotheses.

Until recently, the residual speech production abilities of speakers with dysarthria have typically been disregarded as having communicative function. Although these results are still quite preliminary given the sample size, they suggest that speakers with dysarthria could be taught to exploit their prosodic control capabilities for conveying a host of communicative intents. Given that dysarthric and healthy control productions shared similar acoustic cues, it seems that intervention aimed at enhancing latent prosodic control could be shaped to maintain relatively natural stress patterns. It will be important, however, to monitor the degree and extent to which extraneous and redundant cues are used by speakers with dysarthria to achieve contrastive stress so as to avoid the "bizarre and unnatural" patterns reported by others (cf. Liss & Weismer, 1994; Yorkston et al., 1984). The results from this and previous studies of prosody in dysarthria (cf. Patel 2002a, 2002b, 2003, Vance 1994; Yorkston et. al., 1984) suggest that addressing prosodic control at the earliest stages of intervention may provide the scaffolding on which to build improved speech intelligibility. A larger study with up to 15 speakers with dysarthria and 15 healthy controls, a more diverse set of stimuli, and additional acoustic and perceptual measures is underway.

Acknowledgments This study was conducted in the Department of Speech Language Pathology at Teachers College Columbia University. My gratitude goes to the participants for their time and enthusiasm for the study, to Lisa Wayne for helping with the data collection, to Deepak Chakravadhanula and Elyes Yaich for developing various programs that facilitated data labeling and analysis, and to Howard Cabral for his contributions to the statistical analysis.

Address correspondence to Rupal Patel, Northeastern University, Bouve College of Health Sciences, Department of Speech Language Pathology and Audiology, 360 Huntington Ave., Rm. 102 FR, Boston, MA 02115 USA.
e-mail: r.patel@neu.edu

REFERENCES

Boersma, P., & Weenik, D. (2000). Praat, A system for doing phonetics by computer (Version 4.2). [Computer software]. Available at www.praat.org

Bolinger, D. (1961). Contrastive accent and contrastive stress. *Language, 37,* 83–96.

Darley, F., Aronson, A., & Brown, J. (1969). Differential diagnostic patterns of dysarthria. *Journal of Speech and Hearing Research, 12,* 246–269.

Fry, D. (1955). Duration and intensity as physical correlates of linguistic stress. *Journal of the Acoustical Society of America, 27*, 765–768.

Hardy, J. (1983). *Cerebral palsy.* Englewood Cliffs, NJ: Prentice-Hall.

Howell, P. (1993). Cue trading in the production and perception of vowel stress. *Journal of the Acoustical Society of America, 94*, 2063–2073.

Le Dorze, G., Ouellet, L., & Ryalls, J. (1994). Intonation and speech rate in dysarthric speech. *Journal of Communication Disorders, 27,* 1–18.

Lehiste, I. (1970). *Suprasegmentals.* Cambridge: MIT Press.

Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *Journal of the Acoustical Society of America, 32,* 451–454.

Liss, J., & Weismer, G. (1992). Qualitative acoustic analysis in the study of motor speech disorders. *Journal of the Acoustical Society of America, 92,* 2984–2987.

Liss, J., & Weismer, G. (1994). Selected acoustic characteristics of contrastive stress in control geriatric, apraxic and ataxic dysarthric speakers. *Clinical Linguistics & Phonetics, 8,* 43–66.

Morton, J., & Jassem, W. (1965). Acoustic correlates of stress. *Language & Speech, 8,* 159–181.

Murry, T. (1983). The production of stress in three types of dysarthric speech. In W. Berry (Ed.), *Clinical dysarthria* (pp. 69–84). Austin, TX: Pro-Ed.

Netsell, R. (1973). Speech physiology. In F. Minifie, T. J. Hixon, & F. Williams (Eds.), *Normal aspects of speech, hearing, and language* (pp. 211–234). Englewood Cliffs, NJ: Prentice-Hall.

Patel, R. (2002a). Phonatory control in adults with cerebral palsy and severe dysarthria. *Alternative & Augmentative Communication, 18,* 2–10.

Patel, R. (2002b). Prosodic control in severe dysarthria: Preserved ability to mark the question-statement contrast. *Journal of Speech, Language & Hearing Research, 45*, 858–870.

Patel, R. (2003). Acoustic characteristics of the question-statement contrast in severe dysarthria due to cerebral palsy. *Journal of Speech, Language & Hearing Research, 46*, 1401–1415.

Vance, J. (1994). Prosodic deviation in dysarthria: A case study. *European Journal of Disorders of Communication, 29*, 61–76.

Weismer, G., & Ingrisano, D. (1979). Phrasal-level timing patterns in English: Effects of emphatic stress location and speaking rate. *Journal of Speech & Hearing Research, 22,* 516–33.

Whitehill, T., Ciocca, V., & Lam, S. (2001). Fundamental frequency control in connected speech in Cantonese speakers with dysarthria. In B. Maassen, W. Hulstijn, R. Kent, H. Peters, & P. van Lieshout (Eds.), *Speech motor control in normal and disordered speech* (pp. 228–231). Nijmegen, Netherlands: University of Nijmegen Press.

Yorkston, K. M., & Beukelman, D. R. (1981). *The assessment of intelligibility of dysarthric speakers.* Austin, TX: Pro-Ed.

Yorkston, K. M., Beukelman, D. R., Minifie, F., & Sapir, S. (1984). Assessment of stress patterning. In M. McNeil, J. Rosenbek, & A. Aronson (Eds.), *The dysarthria: Physiology, scoustics, perception, management* (pp. 131–162). Austin, TX: Pro-Ed.