Effect of prosodic manipulation on articulatory kinematics and second formant trajectories in children

Kristen M. Allison,^{1,a)} Sina Salehi,¹ and Jordan R. Green²

¹*Department of Communication Sciences and Disorders, Northeastern University, 360 Huntington Avenue Boston, Massachusetts 02115, USA*

²*Department of Communication Sciences and Disorders, MGH Institute of Health Professions, 36 1st Avenue, Charlestown, Massachusetts 02129, USA*

ABSTRACT:

This study investigated effects of rate reduction and emphatic stress cues on second formant (F2) trajectories and articulatory movements during diphthong production in 11 typically developing school-aged children. F2 extent increased in slow and emphatic stress conditions, and tongue and jaw displacement increased in the emphatic stress condition compared to habitual speech. Tongue displacement significantly predicted F2 extent across speaking conditions. Results suggest that slow rate and emphatic stress cues induce articulatory and acoustic changes in children that may enhance clarity of the acoustic signal. Potential clinical implications for improving speech in children with dysarthria are discussed. © 2020 Acoustical Society of America. <https://doi.org/10.1121/10.0000670>

(Received 2 January 2020; accepted 9 January 2020; published online 5 February 2020)

[Editor: James F. Lynch]

Pages: 769–776

I. INTRODUCTION

When speakers manipulate their rate and prosody, they change multiple parameters of their speech production, including their respiratory patterns, phonation, and articulatory movements.^{1–3} Changes in these speech production parameters have perceptual influences on the clarity, naturalness, and intelligibility of speech.⁴ Because of these perceptual influences, rate and prosodic manipulation strategies have been studied and used clinically as therapeutic techniques for increasing intelligibility in speakers with dysarthria.^{3,5,6} In healthy adult speakers, rate and prosodic manipulation studies have yielded important information about the relations between articulatory adjustments and speech acoustics across a range of speaking styles.^{2,7,8} Across studies, results have shown that prosodic manipulations not only induce changes in the target prosodic dimension (e.g., slowed rate, increased loudness) but also increase acoustic distinctiveness of speech sounds.^{3,9,10} This is important because acoustic distinctiveness is strongly related to speech intelligibility in speakers with dysarthria.^{11,12} Determining how articulatory movement enhances acoustic distinctiveness is important for understanding the mechanisms underlying the success of prosodic manipulation strategies for improving intelligibility in speakers with dysarthria; however, this link is not well-understood and has rarely been studied in children. Because adult prosodic manipulation strategies are being adapted to children with dysarthria,^{13–15} knowledge about how children respond to these manipulations is essential for delineating the ages for which prosodic modulation strategies are developmentally appropriate.

Children's speech differs from adult speech in a number of important ways. Speech motor skills develop over a

protracted period lasting through adolescence.¹⁶ The shape of the vocal tract develops nonlinearly throughout childhood and adolescence, influencing its resonant characteristics and thus affecting formant frequencies.¹⁷ Formant frequencies and articulatory movements are also more variable across repeated productions of the same target in children than adults,^{18,19} presumably due to immature speech motor control. In addition, children's control of speaking rate and prosody changes with age. Habitual speaking rate increases^{20,21} and timing of segmental durations becomes more refined throughout childhood, up to at least age 11.²² The cues children use to mark prosodic contrasts also change with age; between ages 4 and 11, children show refinement in their use of acoustic cues (i.e., pitch, duration, and intensity)²³ and articulatory movements²⁴ to differentiate prosodic intent. Kinematic and acoustic cues used to mark lexical stress also differ between children and adults.^{25,26} Although prior work has provided important information regarding development of prosodic control, little is known regarding the effect of prosodic manipulations on acoustic distinctiveness of speech sounds in children. Because of the developmental differences in anatomy and speech motor control, children may not show the same benefits to articulatory precision from prosodic modulation strategies as do adults. Understanding how children adjust their articulatory movements in response to these cues, and the resulting effect on acoustics, is important for assessing the developmental appropriateness and potential clinical utility of these strategies for children with dysarthria.

A. Prosodic manipulations and articulatory precision

Several rate and prosody manipulation strategies have been shown to influence articulatory precision in both healthy adult speakers and adults with dysarthria. The most

^{a)}Electronic mail: k.allison@northeastern.edu

common strategies are *global* speech modifications, such as rate reduction, increasing loudness, or using “clear speech.”⁵ Other strategies, such as emphatic stress, are *local* speech modification strategies targeted toward modification of particular words in the utterance.⁶ Both global and local prosodic modification strategies have been shown to induce increased vowel distinctiveness relative to habitual speech for adults with dysarthria.^{3,6,27} In adults, rate reduction techniques have been shown to influence the speed and range of articulatory movements²⁸ and may allow speakers more time to achieve precise articulatory configurations, thus increasing vowel distinctiveness. Increased loudness has also been shown to influence articulatory displacement and speed^{1,29} and induce greater vowel distinctiveness.^{9,27} Loudness cues are hypothesized to have carryover effects on articulatory precision due to increased overall drive to the speech motor system.⁹ Clear speech, a strategy that induces hyperarticulation, is also associated with increased acoustic vowel space.²⁷ Emphatic stress has also been shown to induce increases in measures of vowel distinctiveness.⁶ In order to mark emphatic stress, speakers modify prosody through increased duration, loudness, and pitch of stressed words. Thus, similar mechanisms may underlie the effectiveness of emphatic stress for increasing vowel distinctiveness and improving intelligibility.

Much less is known about how children respond to prosodic manipulation cues. A recent body of work has demonstrated efficacy of Lee Silverman Voice Treatment (LSVT), which focuses on increasing loudness, for improving intelligibility in children with dysarthria secondary to cerebral palsy (CP).^{13,14,30} Strategies focused on use of a “big mouth”¹⁴ and slow rate¹⁵ have also shown preliminary evidence for improving intelligibility for some children with dysarthria secondary to CP. One study demonstrated improved vowel intelligibility in two children with motor speech impairment in an emphatic stress condition; however, the manipulation did not show clear effects on vowel acoustics.³¹ All studies examining prosodic modulation strategies in children with motor speech disorders have reported variable results across individuals,^{15,32} and the factors relating to successful use of strategies in children remain unclear. Understanding the effect of these strategies on articulatory movement and speech acoustics in typically developing children is needed to provide a developmentally appropriate frame of reference for comparing how children with dysarthria respond to these strategies.

B. Articulatory-acoustic relations

To test how changes in articulation affect speech acoustics, prosodic manipulations are often used to elicit a wide variety of articulatory changes within a participant. Movement of the vocal tract during speech shapes the acoustic signal. Tongue movement in particular has a large influence on formant frequencies;³³ however, there is not a direct one-to-one mapping between articulatory movements and acoustics, as speakers can produce the same speech sound through different articulatory configurations, consistent with the

principle of motor equivalence.^{34,35} Prior work has shown variability in the strength of articulatory-acoustic relations in adult speakers with and without motor speech disorders.^{2,8,36–38} Second formant (F2) trajectories are of particular interest because F2 is sensitive to changes in tongue movement, and the extent and slope of F2 trajectories relate to speech intelligibility for speakers with dysarthria.^{12,39,40} F2 trajectories are often studied in the context of diphthongs, which involve large F2 transitions between the beginning and end of the vowel. Thus, F2 extent during this transition provides an index of the distinctiveness between the initial and final sounds of the diphthong, and relates to the clarity with which the diphthong is perceived.^{41,42} In studies of adult speakers, F2 slope has been shown to correlate with kinematic measures of lingual speed in both healthy adults and speakers with dysarthria,^{8,36,43} and F2 extent has been shown to correlate with tongue displacement,⁴³ suggesting moderate-to-strong relations between these articulatory and acoustic parameters. Less is known about the strength of these relations in children; however, given the nonlinear changes in formants and the refinement of articulatory movements throughout childhood and adolescence, it is possible that articulatory-acoustic relations in children could differ from adults. In addition, examining the relationship between articulatory adjustments and acoustic changes in response to prosodic manipulation cues is needed to understand developmental factors influencing the potential therapeutic efficacy of these strategies for children. Because the tongue is anatomically coupled to the jaw, the jaw contributes to tongue movement during speech. Thus, examining both jaw and tongue kinematics is important for understanding how articulation is influenced by prosodic manipulations.

The objective of this study was to examine the effect of two prosodic modulation techniques (i.e., slow rate and emphatic stress) on tongue and jaw kinematics and second formant trajectories in typically developing school-aged children, and to assess the degree to which changes in tongue displacement relate to changes in second formant extent across speaking conditions. We focused specifically on these two prosodic modulations because they have received the least attention in the pediatric literature, but have important clinical potential based on prior work in adults. The following research questions were addressed:

- (1) How do slow rate and emphatic stress cues affect F2 extent and F2 slope in school-aged children?
- (2) How do slow rate and emphatic stress cues affect kinematics of tongue and jaw movement in school-aged children?
- (3) What is the relation between range of tongue movement and F2 extent across speaking conditions?

II. METHODS

A. Participants and speaking tasks

Eleven children (four males, seven females) between the ages of 7 and 12 yr (mean = 9 yr, 3 months, SD = 1 yr;

9 months) with no history of speech, language or learning difficulties participated in the study. This age range was chosen because by age 7, phonological development is largely complete and children use duration adjustments to mark prosodic contrasts.²³ In addition, this is an age range in which speech modification strategies may be most appropriate for children with dysarthria, as they are old enough to execute strategies and begin to self-monitor their speech. All children passed a hearing screening, a language screening (i.e., the *Clinical Evaluation of Language Fundamentals Screening Test*, 4th ed.⁴⁴) and achieved a standard score in the normal range on the Goldman-Fristoe Test of Articulation-2.⁴⁵ All children were native English speakers from New England.

Children were asked to repeat a set of sentences, one time each, under three speaking conditions: habitual speech, slow rate, and emphatic stress. The protocol included 14 sentences per condition, four of which contained a monosyllabic target word with the /aI/ diphthong and are included in the present investigation (i.e., “You’ll see a *kite* up there,” “Fox will *bite* a hat,” “He should *type* a note,” and “Give the *pipe* a bang.”). In each speaking condition, children were asked to repeat the target sentences following a pre-recorded model in order to ensure comprehension and consistent interpretation of the task. Stimuli were presented to children in a PowerPoint presentation with a written transcript of the sentence and an accompanying picture to maintain engagement. In each speaking condition, the pre-recorded models demonstrated the desired prosodic modification (i.e., in the “slow” condition, the models presented to the child were spoken at approximately half the speaking rate of those for the habitual condition; in the “stress” condition, models were presented at a typical rate but with emphatic stress placed on the target word). At the beginning of each condition, children were given verbal instructions about how to produce the sentences and were given two practice sentences using the desired prosodic modification prior to beginning the task. All participants began with the habitual speech condition. The order of the slow and stress conditions was counterbalanced across participants.

B. Instrumentation and data acquisition

Acoustic and kinematic data were simultaneously recorded during sessions. Acoustic data were recorded using a Williams Sound MIC 100 (Williams Sound, MN) headset microphone placed approximately four inches from the child’s mouth at a sampling rate of 44.1 kHz. An electromagnetic articulography (EMA) system (Wave, NDI, Canada) was used to record kinematic data at a sampling rate of 100 Hz. EMA data were collected for sensors placed on the forehead, upper lip, lower lip, jaw, tongue tip (approximately 1 cm from the tongue tip) and tongue body (approximately 4 cm from the tongue tip). One child did not tolerate placement of the tongue body sensor, and for another child, the tongue body sensor fell off prior to the emphatic stress condition. Therefore, tongue kinematic data

were available for ten children in the habitual and slow conditions, and nine children in the stress condition.

C. Acoustic analysis

Acoustic data were analyzed in TF32.⁴⁶ Linear predictive coding (LPC) was used to estimate second formant (F2) trajectories, and hand corrected as needed in cases of LPC tracking errors. The extent, duration, and slope of the second formant (F2) were measured for the four /aI/ diphthongs produced in each speaking condition. Beginning and endpoints of F2 trajectories were determined using the 20/20 rule⁴² (i.e., the points at which F2 begins or ceases to change by at least 20 Hz over 20 ms). F2 extent was quantified as the difference between the ending F2 frequency and the initial F2 frequency; F2 slope was quantified as the F2 extent divided by duration of the F2 trajectory (in ms). Data from two participants (18% of the stimuli) were measured by a second rater and re-measured by the first rater for reliability. Inter-judge reliability calculations showed good agreement between the first and second raters’ measurements for F2 transition duration (Pearson’s $r = 0.96$, mean difference = 16 ms) and F2 extent (Pearson’s $r = 0.91$, mean difference = 140 Hz). Intra-judge reliability calculations also showed good agreement between the first and second measurements of the first rater for F2 transition duration (Pearson’s $r = 0.94$, mean difference = 23 ms) and F2 extent (Pearson’s $r = 0.94$, mean difference = 114 Hz). This reliability is consistent with prior research and within an acceptable range.^{11,47}

D. Kinematic analysis

Articulatory kinematic variables were extracted from the three dimensional (3D) movement time-series using a MATLAB-based program, SMASH.⁴⁸ To isolate articulatory movement associated with diphthong production for each target word, movement traces were parsed from the peak in the tongue body sensor signal associated with maximum opening for the /a/ at the beginning of the diphthong, to the trough in the tongue body sensor signal associated with most closed tongue position for the /I/ at the end of the diphthong and final consonant (see Fig. 1). This parsing method enables reliable measurement of the complete target articulatory gesture across participants. For the purpose of this study, only movements of the tongue body sensor and jaw were analyzed. Head movement was automatically subtracted from the tongue and jaw movement signals in the EMA software. Tongue movement was not decoupled from jaw movement; thus, kinematic measures of the tongue body sensor reflect the combined movement of the tongue and jaw. The displacement and average speed of both the tongue body and jaw sensors were measured for the four productions of the /aI/ diphthong in each speaking condition. Values were averaged across the four productions of the /aI/ diphthong to yield an average displacement and speed of the jaw and tongue body markers for each participant. Because kinematic measures were algorithmically determined, reliability is expected to be 100%.

FIG. 1. (Color online) Kinematic parsing for /aI/ diphthong. Example shows the word “kite” for one participant. Dotted lines indicate the beginning and end-points of tongue gesture associated with diphthong production and where the kinematic signal was parsed. These same landmarks were used to parse diphthongs in all four target words.

E. Statistical analysis

Linear mixed effects (LME) models were conducted using R^{49} and *lme4*⁵⁰ to determine the effect of speaking condition on the seven acoustic and kinematic measures: F2 extent, duration, F2 slope, tongue body range, tongue body average speed, jaw range, and jaw average speed. In each model, subjects were included as random intercepts to account for individual variability and speaking condition was included as a fixed effect with the habitual condition as the reference group. Likelihood ratio tests were used to compare successively more complex models containing age and/or sex to determine whether including those variables improved model fit. To determine the most parsimonious model that best fit the data, factors were added to the model until including an additional factor no longer significantly improved model fit.⁵¹ Pairwise differences between conditions were obtained using the multcomp package in R .⁵² A linear regression model was used to evaluate the degree to which tongue displacement predicted F2 extent across speaking conditions. The effects of age and sex on acoustic and kinematic measures were descriptively examined; Pearson’s product moment correlations were used to assess the relationship between participant age and acoustic/kinematic measures.

III. RESULTS

Results of Pearson’s correlations showed that F2 extent and F2 slope were moderately negatively associated with age ($r = -0.45$ and -0.48 , respectively), but F2 transition duration and kinematic variables were not correlated with age. Descriptive analysis showed that displacement and average speed of the jaw and tongue body were lower among boys than girls, but F2 measures were similar across sexes. Results of mixed effects models are shown in Table I. For all seven acoustic and kinematic outcome measures, including speaking condition significantly improved model fit compared to a null model. Including age as a factor significantly improved model fit for F2 extent, including age and sex optimized model fit for F2 slope, and including sex optimized model fit for the four kinematic variables (i.e., tongue body displacement and speed, jaw displacement and speed). The model for duration was not significantly improved by including age or sex. Interactions between age x condition and sex x condition did not improve model fits for F2 extent or kinematic variables.

Pairwise differences in acoustic and kinematic measures between the three speaking conditions are displayed in Fig. 2. Acoustic results showed significant differences in

TABLE I. Results of the best-fitting LME models estimating the effect of speaking condition on F2 extent, duration, F2 slope, tongue body range and speed, and jaw range and speed.

Variable	F2 extent			Duration			F2 Slope			Tongue body displacement			Tongue body average speed			Jaw displacement			Jaw average speed		
	Estimate (SE)	t		Estimate (SE)	t		Estimate (SE)	t		Estimate (SE)	t		Estimate (SE)	t		Estimate (SE)	t		Estimate (SE)	t	
(Intercept)	1416(183)	7.7		98(9)	10.3		16.8(1.1)	15.4		10.8(0.8)	13.4		74(3.5)	21.3		3.7(0.3)	11.5		31(1.9)	16.7	
Slow	207(50)	4.1		104(12)	8.9		-3.4(0.5)	-7.6		0.8(0.7)	1.3		-23(2.4)	-9.9		0.2(0.3)	0.5		-12(2.4)	-4.8	
Stress	271(50)	5.4		36(12)	3.0		-0.5(0.5)	-1.1		1.9(0.7)	2.7		-6(2.4)	-2.6		1.2(0.3)	3.7		-0.1(2.5)	-0.1	
Age	-58(20)	-2.9		—	—		-0.8(0.1)	-6.6		—	—		—	—		—	—		—	—	
Sex (M)	—	—		—	—		-2.0(0.4)	-5.1		-3.5(1.1)	-3.1		-23(5.1)	-4.5		-2.0(0.4)	-4.8		-14(2.1)	-6.8	

vowel duration across all three speaking conditions, with the longest durations in the slow condition and longer durations in the stress condition than the habitual condition. These findings confirm that children were successfully making the intended speaking modifications. Children's F2 extents were significantly larger in the slow and stress conditions than the habitual condition, but the slow and stress conditions did not significantly differ from each other. F2 slope was significantly reduced in the slow condition, but did not differ between the habitual and stress conditions. Kinematic results indicated that children had significantly larger tongue and jaw displacements in the stress condition than in the habitual condition. Jaw displacement in the slow condition was also significantly reduced compared to the stress condition. Speed of tongue and jaw movement were significantly reduced in the slow condition compared to the habitual and stress conditions. Tongue speed was significantly faster in the habitual condition compared to the stress condition, but this difference was not found in jaw speed.

Figure 3 shows the relation between each participant's tongue body displacement and F2 extent for each speaking condition. Results of a linear regression model showed that children's tongue body displacement significantly predicted their F2 extent across speaking conditions ($B = 51.4$, $p < 0.001$, $r^2 = 0.42$).

IV. DISCUSSION

Results of this study suggest that typically developing school-aged children modify the range and speed of their

articulatory movements in response to rate reduction and emphatic stress cues, and that both of these strategies induce increases in F2 extent during diphthong production, suggesting increased acoustic distinctiveness. Children's change in tongue displacement was related to their change in F2 extent between conditions, supporting the correspondence between articulatory adjustments and acoustic distinctiveness in children. Although children's age and sex influenced their F2 extents and articulatory displacements, these factors did not modify the effect of speaking condition on acoustic and kinematic variables. These findings are discussed in detail below.

A. Effect of prosodic manipulation on acoustics

Children in this study adjusted both the duration and extent of their F2 transitions in response to prosodic modulation cues. F2 transition durations varied as expected by speaking condition: the slow condition elicited the longest durations, and the stress condition elicited durations that were significantly longer than the habitual condition but shorter than the slow condition. These findings suggest that typically developing school-aged children were able to adjust their vowel durations to execute the desired prosodic modulations. In addition, the increased F2 transition duration in the stress condition suggests that children used increased duration as a cue to mark emphatic stress, consistent with prior research showing that by age 7, children use duration adjustments to mark prosodic contrasts.²³ The larger F2 extents in the slow and stress conditions compared

FIG. 2. (Color online) Bar graphs displaying differences between habitual, slow, and emphatic stress speaking conditions on: (a) F2 transition extent, duration, and slope, and (b) tongue and jaw displacement and speed. Error bars represent standard error.

FIG. 3. (Color online) Scatterplot showing the relation between tongue body displacement and F2 extent across speaking conditions.

to the habitual condition indicate greater acoustic distinctiveness between the /a/ and /I/ components of the /aI/ diphthong. Although not examined in this study, the larger F2 extent may support increased vowel clarity. This finding is consistent with prior studies showing increased vowel distinctiveness in response to prosodic modulation cues in adults with dysarthria.²⁷ Because increased F2 extent has been linked with increased intelligibility in children with dysarthria,¹¹ these larger extents suggest that rate reduction and emphatic stress cues have the potential to increase intelligibility in school-aged children with dysarthria.

B. Effect of prosodic manipulation on articulatory kinematics

Results also showed that children modified the displacement and speed of their jaw and tongue movements in response to slow rate and emphatic stress cues. The decreased speed of tongue and jaw movement in the slow condition indicate that children slowed their articulatory movements in response to rate reduction cues, as expected. The average speed of children's tongue movements significantly differed between the stress and habitual conditions, but this effect was not significant for jaw speed, despite the increased duration of F2 transitions. The latter finding is consistent with the lack of difference in F2 slope between the habitual and stress conditions. Tongue and jaw displacements were greatest in the stress condition, suggesting children increased the range of their articulatory movements most in this speaking condition. Collectively, these data suggest that in response to rate reduction cues, children slowed their articulatory movements with little change in range of articulatory movement, but in response to emphatic stress cues, children maintained or slightly decreased the speed of their articulatory movements and exaggerated their articulatory gestures by increasing tongue and jaw displacement.

These findings are consistent with prior research on adult speakers showing increased displacement and

decreased speed of articulatory movements in slow rate speaking conditions compared to habitual speech.²⁸ Previous studies on articulatory kinematics of emphatic stress in adults have shown increased articulatory displacements and increased peak velocities in emphatic stress conditions.^{53,54} Although less is known about the effects of stress on articulatory movement in children, one prior study by Grigos and Patel showed that children between ages 4 and 11 demonstrated differences in articulatory displacement in declarative vs interrogative sentences, but no differences in velocity of articulatory movement between prosodic conditions.²⁴ Although emphatic stress differs from the contrasting intonation patterns in Grigos and Patel, results of the present study are consistent with their findings that children use increased articulatory displacement to make prosodic contrasts. The significant difference in tongue speed observed in the present study may indicate a difference associated with emphatic stress specifically. The difference between present findings and previous findings from adult studies may be due to developmental differences in children's ability to control the speed of articulatory movements compared to adults; however, it may also be due to methodological differences in the phonetic contexts used between studies.

C. Articulatory-acoustic relations

The strong correlation between change in F2 extent and change in tongue displacement across speaking conditions provides support for the relation between range of articulatory movement and F2 transition extent in children, at least in the context of diphthong production. Although not examined in this study, these results suggest that increased tongue displacement may support improved acoustic clarity of the speech signal in children. This interpretation is predicated on prior research suggesting a strong relationship between acoustic and articulatory specification in adult talkers.^{2,7} Furthermore, this finding indicates that children who increase their range of articulatory movement most in the slow and stress conditions relative to the habitual condition have the largest gains in F2 extent, suggesting the greatest benefits to vowel clarity. The strong association between tongue displacement and F2 extent also provides support for the hypothesis that increases in articulatory excursion contribute to improved phoneme distinctiveness. In addition, it suggests that the children who increase their articulatory excursions in response to prosodic modulation cues may achieve the most benefit from them; although overly exaggerated excursions could distort the acoustic signal.

Children with dysarthria have reduced acoustic vowel space area and reduced F2 extents in diphthong production compared to typically developing children;^{55–57} however, their articulatory displacements and speeds in utterances can be increased compared to typically developing children.^{58,59} These findings have been interpreted to suggest that children with dysarthria have difficulty controlling the force and timing of their speech movements. An important direction for

future research is to determine whether children with dysarthria further increase their articulatory excursions in response to prosodic modulation cues, and the degree to which resulting articulatory changes are associated with improved acoustic distinctiveness and intelligibility.

D. Limitations and conclusions

Findings from this study should be considered preliminary, due to the small sample size, only one repetition of each sentence, and because acoustic distinctiveness was limited to analysis of a specific diphthong. Further validation of the effects of prosodic modulation strategies on vowel acoustics is needed in larger groups of children, at different ages, and using additional measures of vowel dispersion and distinctiveness. These findings support the need for future research studying the effects of rate reduction and emphatic stress as strategies for increasing speech intelligibility in children with dysarthria. In addition, future work is needed to compare rate and emphatic stress to other prosodic manipulation strategies in children, including clear and loud speech.

ACKNOWLEDGMENTS

This work was supported by the following grants from the National Institutes of Health (NIH) Institute on Deafness and other Communication Disorders: Grant No. F32DC016484-01 awarded to K.A. and Grant No. K24DC016312 awarded to J.G. The authors would like to thank Brian Richburg for his assistance with this project as well as the participants and their parents.

- ¹C. Dromey, L. O. Ramig, and A. B. Johnson, "Phonatory and articulatory changes associated with increased vocal intensity in Parkinson disease: A case study," *J. Speech Hear. Res.* **38**(4), 751–764 (1995).
- ²A. S. Mefferd and J. R. Green, "Articulatory-to-acoustic relations in response to speaking rate and loudness manipulations," *J. Speech, Lang. Hear. Res.* **53**, 1206–1219 (2010).
- ³K. Tjaden and G. E. Wilding, "Rate and loudness manipulations in dysarthria: Acoustic and perceptual findings," *J. Speech Lang. Hear. Res.* **47**(4), 766–783 (2004).
- ⁴K. Tjaden, J. E. Sussman, and G. E. Wilding, "Impact of clear, loud, and slow speech on scaled intelligibility and speech severity in Parkinson's disease and multiple sclerosis," *J. Speech, Lang. Hear. Res.* **57**(3), 779–792 (2014).
- ⁵K. M. Yorkston, M. E. Hakel, D. R. Beukelman, and S. Fager, "Evidence for effectiveness of treatment of loudness, rate, or prosody in dysarthria: A systematic review," *ANCDs Bull Board* **15**(2), xi–xxxvi (2007).
- ⁶K. P. Connaghan and R. Patel, "The impact of contrastive stress on vowel acoustics and intelligibility in dysarthria," *J. Speech Lang. Hear. Res.* **60**, 38–50 (2017).
- ⁷A. S. Mefferd, "Tongue- and jaw-specific contributions to acoustic vowel contrast changes in the diphthong /ai/ in response to slow, loud, and clear speech," *J. Speech, Lang. Hear. Res.* **60**(11), 3144–3158 (2017).
- ⁸A. Thompson and Y. Kim, "Relation of second formant trajectories to tongue kinematics," *J. Acoust. Soc. Am.* **145**(4), EL323–EL328 (2019).
- ⁹S. Sapir, J. L. Spielman, L. O. Ramig, B. H. Story, and C. Fox, "Effects of intensive voice treatment (the Lee Silverman Voice Treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: Acoustic and perceptual findings," *J. Speech Lang. Hear. Res.* **50**, 899–913 (2007).
- ¹⁰G. S. Turner, K. Tjaden, and G. Weismer, "The influence of speaking rate on vowel space and speech intelligibility for individuals with amyotrophic lateral sclerosis," *J. Speech Hear. Res.* **38**(5), 1001–1013 (1995).
- ¹¹J. Lee, K. C. Hustad, and G. Weismer, "Predicting speech intelligibility with multiple speech subsystem approach in children with cerebral palsy," *J. Speech Lang. Hear. Res.* **57**(5), 1666–1678 (2014).
- ¹²Y. Kim, R. D. Kent, and G. Weismer, "An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria," *J. Speech Lang. Hear. Res.* **54**(2), 417–429 (2011).
- ¹³C. A. Boliek and C. M. Fox, "Therapeutic effects of intensive voice treatment (LSVT LOUD®) for children with spastic cerebral palsy and dysarthria: A phase I treatment validation study," *Int. J. Speech Lang. Pathol. [Internet]* **19**(6), 601–615 (2017).
- ¹⁴E. S. Levy, L. O. Ramig, and S. M. Camarata, "The effects of two speech interventions on speech function in pediatric dysarthria," *J. Medical-Speech Lang. Pathol.* **20**(4), 82–87 (2013).
- ¹⁵A. Sakash, T. Mahr, P. Natzke, and K. C. Hustad, "Effects of rate manipulation on intelligibility in children with cerebral palsy," *Am. J. Speech-Lang. Pathol.* **23**, 1–15 (2019).
- ¹⁶A. Smith and H. N. Zelaznik, "Development of functional synergies for speech motor coordination in childhood and adolescence," *Dev. Psychobiol.* **45**(1), 22–33 (2004).
- ¹⁷H. K. Vorperian and R. D. Kent, "Vowel acoustic space development in children: A synthesis of acoustic and anatomic data," *J. Speech Lang. Hear. Res.* **50**(6), 1510–1545 (2007).
- ¹⁸S. Lee, A. Potamianos, and S. Narayanan, "Acoustics of children's speech: Developmental changes of temporal and spectral parameters," *J. Acoust. Soc. Am.* **105**(3), 1455–1468 (1999).
- ¹⁹M. I. Grigos, "Changes in articulator movement variability during phonemic development: A longitudinal study," *J. Speech Lang. Hear. Res.* **52**(1), 164–177 (2009).
- ²⁰K. J. Logan, C. T. Byrd, E. M. Mazzocchi, and R. B. Gillam, "Speaking rate characteristics of elementary-school-aged children who do and do not stutter," *J. Commun. Disord.* **44**(1), 130–147 (2011).
- ²¹J. F. Walker, L. M. Archibald, S. R. Cherniak, and V. G. Fish, "Articulation rate in 3- and 5-year-old children," *J. Speech Hear. Res.* **35**(1), 4–13 (1992).
- ²²S. Hawkins, "On the development of motor control in speech: Evidence from studies of temporal coordination," in *Speech and Language: Advances in Basic Research and Practice* (Academic, New York, 1984), pp. 317–374.
- ²³R. Patel and M. I. Grigos, "Acoustic characterization of the question-statement contrast in 4, 7 and 11 year-old children," *Speech Commun.* **48**(10), 1308–1318 (2006).
- ²⁴M. I. Grigos and R. Patel, "Articulator movement associated with the development of prosodic control in children," *J. Speech, Lang. Hear. Res.* **50**(1), 119–1130 (2007).
- ²⁵L. Goffman and A. Smith, "Development and phonetic differentiation of speech movement patterns," *J. Exp. Psychol. Hum. Percept. Perform.* **25**(3), 649–660 (1999).
- ²⁶L. Goffman, "Kinematic differentiation of prosodic categories in normal," *J. Speech Lang. Hear. Res.* **2004**, 1088.
- ²⁷K. Tjaden, J. Lam, and G. Wilding, "Vowel acoustics in Parkinson's disease and multiple sclerosis: Comparison of clear, loud, and slow speaking conditions," *J. Speech, Lang. Hear. Res.* **56**(5), 1485–1502 (2013).
- ²⁸C. Dromey and L. O. Ramig, "Intentional changes in sound pressure level and rate," *J. Speech, Lang. Hear. Res.* **41**(5), 1003–1018 (1998).
- ²⁹M. Darling and J. E. Huber, "Changes to articulatory kinematics in response to loudness cues in individuals with Parkinson's disease," *J. Speech Lang. Hear. Res.* **54**, 1247–1260 (2011).
- ³⁰R. Bakhtiari, J. Cummine, A. Reed, C. M. Fox, B. Chouinard, I. Cribben, and C. A. Boliek, "Changes in brain activity following intensive voice treatment in children with cerebral palsy," *Hum. Brain Mapp.* **38**(9), 4413–4429 (2017).
- ³¹K. P. Connaghan and R. Patel, "Impact of prosodic strategies on vowel intelligibility in childhood motor speech impairment," *J. Med. Speech-Lang. Pathol.* **20**(4), 133–139 (2013).
- ³²C. A. Boliek and C. M. Fox, "Individual and environmental contributions to treatment outcomes following a neuroplasticity-principled speech treatment (LSVT LOUD) in children with dysarthria secondary to cerebral palsy: A case study review," *Int. J. Speech Lang. Pathol.* **16**(4), 372–385 (2014).
- ³³G. Fant, *Acoustic Theory of Speech Production* (Walter de Gruyter, Berlin, 1970), No. 2.

- ³⁴J. S. Perkell, M. L. Matthies, M. A. Svirsky, and M. I. Jordan, "Trading relations between tongue-body raising and lip rounding in production of the vowel /u/: A pilot 'motor equivalence' study," *J. Acoust. Soc. Am.* **93**(5), 2948–2961 (1993).
- ³⁵J. Perkell, M. Matthies, H. Lane, F. Guenther, R. Wilhelms-Tricarico, J. Wozniak, and P. Guiod, "Speech motor control: Acoustic goals, saturation effects, auditory feedback and internal models," *Speech Commun.* **22**(2–3), 227–250 (1997).
- ³⁶P. Rong, T. Loucks, H. Kim, and M. Hasegawa-Johnson, "Relationship between kinematics, F2 slope and speech intelligibility in dysarthria due to cerebral palsy," *Clin. Linguist. Phon.* **26**(9), 806–822 (2012).
- ³⁷J. Lee, S. Shaiman, and G. Weismer, "Relationship between tongue positions and formant frequencies in female speakers," *J. Acoust. Soc. Am.* **139**(1), 426–440 (2016).
- ³⁸C. Dromey, G. O. Jang, and K. Hollis, "Assessing correlations between lingual movements and formants," *Speech Commun.* **55**(2), 315–328 (2013).
- ³⁹Y. Kim, G. Weismer, R. D. Kent, and J. R. Duffy, "Statistical models of F2 slope in relation to severity of dysarthria," *Folia Phoniatr. Logop.* **61**(6), 329–335 (2009).
- ⁴⁰G. Weismer and R. Martin, "Acoustic and perceptual approaches to the study of intelligibility," in *Intelligibility in Speech Disorders*, edited by R. D. Kent (John Benjamins, Philadelphia, PA, 1992).
- ⁴¹G. Weismer, J. Y. Jeng, J. S. Laues, R. D. Kent, and J. F. Kent, "Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders," *Folia Phoniatr. Logop.* **53**(1), 1–18 (2001).
- ⁴²G. Weismer, R. Martin, R. D. Kent, and J. F. Kent, "Formant trajectory characteristics of males with amyotrophic lateral sclerosis," *J. Acoust. Soc. Am.* **91**(2), 1085–1098 (1992).
- ⁴³Y. Yunusova, J. R. Green, L. Greenwood, J. Wang, G. L. Pattee, and L. Zinman, "Tongue movements and their acoustic consequences in amyotrophic lateral sclerosis," *Folia Phoniatr. Logop.* **64**(2), 94–102 (2012).
- ⁴⁴E. Semel, E. H. Wiig, and W. A. Secord, *Clinical Evaluation of Language Fundamentals Screening Test*, 4th ed. (Pearson Assessments, Bloomington, MN, 2004).
- ⁴⁵R. Goldman and M. Fristoe, *Goldman-Fristoe Test of Articulation*, 2nd ed. (Pearson, San Antonio, 2000).
- ⁴⁶P. Milenkovic, "TF32," an acoustic analysis software, University of Wisconsin-Madison, 2002.
- ⁴⁷K. M. Allison and K. C. Hustad, "Acoustic predictors of pediatric dysarthria in cerebral palsy," *J. Speech, Lang. Hear. Res.* **61**(3), 462–478 (2018).
- ⁴⁸J. R. Green, J. Wang, and D. L. Wilson, "SMASH: A tool for articulatory data processing and analysis," *Interspeech* 2013, Sept., pp. 1331–1335.
- ⁴⁹R Core Team, "R: A language and environment for statistical computing" Vienna, Austria: R Foundation for Statistical Computing (2017).
- ⁵⁰D. Bates, M. Maechler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *J. Stat. Softw.* **67**(1), 1–48 (2015).
- ⁵¹J. D. Singer and J. B. Willett, *Applied Longitudinal Data Analysis* (Oxford University Press, New York, 2003).
- ⁵²T. Hothorn, F. Bretz, and P. Westfall, "Simultaneous inference in general parametric models," *Biometrical J.* **50**(3), 346–363 (2008).
- ⁵³J. A. S. Kelso, E. Vatikiotis-Bateson, E. L. Saltzman, and B. Kay, "A qualitative dynamic analysis of reiterant speech production: Phase portraits, kinematics, and dynamic modeling," *J. Acoust. Soc. Am.* **77**(1), 266–280 (1985).
- ⁵⁴D. J. Ostry, E. Keller, and A. Parush, "Similarities in the control of the speech articulators and the limbs: Kinematics of tongue dorsum movement in speech," *J. Exp. Psychol. Hum. Percept. Perform.* **9**(4), 622–636 (1983).
- ⁵⁵K. M. Allison, L. Annear, M. Policcicio, and K. C. Hustad, "Range and precision of formant movement in pediatric dysarthria," *J. Speech, Lang. Hear. Res.* **60**(7), 1864–1876 (2017).
- ⁵⁶J. Lee and K. C. Hustad, "A preliminary investigation of longitudinal changes in speech production over 18 months in young children with cerebral palsy," *Folia Phoniatr. Logop.* **65**(1), 32–39 (2013).
- ⁵⁷C. M. Higgins and M. M. Hodge, "Vowel area and intelligibility in children with and without dysarthria," *J. Med. Speech Lang. Pathol.* **10**(4), 271–277 (2002).
- ⁵⁸I. S. Nip, "Kinematic characteristics of speaking rate in individuals with cerebral palsy: A preliminary study," *J. Med. Speech Lang. Pathol.* **20**(4), 88–94 (2013).
- ⁵⁹I. S. B. Nip, C. Arias, K. Moriat, and H. Richardson, "Initial observations of lingual movement characteristics of children with cerebral palsy," *J. Speech, Lang. Hear. Res.* **60**(6S), 1780–1790 (2017).