Effect of Speaking Rate on Comprehension of Prosodic Intent in Dysarthria

Rupal Patel, Ph.D.

Department of Speech-Language Pathology and Audiology, Northeastern University

Harvard-MIT Division of Health Sciences & Technology—Program in Speech and Hearing Bioscience and Technology

Katherine Alexander, M.S.

Department of Speech Language Pathology and Audiology, Northeastern University

Keywords: rate; prosody; dysarthria; perceptual; acoustics; contrastive stress; question–statement contrast

This study sought to examine whether reduced speaking rate impacts listener comprehension of prosodic information in utterances produced by speakers with dysarthria (N = 7) and healthy controls (N = 7). Forty-two English-speaking adults were recruited for the listening task in the present study. Spoken samples were acquired in a previous production study in which speakers engaged in a naturalistic task in which they produced utterances with contrastive stress, as questions, or affirmative statements at habitual and slow speaking rates. Listeners heard conversational exchanges extracted from the original production study and were asked to classify the speaker's prosodic intent. Results indicated that listener accuracy was higher at habitual rate for all three prosodic intentions produced by speakers with dysarthria. Contrastive tokens produced by speakers with dysarthria were especially difficult to discern at slow rate. In contrast, reduced rate had little impact on classifying the productions of healthy controls. These findings suggest that although rate reduction is commonly used to improve intelligibility in dysarthria, it may hamper the speaker's ability to signal some prosodic contrasts. Interventions that emphasize maintaining prosodic contrasts at slowed rates may mitigate the trade-off between prosodic and segmental intelligibility.

Interventions aimed at improving intelligibility in dysarthria have typically focused on optimizing segmental contrasts. Rate reduction is commonly used to improve intelligibility and enhance overall communicative effectiveness (Weismer, 2007; Yorkston, Beukelman, & Bell, 1988). It is thought that slow rate alters the movements associated with speech production, thereby allowing speakers more time to achieve articulatory targets (Dromey & Ramig, 1998; Tjaden & Weismer, 1998; Tjaden & Wilding, 2005; Turner, Tjaden, & Weismer, 1995; Yorkston, Hammen, Beukelman, & Traynor, 1990). In healthy speakers, reduced rate has been associated with increased duration of lower lip and tongue tip movements and thus improved segmental accuracy (Adams, Weismer, & Kent, 1993; Dromey & Ramig, 1998). Similarly, in speakers with dysarthria, slow rate has been shown to reduce spatiotemporal variability (McHenry, 2003) and increase vowel

Journal of Medical Speech-Language Pathology
Volume 18, Number 4, pp. 109–114

space (Turner et al., 1995), thereby improving intelligibility.

Although rate reduction improves segmental clarity, it has less desirable consequences on speech prosody and naturalness (Dromey & Ramig, 1998; Tjaden & Weismer, 1998; Yorkston et al., 1990). Variations in prosody consist of changes in fundamental frequency (F0), intensity, and duration. Prosody plays an important role in sentence comprehension and indication of sentence focus or stress and is highly intertwined with segmental aspects of speech (Eady & Cooper, 1988; Lehiste, 1970). Thus it is not surprising that slow rate in healthy speakers has been associated with restricted F0 variability (Dromey & Ramig, 1998) and is perceived as less natural and monotonous (Schaeffer & Eichorn, 2001). While rate reduction may be beneficial in counteracting prosodic disturbances noted in hypokinetic and/or ataxic dysarthria, it may inadvertently restrict the ability to leverage residual prosodic abilities noted in other dysarthria subtypes (Le Dorze, Oullet, & Ryalls, 1994; Patel, 2002, 2003; Patel & Campellone, 2009; Wang, Kent, Duffy, & Thomas, 2005; Yorkston, Beukelman, Minifie, & Sapir, 1984).

Our group (Patel and Campellone) has collected a database of prosodic contrasts produced at slow and habitual rates from speakers with dysarthria due to cerebral palsy and healthy controls. Acoustic analyses indicated that differences in F0 and intensity range between contrastive stress, affirmative, and question tokens at habitual rate were reduced at slow rate especially for speakers with dysarthria. The present study aimed at determining the impact of reduced speaking rate on listener ability to identify prosodic intent in conversational dialogues between speakers with and without dysarthria extracted from the original production study.

METHOD

A subset of tokens from our previously collected database of short conversational dialogues of seven speakers with dysarthria due to cerebral palsy (mean age 34.9 years; 7M; spastic, mixed spastic and ataxic, mixed spastic and flaccid) and 7 age- and gender-matched healthy controls (mean age 36.9 years) served as stimuli for the present experiment. The dataset was comprised of recordings of speakers playing a modified game of "Sorry!" Speakers alternated taking roles of the "Director" and "Mover" and played four game rounds, two at habitual rate and two at half of the habitual rate or slow. Dialogues consisted of the Director instructing the mover to "move the [color] [shape] X spaces", the Mover asking a clarifying question (Q) in the form "the [color] [shape]?" and the Director producing either an affirmative statement "Yes, the [color] [shape]" (A) or a contrastive statement "No, [COLOR][shape]" (C).

Listener Participants. Forty-two English speakers (mean age = 25.3; 17M, 25F) with no known history of speech, language, or cognitive impairment were recruited as listeners. All listeners passed a hearing screening with thresholds at or below 25 dB at 0.5, 1, 2, and 4 kHz.

Procedure. The listening experiment was conducted in a sound-treated booth and stimuli were presented via headphones (AKG 240). A custom graphical interface was used to play sound samples and record listener responses. Listeners heard excerpts of conversational dialogues between a Director and a Mover extracted from the production study. Speakers with dysarthria and healthy controls took turns playing the Director or Mover. The perceptual task consisted of completing a dialogue sequence by selecting the Q, A, or C token from a series of three segmentally identical sound samples (Figure 1). Each dialogue frame consisted of three panels. In the top panel, listeners heard the Director instruct the Mover to move a game piece (square or star), without specifying its color. In the second panel, the Mover asks for clarification on which color piece needs to be moved. Thus listeners heard three possible tokens (C, A, Q) and were asked to choose the appropriate sample (Q). In the bottom panel, the Director then provided feedback either affirming or contrasting the color of the piece. Listeners again heard three new randomly ordered tokens (A, C, Q) and were asked to choose the appropriate token (either A or C). Speech tokens were randomly ordered for each frame. Visual cues in the form of each player's "cards" were provided to indicate intended actions. Listeners heard 106 tokens: 96 unique tokens (32 of each type, C, Q, and A) and 10 reliability tokens.

RESULTS

A split plot factorial design was used to account for individual differences among speakers with dysathria to signal prosodic contrasts. The 7 speaker pairs formed the main plot, while the 6 listeners who heard each speaker pair formed split plots.

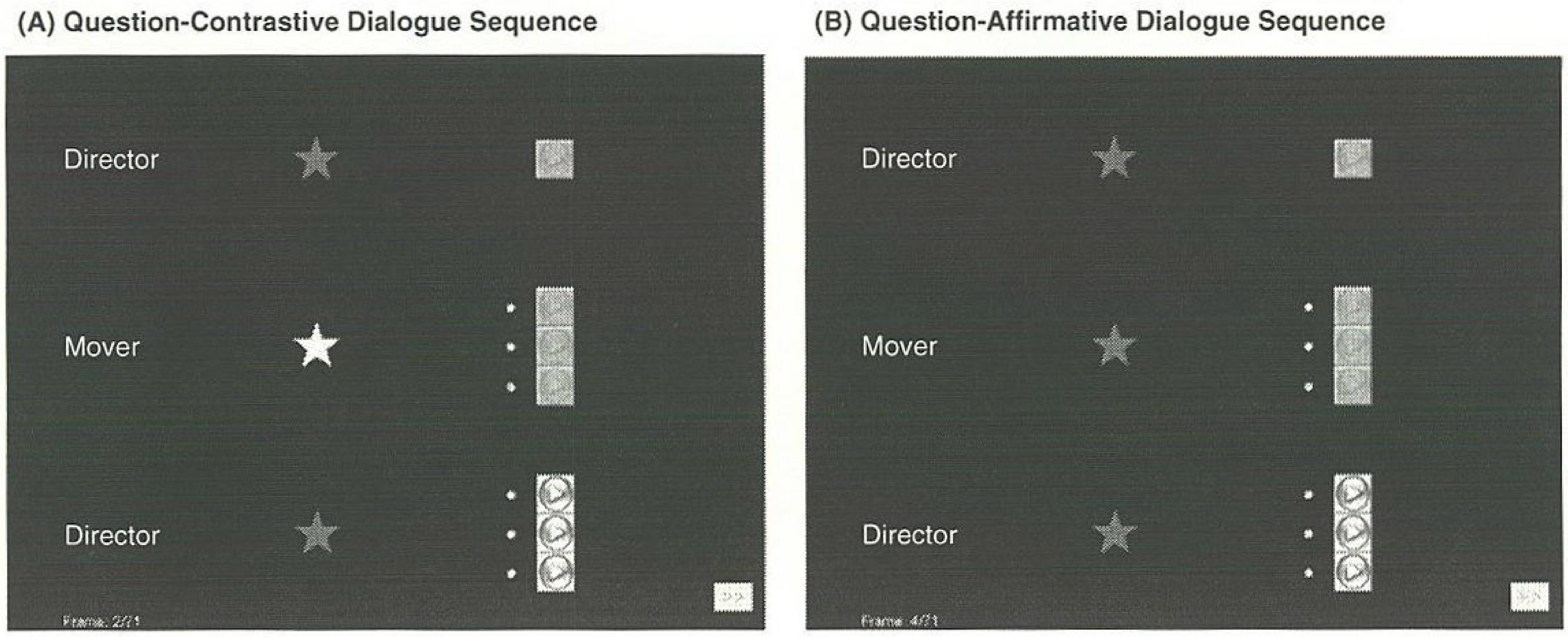

Figure 1. Screenshots of the Experimental Interface: (A) The Dialogue Sequence Requires Listeners to Identify a Question in the Middle Panel Followed by a Contrastive Token in the Bottom Panel; (B) Again Listeners Identify a Question in the Middle Panel Followed by an Affirmative Token in the Bottom Panel

To examine differences in listener accuracy, an analysis of variance (ANOVA) was performed using SPSS (Version 15.0) with three within subject factors of prosodic condition (C, A, Q), rate (slow and habitual) and speaker type (speakers with dysarthria, healthy control), and one between subjects factor of speaker pair (7 speaker pairs). An alpha level of 0.05 was used to test all main effects and adjusted to account for multiple post-hoc comparisons.

Statistically significant main effects were found for condition ($p < .0001$) and speaker pair ($p < .0001$). Figure 2 illustrates differences in listener accuracy by prosodic condition (C, A, Q) and speaker group. While questions were easiest to discern at both rates for both speaker groups, listeners had more difficulty discriminating between C and A, particularly for samples produced by speakers with dysarthria at slow rate (Figure 2A). Differences in listener accuracy between speaker pairs were anticipated and accounted for by the split plot design. Significant two-way interactions were noted between speaker type*condition ($p < .0001$), speaker type*rate ($p = .039$) and condition*speaker pair ($p = .015$). The following three- and four-way interactions were also found to be significant: speaker type*rate*speaker pair ($p = .018$), speaker type*condition*speaker pair ($p < .0001$), rate*condition* speaker pair ($p = .042$), speaker type*rate*condition ($p = .038$), and speaker type*rate* condition*speaker pair ($p = .026$). Thus listener's ability to discern conditions (C, A, Q) varied depending on which speaker they heard and whether the speaker had dysarthria or not. Listeners were more accurate at identifying C produced by speakers with dysarthria than healthy controls at habitual rate (compare Figure 2A and B). Also noteworthy are the differences in listener accuracies within speaker pairs across conditions and rates (see Figure 2C and D).

DISCUSSION

This study was premised on the notion that rate reduction may have adverse consequences on prosodic modulation. Results from the current study indicate that at reduced rate, listener's ability to discern prosodic contrasts was decreased for speakers with dysarthria. While the proportion of accurate responses for each prosodic condition differed for speakers with dysarthria at habitual rate, the ability to discern between A and C tokens was reduced at slow rate. Error analyses revealed that C tokens were often confused for A at slow rate. These perceptual findings mirror the lack of acoustic contrasts noted between A and C tokens produced by speakers with dysarthria at slow rate in the production data.

To simulate a natural conversational dialogue, the perceptual task had an implicit order pattern of prosodic conditions. In each dialogue frame, listeners first identified a Q token and then either the C or A token based on the conversational context. Thus while listeners were not explicitly told

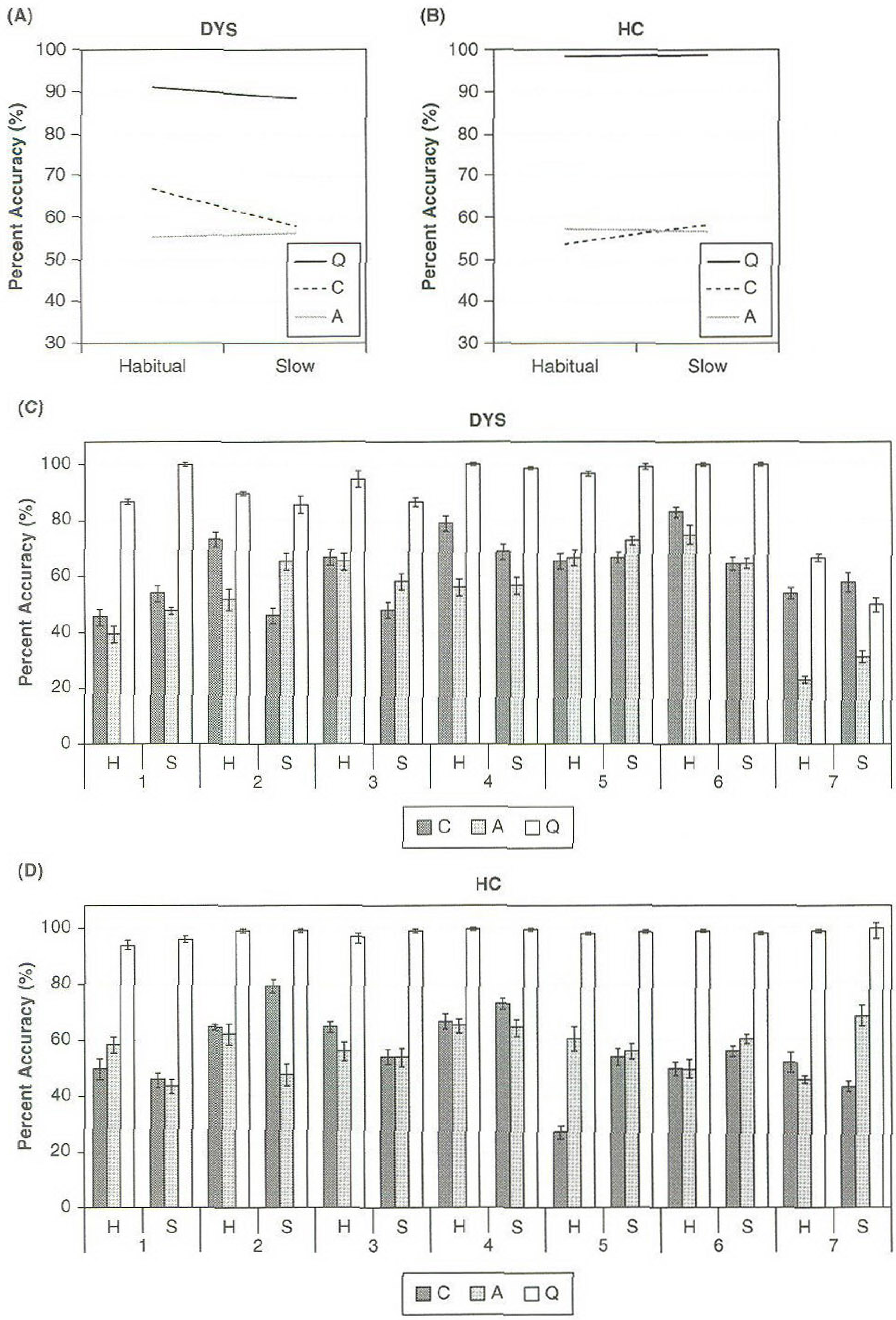

Figure 2. Listener Accuracies by Speaker Group (Top Two Graphs; A. Speakers with Dysarthria [DYS], and B. Healthy Controls [HC]) as well as by Individual Speakers (Bottom Two Graphs) Across Prosodic Conditions (C = Contrastive, A = Affirmative, Q = Question) and Rates (H= Habitual; S = Slow)

to select Q tokens from the middle panel, and C or A tokens from the bottom panel, they learned this pattern over the course of the experiment. For some speakers with dysarthria (DYS1, DYS 2, DYS 7), listeners rarely selected Q tokens for the bottom panel (i.e. when only A or C was appropriate), yet they often mistakenly selected A or C tokens from the middle panel (i.e., when only Q was appropriate). These observations suggest that while listeners may have known they were looking for a Q in the middle panel, the acoustic contrasts between conditions may have been dampened particularly at slow rate.

Reduced rate impacted healthy controls to a lesser extent than speakers with dysarthria. Interestingly, at habitual rate, listeners were more accurate at identifying C tokens produced by speakers with dysarthria compared to healthy controls. Perhaps speakers with dysarthria were using redundant and exaggerated acoustic cues to signal C at habitual rate (Howell, 1993; Patel, 2003; Patel & Campellone, 2009, Yorkston et al., 1984). Yet this apparent "advantage" was diminished as speakers with dysarthria implemented the slow speaking rate strategy.

In summary, the present findings suggest that although traditional dysarthria therapies may be effective in increasing segmental clarity, prosodic contrasts may become less distinct at reduced rate and thus more difficult to discern. While the present findings need to be extended to larger sample and more diverse group of speakers with dysarthria (subtype, gender, and age), they highlight the interaction between prosodic and segmental aspects of speech. Novel interventions that focus on improving segmental intelligibility while harnessing residual prosodic control to optimize communication success are warranted.

Acknowledgments This work was funded in part by the American Speech-Language-Hearing Association SPARC award (Students Preparing for Academic and Research Careers). We thank Jon Brumberg for implementing the perceptual interface and the participants for their time.

Address Correspondence to *Rupal Patel, Ph.D., Department of Speech Language Pathology and Audiology, Northeastern University, 360 Huntington Ave., Rm 102 FR, Boston, MA 02115*
e-mail: r.patel@neu.edu

REFERENCES

Adams, S. G., Wesimer, G., & Kent, R. D. (1993). Speaking rate and speech movement velocity profiles. *Journal of Speech and Hearing Research, 36*, 41–54.

Dromey, C., & Ramig, L. (1998). Intentional changes in sound pressure level and rate: Their impact of respiration, phonation and articulation. *Journal of Speech, Language, and Hearing Research, 41*, 1003–1018.

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

Howell, P. (1993). Cue trading in the production and perception of vowel stress. *Journal of the Acoustical Society of America, 94*(4), 2063–2073.

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

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

McHenry, M. (2003). The effect of pacing strategies on the variability of speech movement sequences in dysarthria. *Journal of Speech, Language, and Hearing Research, 46*, 702–710.

Patel, R. (2002). Prosodic control in severe dysarthria: Preserved ability to mark the question-statement contrast. *Journal of Speech, Language, and Hearing Research, 45*(5), 858–870.

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., & Campellone, P. (2009). Acoustic and perceptual cues to contrastive stress in dysarthria. *Journal of Speech, Language, and Hearing Research, 52*(1), 206–222.

Schaeffer, N., & Eichorn, N. (2001). The effects of differential vowel prolongations on the perception of speech naturalness. *Journal of Fluency Disorder, 26*, 335–348.

Tjaden, K., & Weismer, G. (1998). Speaking-rate-induced variability in F2 trajectories. *Journal of Speech, Language, and Hearing Research, 41*, 976–989.

Tjaden, K., & Wilding, G. (2005). Effect of rate reduction and increased loudness on acoustic measures of anticipatory coarticulation in multiple sclerosis and Parkinson's disease. *Journal of Speech, Language, and Hearing Research, 48*, 261–277.

Turner, G., Tjaden, K., & Weismer, G. (1995). The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis. *Journal of Speech and Hearing Research, 38*, 1001–1013.

Wang, Y.-T., Kent, R. D., Duffy, J. R., & Thomas, J. E. (2005). Dysarthria associated with traumatic brain injury: Speaking rate and emphatic stress. *Journal of Communication Disorders, 38,* 231–260.

Weismer, G. (2007). *Motor speech disorders*. San Diego, CA: Plural Publishing.

Yorkston, K., Beukelman, D., & Bell, K. (1988). *Clinical management of dysarthric speakers*. Boston, MA: College Hill Press.

Yorkston, K., Beukelman, D., Minifie, F., & Sapir, S. (1984). Assessment of stress patterning. In McNeil, M., Rosenbek, J., & Aronson, A. (Eds.). *The dysarthrias: Acoustics, perception, management* (pp. 131–162). San Diego, CA: College-Hill Press.

Yorkston, K., Hammen, V., Beukelman, D., & Traynor, C. (1990). The effect of rate control on the intelligibility and naturalness of dysarthric speech. *Journal of Speech and Hearing Disorders, 55*, 550–560.