Clinical Linguistics & Phonetics, Month 2007; 000(000): 1–16

Influence of familiarity on identifying prosodic vocalizations produced by children with severe dysarthria

RUPAL PATEL, & BETHANY SCHROEDER

Northeastern University, Boston, MA, USA

(Received 22 September 2006; accepted 29 June 2007)

Abstract

Familiarity is thought to aid listeners in decoding disordered speech; however, as the speech signal degrades, the ''familiarity advantage'' becomes less beneficial. Despite highly unintelligible speech sound production, many children with dysarthria vocalize when interacting with familiar caregivers. Perhaps listeners can understand these vocalizations by cuing into prosodic consistencies in their child's productions. This paper examined whether familiarity influenced the identification of sustained vowels that varied in pitch, duration, and pitch–duration combinations, produced by 3 children with severe dysarthria due to cerebral palsy. Thirty-six listeners participated in the study. For each speaker, there were 2 familiar listeners (FAM), 5 experienced listeners (EXP), and 5 unfamiliar/inexperienced listeners (INX). Results indicated that familiarity did not impact identification of prosodic contrasts. In fact, all 3 listener groups were highly accurate in identifying duration, somewhat less successful at identifying pitch, and least accurate in identifying combinations of pitch and duration. Influences of speaker–listener variables on familiarity are discussed.

Keywords: *Familiarity, children, prosody, dysarthria, intelligibility*

Introduction

Dysarthric speech is characterized by imprecise consonants, distorted vowels, and prosodic disturbances (Darley, Aronson, & Brown, 1975). However, even in the case of children with severe dysarthria, it is not uncommon for caregivers to be able to ''translate'' their child's vocalizations. Anecdotal evidence suggests that caregivers and siblings have greater facility in understanding a family member with dysarthria than naïve listeners. Several studies have also documented that listener familiarity enhances intelligibility judgements of dysarthric speech (Tjaden & Liss, 1995a; DePaul & Kent, 2000; Liss, Spitzer, Caviness, & Adler, 2002; Hustad & Cahill, 2003). Familiarity has been defined as having previous contact or exposure with an acoustic signal (Spitzer, Liss, Caviness, & Adler, 2000). Familiarity, however, is a multifaceted concept that is affected by many variables (Liss et al., 2002) including personal experience with the speaker, previous exposure to the disorder, and training or adaptation to the speech signal.

Correspondence: 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

ISSN 0269-9206 print/ISSN 1464-5076 online
DOI: 10.1080/02699200701559476

As the speech signal becomes increasingly degraded, however, this ''familiarity advantage'' may be less evident (Dongilli, 1994; DePaul & Kent, 2000). Perhaps this degradation is partly due to the fact that listener strategies for decoding dysarthric speech may be focused primarily on speech sound segments, which may be less informative than other acoustic cues. Meaning is conveyed not only through segmental aspects of speech, but also through the prosodic features, namely pitch, duration, and loudness (cf. Netsell, 1973; Lehiste, 1976; Bolinger, 1989; Shattuck-Hufnagel & Turk, 1996). Recent findings suggest that, despite impoverished speech sound intelligibility, some individuals with severe dysarthria can convey prosodic contrasts (Patel, 2002, 2003, 2004; Ciocca, Whitehill, & Joan, 2004; Puyuelo & Rondal, 2005; Patel & Salata, 2006). The question then arises, if familiar listeners are asked to identify dysarthric vocalizations that vary along prosodic dimensions alone, will they be more proficient than experienced listeners and/or inexperienced/unfamiliar listeners?

Impact of familiarity on segmental intelligibility

The role of familiarity in decoding words and sentences produced by individuals with dysarthria has received considerable attention. In these studies, familiarity has been broadly defined to include personal experience with the speaker, academic or clinical training in the disorder area, and exposure to the specific speech stimuli. Although Flipsen (1995) found that productions of children with speech delay were more easily identified by their mothers than by their fathers or unfamiliar listeners, this advantage may be less robust in dysarthria. Kwiatkowski and Shriberg (1992) studied whether familiar communication partners were, in fact, better able to discern among dysarthric productions or if this intelligibility bias was merely their perception. They found that caregiver estimates of speech intelligibility were significantly higher than actual word and sentence transcription scores. Thus, the authors concluded that caregivers may not have a particular advantage in decoding their child's productions when utterances are decontextualized from everyday routines and preferences.

Much research has also focused on the effect of previous experience with the disorder on intelligibility (cf. McGarr, 1983; Mencke, Ochsner, & Testut, 1983; Boothroyd, 1985; Kreiman, Gerratt, & Precoda, 1990; Ellis & Fucci, 1991, 1992, 2000; Most, Weisel, & Lev-Matezky, 1996; Klimacka, Patterson, & Patterson, 2001; Lewis, Watterson, & Houghton, 2003). The impact of familiarity, however, has varied within and across disorder populations. Ellis and Fucci (1991, 1992, 2000) found no difference between experienced and inexperienced listeners' ability to rate poorly articulated speech. In contrast, McGarr (1983) found that experienced listeners achieved significantly higher scores on single-word and sentence intelligibility tests of deaf speech compared to inexperienced listeners.

Familiarization procedures, which provide a listener with previous exposure to a particular speaker or voice, have proven effective in increasing intelligibility of dysarthric speech (Beukelman & Yorkston, 1980; Tjaden & Liss, 1995b; King & Gallegos-Santillan, 1999; DePaul & Kent, 2000; Spitzer et al., 2000; Liss et al., 2002; Hustad & Cahill, 2003; D'Innocenzo, Tjaden, & Greenman, 2006). Liss et al. (2002) examined the effect of familiarization training on intelligibility of adult speakers with ataxic or hypokinetic dysarthria. Sentence transcription accuracy of a control group, who had no previous experience with dysarthric speech, was compared to two groups who each received training with either hypokinetic or ataxic dysarthria. During the familiarization phase, listeners heard dysarthric utterances while following a written transcript. There was a significant difference in intelligibility scores of listeners who received familiarization training vs the

control group. The authors suggest that there is a perceptual benefit of being familiarized with dysarthric speech before interacting or listening to a speaker with dysarthria. Given individual differences among speakers with dysarthria, familiarizing listeners to a particular target speaker's idiosyncratic patterns may be more effective than familiarization to dysarthric speech in general. While this notion is supported by evidence for a ''maternal advantage'' when decoding speaker-specific phonological errors (Powell, Elbert, & Forrest, 1990; Flipsen, 1995), it requires further inquiry in terms of decoding prosodic contrasts in dysarthric speech.

Even when familiarity was shown to aid listener judgements, the benefits of familiarization appeared to subside as the signal became exceedingly impaired (Mencke et al., 1983; DePaul & Kent, 2000). DePaul and Kent (2000) studied the effect of familiarity on the intelligibility of an adult with Amyotrophic Lateral Sclerosis (ALS), a progressive neuromotor disorder. They compared intelligibility ratings of a familiar listener (spouse), a semi-trained listener group consisting of students in the communication disorders field, and an untrained listener group. Although the familiar listener initially outperformed the other groups, her accuracy fell below that of the other listeners as the disease progressed and the speech impairment increased in severity. Although Lindblom's (1990) mutuality model suggests that listeners will make use of extrinsic cues even when the speech signal is degraded, Dongilli (1994) found that if severity is too high, it may not be possible to make use of any aspects of the signal.

Prosodic control in dysarthria

While familiarity may play an important role in deciphering words and sentences, little is known about how it influences the perception of vocalizations that vary in prosody. Several studies of adults with dysarthria suggest that prosodic control may be preserved despite severely degraded speech sound intelligibility (Le Dorze, Ouellet, & Ryalls, 1994; Vance, 1994; Patel, 1999, 2002, 2003; Ciocca et al., 2004). Patel (2002) studied pitch and duration control for sustained vowel productions in 8 adults with severe dysarthria due to cerebral palsy. While the speakers varied in their ability to produce distinct levels of pitch (high, medium, and low), all 8 speakers were able to produce distinct levels of duration (short, medium, and long). If children with severe dysarthria can also convey these prosodic contrasts, then it may be possible that familiar listeners will be able to attune to these cues. This study aimed to determine whether familiarity influences the identification of prosodically varying vocalizations produced by 3 children with severe dysarthria due to cerebral palsy. In particular, the study aimed to isolate which aspect of familiarity, experience with a given speaker vs experience with the speech disorder, was most advantageous in identifying prosodic contrasts and whether the pattern was similar across all three speakers.

Method

Speaker database

Data collected from a previous study yielded a database of children with severe dysarthria (DYS) secondary to cerebral palsy producing sustained vowels at 3 pitch levels, 3 duration levels, and 9 combinations of pitch and duration (Patel & Salata, 2006). Recordings from a sub-set of 3 female speakers ranging from 6–13 years of age (mean = 9 years, 7 months)

were used in the present study (see Table I for demographic information). All 3 children had hearing and visual function within normal limits and were primarily non-verbal communicators who used multiple modes of communication. For all children, English was the primary language of communication. Parental report was used to gather a detailed case history and to determine the primary medical diagnosis of cerebral palsy. Although no standardized tests were administered, informal observation, parental report, and information gathered from the referring speech-language pathologist indicated that all 3 children had adequate cognitive and linguistic abilities for completing the task. The type of dysarthria was determined by the referring speech-language pathologist and later confirmed by an additional certified speech-language pathologist.

Speech recordings were collected in a sound-treated audiometric booth using a unidirectional head-mounted cardioid dynamic microphone (Shure, SM10A) and saved directly to a computer at a sampling rate of 22 050 Hz. Each child was seated in front of a computer screen that displayed 3 interactive computer games used to elicit vocalizations that varied in pitch, duration, and a combination of pitch and duration. The software simultaneously collected the child's vocalizations and his/her caregiver's perception of those vocalizations. The child's task was to help the computer[1] move various characters on the screen by producing distinct levels of pitch, duration, or combination of pitch and duration. For the pitch game, the child was asked to sustain the vowel /a/ at 3 distinct levels of pitch (i.e. high, medium, and low). For the duration game, the child was asked to produce the same vowel at 3 distinct levels of duration (i.e. short, medium, and long). The order of the pitch and duration games was counterbalanced across speakers. The last game required the child to sustain production of the vowel /a/ at 9 distinct combinations of pitch and duration (high and long, low and short, etc.). The caregiver was seated on the other side of an audiometric booth and simulated the computer's intelligence by selecting the character she believed her child intended to move. Caregiver classifications resulted in a movement of the selected character on both the child's and parent's screens. Before each phase, 3 practice trials were elicited at each of the distinct levels (i.e. 3 levels of pitch, 3 levels of duration and 9 levels of pitch–duration) to ensure that the child understood the task. These demonstration trials followed the same procedures as the experimental trial, except that the results were not analysed.

Recordings from each speaker were first segmented into individual trials using Adobe® Audition®. No further manipulations were made to the sound files for the listening task. The Praat speech analysis software package (Boersma & Weenik, 2000) was used to

Table I. Description of the 3 speakers with severe dysarthria.

Variable	S1	S2	S3
Age	6 years	13 years	9 years
Medical diagnosis	spastic cerebral palsy	spastic/athetoid cerebral palsy	spastic cerebral palsy
Speech diagnosis	Spastic dysarthria	Spastic dysarthria	Spastic dysarthria
Primary mode of communication	Multi-modal AAC: • picture communication book • speech w/ familiar partners	Multi-modal AAC: • vocalizations • eye gaze	Multi-modal AAC: • AAC device • sign language • vocalizations
AAC system	None	None	Vanguard
Length of time w/ AAC system	N/A	N/A	1 year

calculate the average fundamental frequency and duration of each vocalization. A summary of each speaker's (S1, S2, S3) mean fundamental frequency and duration at each distinct level for the data sub-set used in the present listening experiment are provided in Table II for the pitch and duration games, and in Table III for the combination game.

Participants

Thirty-six monolingual speakers of American English (mean age = 28.25; SD = 9.29) participated in this study. A questionnaire was used to group listeners according to their level of familiarity with the speaker and/or experience with dysarthria. Three versions of the questionnaire were developed. While there was overlap amongst the questionnaires, each version probed specific information that was used to determine eligibility and classification into the 3 listener groups. For example, the familiar group was asked how long they had known the child, and how often they orally communicated with the child to assure they were not only familiar with the child, but also familiar with hearing his or her voice. The aim was to recruit familiar communication partners who had similar levels of interaction with the speaker. The experienced group was probed about their experience with listening to and treating both disordered speech in general, as well as dysarthric speech in particular. Last, the inexperienced group was asked whether they had any experience with dysarthria and/or speech disorders in general, whether they had any exposure in school or in their work environment to speech impairment, and whether they had any family members, co-workers, friends, etc., who had a disability or known speech disorder.

For each speaker, 2 familiar listeners (FAM), 5 experienced listeners (EXP), and 5 unfamiliar and inexperienced listeners (INX) were recruited. Familiar listeners consisted of a parent, speech-language pathologist, one-on-one aid, and/or teacher of the child with dysarthria. The experienced listener group was comprised of second-year speech-language pathology graduate students, all of whom had previous clinical experience listening to both disordered speech as a whole and more specifically to dysarthric speech. The unfamiliar, inexperienced group consisted of listeners who had no previous exposure to the children in the study or to disordered speech. Listeners who met the criteria for the EXP and INX groups were randomly assigned to one of the three speakers. All participants had adequate visual acuity and hearing thresholds that fell below 25 dB in at least one ear for 250, 500, 1000, 2000, 4000, and 8000 Hz.

Listener task

The listening task was completed in a quiet room in our laboratory or at the participant's home. Listeners engaged in a series of 3 interactive computer games, which were similar to those used to elicit the child's vocalizations. The games were used to present the auditory

Table II. Production results by speaker for the pitch and duration control; mean fundamental frequency (SD) for each pitch level requested and mean vowel length in seconds (SD) for each duration level requested.

	Fundamental frequency (Hz)			Duration (sec)		
	High	Medium	Low	Long	Medium	Short
S1	338.1 (26.1)	279.2 (17.6)	260.9 (17.9)	3.08 (.47)	1.52 (.31)	.26 (.05)
S2	356.3 (12.3)	323.0 (23.6)	347.0 (13.8)	2.31 (.34)	1.30 (.21)	.26 (.10)
S3	285.4 (13.0)	309.0 (11.9)	122.8 (15.4)	1.58 (.20)	.61 (.09)	.19 (.02)

Table III. Production results by speaker for combined pitch and duration control in terms of mean fundamental frequency (Hz); vowel length (sec) and corresponding standard deviation (SD_{F0}; SD_{dur}) for each of 9 pitch and duration combinations.

	High; Long	High; Med	High; Short	Med; Long	Med; Med	Med; Short	Low; Long	Low; Med	Low; Short
S1	357.9; 1.49	344.5; 0.90	346.8; 0.38	286.5; 1.63	272.1; 0.86	279.9; 0.30	282.4; 1.81	261.4; 0.98	257.2; 0.34
	(37.4; 0.38)	(28.3; 0.19)	(25.3; 0.12)	(8.8; 0.29)	(4.3; 0.08)	(4.56; 0.09)	(7.3; 0.4)	(6.5; 0.15)	(9.5;0.06)
S2	346.4; 2.25	347.4; 1.25	347.5; 0.41	341.2;1.34	300.5;1.02	329.8; 0.34	333.7; 1.67	332.4; 0.78	332.1; 0.39
	(8.2; 0.7)	(6.74; 0.37)	(10.8;.0.16)	(4.9;0.31)	(7.3; 0.18)	(9.39; 0.08)	(5.75;0.99)	(17.1; 0.21)	(18.4; 0.07)
S3	544.6; 2.22	522.6; 1.08	526.4; 0.40	314.4; 1.91	289.5; 0.89	294.3; 0.31	140.9; 3.4	124.2; 1.13	140.9; 0.34
	(24.2;0.69)	(23.5;0.16)	(22.2;0.09)	(26.6;0.38)	(28.2; 0.28)	(25.6; 0.08)	(16.5;0.72)	(8.59;0.26)	(26.4;0.06)

stimuli, guide listeners through the listening task, record listener responses, and to calculate listener accuracy. The experimental task consisted of 3 phases: the pitch phase, the duration phase, and the combination phase. The order of the pitch and duration phases was counterbalanced across listeners. The combination phase, however, was always the last phase of the listening task. During each phase, the listener wore headphones (model AKG 240) and sat in front of a computer screen that displayed the interactive game and guided them through the experimental task. The task consisted of listening to a vocalization and then classifying it into one of 3 distinct levels in both the pitch and duration games and one of 9 distinct levels in the combination game. The task was not timed and listeners were able to replay each vocalization once prior to making their decision. Once they had made their choice, the scene was cleared and a "play" button appeared to signal readiness for the next trial. Three demonstration trials were provided before the pitch and duration games and 5 demonstration trials were presented before the combination game to ensure that participants understood the task and felt comfortable with the computer interface.

The pitch phase consisted of listening to 50 vocalizations and determining whether the child produced a high-, medium-, or low-pitch sound. Listeners heard a total of 15 low-, 15 medium-, and 15 high-pitch tokens, as well as 5 random repeats that were used to determine intra-rater reliability. Figure 1 provides a screenshot of the listener interface for the pitch phase.

The duration phase also consisted of 50 vocalizations that could be classified as short, medium, or long. Listeners heard 15 short-, 15 medium-, and 15 long-duration tokens, as well as 5 random repeats. Finally, the combination phase consisted of vocalizations that

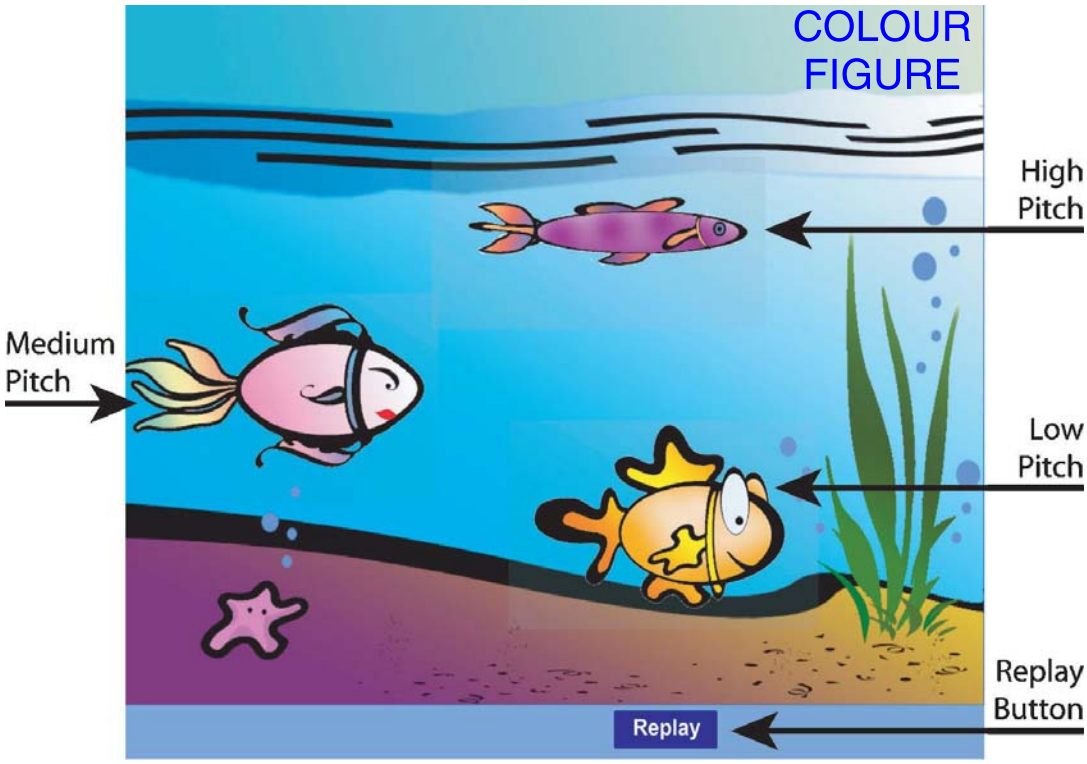

Figure 1. Screen shot of the listener interface for pitch identification. Each fish represents a specific pitch level. The topmost fish corresponds to high pitch, the middle fish to mid-pitch, and the lowest fish to low-pitch. Listeners selected the appropriate fish based upon the pitch level they thought the child produced.

varied in both pitch and duration (see Figure 2 for a screenshot of the listener interface). Listeners heard a total of 82 productions, 8 for each combination of pitch and duration (high and long, low and short, etc.), and 10 random repeats. The results of only those listeners who had greater than 80% reliability on the random repeats across all 3 experimental phases were analysed. Only one listener was excluded based on the reliability criterion.

Results

The results for each speaker are presented separately due to the inherently large individual differences in vocal abilities among speakers. Thus, for a given child, identification accuracy is compared across listener groups (FAM, EXP, INX) for each prosodic task. A criterion reference was used to determine clinically significant differences between listener groups rather than parametric analyses due to the small and unequal sample size in each listener group. While some previous studies have noted 15–20% gains in intelligibility due to familiarity (Liss et al., 2002), others have shown that even 7–8% gains can be clinically meaningful (Flipsen, 1995; D'Innocenzo et al., 2006). To err on the conservative end and to account for variability differences between listeners in each group, a 15% difference in mean accuracy between listener groups was deemed clinically significant in the present study.

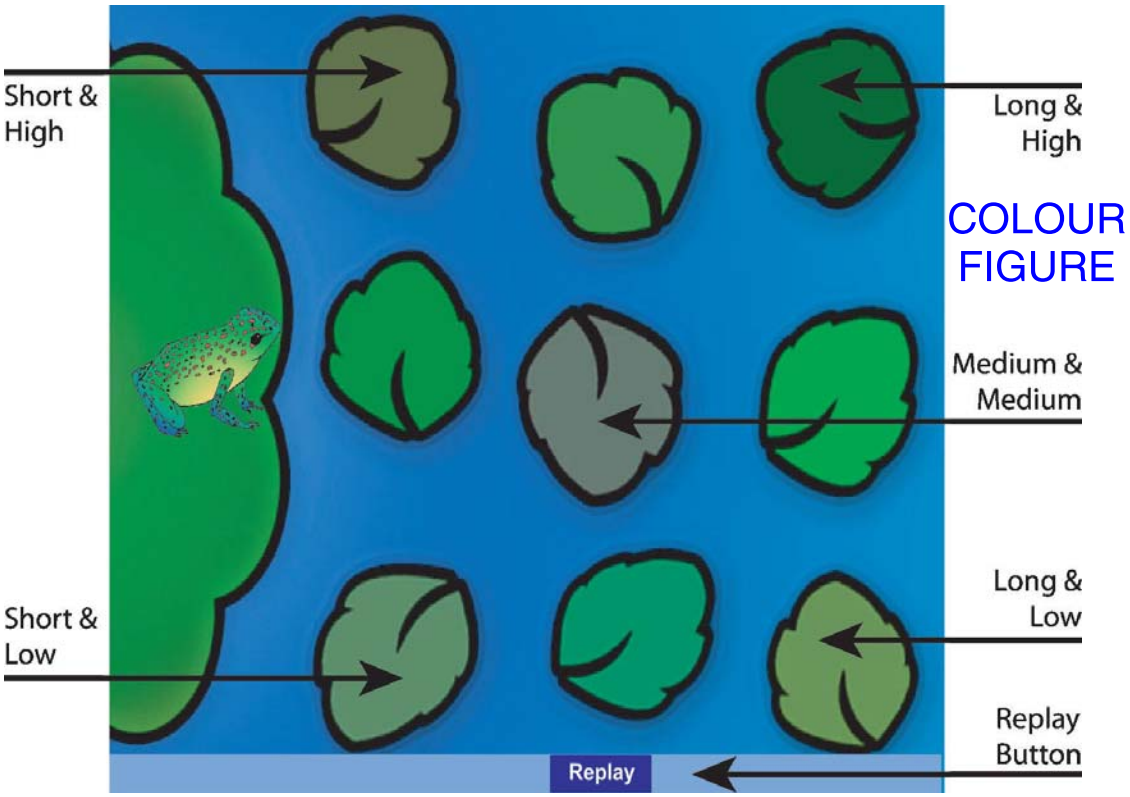

Figure 2. Screen shot of the listener interface for identifying pitch and duration combinations. Each of the 9 lily pads represents a unique pitch and duration combination with the rows corresponding to pitch and the columns corresponding to duration. Thus, the lily pad in the upper right corner represents the high-long combination, while the lower left corner represents the low-short combination. Listeners selected the appropriate lily pad based on the pitch and duration combination they thought the child produced.

Table IV. Mean percentage accuracy (SD) for identifying pitch, duration, and combination of pitch and duration within each listener group per speaker.

	Pitch			Duration			Combination		
	FAM	EXP	INX	FAM	EXP	INX	FAM	EXP	INX
S1	62.2 (4.2)	85.7 (4.0)	73.3 (6.1)	92.2 (.7)	92.4 (1.1)	95.1 (1.6)	54.9 (2.8)	47.5 (6.5)	44.4 (5.2)
S2	56.6 (.7)	59.1 (2.1)	55.6 (3.3)	75.5 (1.4)	74.7 (3.4)	79.1 (2.5)	26.4 (4.2)	28.3 (2.3)	31.1 (2.3)
S3	77.7 (.0)	73.3 (5.1)	67.1 (6.8)	100 (.0)	93.8 (2.0)	97.8 (1.0)	77.1 (3.5)	75.6 (2.3)	71.9 (4.4)

Speaker 1

Summary statistics—broken down by speaker, listener group, and prosodic task—are provided in Table IV. In addition, Figures 3–5 illustrate the accuracy levels per distinct category in the pitch, duration, and combined pitch and duration tasks, respectively.

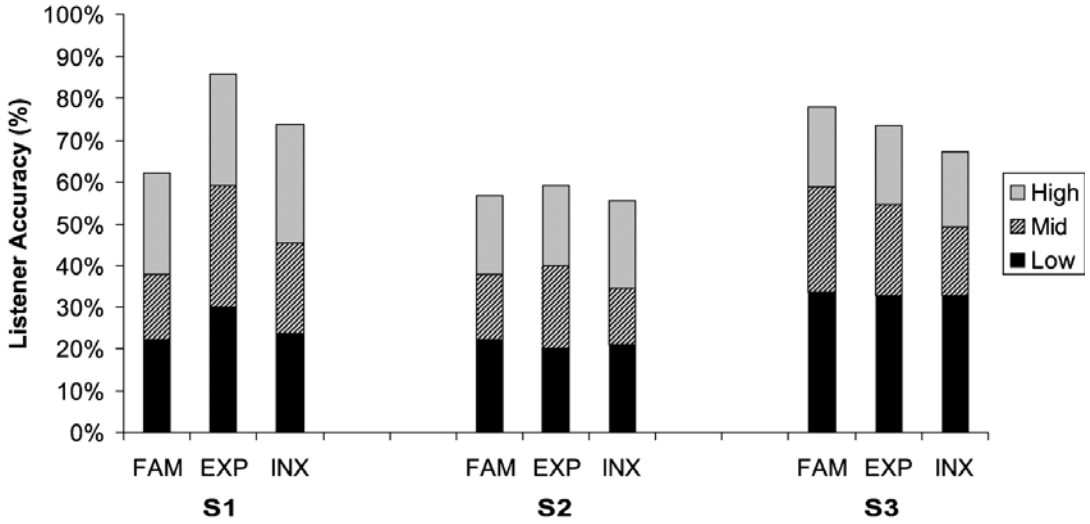

Figure 3. Mean identification accuracy for all 3 levels of pitch across familiar (FAM), experienced (EXP), and unfamiliar/inexperienced (INX) listeners.

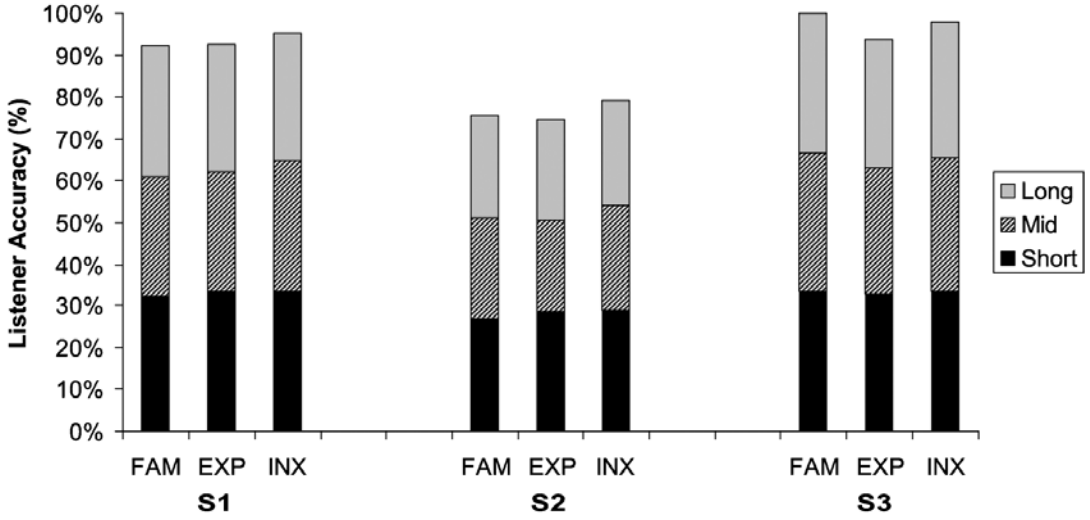

Figure 4. Mean identification accuracy for all 3 levels of duration across listener groups.

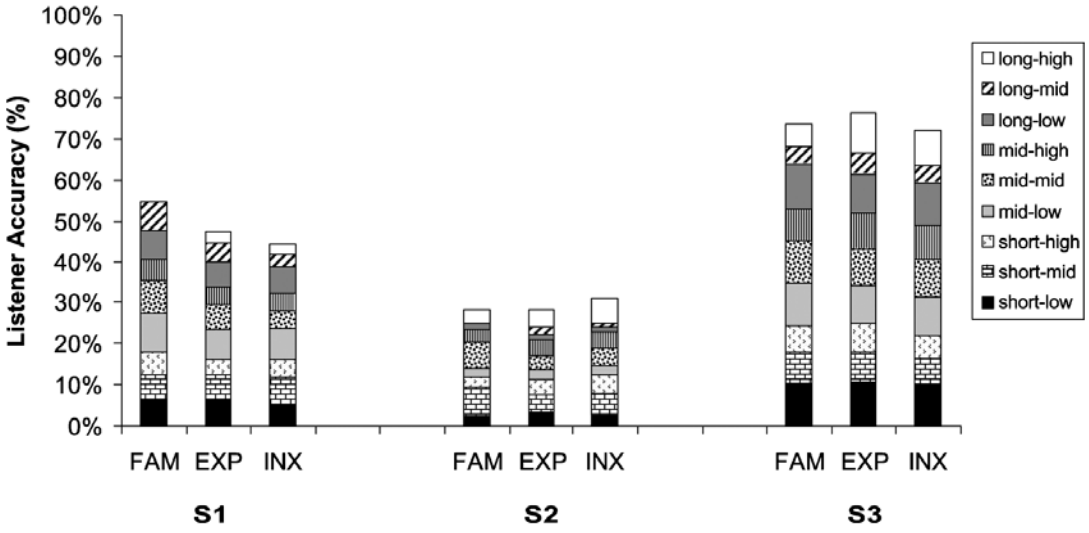

Figure 5. Mean identification accuracy for all 9 combinations of pitch and duration across listener groups.

Overall, listeners who heard S1's productions were most accurate at identifying duration contrasts and least accurate at identifying combinations of pitch and duration.

In terms of pitch identification, FAM listeners were least accurate (62.2%), whereas EXP listeners were most accurate (85.7%) at discerning vocalizations produced by S1. While the difference between the FAM and EXP groups met the 15% significance criterion, mean accuracies did not differ between the EXP and INX group or the FAM and INX groups. Figure 3 indicates that all 3 listener groups had considerable difficulty identifying the mid- and low-pitch categories. These perceptual results reflect the production results reported in Table II in which the distributions of the mid- and low-pitch productions appear to overlap.

All 3 listener groups were highly adept at identifying duration contrasts produced by S1 (mean range 92.2% for FAM to 95.1% for INX). As a result, there were no apparent differences among listener groups. Moreover, there were no notable differences in listener accuracies for the long-, medium-, and short-duration categories (see Figure 4). Table II indicates that S1 was able to produce 3 distinct duration categories with average short vowels lasting .26 seconds, mid vowels being sustained for 1.52 seconds, and long vowels extending to 3.08 seconds.

There were no apparent differences among listener groups in identifying combinations of pitch and duration with accuracies ranging from 44.4% for INX, 47.5% for EXP, to 54.7% for FAM, with the former two groups displaying more variability than the FAM group. Listeners in all 3 groups had difficulty identifying combinations with high pitch; in particular, FAM listeners were not able to identify any long-high vocalizations (see Figure 5). The production results reported in Table III indicate that the distributions of all three pitch categories overlapped and were tightly clustered and thus may have been challenging for listeners to decode.

Speaker 2

In comparison to those who heard S1 and S3, listeners who heard S2 were the least accurate, regardless of familiarity level or prosodic task. Pitch identification did not differ by familiarity, with mean accuracies ranging from 56.6% for FAM, 59.1 % for EXP, and

55.6% for INX. Figure 3 further illustrates that listeners had difficulty identifying all 3 pitch categories. These findings are supported by the production results, which indicate overlapping frequency distributions for S2's low-, medium-, and high-pitch vocalizations (see Table II).

Although accuracy was higher for identifying S2's duration categories compared to pitch contrasts, there were no apparent differences between listener groups. Furthermore, no particular duration category was more difficult to discern than others. In fact, the production results would predict higher accuracy levels given relatively distinct distributions for S2's short, medium, and long productions.

Listener accuracy levels across groups approached chance performance for the combined pitch and duration task. Once again there were no differences among listener groups or between pitch and duration categories. All listener groups had particular difficulty identifying long-mid, long-low, and mid-low vocalizations. In general, across all 9 pitch and duration combinations, listener accuracy scores were low and the variability within listener groups was high. S2's inability to produce distinct pitch differences appears to impact listener performance (see Table III and Figure 5).

Speaker 3

As a group, listeners who heard S3's productions performed better on all 3 prosodic tasks than those who heard S1 and S2. There were, however, no differences among listener groups on all 3 tasks. Pitch identification accuracy varied from 77.7% for FAM, 73.3% for EXP, to 67.1% for INX. Noteworthy is the variability among listeners in the EXP and INX groups. Regardless of familiarity level, listeners were least accurate in identifying high pitch and most accurate in identifying low pitch (see Figure 3). Table II indicates that this may be due, at least in part, to overlapping distributions of S3's high and medium pitch categories.

All three listener groups identified duration categories with greater than 93% accuracy. Given that S2 [] produced distinct non-overlapping duration categories (see Table II), it is 1 not surprising that there were no differences between listener groups and among the three duration levels (see Figure 4).

Listener accuracy scores for identifying pitch and duration combinations were considerably higher for S3, as compared to S1 and S2. Although FAM listeners were 77.1% accurate, this advantage was only less than 2% greater than the EXP group and ~ 5% greater than the INX group. All listeners had particular difficulty identifying short-high and long-mid vocalizations (see Figure 5). The production results indicate that while S2's [] fundamental frequency range on the pitch task was only 160 Hz; it was closer to 420 Hz 2 for the combined pitch and duration task (see Table III). This improvement in prosodic differentiation may account for the relatively high levels of listener accuracy.

Discussion

Within this limited sample, listener familiarity with the speaker or with speech impairment did not prove advantageous for deciphering among prosodic variations. In fact, all 3 listener groups were highly accurate in identifying duration changes and to a lesser degree pitch changes produced by the 3 speakers. While listener accuracies varied by speaker, there was no evidence of a familiarity advantage for any of three speakers. This lack of a clinically-meaningful difference among listener groups may have been due to the fact that familiarity, by definition, relates to previous exposure to the specific acoustic signal (Spitzer et al.,

2000). Although familiar listeners in the present study had previous exposure to the children's vocalizations, the task of producing sustained vowels at varying levels of prosody may not have been part of the child's everyday vocalization repertoire. Thus, familiar listeners were at no particular advantage for the tasks studied herein. Perhaps if they had access to visual cues, or additional communicative signals, a familiarity advantage may have been evident. Alternatively, as Kwiatkowski and Shriberg (1992) noted, when dysarthric vocalizations are decontextualized and listeners must rely on the acoustic cues alone, the familiarity advantage is negligible.

In addition to listener variables, it is likely that identification accuracy was impacted by the level of prosodic control available to the speakers. For example, listeners who heard S2 were least accurate on all 3 prosodic tasks compared to those who heard S1 and S3, regardless of the level of familiarity. This implies that S2 may not have been reliable in signalling the prosodic contrasts, a claim that is supported by the production data. Case history also suggests that this speaker was the most severely impaired of the group. Consistent with Dongilli's (1994) findings, it appears that when severity increases, it may be more difficult to make use of any aspects of the speech signal.

Task variables may also have impacted the results. The results suggest that, regardless of familiarity, all listener groups were highly successful in identifying duration across all speakers. Perhaps because duration changes were the easiest for children to produce, they were also the most concrete category for listeners to distinguish. In contrast, the pitch and combined pitch and duration tasks may not have been sensitive to level of familiarity for the opposite reason. The motor complexity of these tasks may have led to poor listener performance across all 3 familiarity groups. The production data from the present study and previous work with adults with dysarthria due to cerebral palsy (Patel, 2002) suggest that speakers are less reliable at producing distinct pitch categories.

Interactions among speaker and listener variables are evident, in that listeners were most accurate in identifying short duration sounds. Given the respiratory control difficulties associated with dysarthria (Blumberg, 1955; Hardy, 1983), it is expected that long duration sounds may be more difficult to produce and, thus, more difficult to accurately identify. Even if a child were able to produce long-duration vowel sounds that differed from short- and medium-duration sounds, insufficient respiratory support and vocal fatigue may have led to greater variance in the length of the longer vowels. The production data support this claim. Similarly, given that narrowed fundamental frequency range has often been noted as a hallmark of dysarthria (Hardy, 1983; Jacques, Rastatter, & Sullivan, 1985; Schlenck, Bettrich, & Willmes, 1993; Wit, Maassen, Gabreels, & Thoonen, 1993), children may have had difficulty producing 3 distinct pitch categories. The results indicated that listeners often confused medium-pitch vocalizations for high- and low-pitch sounds. The production results illustrate that all 3 speakers were able to produce 2 distinct pitch categories at most. In other words, adjacent pitch categories (i.e. high and medium; medium and low) often overlapped, thus listener performance appears consistent with the cues available in the acoustics. These findings are also supported by previous findings in adults with severe dysarthria due to cerebral palsy, where it was noted that the frequency range of the medium pitch category often overlapped with the high and low sounds (Patel, 2002).

In contrast to S1 and S2 listeners, S3 listeners were more accurate in identifying pitch and duration combinations compared to pitch contrasts alone. Given that the combination task was always completed last, S3 may have learned to differentiate between her productions during the combination phase of the speaking experiment. This claim is

supported by the production data, which indicate that S3 used a wider fundamental frequency range in the combination task than in the pitch alone task. Additionally, this increased precision in marking prosodic contrasts may have facilitated S3 listeners in distinguishing among combined pitch and duration categories. In either case, there appears to be some degree of learning in the production and/or classification of prosodically varying vocalizations.

Additional evidence of learning was noted in terms of performance on the reliability tokens. Occasionally, listeners across all 3 speakers were inaccurate in identifying the first reliability token, but correct in identifying the second token. Given that all listeners received visual feedback regarding the correctness of each perceptual judgement, this information may have contributed to learning the child's idiosyncratic prosodic contrasts, regardless of previous contact with the speaker. This finding is congruent with previous studies that have shown that repeated exposure to dysarthric productions is a form of familiarization that improves intelligibility (Flipsen, 1995; Tjaden & Liss, 1995b; King & Gallegos-Santillan, 1999; DePaul & Kent, 2000; Spitzer et al., 2000; Liss et al., 2002; Hustad & Cahill, 2003; D'Innocenzo et al., 2006).

Within all listener groups, particularly the EXP and INX groups, there was a wide range of variability in identification accuracy. Variability may be higher due to differing levels of experience among listeners within the EXP group and varying levels of ''noviceness'' in the INX group. Although all EXP listeners were second-year speech-language pathology graduate students, and all reported having experience with speakers with dysarthria, some participants may have had greater exposure to clients with severe impairments compared to other listeners in the group. The extent and nature of the clinical experiences among the EXP group may also have varied. Similarly, while all listeners in the INX group reported having no previous experience in listening to dysarthric or disordered speech, some participants may have had more experience than they realized, given their place of employment, or exposure to family members or friends with even mild speech impairments. While the operational definitions of familiarity followed the convention of previous studies, more stringent criteria for assessing familiarity, experience, and ''noviceness'' may be necessary in future work.

Limitations and future directions

The present study did not find a familiarity advantage for deciphering among prosodically varying vocalizations produced by children with severe dysarthria. Further inquiry is warranted to substantiate these findings. In future extensions of this work, increasing the sample size of the EXP and INX groups may help to reduce within group variance and potentially equalize variance between all 3 listeners groups. Additionally, more detailed criteria for establishing level of familiarity may be necessary to obtain a more homogenous sample of listeners within each group. Similarly, it would be ideal to increase the sample of familiar listeners, perhaps by including siblings; however, practical limitations may make this extension less feasible. Attaining more information about previous exposure to speech impairments would be beneficial for either creating additional listener groups or for excluding listeners from the EXP or INX groups. While a questionnaire was used to confirm listener level of familiarity, it may not have been sufficiently sensitive. For example, musical training or having children of their own may contribute to improved listener success in identifying prosodically varying vocalizations. In addition to listener variables, heterogeneity in the prosodic control abilities of the children made it difficult to identify

trends across listener groups if they in fact existed. Selecting speakers with similar prosodic abilities may be helpful for studying the effect of familiarity across speakers. Last, the naturalness of the prosodic tasks was sacrificed at the expense of experimental control. In the future, gathering the child's vocalization repertoire and then requesting prosodic manipulations within that repertoire may create a more natural speaking and listening task in which familiarity plays a role. Such an extension may help to tease apart the relative contribution of prosodic, phonetic, pragmatic, and non-verbal cues on communication effectiveness.

Conclusion

Familiarity with the speaker or with the speech disorder did not appear to impact listener identification of prosodic contrasts produced by 3 children with severe dysarthria due to cerebral palsy. In fact, scores earned by familiar listeners, experienced listeners, and inexperienced/unfamiliar listeners were similar across all 3 prosodic control tasks for any given speaker. Although accuracy varied across speakers, a general trend was that listeners were all highly accurate in identifying duration, somewhat less successful at identifying pitch, and least accurate at identifying combined pitch and duration contrasts. Thus, despite severely impaired speech sound intelligibility these children were able to signal some prosodic contrasts such that listeners could attune to acoustic consistencies, even with relatively little exposure. These findings suggest that interventions aimed at harnessing residual prosodic control in children with severe dysarthria may be highly efficacious. Perhaps familiarity will prove to be more advantageous once the child begins to use these prosodic contrasts in daily communication. Furthermore, listener familiarization procedures combined with speaker-specific interventions aimed at improving prosodic control may yield improved intelligibility than either intervention alone.

Acknowledgements

We would like to express our gratitude to Elyes Yaich and Chris Cieslik for illustrating and implementing the interactive games, Howard Cabral for his assistance with the statistical analysis as well as to the participants for their time and enthusiasm for the study. This work was supported by the Charles H. Hood Foundation.

Note

1. Children were told they were playing with a computer to increase motivation and minimize frustration toward the caregiver for incorrect trials.

References

Adobe® Audition® (Version 1.5) [Computer software]. (2004) San Jose, CA: Adobe Systems Inc.

Beukelman, D. R., & Yorkston, K. M. (1980). Influence of passage familiarity on intelligibility estimates of dysarthric speech. *Journal of Communication Disorders*, *13*, 33–31.

Blumberg, M. (1955). Respiration and speech in the cerebral palsied child. *American Journal of Disorders of Childhood*, *89*, 48–53.

Boersma, P., & Weenik, D. (2000). Praat, A system for doing phonetics by computer (Version 3.4) [Computer software]. Technical Report 132, Institute of Phonetic Sciences of the University of Amsterdam. Available online at: http://www.praat.org.

Bolinger, D. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford, CA: Stanford University Press.

Boothroyd, A. (1985). Evaluation of speech production of the hearing impaired: Some benefits of forced-choice testing. *Journal of Speech and Hearing Research*, *28*, 185–196.

Ciocca, V., Whitehill, T., & Joan, M. K. Y. (2004). The impact of cerebral palsy on the intelligibility of pitch-based linguistic contrasts. *Journal of Physiological Anthropology and Applied Human Science*, *23*, 283–287.

Darley, F., Aronson, A. E., & Brown, J. R. (1975). *Motor speech disorders*. Philadelphia: W. B. Saunders Company.

DePaul, R., & Kent, R. D. (2000). A longitudinal case study of ALS: effects of listener familiarity and proficiency on intelligibility judgments. *American Journal of Speech-Language Pathology*, *9*, 230–240.

D'Innocenzo, J., Tjaden, K., & Greenman, G. (2006). Intelligibility in dysarthria: effects of listener familiarity and speaking condition. *Clinical Linguistics and Phonetics*, *20*, 659–675.

Dongilli, P. Jr. (1994). Semantic context and speech intelligibility. In J. Till, K. Yorkston, & D. Beukelman (Eds.), *Motor speech disorders: Advances in assessment and treatment* (pp. 175–191). Baltimore: Paul H. Brookes.

Ellis, L. W., & Fucci, D. J. (1991). Magnitude estimation scaling of speech intelligibility: effects of listeners' experience and semantic-syntactic context. *Perceptual and Motor Skills*, *73*, 295–305.

Ellis, L. W., & Fucci, D. J. (1992). Effects of listeners' experience on two measures of intelligibility. *Perceptual and Motor Skills*, *74*, 1099–1104.

Ellis, L. W., & Fucci, D. J. (2000). Sophisticated and naïve listeners' magnitude estimation scaling judgments of speech intelligibility and speech annoyance. *Perceptual and Motor Skills*, *90*, 1231–1234.

Flipsen, P. (1995). Speaker-listener familiarity: parents as judges of delayed speech intelligibility. *Journal of Communication Disorders*, *28*, 3–19.

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

Hustad, K. C., & Cahill, M. A. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology*, *12*, 198–208.

Jacques, R. D., Rastatter, M., & Sullivan, J. (1985). Some effects of congenital spasticity on fundamental frequency. *Perceptual Motor Skills*, *61*, 75–80.

King, J. M., & Gallegos-Santillan, P. (1999). Strategy use by speakers with dysarthria and both familiar and unfamiliar conversational partners. *Journal of Medical Speech Language Pathology*, *7*, 113–116.

Klimacka, L., Patterson, A., & Patterson, R. (2001). Listening to deaf speech: does experience count? *International Journal of Language and Communication Disorders*, *36*, 210–215.

Kreiman, J., Gerratt, B. R., & Precoda, K. (1990). Listener experience and perception of voice quality. *Journal of Speech and Hearing Research*, *33*, 103–115.

Kwiatkowski, J., & Shriberg, L. D. (1992). Intelligibility assessment in developmental phonological disorders: accuracy of caregiver gloss. *Journal of Speech and Hearing Research*, *35*, 1095–1104.

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

Lehiste, I. (1976). Suprasegmental features of speech. In N. J. Lass (Ed.), *Contemporary issues in experimental phonetics* (pp. 225–239). New York: Academic Press.

Lewis, K. E., Watterson, T. L., & Houghton, S. M. (2003). The influence of listener experience and academic training on ratings of nasality. *Journal of Communication Disorders*, *36*, 49–58.

Lindblom, B. (1990). On the communication process: speaker–listener interaction and the development of speech. *Augmentative and Alternative Communication*, *6*, 220–230.

Liss, J. M., Spitzer, S. M., Caviness, J. N., & Adler, C. (2002). The effects of familiarization on intelligibility and segmentation in hypokinetic and ataxic dysarthria. *Journal of the Acoustical Society of America*, *112*, 3022–3030.

McGarr, N. S. (1983). The intelligibility of deaf speech to experienced and inexperienced listeners. *Journal of Speech and Hearing Research*, *26*, 451–458.

Mencke, E. O., Ochsner, G. J., & Testut, E. W. (1983). Listener judges and the speech intelligibility of deaf children. *Journal of Communication Disorders*, *16*, 175–180.

Most, T., Weisel, A., & Lev-Matezky, A. (1996). Speech intelligibility and the evaluation of personal qualities by experienced and inexperienced listeners. *Volta Review*, *98*, 181–190.

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. (1999). Prosody conveys information in severely impaired speech. *Proceedings of the European Speech Community Association (ESCA) Workshop of Dialogue and Prosody* 129–135, Veldhoven, The Netherlands.

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

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

Patel, R. (2004). The acoustics of contrastive prosody in adults with cerebral palsy. *Journal of Medical Speech-Language Pathology*, *12*, 189–193.

Patel, R., & Salata, A. (2006). Using computer games to mediate caregiver-child communication for children with severe dysarthria. *Journal of Medical Speech Language Pathology*, *14*, 279–284.

Powell, T. W., Elbert, M., & Forrest, K. (1990). Listeners' familiarity and perception of an idiosyncratic phonological pattern. *Perceptual and Motor Skills*, *70*, 859–864.

Puyuelo, M., & Rondal, J. A. (2005). Speech rehabilitation in 10 Spanish-speaking children with severe cerebral palsy: a 4-year longitudinal study. *Pediatric Rehabilitation*, *8*, 113–116.

Schlenck, K., Bettrich, R., & Willmes, K. (1993). Aspects of disturbed prosody in dysarthria. *Clinical Linguistics and Phonetics*, *7*, 119–128.

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

Spitzer, S. M., Liss, J. M., Caviness, J. N., & Adler, C. (2000). An exploration of familiarization effects in the perception of hypokinetic and ataxic dysarthric speech. *Journal of Medical Speech-Language Pathology*, *8*, 285–293.

Tjaden, K. K., & Liss, J. M. (1995a). The role of listener familiarity in the perception of dysarthric speech. *Clinical Linguistics & Phonetics*, *9*, 139–154.

Tjaden, K. K., & Liss, J. M. (1995b). The influence of familiarity on judgments of treated speech. *American Journal of Speech-Language Pathology*, *4*, 39–48.

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

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.