Impact of Prosodic Strategies on Vowel Intelligibility in Childhood Motor Speech Impairment

Kathryn P. Connaghan, Ph.D.

Department of Speech Language Pathology & Audiology, Northeastern University, Boston, Massachusetts

Rupal Patel, Ph.D.

Department of Speech Language Pathology & Audiology and College of Computer and Information Sciences, Northeastern University, Boston, Massachusetts

Keywords: children, prosody, rate reduction, loudness modification, intelligibility

---

Vowel production contributes to the intelligibility of individuals with motor speech disorders. In adults with dysarthria, rate reduction and increased loudness have been shown to improve intelligibility partly because of changes in vowel production. The impact of prosodic modulation on vowel acoustics and intelligibility in children with motor speech impairment (MSI), however, is unclear. The current investigation aimed to provide perceptual and acoustic data from three children (ages 3 years, 6 months to 8 years, 6 months) with motor programming deficits producing utterances in four conditions: (1) habitual, (2) increased loudness, (3) slowed rate, and (4) emphatic stress. Additional acoustic data were collected from three healthy age-matched control participants (ages 4 years, 1 month to 8 years, 10 months). For each condition, utterances containing three corner vowels (/i/, /a/, /u/) were elicited using an audio-visualization technique. Perceptual intelligibility was based on judgments of 12 unfamiliar listeners. The impact of prosodic strategies varied across speakers, with two children with MSI demonstrating increased vowel intelligibility in at least one modulation condition. Acoustic analyses included first (F1) and second (F2) formant extraction and comparison of variance across conditions. F2 variance was greater for the children with MSI. These preliminary findings shed light on the interaction between prosodic modulation and intelligibility in childhood MSI and warrant further investigation.

INTRODUCTION

Prosodic modulation strategies are commonly used to improve intelligibility in motor speech disorders (dysarthria and apraxia of speech). Global prosodic strategies entail modulating a prosodic feature (e.g., rate, intensity) across the entire utterance as a means to improve speech intelligibility (Yorkston, Hakel, Beukelman, & Fager, 2007). In contrast, local strategies focus on specific words or phrases, such as placing emphatic stress on a linguistically salient word. A common goal of these strategies is to increase vowel space area (VSA; Tjaden & Wilding, 2004) as reduced VSA caused by vowel centralization has been documented in a number of populations with dysarthria (Bunton & Leddy, 2011; Liu, Tsao, & Kuhl, 2005; Weismer, Jeng, Laures, Kent, & Kent, 2001). Although VSA is positively correlated with intelligibility (Bunton & Leddy, 2011; Liu et al., 2005;

Journal of Medical Speech-Language Pathology
Volume 20, Number 4, pp. 133–139

Higgins & Hodge, 2002; Hustad, Gorton, & Lee, 2010; Weismer et al., 2001), there are differing reports on the strength of the relationship (Higgins & Hodge, 2002; Tjaden & Wilding, 2004; Weismer et al., 2001).

Moreover, there is a paucity of data demonstrating a link between prosodic modulation and intelligibility in childhood motor speech disorders. Although this link may be clearer in children with dysarthria who have reduced VSA (Higgins & Hodge, 2002; Hustad et al., 2010), it is more tenuous in childhood apraxia of speech (CAS) in which VSA is not particularly restricted. Nonetheless, children with CAS demonstrate deficits such as vowel errors (Davis, Jacks, & Marquardt, 2005), inconsistency of vowel productions (Smith, Marquardt, Cannito, & Davis, 1994), and decreased vowel distinctiveness (Nijland, Maasen, Van der Meulen, Gabreels, Kraaimaat, & Schreuder, 2002), which may be amenable to remediation through prosodic strategies. For example, a global strategy such as rate reduction may increase the duration of the speech movements, thereby facilitating feedback processing, which has been shown to be impaired in CAS (Terband & Maasen, 2010; Terband, Maasen, Guenther, & Brumberg, 2009). In contrast, a local strategy such as emphatic stress may facilitate appropriate rhythm and temporal dynamics to address the prosodic deficits commonly noted in CAS. The current investigation sought to determine whether (1) vowel acoustics and (2) listener judgments of vowel intelligibility differ across the prosodic modulation and habitual speech conditions for children with motor speech impairment (MSI).

METHOD

Speaking Task

Six children (ages 3 year, 6 months to 8 years, 10 months) were recruited for the speech production portion of the investigation. Three participants were referred with diagnoses of CAS. Presence of MSI was confirmed by the investigators, certified speech-language pathologists, using a battery of standardized and informal assessment tools. Three age-matched (within 6 months) healthy control (HC) participants with no reported or obvious speech, language, or developmental deficits completed the speaking task to provide a normative comparison for the formant analysis. Participant characteristics can be found in Table 1.

Participants were recorded in a quiet room in their home or a sound-treated booth using an AKG C 520 condenser microphone, placed approximately 3 cm from the left corner of the

TABLE 1. Participant Characteristics

Participant	Age/Sex	Reported Characteristics	Observed Characteristics
MSI-1	3;6/M	Referral diagnosis of CAS	Limited vowel and consonant inventory; vowel distortions; verbal sequencing difficulties; inconsistent errors; GFTA-2 below the first percentile; structural-functional examination WNL
MSI-2	5;7/M	Referral diagnosis of CAS and phonological disorder; generalized motor planning impairment	Speech distortions and idiosyncratic substitutions; verbal and nonverbal oral sequencing difficulties; phonological impairment; GFTA-2 at the 10th percentile
MSI-3	8;5/M	Referral diagnosis of CAS; receptive and expressive language and developmental delays; generalized motor planning impairment	Slow rate of speech; inconsistent errors, speech distortions; verbal and nonverbal oral sequencing difficulties; GFTA-2 at the 18th percentile
HC-1	4;1/F		
HC-2	5;6/F		
HC-3	8;10/M		

CAS = childhood apraxia of speech; GFTA-2 = Goldman-Fristoe Test of Articulation – 2 (Goldman & Fristoe, 2000): F = female; M = male; WNT = within normal limits mouth. Audio signals were pre-amplified with a Presonus AudioBox USB before recording to a Thinkpad X201 laptop computer at a sampling rate of 44,100 Hz using Audacity 1.2.6 software. After a brief training, participants were recorded producing 24 sentences that were elicited using an audio-visualization technique (Patel & Furr, 2011), which provided a recording of an adult female model and corresponding visualization of pitch, loudness, and duration modulation within the utterance. Given differing literacy skills, iconic and orthographic labels were provided for each word in the sentence. Stimuli included six sentences ranging in length from four to five monosyllabic words, with two of the words containing one of the corner vowels /i/, /a/, or /u/. Each sentence was produced in four different conditions: LOUD, SLOW, emphatic stress (STRESS), or habitual speech (HAB), for a total of 24 sentence (48 vowel) productions per participant.

Listening Task

Sixteen students in the speech-language pathology program at Northeastern University (mean age = 25 years) with no known speech, language, or hearing concerns were recruited as unfamiliar yet knowledgeable listeners who could provide reliable transcriptions of children with MSI. Sound files were delivered through a Thinkpad X201 laptop computer via headphones (K240 Studio AKG) in a sound-treated booth. Twelve listeners orthographically transcribed a subset of sentences (25%) from each MSI participant such that three listeners heard the same word in the same condition for each MSI speaker. Although each listener heard productions from all three children with MSI, they heard each word only once per child, and listening tokens were counterbalanced by speaker across listeners. Immediately following each sentence transcription, the listeners confirmed the vowel within each word by circling it on a list. For validation of stress placement, each of the twelve listeners indicated the stress placement for one of the MSI participants (four listeners per MSI participant); the remaining four listeners validated stress placement of HC productions.

Analyses

Recordings were analyzed offline using Praat (Boersma & Weenink, 2009). To confirm that speakers slowed rate and increased loudness, utterance duration was compared across HAB and SLOW conditions, and peak word intensity was compared across HAB and LOUD conditions. Participants MSI-3 and HC-2 did not meet the criteria of 25% increase in intensity or a statistically significant difference between HAB and LOUD conditions. Additionally, vowel acoustics were measured in all conditions by extracting the mean F1 and F2 from a 30-msec window centered on the vowel nucleus (Higgins & Hodge, 2002). Formants were then Bark transformed (Zwicker & Terhardt, 1980), plotted on the F1—F2 plane, and analyzed using the coefficient of variation (COV) to identify changes in formant variation across groups and conditions. Interjudge reliability of acoustic analysis was conducted on 10% of tokens ($r = 0.92$).

Perceptual analyses also followed a two-step process. First, only tokens judged as the "stressed" word within the utterance by three of the four listeners were included for analysis in the STRESS condition. Second, vowel intelligibility was calculated for each MSI participant across conditions. Each token was judged to be intelligible if at least two of the three listeners correctly identified the target vowel.

RESULTS

The F1—F2 Bark-transformed MSI plots (*filled*) and intelligibility scores across the four conditions are shown in Figure 1. Each plot includes the data from the age-matched HC (*unfilled*) to allow for normative comparison. The perceptual analysis of vowels revealed increases in intelligibility for two participants, MSI-1 and MSI-3. This finding is exemplified by the increase in the vowel intelligibility of MSI-3, which increased from 67% in HAB to 100% in STRESS. Notably, MSI-1 did not produce an intelligible /a/ across any of the conditions. The COV for F1 and F2 (Table 2) was calculated for each vowel across conditions as a measure of variability and ranged from 0.00 to 0.25 for the HC group and 0.01 to 0.69 for the MSI group. Four separate repeated measures analyses of variance were conducted comparing the COV of either F1 or F2 between the groups (MSI vs. HC) and across either speaking condition (e.g., HAB, STRESS) or vowel (/i/ vs. /a/ vs. /u/). For the F1 analyses, no differences were found in either of the condition or vowel analyses ($p > .05$). For F2, the MSI group produced significantly higher COVs across vowels than the HC group ($F[1,19] = 7.02; p = .02$).

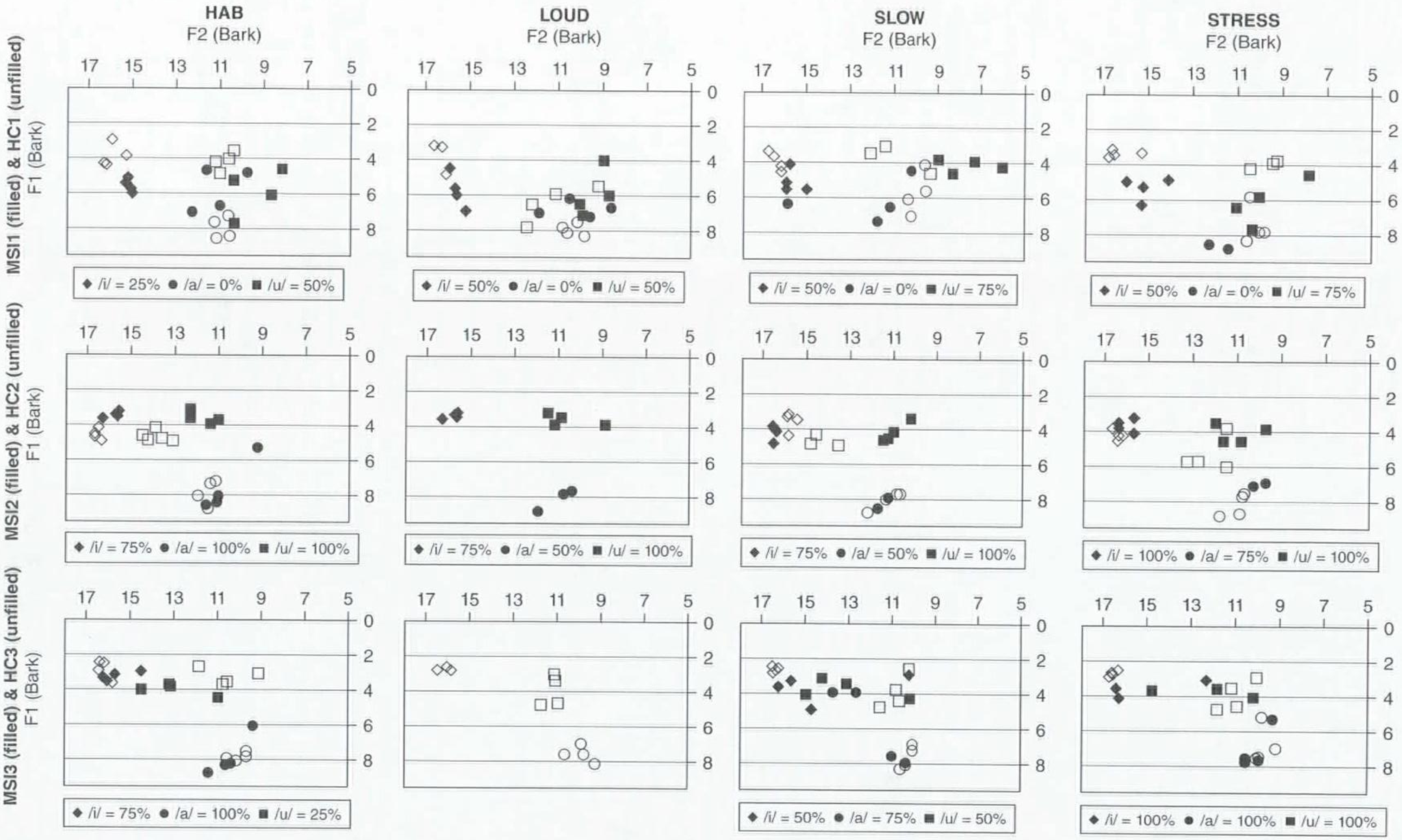

Figure 1. The F_1–F_2 plots for each participant with motor speech impairment (MSI) (*filled*) and his healthy age-matched control participant (HC) (*unfilled*) across the four speaking conditions (habitual speech [HAB], SLOW, emphatic stress [STRESS], and LOUD). The formants are Bark transformed with F2 and F1 on the *x*- and *y*-axes, respectively. The data-point shapes correspond to different vowels (*triangle* = /i/; *circle* = /a/; *square* = /u/). Beneath each plot is the vowel intelligibility scores for the MSI participants across conditions.

TABLE 2. Coefficient of Variation (standard deviation/mean) of Bark transformed F1 and F2

		HAB			SLOW			STRESS			LOUD		
		/i/	/a/	/u/	/i/	/a/	/u/	/i/	/a/	/u/	/i/	/a/	/u/
MSI-1	**F1**	0.07	0.19	0.23	0.14	0.20	0.09	0.12	0.09	0.21	0.22	0.07	0.23
	F2	0.01	0.08	0.12	0.03	0.20	0.17	0.05	0.07	0.14	0.02	0.13	0.07
HC1	**F1**	0.17	0.07	0.10	0.14	0.21	0.20	0.05	0.14	0.07	0.21	0.05	0.16
	F2	0.03	0.03	0.03	0.02	0.05	0.13	0.04	0.04	0.13	0.02	0.04	0.13
MSI-2	**F1**	0.05	0.20	0.10	0.15	0.06	0.14	0.11	0.02	0.13	0.05	0.09	0.10
	F2	0.02	0.10	0.02	0.01	0.03	0.05	0.02	0.24	0.09	0.06	0.06	0.11
HC2	**F1**	0.07	0.09	0.07	0.10	0.06	0.10	0.07	0.08	0.19			
	F2	0.01	0.03	0.03	0.00	0.06	0.05	0.01	0.05	0.07			
MSI-3	**F1**	0.07	0.15	0.09	0.24	0.38	0.14	0.68	0.17	0.08			
	F2	0.05	0.08	0.11	0.19	0.13	0.16	0.69	0.06	0.19			
HC3	**F1**	0.19	0.03	0.14	0.06	0.08	0.25	0.05	0.16	0.22	0.03	0.06	0.22
	F2	0.02	0.04	0.11	0.01	0.03	0.05	0.01	0.04	0.07	0.02	0.06	0.03

F1 = first formant; F2 = second formant; HAB = habitual speech; HC = healthy control participant; MSI = participant with motor speech impairment; STRESS = emphatic stress.

DISCUSSION AND CONCLUSIONS

The aim of this investigation was to provide preliminary data describing the impact of prosodic modulations on vowel acoustics and intelligibility in childhood MSI. Although the results were mixed across participants, two MSI participants increased intelligibility when using at least one strategy, warranting further investigation to fully explore the viability of these strategies for intervention and the mechanisms by which they influence speech production.

A noteworthy theme across both the acoustic and perceptual findings was individual speaker differences. Improvements to the intelligibility of MSI-1, the youngest and most severely impaired participant, were generally consistent across modulations, with the greatest gain observed for /i/ and no benefit gained for /a/. MSI-3 demonstrated a substantial increase to 100% vowel intelligibility in the STRESS condition yet no net benefit from reducing his rate. In contrast, the vowel intelligibility of MSI-2 was not enhanced when using any of the prosodic strategies and in fact decreased compared with the HAB condition. This finding may reflect his relatively high HAB intelligibility or be suggestive of differences in type or subtype of the motor impairment characterizing these participants.

Consistent with previous findings (Smith et al., 1994), the coefficients of variation were greater for the children with MSI than HC for at least the vowel analysis. The fact that it was significant for F2 only suggests greater variance in movements requiring tongue advancement or retraction than changes to movement of tongue height. Participant MSI-3 demonstrated a substantial range of COV particularly in STRESS, in which he also attained the highest vowel intelligibility, indicating that high variability did not affect listeners' perceptions of his vowel production in that condition. Interestingly, although alterations in movement stability have been shown with reduced speaking rate and increased loudness (see Mefferd & Green, 2010, for findings on healthy adults; see Tjaden & Wilding, 2004 for speakers with dysarthria), no such differences across conditions were found this sample.

The mechanism by which prosodic modulation may have enhanced intelligibility in MSI-1 and MSI-3 remains unclear. Given that children with CAS are thought to rely excessively on feedback (Terband & Maasen, 2010; Terband et al., 2009), perhaps modulations that increased feedback (e.g., auditory, kinesthetic) or time to process feedback was beneficial. Additionally, rate reduction, loudness variation, or emphatic stress may have targeted the very nature of the prosodic deficits that are core to CAS (American Speech-Language-Hearing Association, 2007).

The sample size of the current study was insufficient for differentiating the impact of local versus global prosodic strategies on intelligibility. It is unlikely that there is a one-size-fits all strategy given the heterogeneity across individuals and subgroups of motor speech disorders. For instance, global strategies may be less cognitively demanding and therefore more appropriate for individuals with resource allocation limitations. On the other hand, a local strategy may be well suited to individuals with limited physiological support who may fatigue when implementing a global strategy for prolonged periods. Focusing on prosody at the local level has the added benefit of assisting the listener by highlighting salient words (Liss, 2007).

This investigation provides initial findings toward our understanding of the role of prosodic modulation in the treatment of childhood MSI. Future studies that include a larger sample encompassing a broader range of impairments and severities are warranted. These efforts will provide the opportunity to test hypotheses regarding the underlying nature of childhood MSIs and inform the development of novel interventions to enhance current treatments and impact clinical practice.

Acknowledgments We would like to express our gratitude to the participants and their families. We also appreciate the Childhood Apraxia of Speech Association of North America (CASANA) for assistance with participant recruitment. Finally, we thank the members of the Communication Design and Analysis Laboratory (CadLab) at Northeastern University for their help with data collection and analysis.

Address Correspondence to *Kathryn P. Connaghan, Ph.D., Department of Speech Language Pathology and Audiology, Northeastern University, 360 Huntington Ave., Rm. 100FR, Boston, MA 02115*
e-mail: kconnaghan.neu@gmail.com

REFERENCES

American Speech-Language-Hearing Association. (2007). *Childhood apraxia of speech* [technical report]. Retrieved from http://www.asha.org/policy.

Boersma, P., & Weenink, D. (2009). PRAAT: *A system for doing phonetics by computer* [computer program]. Retrieved from http://www.praat.org/.

Bunton, K., & Leddy, M. (2011). An evaluation of articulatory working space area in vowel production of adults with Down syndrome. *Clinical Linguistics & Phonetics, 25,* 321–334.

Davis, B. L., Jacks, A., & Marquardt, T. P. (2005). Vowel patterns in developmental apraxia of speech: three longitudinal case studies. *Clinical Linguistics & Phonetics, 19,* 249–274.

Goldman, R., & Fristoe, M. (2000). Goldman-Fristoe Test of Articulation–Second Edition. Circle Pines, MN: American Guidance Service.

Higgins, C. M., & Hodge, M. M. (2002). Vowel area and intelligibility in children with and without dysarthria. *Journal of Medical Speech Language Pathology, 10,* 271–278.

Hustad, K. C., Gorton, K., & Lee, J. (2010). Classification of speech and language profiles in 4-year-old children with cerebral palsy: A prospective preliminary study. *Journal of Speech, Language, & Hearing Research, 53,* 1496–1513.

Liss, J. M. (2007). The role of speech perception in motor speech disorders. In G. Weismer (Ed.), *Motor speech disorders* (pp. 187–219). San Diego: Plural Publishing.

Liu, H. M., Tsao, F. M., & Kuhl, P. K. (2005). The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *Journal of the Acoustical Society of America, 117,* 3879–3889.

Mefferd, A., & Green, J. R. (2010). Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations. *Journal of Speech, Language, & Hearing Research, 53,* 1206–1219.

Nijland, L., Maasen, B., Van der Meulen, S., Gabreels, F., Kraaimaat, F., & Schreuder, R. (2002). Coarticulation patterns in children with developmental apraxia of speech. *Clinical Linguistics & Phonetics, 16,* 461–483.

Patel, R., & Furr, W. (2011). *ReadN'Karaoke: Visualizing prosody in children's books for expressive oral reading*. Proceedings of the Conference on Human Factors in Computing Systems, Vancouver, Canada.

Smith, B., Marquardt, T. P., Cannito, M. P., & Davis, B. L. (1994). Vowel variability in developmental apraxia of speech (pp. 81–89). In J. A. Till, K. M. Yorkston, & D. R. Beukelman (Eds.), *Motor speech disorders: Advances in assessment and treatment*. Baltimore: Paul H. Brookes Publishing.

Terband, H., & Maasen, B. (2010). Speech motor development in children with apraxia of speech: Generating testable hypotheses by neurocomputational modeling. *Folia Phoniatrica et Logopedia, 62,* 134–142.

Terband, H., Maasen, B., Guenther, F. H., & Brumberg, J. (2009). Computational neural modeling of speech motor control in childhood apraxia of speech (CAS).

Journal of Speech, Language, and Hearing Research, 52, 1595–1609.

Tjaden, K., & Wilding, G.E. (2004). Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings. *Journal of Speech, Language, & Hearing Research, 47,* 766–783

Weismer, G., Jeng, J. -Y., Laures, J., Kent, R., & Kent, J. (2001). Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. *Folia Phoniatrica et Logopaedica, 53,* 1–18.

Yorkston, K. M., Hakel, M., Beukelman, D. R., & Fager, S. (2007). Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review. *Journal of Medical Speech-Language Pathology, 15,* xi–xxxvi.

Zwicker, E., & Terhardt, E. (1980). Analytical expressions for critical band rate and critical bandwidth as a function of frequency. *Journal of the Acoustical Society of America, 68,* 1523–1524.