Acoustic Characteristics of the Question–Statement Contrast in Severe Dysarthria Due to Cerebral Palsy

Rupal Patel*
Teachers College,
Columbia University,
New York, NY

Studies of prosodic control in severe dysarthria (DYS) have focused on differences between impaired and nonimpaired speech in terms of the range and variation of fundamental frequency (F0), intensity, and duration. Whether individuals with severe DYS *can* adequately signal prosodic contrasts and *which* acoustic cues they use to do so has received far less attention. This article focused on the question–statement contrast. In nonimpaired speech, this contrast is believed to be cued primarily by F0, although some researchers have argued that duration also plays a role. This study examined how 8 speakers with severe DYS due to cerebral palsy signaled the question–statement contrast for a set of 10 short phrases. An additional 8 healthy controls (HCs) produced the same set of phrases as questions and statements. To analyze the speech recordings, peak F0 (F0peak), average F0 (F0ave), slope of F0 (F0slope), peak intensity (INTpeak), average intensity (INTave), slope of intensity (INTslope), and duration measures were calculated for each syllable (S1, S2, S3) within each phrase. Acoustic analyses revealed that speakers with DYS and HCs used F0, duration, and to a lesser degree, intensity cues to signal the contrast. Moreover, productions by speakers with DYS had longer and louder S3 for questions compared to productions by HCs, suggesting that speakers with DYS may have been compensating for their reduced ability to control F0 by exploiting their residual control of loudness and duration. Data from a previous perceptual study (R. Patel, 2002b) with the same speakers with DYS were used to analyze the relationship between acoustic characteristics and listener perceptions of their productions. Logistic regression analysis revealed that S1_F0ave; S2_duration; and S3_duration, S3_F0peak, S3_F0slope, S3_INTave, and S3_INTslope were significant predictors of the perceived prosodic contrast. Identifying acoustic consistencies in prosodic control among speakers with DYS provides the impetus to build vocalization recognition algorithms that are capable of processing dysarthric speech for use in assistive communication aids. These findings suggest that speakers with DYS may also benefit from intervention aimed at improving prosodic control such that these contrasts may be exploited for communication.

KEY WORDS: dysarthria, cerebral palsy, prosody, acoustics, fundamental frequency, question–statement, pitch, intonation, loudness, syllable duration

Individuals with severe dysarthria (DYS) rely on augmentative and alternative communication systems, such as a voice output communication aid, to express their thoughts and needs. At present, these systems can only be accessed through pointing or scanning interfaces. Despite poor speech intelligibility, many users prefer to use their residual

*Currently at Northeastern University, Boston

speech abilities when interacting with familiar caregivers. Speech is a natural, socially satisfying, and efficient method of interaction. Enabling users to access their communication aids using vocalizations may improve communication efficiency, improve social and interpersonal aspects of interaction, and enhance overall communication satisfaction. This work is motivated in part by a long-term goal to develop a dedicated speech recognition system that is tuned to the speech production abilities of individuals with severe DYS. Furthermore, these findings may provide the basis for developing intervention strategies that harness residual prosodic control capabilities for communicative function. As a first step toward these goals, it is essential to identify information-bearing acoustic characteristics of dysarthric speech.

Numerous studies have examined the perceptual and acoustic properties of segmental cues in DYS (cf. Ansel & Kent, 1992; Darley, Aronson, & Brown, 1969; Hardy, 1983; Platt, Andrews, & Howie, 1980; Platt, Andrews, Young, & Quinn, 1980; Rosenbek & La Pointe, 1978). The study of prosodic control, however, has been less extensive and has tended to focus on differences between speakers with and without DYS (cf. Canter, 1963; Darley et al., 1969; Hardy, 1983; Le Dorze, Ouellet, & Ryalls, 1994; Rosenbek & La Pointe, 1978; Wit, Maassen, Gabreels, & Thoonen, 1993; Yorkston, Beukelman, Minifie, & Sapir, 1984). Relative to nonimpaired speech, DYS has been characterized by one or more of the following prosodic qualities: monopitch, monoloudness, and abnormally fast or slow rate. Impairments in range and flexibility of control of F0, intensity, and excessively fast or slow rate have been noted across various etiologies of DYS, including congenital disorders such as cerebral palsy (cf. Hardy, 1983; Irwin, 1955; Murry, 1983; Rosenbek & La Pointe, 1978), acquired disorders (cf. Baum & Pell, 1997; Canter, 1963; Danly & Shapiro, 1982; Robin, Klouda, & Hug, 1991), and hearing impairment (cf. Allen & Arndorfer, 2000; Hood & Dixon, 1969; Rosenhouse, 1986). These deviations of prosody have also been noted in languages other than English, including French (Le Dorze et al., 1994) and Cantonese (Whitehill, Ciocca, & Lam, 2001), and across severity levels of DYS (cf. Patel 2002a, 2002b; Vance, 1994; Yorkston et al., 1984). Narrowed range of F0 and intensity variation, however, do not preclude communicative use of that range. It is plausible that speakers with DYS can learn to exploit their residual range to consistently signal prosodic contrasts.

Alterations of prosody may also impact speech intelligibility. Studies have shown that flattening F0 contours reduced speech intelligibility of nonimpaired speech (Laures & Weismer, 1999; Wingfield, Lombardi, & Sokol, 1984) and dysarthric speech (Bunton, Weismer, & Kent, 2000). Thus, identifying and harnessing residual prosodic control abilities of speakers with DYS may improve overall communication efficiency.

Question–Statement Contrast

Although many researchers believe that a rising F0 contour is the primary acoustic cue for marking the question–statement contrast (cf. Cruttenden, 1986; Eady & Cooper 1986; Hadding-Koch & Studdert-Kennedy, 1964; Majewski & Blasdell, 1969; O'Shaughnessy, 1979; Shattuck-Hufnagel & Turk, 1996), others have suggested that duration is a key factor for marking linguistic stress (Fry, 1955; Lieberman, 1967; Morton & Jassem, 1965).

Some researchers have argued that intonational patterns only act as linguistic conventions, and that these patterns are highly influenced by contextual factors (Howell, 1993; Välimaa-Blum, 2001). Lieberman (1960, 1967), Howell, and other proponents of cue-trading relations have suggested that speakers can vary the relative salience of prosodic markers within equivalent cue combinations to communicate the perception of stress. Howell found that although different speakers mark stress using different sequences of acoustic features (i.e., some may use F0, others may use loudness as a prominent feature, etc.), listeners were able to tune to the speakers' idiosyncratic stress pattern even if they differed from their own method of signaling stress. In contrast, studies of listener perception of synthetic speech have shown that although intensity and duration can signal prosodic contrasts, F0 is the dominant cue and conflicting F0 evidence can override other cues (Denes, 1959; Denes & Milton-Williams, 1962).

Regardless of whether F0 is the prominent cue, if prosodic cues can be "traded," it may be possible for speakers with DYS who have poor range and variation of F0 to signal the question–statement contrast using alternate cues such as intensity and duration. Some researchers have posited that speakers with DYS may exploit cue-trading equivalencies to remap control of physiologically constrained prosodic parameters onto other features that are still within their control (Brewster, 1989; Patel 2002a; Vance, 1994). Speakers may substitute cues within a prosodic parameter such as using falling tones in place of rising tones or they may trade between prosodic parameters such as increasing loudness in place of heightening pitch (Vance, 1994). It is also possible that speakers with DYS do not differ from healthy control (HCs) in which cues they use, but differ only in the extent to which they use these cues. Along these lines, hearing impaired speakers use the same cues as nonimpaired speakers to mark prosodic contrasts (namely F0, duration, and intensity), but their productions are less pronounced than those of HCs (Allen & Arndorfer, 2000).

In previous work, speakers with DYS due to cerebral palsy were able to control pitch and duration for sustained vowel productions (Patel, 2002a). Moreover, the same group of speakers was also able to successfully signal the question–statement contrast such that human listeners could classify their productions with accuracy levels ranging from 81% to 98% (Patel, 2002b). To shed light on which acoustic features may have impacted listener judgments, listeners were asked to classify an additional set of stimuli in which combinations of pitch and duration cues were removed. Flattening the F0 contour dramatically reduced listener performance, whereas removal of duration cues had little, if any, impact. It remained unclear, however, how speakers conveyed the contrast. Did they raise F0 throughout the utterance? Did they drop F0 instead of raising F0? Did they use the same F0 pattern as speakers without DYS? Did they alter intensity or syllable duration instead of, or in addition to, changes in F0? It was also unclear whether speakers with DYS only altered F0 or if they also altered syllable duration and intensity but listeners did not make use of this information given the prominence of F0 cues. In the present study, two different analyses were conducted to answer these questions: (a) an acoustic analysis of dysarthric productions from Patel (2002b) replicated with 8 HCs and (b) a regression analysis to relate the acoustic findings of this study with the perceptual findings of Patel (2002b).

The specific experimental questions of the present study included the following: How do speakers with DYS use F0, duration, and intensity to signal the question–statement contrast? How do these patterns differ from how HCs signal the contrast? How well can listener perceptions of the question–statement contrast produced by speakers with DYS be predicted by various combinations of acoustic cues?

Method

Participants

Speakers With DYS

The DYS dataset used in the present study was first described in Patel (2002b). The dataset consisted of question–statement productions of 8 individuals with primarily spastic DYS due to cerebral palsy. Speakers ranged in age from 27 to 44 years (*M* = 36 years) and were monolingual speakers of English. Speakers met a set of four selection criteria. First, all speakers had a primary speech diagnosis of DYS. The referring clinicians' assessment was confirmed by the investigator using a battery of formal and informal evaluations that examined the nature and degree of speech motor impairment. Second, a modified version of the Assessment of Intelligibility of Dysarthric Speech (Yorkston & Beukelman, 1981) was administered to determine the level of severity of DYS. To minimize speaker fatigue, the modified assessment consisted of 25 rather than 50 isolated word productions. The average of three unfamiliar raters' scores was used as a measure of the speaker's intelligibility. All of the speakers' speech intelligibility ratings fell below 25%, underscoring the severe extent of speech impairment in this group. Third, speakers were required to pass a pure-tone audiometric evaluation with thresholds at or below 25 dB HL in at least one ear. Fourth, all speakers demonstrated grossly adequate receptive language and cognitive skills necessary for completing the experimental task (see Table 1 for a description of speakers with DYS).

HCs

The Patel (2002b) protocol was replicated with 8 normal-hearing, monolingual speakers of English between the ages of 21 and 37 years (*M* = 30 years). The HC speakers were matched in gender with the DYS dataset, which included 2 women and 6 men. All HCs passed an audiometric screening evaluation with average pure-tone thresholds (at 500, 1000, and 2000 Hz) at or below 25 dB HL in at least one ear.

Materials and Apparatus

Speech recordings were collected in a sound-treated audiometric booth using a digital audiotape recorder (Sony, PCM-2300) and a unidirectional head-mounted cardioid dynamic microphone (Shure, SM10A) placed 2 cm from the left corner of the speaker's mouth.

Procedure

Each speaker was recorded while he or she produced 10 phrases spoken as a question and the same 10 phrases spoken as a statement. All phrases were three syllables in length to facilitate dysarthric productions in light of poor breath support and coordination (for a description of the stimuli selection, see Patel, 2002b). The Appendix provides the phrase list and contextual scenarios used to elicit the contrast. Each phrase was produced as a statement five times and as a question five times, resulting in a total of 100 recordings (hereafter referred to as tokens) per speaker. To control for order effects, phrase type and sentence type orders were randomized across speakers. To be consistent with Patel (2002b), a randomly selected subset of 60 utterances (30 question tokens and 30 statement tokens) per HC speaker was used for acoustic analyses.

Acoustic Analyses

In total, 480 utterances by speakers with DYS and 480 utterances by HCs were acoustically analyzed. All utterances were sampled at 22050 Hz. Prior to analysis,

Table 1. Description of speakers with dysarthria.

Speaker	Age (years)	Gender	Speech intelligibility	Mode(s) of communication	Motor control
D1	27	M	12%	Gesture, vocalizations, sign language	Able to sign, unable to write Uses a wheelchair
D2	33	F	18%	Head pointer, vocalizations	Only head control Uses a wheelchair
D3	35	M	20%	Communication board with bliss symbols, vocalizations	Pointing gestures Uses a wheelchair
D4	32	M	22%	Picture symbols, vocalizations	Pointing gestures Uses a wheelchair
D5	44	F	16%	Alphanumeric communication board, vocalizations	Pointing gestures Uses a wheelchair
D6	36	M	22%	Alphanumeric communication board, some vocalizations	Pointing gestures Ambulatory
D7	40	M	18%	Alphanumeric and phrase communication board	Pointing gestures Uses a wheelchair
D8	44	M	24%	Alphanumeric board, vocalizations with familiar people	Pointing gestures Ambulatory

in order to obtain measures of relative intensity, all recordings were rescaled such that peak intensity was matched across all utterances and speakers. The Praat speech analysis software package (Boersma & Weenik, 2000) was used to calculate seven acoustic features: syllable duration, peak F0 (F0peak), average F0 (F0ave), slope of F0 (F0slope), peak intensity (INTpeak), average intensity (INTave), and slope of intensity (INTslope).

Duration

The beginning and end of each syllable within each phrase was marked by listening to the waveform and using the intensity envelope as a guide. Despite efforts to minimize respiratory demands by limiting phrase length to three syllables, many speakers with DYS were unable to produce the phrase in one breath group, resulting in pauses between syllables. In order to remove inspiration or termination fall-off cues that would otherwise confound the results, a software routine was used to systematically clip 0.01 s from the beginning and end of each syllable. If two adjacent syllables were produced in one breath group, the beginning of one syllable was interpreted as the end of the previous syllable and clipping was not performed on that segment. Three duration measures were calculated for each phrase: S1_dur, S2_dur, and S3_dur, the duration of Syllables 1, 2, and 3, respectively.

F0

The Praat system generated F0 values (Hz) for each syllable within each phrase. Manual correction of the automatically generated F0 values was required on 28 of the 480 utterances by speakers with DYS and on 4 of the 480 utterances by HCs because the pitch-tracking algorithm reported octave jumps that could not be verified auditorily. Manually adjusting the upper and lower F0 limits and frame duration parameters in Praat typically led to improved F0 tracking. These new F0 values were verified through visual and auditory inspection and confirmation using direct calculation of the pitch period from the waveform. Praat-derived F0 values that continued to be judged as errors (this included only 7 dysarthric productions in total) were replaced by manually derived values obtained from the waveform. In total, nine F0 measures were calculated for each phrase: (a) S1_F0peak, S2_F0peak, and S3_F0peak, the highest F0 value within each syllable; (b) S1_F0ave, S2_F0ave, and S3_F0ave, the average of F0 values within each syllable; and (c) S1_F0slope, S2_F0slope, and S3_F0slope, the rate of change in F0 over syllable duration.

Intensity

The Praat system generated a series of relative intensity values (dB) across the duration of each syllable within each phrase. Similar to F0, nine intensity measures were calculated for each phrase: (a) S1_INTpeak, S2_INTpeak, and S3_INTpeak, the highest intensity value within each syllable; (b) S1_INTave, S2_INTave, and S3_INTave, the average of intensity values within each syllable; and (c) S1_INTslope, S2_INTslope, and S3_INTslope, the rate of change in intensity over syllable duration.

Reliability of Acoustic Measures

Intrajudge reliability was assessed using a randomly selected sample of 10% of the dysarthric and HC productions. Syllable-start and syllable-end points were manually relabeled 10 weeks after the original analysis. Intrajudge reliability of syllable duration measures across these two points in time was $r = .989$ ($M = 0.008$ s, $SD = 0.011$ s). Based on the new duration labels, all intensity and F0 values for this sample were recalculated. The mean difference between the first and second measurement was 1.1 Hz ($SD = 1.8$ Hz) for F0peak, 1.7 Hz ($SD = 3.8$ Hz) for F0ave, –40.2 Hz/s ($SD = 54.8$ Hz) for F0slope, –0.91 dB ($SD = 1.4$ dB) for INTpeak, –1.2 dB ($SD = 3.1$ dB) for INTave, and –4.2 dB/s ($SD = 12.7$ dB/s) for INTslope.

Listener Perception Judgments

Logistic regression analyses were used to understand the relationship between listener perception and acoustic consistencies in dysarthric speech. Given the binary nature of the listener response variable (i.e., token judged as a question or a statement), logistic regression was used. This analysis was only performed on the DYS dataset, because the goal of this study was to better understand how speakers with DYS signal the prosodic contrast.

Listener perceptual judgements were taken from Patel (2002b). For each speaker with DYS, 6 independent normal-hearing listeners who were unfamiliar with dysarthric speech classified 60 utterances (30 question tokens and 30 statement tokens) as either a question or a statement. In the present analysis, an utterance was considered to be accurately judged only if more that 50% of the 6 listeners (i.e., 4 or more listeners) had correctly classified that token. Using this perceptual accuracy criterion, each utterance (60 utterances × 8 speakers, for a total of 480 utterances) was placed into one of two categories, judged as (a) a question or (b) a statement. This classification resulted in 231 tokens perceived as questions and 249 tokens perceived as statements.

Results

Acoustic Consistencies Within Question–Statement Productions

Acoustic analyses were conducted on all 480 productions by speakers with DYS and on all 480 productions by HCs regardless of whether unfamiliar listeners accurately identified the speaker's intention. Although some studies have analyzed only correctly identified tokens, all tokens were analyzed in this study because it better addressed the goal of building a speech recognition algorithm capable of processing dysarthric utterances. Such a system would not have access to a priori knowledge about correctly perceived tokens.

Separate repeated-measures analyses of variance were conducted for each of the seven dependant measures. In each analysis, the effect of one between-subject factor (group: speakers with DYS vs. HCs) and two within-subject factors were examined. Factor 1 represented the sentence type and had two levels (question, statement). Factor 2 represented syllables and had three levels (Syllable 1 [S1], Syllable 2 [S2], and Syllable 3 [S3]). The response variables were all continuous—fundamental frequency in hertz, intensity in decibels, and duration in seconds. The F statistic was used to test the null hypothesis, with $\alpha = .05$. Considering each acoustic parameter separately, up to 12 pairwise contrasts were conducted to examine differences within and across syllables, sentence types, and speaker groups. To account for multiple comparisons in these post hoc tests, a Bonferroni correction factor was applied and any $p < .004$ was considered to be statistically significant.

Syllable Duration

Statistically significant main effects in syllable duration were found for group, $F(1, 14) = 165.2$, $p < .0001$; sentence type, $F(1, 14) = 45.8$, $p < .0001$; and syllable, $F(2, 28) = 103.6$, $p < .0001$ (see Figure 1). All two-way and three-way interactions were also statistically significant ($p < .0001$). Although both groups elongated S3 for questions compared to statements, questions by speakers with DYS were more than four times longer than questions by HCs ($p < .0001$; see Table 2 for mean syllable duration values for each speaker with DYS and each HC speaker). For the DYS group, the mean difference in S3 duration between questions and statements was 0.36 s ($p < .0001$), while for HCs the mean difference was 0.08 s ($p = .0008$). Speaker groups also differed in the change in duration from S2 to S3 for questions ($p = .0003$). Speakers with DYS increased duration from S2 to S3 for questions by 0.36 s ($p < .0001$), while HCs increased duration by 0.14 s ($p = .0002$). Speakers with DYS also lengthened S2 for questions compared to statements by 0.06 s ($p = .002$), which HCs did not do.

F0

Three F0 measures were examined for each syllable within each phrase: F0peak, F0ave, and F0slope. In each analysis, gender was included as a second between-subjects factor to account for inherent differences in F0 among female and male speakers. Note that although F0peak and F0ave are related and thus changes in one parameter are likely to lead to similar changes in the

Figure 1. Mean syllable duration (s) for each syllable of question and statement tokens produced by healthy controls (HC) and speakers with dysarthria (DYS).

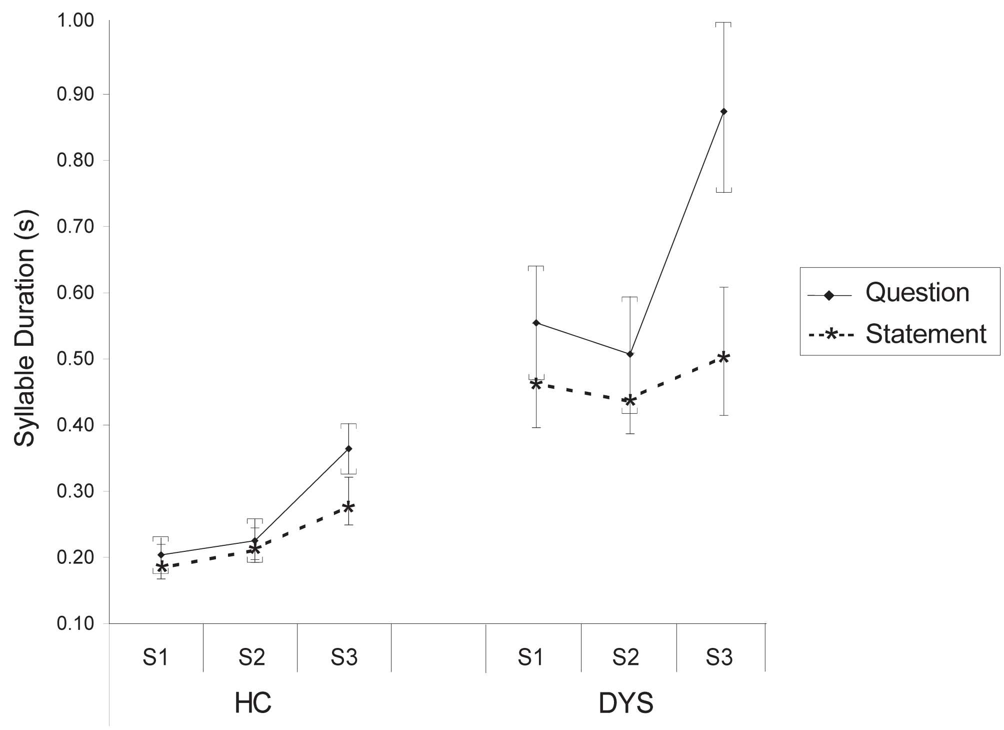

Table 2. Mean syllable duration (s) of Syllable 1 (S1), Syllable 2 (S2), and Syllable 3 (S3) for question and statement tokens, by speaker, for healthy controls (H1–H8) and speakers with dysarthria (D1–D8).

	Questions			Statements		
Speaker	**S1**	**S2**	**S3**	**S1**	**S2**	**S3**
H1	0.24	0.28	0.42	0.21	0.25	0.32
H2	0.19	0.20	0.35	0.19	0.20	0.27
H3	0.18	0.18	0.37	0.17	0.19	0.32
H4	0.18	0.20	0.37	0.17	0.20	0.29
H5	0.17	0.19	0.33	0.14	0.18	0.21
H6	0.20	0.23	0.33	0.18	0.22	0.25
H7	0.16	0.19	0.29	0.17	0.20	0.26
H8	0.22	0.25	0.37	0.22	0.23	0.28
Group	0.19 (0.03)	0.21 (0.03)	0.35 (0.04)	0.18 (0.03)	0.21 (0.02)	0.27 (0.04)
D1	0.50	0.51	0.91	0.48	0.44	0.47
D2	0.48	0.40	0.69	0.36	0.37	0.34
D3	0.41	0.42	0.85	0.44	0.40	0.64
D4	0.61	0.40	0.78	0.38	0.35	0.46
D5	0.47	0.54	0.80	0.45	0.48	0.54
D6	0.64	0.60	1.04	0.48	0.45	0.43
D7	0.60	0.62	1.03	0.51	0.53	0.50
D8	0.64	0.49	0.81	0.61	0.48	0.62
Group	0.54 (0.09)	0.50 (0.09)	0.86 (0.12)	0.46 (0.08)	0.44 (0.06)	0.50 (0.09)

Note. Standard deviations are in parentheses.

other, differences between F0peak and F0ave may be informative about the F0 contour.

Statistically significant main effects in F0peak were found for gender, $F(1, 12) = 6.85$, $p = .02$; sentence type, $F(1, 12) = 52.5$, $p < .0001$; and syllable, $F(2, 24) = 16.0$, $p < .0001$ (see Figure 2). The interaction between sentence type and syllable was significant, $F(2, 24) = 37.9$, $p < .0001$; however, all other interactions were not significant. Although speakers with DYS had higher F0peak values for question and statement tokens compared to HCs, this difference was not statistically significant, $F(1, 14) = 4.1$, $p = .06$. In terms of gender differences, female speakers in both groups had higher F0peak values for both question and statement tokens compared to male speakers; however, none of the interactions with gender were statistically significant (see Table 3 for mean F0peak values for each speaker with DYS and each HC speaker). In other words, male and female speakers in both groups used similar patterns of F0peak to mark the question–statement contrast. For statement tokens, F0peak was steady across S1, S2, and S3, while question tokens were produced with steadily increased F0peak from S1 to S2 and marked heightening of S3. There was no difference between groups for mean difference in S3_F0peak across sentence type or in the extent of raising F0peak from S2 to S3 for questions.

Statistically significant main effects in F0ave were found for group, $F(1, 12) = 7.1$, $p = 0.02$; gender, $F(1, 12) = 7.1$, $p = .02$; sentence type, $F(1, 12) = 71.8$, $p < .0001$; and syllable, $F(2, 24) = 8.3$, $p = .002$ (see Figure 3). Interactions between sentence type and gender, $F(1, 12) = 7.9$, $p = .02$; syllable and group, $F(2, 24) = 9.1$, $p = .001$; and sentence type and syllable, $F(2, 24) = 18.4$, $p < .0001$, were all statistically significant. All other interactions between group, gender, sentence type, and syllable were not significant. Female speakers in both speaker groups used a higher overall F0ave and larger difference in F0ave between question and statement tokens than did male speakers (see Table 4 for mean F0ave values for each speaker with DYS and each HC speaker). With regard to group differences, speakers with DYS used higher F0ave for question and statement tokens compared to HCs ($p = .002$). Although the increase in S3_F0ave for questions was statistically significant for speakers with DYS (mean difference = 106.0 Hz, $p = .004$), HCs only increased F0peak by 47.0 Hz ($p = .03$). The degree of increase in F0ave from S2 to S3 for questions was also higher among speakers with DYS (25%) compared to HCs (19%); however, this difference was not statistically significant ($p = .09$).

Statistically significant main effects for F0slope were found for gender, $F(1, 12) = 5.8$, $p = .03$; sentence type, $F(1, 12) = 50.0$, $p < .0001$; and syllable, $F(2, 24) = 3.9$, $p = .03$ (see Figure 4). Interactions between sentence type and group, $F(2, 24) = 5.7$, $p = .03$; syllable and group, $F(2, 24) = 4.7$, $p = .02$; sentence type and syllable, $F(2, 24) = 12.4$, $p = .0002$; and sentence type and syllable and group, $F(2, 24) = 3.8$, $p = .04$, were also significant. All other main effects and interactions were

Figure 2. Mean peak fundamental frequency (Hz) for each syllable of question and statement tokens produced by healthy controls (HC) and speakers with dysarthria (DYS).

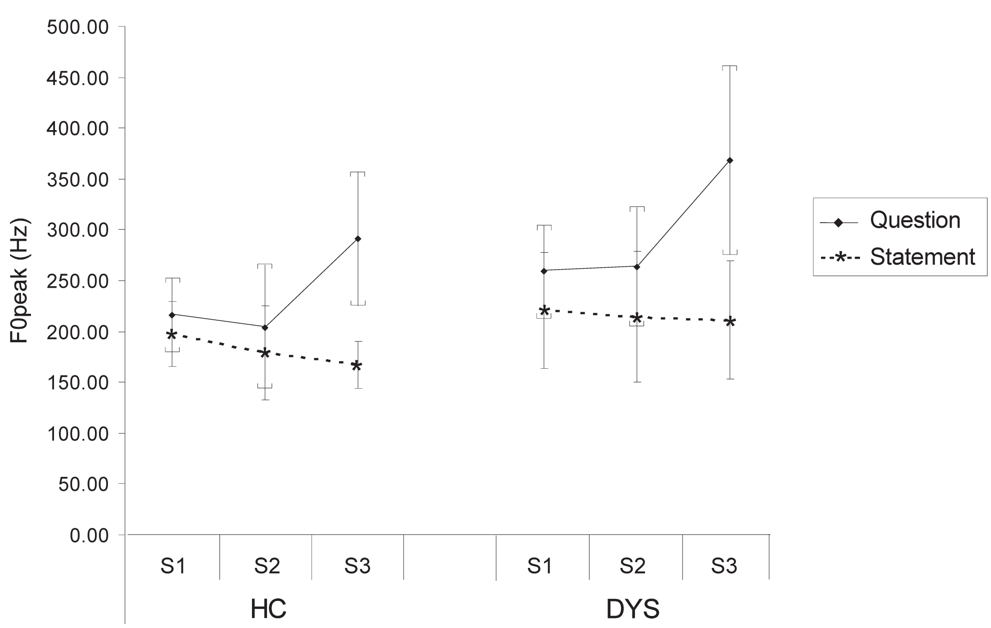

Table 3. Mean F0peak (Hz) of Syllable 1 (S1), Syllable 2 (S2), and Syllable 3 (S3) for question and statement tokens, by speaker, for healthy controls (H1–H8) and speakers with dysarthria (D1–D8).

		Questions			Statements		
Speaker	Gender	S1	S2	S3	S1	S2	S3
H1	M	232.6	165.7	293.5	214.5	159.3	168.1
H2	M	190.5	174.9	258.6	193.8	172.8	160.8
H3	M	219.6	205.2	299.6	216.5	192.3	171.5
H4	M	208.1	164.4	217.7	169.1	124.5	121.9
H5	M	199.4	175.9	250.5	172.2	213.4	195.6
H6	F	261.0	282.9	343.6	204.1	222.9	173.1
H7	M	155.3	147.4	241.2	154.1	108.0	191.7
H8	F	261.6	314.1	421.1	253.4	236.5	155.5
Group		216.0 (35.9)	203.8 (61.2)	290.7 (65.8)	197.2 (31.9)	178.7 (46.3)	167.3 (22.9)
D1	M	288.0	290.1	328.0	255.7	278.5	256.5
D2	F	292.5	322.1	393.1	243.6	230.2	213.7
D3	M	250.3	266.8	309.5	224.5	188.5	220.7
D4	M	247.3	212.7	462.7	134.9	119.6	125.6
D5	F	287.64	293.95	456.9	254.3	275.1	278.1
D6	M	307.2	327.8	468.4	309.4	292.3	275.1
D7	M	167.4	158.0	207.5	161.1	150.3	145.7
D8	M	234.9	235.3	322.0	183.9	179.4	173.4
Group		259.5 (45.2)	263.4 (58.2)	368.5 (92.9)	220.9 (57.4)	214.2 (64.5)	211.1 (58.3)

Note. Standard deviations are in parentheses.

not significant. Although female speakers in both groups used higher F0slope values than male speakers to mark question tokens, the pattern of using rising S3_F0slope for questions and using falling S3_F0slope to mark statement tokens was common across genders (see Table 5 for mean F0slope values for each speaker with DYS and each HC speaker). For HCs, S3_F0slope in questions and statements contrasted by 511.8 Hz/s (p = .0002), whereas productions by speakers with DYS only differed by 206.3 Hz/s (p = .001). This difference between groups, however, was not found to be statistically significant because both groups had very large variance in F0slope. The pattern of change in F0slope was different between speaker groups. Although speakers with DYS marked questions with a gradually rising S3_F0slope and statements with a sharply falling S3_F0slope, HCs used a sharply rising S3_F0slope for questions and gradually falling S3_F0slope for statements.

Intensity

Three intensity measures were examined for each syllable within each phrase: INTpeak, INTave, and INTslope. Table 6 shows mean values for each intensity measure by syllable and by group. A statistically significant main effect in INTpeak was found only for sentence type, $F(1, 14) = 13.8$, $p = .002$. Only the interactions between sentence type and group, $F(1, 14) = 5.41$, $p = .04$, and sentence type and syllable, $F(2, 28) = 43.5$, $p < .0001$, were significant. Although speaker groups did not differ significantly in overall INTpeak, they used different patterns of INTpeak to mark the question–statement contrast (see Table 6). Although speakers with DYS increased S3_INTpeak for questions by 6.8 dB, HCs only increased S3_INTpeak by 3.0 dB. Differences in S3_INTpeak across groups and within a group were not statistically significant.

A statistically significant main effect in INTave was also found only for sentence type, $F(1, 14) = 14.3$, $p < .002$. Only interactions between sentence type and group, $F(1, 14) = 5.28$, $p = .04$, and sentence type and syllable were significant, $F(2, 28) = 58.5$, $p < .0001$. Both groups had similar overall INTave values, yet they used different patterns of INTave to mark the question–statement contrast (see Table 6). Speakers with DYS increased S3_INTave for questions by 5.2 dB, whereas HCs only increased S3_INTave by 3.4 dB. Differences across groups and within a group, however, were not statistically significant.

A statistically significant main effect in INTslope was found only for syllable, $F(1, 14) = 56.0$, $p < .0001$. Interactions between sentence type and group, $F(1, 14) = 7.8$, $p = .01$; syllable and group, $F(2, 28) = 23.7$, $p < .0001$; and sentence type and syllable, $F(2, 28) = 5.1$, $p < .01$, were also significant. In both groups, INTslope was positive in S1 and negative in S2 and S3 for both questions and statements (see Table 6). Both sentence types

Figure 3. Mean of the average fundamental frequency (Hz) for each syllable of question and statement tokens produced by healthy controls (HC) and speakers with dysarthria (DYS).

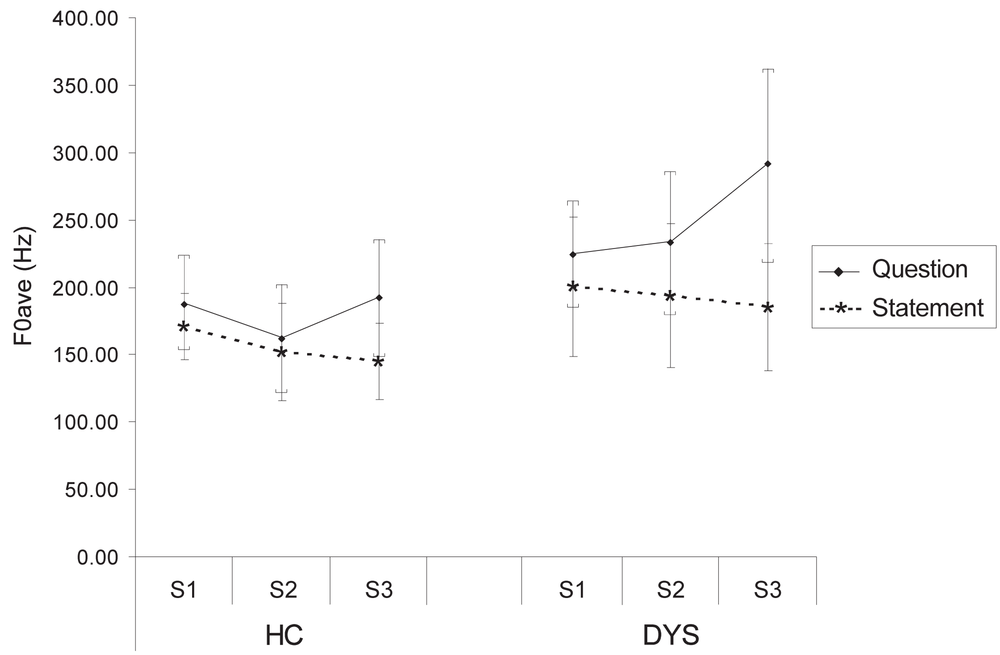

Table 4. Mean F0ave (Hz) of Syllable 1 (S1), Syllable 2 (S2), and Syllable 3 (S3) for question and statement tokens, by speaker, for healthy controls (H1–H8) and speakers with dysarthria (D1–D8).

		Questions			Statements		
Speaker	**Gender**	**S1**	**S2**	**S3**	**S1**	**S2**	**S3**
H1	M	183.6	129.5	174.1	174.0	136.4	128.2
H2	M	166.7	147.9	154.8	163.0	147.2	128.7
H3	M	196.2	169.9	219.8	187.2	177.7	148.1
H4	M	165.5	132.4	160.4	140.9	110.2	99.6
H5	M	171.9	143.4	180.1	174.6	169.3	188.4
H6	F	243.4	238.1	244.5	195.2	190.5	151.9
H7	M	139.1	125.2	141.7	132.1	95.7	175.8
H8	F	234.2	205.3	261.8	200.2	189.1	139.9
Group		187.6 (35.7)	161.5 (40.6)	192.1 (44.3)	170.9 (24.5)	152.0 (35.9)	145.1 (28.1)
D1	M	266.0	267.7	283.1	231.3	231.2	215.1
D2	F	253.6	290.4	327.2	230.0	215.6	199.3
D3	M	211.0	254.5	278.9	197.8	185.5	201.6
D4	M	210.9	167.5	250.8	121.3	109.3	114.4
D5	F	263.4	267.8	377.8	237.3	242.9	242.0
D6	M	243.4	258.3	388.3	271.2	262.3	228.9
D7	M	151.3	142.0	175.3	144.9	136.9	132.7
D8	M	197.4	215.3	249.0	169.2	168.7	148.4
Group		224.6 (39.5)	233.0 (53.0)	291.3 (70.9)	200.4 (51.5)	194.1 (53.5)	185.3 (47.2)

Note. Standard deviations are in parentheses.

were marked by a falling INTslope contour by speakers with DYS and a fall–rise contour by HC speakers.

Relationship Between Listener Perception and Acoustic Consistencies

Logistic regression was used to understand the relationship between listener perception judgments and acoustic consistencies in productions by speakers with DYS. Given the potential lack of statistical independence among the 480 utterances, multiple logistic regression analyses were performed for clustered data using generalized estimating equations (GEE; Zeger & Liang, 1986).

Before fitting the model, a stepwise logistic regression analysis was conducted with all 21 acoustic predictor variables (i.e., seven acoustic parameters for each syllable), assuming statistical independence among the observations and using a conservative entry/exit criterion of $p < .2$ while forcing gender as a predictor (given that F0 values vary with gender). In this way, the potential for multicollinearity in the final model was addressed by including only those acoustic parameters that were most likely to be predictive of listener perception. The acoustic parameters retained for the final GEE-based analysis were S1_F0ave, S1_F0slope, S2_dur, S3_dur, S3_F0peak, S3_F0slope, S3_INTave, and S3_INTslope. The dependant measure was the listeners' perceptual judgement (question vs. statement). Small values in the working correlation matrix ($r < .05$) of the final model allowed for assuming an independence working correlation in the estimation of the regression parameters and standard errors. All two-way interactions between gender and each acoustic parameter were tested in the final model and none were found to be significant at the .05 level.

The results of the final logistic regression are shown in Table 7. Listener perception of a question was statistically significantly more likely when utterances by speakers with DYS were of the following acoustic characteristics: lower S1_F0ave, longer S2_dur, and longer S3_dur, higher S3_F0peak, higher S3_F0slope, higher S3_INTave, and higher S3_INTslope. Meaningful unit differences used for the odds ratios were calculated based on observed means and standard deviations as well as clinical judgement. A nonsignificant main effect was found for S1_F0slope (p = .12). A main effect of gender was also not statistically significant (p = .77) and did not alter the odds ratio estimates of the acoustic parameters compared to a model in which its main effect was excluded.

To assess the predictive value of the final logistic regression, predicted probabilities were computed for perceiving a question for each observation. These probabilities were then classified as predicted statements and predicted questions. The C statistic (Hanley & McNeil, 1982), a measure of the predictive ability of the model, was .98, indicating that the final logistic model

Figure 4. Mean slope of fundamental frequency (Hz/s) for each syllable of question and statement tokens produced by healthy controls (HC) and speakers with dysarthria (DYS).

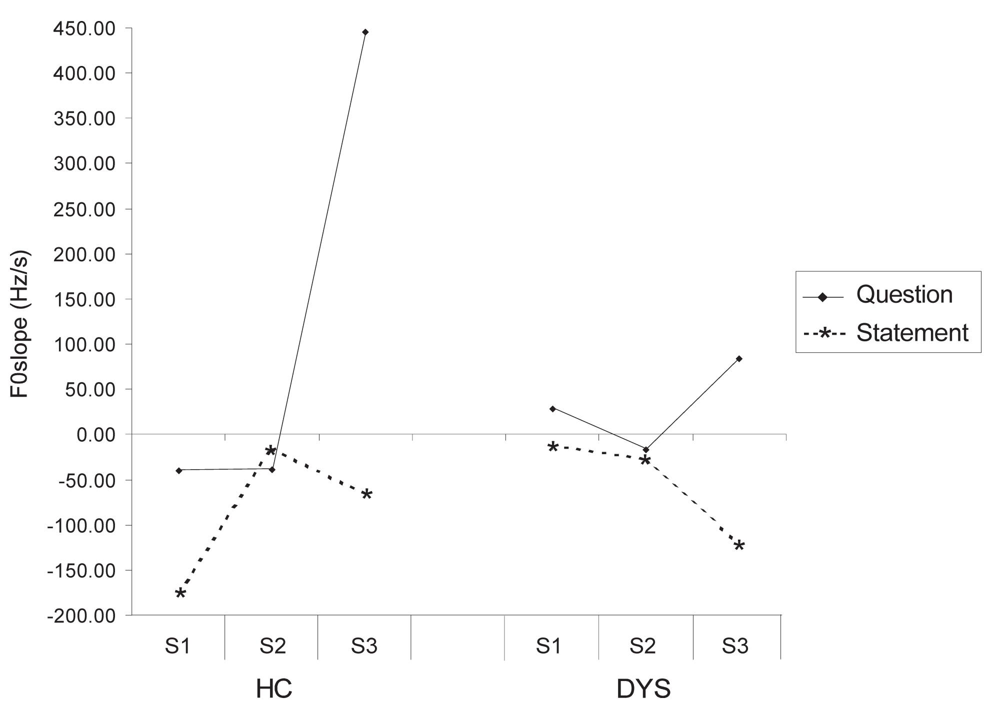

Table 5. Mean F0slope (Hz/s) of Syllable 1 (S1), Syllable 2 (S2), and Syllable 3 (S3) for question and statement tokens, by speaker, for healthy controls (H1–H8) and speakers with dysarthria (D1–D8).

		Questions			Statements		
Speaker	**Gender**	**S1**	**S2**	**S3**	**S1**	**S2**	**S3**
H1	M	292.7	−262.2	460.7	−97.9	−365.8	−80.3
H2	M	224.8	132.2	491.5	−79.9	143.5	−147.0
H3	M	−0.3	−570.5	498.5	34.8	−216.2	−156.2
H4	M	−283.3	8.9	314.2	−381.1	−91.0	−139.5
H5	M	−404.7	−237.9	371.9	−330.4	−167.3	−40.3
H6	F	69.7	73.3	564.1	−249.2	313.2	−100.6
H7	M	−339.1	592.2	45.2	−163.7	−172.2	102.1
H8	F	122.0	−43.8	820.5	−129.0	432.8	33.3
Group		−39.8 (267.9)	−38.5 (341.1)	445.8 (221.3)	−174.6 (138.0)	−15.4 (280.5)	−66.0 (92.8)
D1	M	23.8	−84.5	9.4	−30.7	−82.1	−336.4
D2	F	148.0	48.4	133.1	−27.4	−32.1	−74.8
D3	M	50.7	23.1	−49.2	−28.6	−20.7	−42.4
D4	M	4.4	4.0	250.1	−37.0	6.9	3.4
D5	F	−16.1	−56.0	145.8	−66.4	−70.7	−177.7
D6	M	−88.9	−87.5	100.4	21.2	−8.8	−200.4
D7	M	−12.4	−15.0	21.0	12.6	−14.6	−28.0
D8	M	113.3	27.3	62.4	58.0	3.2	−121.3
Group		27.9 (75.6)	−17.5 (52.6)	84.1 (94.0)	−12.3 (39.7)	−27.4 (32.8)	−122.2 (112.4)

Note. Standard deviations are in parentheses.

was fairly robust. The confusion matrix in Table 8 indicates that 95% of the 480 observations were correctly classified, with a sensitivity of 96%, a specificity of 95%, a positive predictive value (Fleiss, 1981) of 94%, and a negative predictive value (Fleiss, 1981) of 96%.

Discussion

In previous work it was found that speakers with DYS were able to reliably and consistently signal the question–statement contrast to unfamiliar listeners (Patel, 2002b). The present study further examined that dataset to identify acoustic consistencies in productions by speakers with DYS. In addition, the Patel (2002b) study was replicated with 8 HC speakers to investigate whether speakers with DYS were using alternative strategies to mark the contrast in light of their speech motor control deficiencies. Finally, listener perception findings of Patel (2002b) were connected to the acoustic consequences calculated in this study.

Although a rising F0 contour is thought to be the primary acoustic cue for marking the question–statement

Table 6. Mean and standard deviation of INTpeak (dB), INTave (dB), and INTslope (dB/s) of Syllable 1 (S1), Syllable 2 (S2), and Syllable 3 (S3) for question and statement tokens, by speaker, for healthy controls (H1–H8) and speakers with dysarthria (D1–D8).

	INTpeak (dB)		INTave (dB)		INTslope (dB/s)	
	Questions	**Statements**	**Questions**	**Statements**	**Questions**	**Statements**
HC						
S1	68.6 (4.3)	68.6 (4.3)	61.7 (3.8)	61.8 (3.4)	35.4 (12.1)	49.8 (11.1)
S2	68.1 (4.9)	68.1 (4.1)	61.1 (3.8)	61.5 (3.2)	−34.2 (12.6)	−33.8 (17.2)
S3	70.2 (4.4)	67.2 (4.5)	63.8 (4.0)	60.4 (4.5)	−6.4 (11.1)	−14.5 (17.6)
DYS						
S1	69.4 (3.9)	66.6 (2.7)	62.3 (3.4)	60.6 (1.6)	12.8 (18.1)	4.8 (16.9)
S2	69.8 (3.4)	66.4 (4.3)	63.8 (3.7)	60.7 (2.6)	−2.5 (5.0)	−6.4 (4.9)
S3	73.1 (3.4)	66.3 (5.4)	66.1 (4.0)	58.9 (4.6)	−10.3 (9.8)	−20.5 (8.5)

Table 7. Results of the logistic regression of perceptual judgment (question vs. statement) on 10 acoustic measures.

Acoustic parameter (meaningful difference in units)	Relative odds of perceiving an utterance as a question per one unit difference in parameter	*p*
S1_F0ave (50 Hz)	0.54	.0001
S1_F0slope (150 Hz/s)	1.91	.12
S2_Duration (0.3 s)	2.15	.02
S3_Duration (0.3 s)	6.02	<.0001
S3_F0peak (50 Hz)	3.89	<.0001
S3_F0slope (150 Hz/s)	5.29	.0002
S3_INTave (5 dB)	2.94	.005
S3_INTslope (15 dB/s)	0.63	<.0001

contrast (cf. Cruttenden, 1986; Eady & Cooper, 1986), there has been unequivocal data to support whether or not syllable duration is also varied (cf. Denes, 1959; Lieberman, 1967). Results of the acoustic analyses indicate that although HC speakers primarily used F0 and duration cues, speakers with DYS used all three cues of linguistic stress—F0, duration, and intensity—to mark the difference between questions and statements.

For both groups, questions were longer than statements and they were marked by a longer S3. Within questions, syllable lengthening occurred from S2 to S3. It is clear from this acoustic analysis that speakers with DYS were using duration to mark the question–statement contrast. In fact, the contrastive use of duration was greater for speakers with DYS compared to HCs. Perhaps speakers with DYS were remapping control of F0 onto duration, a cue over which they had greater control.

Question tokens produced by speakers with DYS and HCs were also higher in F0peak and F0ave than statements with focused heightening for S3. For questions, a change in F0peak and F0ave occurred from S2 to S3. The change in F0peak and F0ave within question productions and across sentence types in S3 was greater for speakers with DYS compared to HCs. However, when changes in F0peak and F0ave are considered in relative terms (i.e., percentage increase from S2 to S3 in question or between S3 in question vs. statements), the group differences are fairly small, suggesting that both groups were using similar strategies to signal the contrast.

Table 8. Classification of predicted versus observed statements using the logistic regression model.

	Predicted	
	Statement	Question
Observed perception		
Statement	236	13
Question	9	222

Although speakers with DYS due to cerebral palsy are typically characterized as having a reduced range in F0 (cf. Hardy, 1983; Le Dorze et al., 1994; Yorkston et al., 1984), speakers with DYS in this study did not differ considerably from HCs in how they used their F0 range to mark the question–statement contrast. It is likely that these speakers with DYS do have a reduced range in F0 on other tasks such as sustained vowel production; however, such a range may not be necessary for marking communicative contrasts as in question versus statement. In addition, speakers with DYS may be still be able to exert reliable control over a limited range.

F0slope was relatively flat for S1 and S2 but differentiated itself at S3, where it was rising for questions and falling for statements. The extent of raising S3_F0slope for questions was greater for HC speakers compared to speakers with DYS. Interestingly, however, speakers with DYS marked statements with a steeper falling S3_F0slope than did HCs. Perhaps speakers with DYS used this pattern of contrastive S3_F0slope as a compensatory strategy to ensure maximal separation of questions and statements. Physiological constraints imposed by the speech motor control system of people with DYS may make it easier to produce a falling F0 for statements than to make a sharply rising contour for questions. Although raising F0 requires precise adjustments to vocal fold tension using active muscle contraction, it is possible to drop F0 by relaxing vocal fold tensors. The F0peak, F0ave, and F0slope results taken together suggest that F0 range may not be as much of a problem for speakers with DYS as temporal aspects of F0.

Although group differences in INTpeak, INTave, and INTslope were not statistically significant, interactions of Group × Sentence Type and Group × Syllable were significant, suggesting that speakers in the two groups may have been using intensity cues differentially. Although speakers with DYS used a falling INTslope for questions and statements, HCs used a fall–rise INT contour. In addition, HCs did not mark S3 in questions with as high values of INTpeak or INTave compared to speakers with DYS. Changes in INTpeak and INTave were most noticeable from S2 to S3 in question tokens produced by speakers with DYS. Given that these INT differences were not statistically significant, it is possible that they were merely a by-product of the increased F0 used by speakers with DYS. It may also be the case that speakers with DYS were attempting to use INT as an additional cue given that their F0 cues alone may not have been sufficiently contrastive.

This notion of sufficient contrast refers to the effect of acoustic differences on listener perception. The presence of acoustic consistencies does not mean that these differences are salient to listeners. Conversely, listeners may attend to acoustic cues or cue combinations even if the acoustic differences are not statistically significant. Results of the regression analysis of listener perception and acoustic consistencies verify previous findings that questions and statements differ predominantly in the final syllable, S3 (cf. Cruttenden, 1986; Eady & Cooper, 1986). Some S1 and S2 features were also significant, suggesting that various features throughout the phrase influence listener perceptions of dysarthric speech.

Although F0 cues are thought to be the most prominent cues for predicting the contrast in nonimpaired speech, duration, F0, and intensity cues all contributed to accurate listener perception of questions versus statements by speakers with DYS. Listeners were more likely to perceive a token as a question when S1_F0ave was low and S3_F0peak and S3_F0slope were high. Perhaps the use of a higher overall F0ave and F0peak along with a moderate rise in F0slope at the end of questions by speakers with DYS was adequate for accurate listener perception (see Hadding-Koch, 1961, for similar findings in Swedish speakers). In addition to these F0 cues, longer S2_dur, longer S3_dur, louder S3_INTave, and falling S3_INTslope also contributed to the perception of a question.

Teasing apart what listeners were doing from what they thought they were doing was challenging. At the end of the perceptual experiment in Patel (2002b), an open-ended questionnaire was conducted to gather data on what listeners thought they were listening to when making their classifications. Although only 19 of the 48 listeners reported using duration as a cue, the results of the logistic regression indicate that duration was a highly informative cue for S2 and S3 of questions. Additionally, 18 of the 48 listeners had indicated that questions seemed to be louder. Although intensity differences were not statistically significant in the acoustic analyses, results of the logistic regression suggest that listeners were more likely to perceive a token as a question when S3_INTave was high but S3_INTslope was decreasing. Although intensity cues alone may not have been contrastive enough, the combined effect of changes in intensity, duration, and F0 may have led to accurate perception of the question–statement contrast produced by speakers with DYS.

Future Directions

Although the findings of this study are promising in terms of prosodic control in severe dysarthria due to cerebral palsy, several issues require further exploration. First, it will be important to replicate these findings with longer phrases given the known respiratory problems in this group. The results of this study differed from previous work in which longer utterances have been used (e.g., Le Dorze et al., 1994). A second extension of this work would be to examine local contrastive stress to better understand the flexibility and physiological limits of the speech production mechanism in people with DYS. Future studies of prosodic control in DYS may consider additional acoustic cues such as the turning point in F0 and the shape of the full F0 contour (Hadding-Koch & Studdert-Kennedy, 1964; Majewski & Blasdell, 1969). To increase the generalizability of the present findings, it will be important to collect data from speakers with DYS with varying etiologies.

This line of research gives rise to three application areas that would require further exploration. First, acoustic consistencies in prosodic control provide the impetus for building a communication aid that can at least partially be controlled through vocal contrasts. Second, human listeners may also be able to use these prosodic consistencies. Communication partners may thus benefit from training aimed at attending to prosodic aspects of dysarthric speech. Third, the results of the acoustic and perceptual analyses suggest that speakers with DYS may benefit from intervention aimed at further refining control of duration and F0 for communicative function.

Acknowledgment

This research was conducted in the Speech Communication Group at the Research Laboratory of Electronics at the Massachusetts Institute of Technology and in the Department of Speech-Language Pathology at Teachers College, Columbia University. I am grateful to Kenneth Stevens and Stephanie Shattuck-Hufnagel for their invaluable input. I would especially like to thank Deb Roy and Howard Cabral for their contributions to the acoustic and statistical analyses. In addition, this study would not have been possible without the involvement of the participants. This material is based on work supported by National Science Foundation Grant 0083032.

References

Allen, G. D., & Arndorfer, P. M. (2000). Production of sentence-final intonation contours by hearing-impaired children. *Journal of Speech, Language, and Hearing Research, 43*, 441–455.

Ansel, B. M., & Kent, R. D. (1992). Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *Journal of Speech and Hearing Research, 35*, 296–308.

Baum, S., & Pell, M. (1997). Production of affective and linguistic prosody by brain damaged patients. *Aphasiology, 11*, 177–198.

Boersma, P., & Weenik, D. (2000). *Praat, a system for doing phonetics by computer, Version 3.4*. Amsterdam:

University of Amsterdam, Institute of Phonetic Sciences. Available online at http://www.fon.hum.uva.nl/praat

Brewster, K. (1989). The assessment of prosody. In K. Grundy (Ed.), *Linguistics in clinical practice* (pp. 186–202). London: Taylor & Francis.

Bunton, K., Weismer, G., & Kent, R. (2000, February). *The effects of flattened fundamental frequency contours on sentence intelligibility in dysarthric speakers.* Poster presentation at the Tenth Biennial Conference on Motor Speech: Motor Speech Disorders and Speech Motor Control, San Antonio, TX.

Canter, G. J. (1963). Speech characteristics of patients with Parkinson's disease: Intensity, pitch and duration. *Journal of Speech and Hearing Disorders, 28,* 221–228.

Cruttenden, D. (1986). *Intonation*. Cambridge, U.K.: Cambridge University Press.

Danly, M., & Shapiro, B. (1982). Speech prosody in Broca's aphasia. *Brain and Language, 16,* 171–190.

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

Denes, P. (1959). A preliminary investigation of certain aspects of intonation. *Language and Speech, 2,* 106–122.

Denes, P., & Milton-Williams, J. (1962). Further studies in intonation. *Language and Speech, 5,* 1–14.

Eady, S. J., & Cooper, W. E. (1986). Speech intonation and focus location in matched statements and questions. *Journal of the Acoustical Society of America, 80,* 402–415.

Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). New York: Wiley.

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

Hadding-Koch, K. (1961). *Acoustico-phonetic studies in the intonation of southern Swedish.* Research report, Institute of Phonetics, University of Lund, Sweden.

Hadding-Koch, K., & Studdert-Kennedy, M. (1964). An experimental study of some intonation contours. *Phonetica, 11*, 175–185.

Hanley, J. A., & McNeil, B. J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology, 143*(1), 29–36.

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

Hood, R., & Dixon, R. (1969). Physical characteristics of speech rhythm of deaf and normal-hearing speakers. *Journal of Communication Disorders, 2,* 20–28.

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

Irwin, J. (1955). Phonetic equipment of spastic and athetoid children. *Journal of Speech and Hearing Disorders, 20,* 54–67.

Laures, J. S., & Weismer, G. (1999). The effects of a flattened fundamental frequency contour on intelligibility at the sentence level. *Journal of Speech, Language, and Hearing Research, 42,* 1148–1156.

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

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

Lieberman, P. (1967). *Intonation, perception and language*. Cambridge, MA: MIT Press.

Majewski, W., & Blasdell, R. (1969). Influence of fundamental frequency cues on the perception of some synthetic intonation contours. *Journal of the Acoustical Society of America, 45*, 450–457.

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.

O'Shaughnessy, D. (1979). Linguistic features in fundamental frequency patterns. *Journal of Phonetics, 7*, 119–145.

Patel, R. (2002a). Phonatory control in adults with cerebral palsy and severe dysarthria. *Alternative and 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, and Hearing Research, 45,* 858–870.

Platt, L., Andrews, G., & Howie, F. M. (1980). Dysarthria of adult cerebral palsy: II. Phonemic analysis of articulation errors. *Journal of Speech and Hearing Research, 23,* 41–55.

Platt, L., Andrews, G., Young, M., & Quinn, P. (1980). Dysarthria of adult cerebral palsy: I. Intelligibility and articulatory impairment. *Journal of Speech and Hearing Research, 23,* 28–40.

Robin, D. A., Klouda, G. V., & Hug, L. N. (1991). Neurogenic disorders of prosody. In M. P. Cannito & D. Vogel (Eds.), *Treating disordered speech motor control: For clinicians by clinicians* (pp. 241–271). Austin, TX: Pro-Ed.

Rosenbek, J., & La Pointe, L. (1978). The dysarthrias: Description, diagnosis, and treatment. In D. Johns (Ed.), *Clinical management of neurogenic communicative disorders* (pp. 251–310). Boston: Little-Brown.

Rosenhouse, J. (1986). Intonation problems of hearing impaired Hebrew-speaking children. *Language and Speech, 29,* 69–92.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research, 25*(2), 193–247.

Vance, J. E. (1994). Prosodic deviation in dysarthria: A case study. *European Journal of Disorders of Communication, 29*(1), 61–76.

Välimaa-Blum, R. (2001, March 17). Intonation and interrogation in English: It's all a matter of defaults. In *L'Interrogation: Actes du Colloque du 17 Mars 2001* (Proceedings of the Colloquium on Interrogation, March 17, 2001), Association des Linguistes Anglicistes de l'Enseignement Supérieur (ALAES) and Sémantique Syntaxe et Linguistique Anglaises (SESYLIA). Available online at http://www.univ-pau.fr/ANGLAIS/alaes/blum.pdf

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

Wingfield, A., Lombardi, L., & Sokol, S. (1984). Prosodic features and the intelligibility of accelerated speech: Syntactic versus periodic segmentation. *Journal of Speech and Hearing Research, 27,* 128–134.

Wit, J., Maassen, B., Gabreels, F. J. M., & Thoonen, G. (1993). Maximum performance tests in children with developmental spastic dysarthria. *Journal of Speech and Hearing Research, 36,* 452–459.

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, acoustics, perception, management* (pp. 131–162). Austin, TX: Pro-Ed.

Zeger, S., & Liang, L. (1986). Longitudinal data analysis using generalized linear models. *Biometrika, 73*, 13–22.

Received March 27, 2003

Accepted April 30, 2003

DOI: 10.1044/1092-4388(2003/xxx)

Contact author: Rupal Patel, PhD, 360 Huntington Avenue, #102 Forysth Building, Boston, MA 02115. E-mail: r.patel@neu.edu

Appendix A. Target phrases.

Phrase	Abbreviation	Statement context	Question context
She was here (?)	SWH	Bob asks if you have seen Jane. You say…	John tells you that the Queen of England visited your house. You ask…
It was hot (?)	IWH	You just got back from holidays in Florida. Jane asks if the weather was nice. You say…	Jane says her vacation to Alaska was too hot. You ask…
She loves dogs (?)	SLD	Molly asks if Sue likes dogs. You say…	Jane doesn't like animals. Someone tells you she loves dogs. You ask…
Play it loud (?)	PIL	You can't hear the message on your answering machine. You tell Jane to …	It's midnight and someone asks you to turn up the volume on the stereo. You ask…
Use some soap (?)	USS	You ask Jenny to get out a stain from your shirt. She asks how. You say…	You run out of shampoo while washing your hair. Your attendant tells you to just use soap. You ask…
She said no (?)	SSN	Bill proposed to Sue but she refused. Someone asks what she said. You say…	John offered Beth a million dollars. She declined. You ask…
It's Thursday (?)	ITD	Someone asks what day it is. You say…	Someone tells you it's Thursday on the weekend. You ask…
He lives there (?)	HLT	Doug asks you where Chris lives. You point to the blue house and say…	Chris is a millionaire. Paul says he lives in a small run down apartment downtown. You ask…
Pass some salt (?)	PSS	Your food is a little bland. You say…	While eating desert Jim asks you for the salt. You ask…
Give it back (?)	GIB	Someone grabs your pen. You say…	Someone gives you a gift and then asks for it back. You ask…