



Published in final edited form as:

J Med Speech Lang Pathol. 2012 December ; 20(4): .

Relationship Between Prosody and Intelligibility in Children with Dysarthria

Rupal Patel, Ph.D.^{*,A,B}, Katherine C. Hustad, Ph.D.[^], Kathryn P. Connaghan, Ph.D.^{*,A}, and William Furr, MS^{*,B}

^{*}Northeastern University

^ADepartment of Speech Language Pathology & Audiology

^BCollege of Computer and Information Sciences

[^]University of Wisconsin – Madison, Department of Communicative Disorders Waisman Center

Abstract

Exaggerated and redundant prosodic cue use has been noted among adults with dysarthria secondary to cerebral palsy (CP) (Patel, 2004; van Doorn & Sheard, 2001). A possible explanation may be that speakers heighten prosodic contrasts to increase intelligibility. The current work examined whether children with dysarthria due to CP also produce exaggerated prosodic contours and if so, how prosodic cue use in these speakers impacts intelligibility. Acoustic analyses were conducted on a previously collected dataset of 2-7 word utterances produced by fourteen children with CP (7 with dysarthria and 7 without) (Hustad, Gorton & Lee, 2010). The dataset also included sentence-level transcriptions obtained from five listeners per speaker. Word intelligibility scores were derived from these transcripts and used to determine whether prosodic modulation differed for words with high versus low intelligibility. Although mean fundamental frequency (F0) and intensity range were similar across groups, words produced by children with dysarthria were slower and more variable in F0 than the group without dysarthria. Moreover, intelligibility decreased when children with dysarthria increased F0 and duration beyond the range used by children without dysarthria. Thus findings suggest that interventions targeting appropriate prosody may be beneficial in improving intelligibility in children with dysarthria and CP.

Keywords

Cerebral palsy; dysarthria; prosody; children; intelligibility

Cerebral palsy (CP) is a congenital motor disorder with prevalence estimates of approximately 3 per 1000 children (Arneson et al., 2009). Studies suggest that 60% of children with CP may have communication impairments (Bax, Tydeman, & Flodmark, 2006). Although individuals with CP can have a wide range of different underlying speech, language, and cognitive impairments, one common problem is dysarthria. Reduced intelligibility is a hallmark characteristic of dysarthria, and improving intelligibility is often

a key component of intervention (Ansel & Kent, 1992). A number of different variables have been shown to contribute to speech intelligibility. For example, studies have revealed that the size of the vowel space can account for nearly 50% of the variability in intelligibility scores (Higgins & Hodge, 2002; Liu, Tsao, & Kuhl, 2005). In addition, reduced slope of F2 transition, indicative of reduced speed of articulatory motions, has been associated with intelligibility deficits (Kent et al., 1989; Kim, Kent, & Weismer, 2011). A relationship between speech rate and intelligibility has also been demonstrated (Hustad & Sassano, 2002; Kim et al., 2011), and research on children with CP has shown that rate may be a key variable that differentiates children who have speech motor control deficits from those who do not (Hustad, Gorton, & Lee, 2010). One variable that has not been systematically explored with regard to its contribution to intelligibility in dysarthria is prosody (see De Bodt, Hernandez-Dia Huici & Van de Heyning, 2002 & Falk, Chan & Shein, 2012 for recent exceptions).

Studies of adults with CP and dysarthria indicate that many speakers retain the ability to modulate prosody (i.e. pitch, loudness and duration) despite severe speech impairment (Ciocca, Whitehill, & Ng, 2002; Patel, 2003, 2004; Patel & Campellone, 2009; van Doorn & Sheard, 2001; Whitehill, Patel & Lai, 2008). Moreover, unfamiliar listeners can accurately identify the communicative intent even when speakers use atypical prosodic cues such as increasing duration and/or intensity instead of pitch to signal a question (Patel, 2003; Whitehill & Ciocca, 2000). While speakers with CP may be compensating for physiological limitations by relying on non-standard cues, it is possible that the use of multiple redundant prosodic cues enhances intelligibility. The extent and nature of the relationship between prosody and intelligibility in dysarthria requires further inquiry. This study sought to determine whether children with dysarthria secondary to CP modulate word level prosody and if so, how these prosodic changes impact listener judgments of intelligibility.

Method

Dataset of speech recordings and listener judgments

Speech samples for the current study were obtained from an existing dataset, described in detail elsewhere (Hustad, et al., 2010). In brief, the current dataset consisted of 14 children with CP producing 2-7 word sentences from the Test of Children's Speech (TOCS+) (Hodge & Daniels, 2007). All children met the following inclusion criteria: 1) medical diagnosis of CP; 2) language comprehension abilities that were within age level expectations; and 3) hearing abilities within normal limits.

Following previously established criteria (Hustad et al., 2010), 7 of the children with CP had no evidence of speech motor impairment (NSMI) (average intelligibility = 79.26%), and 7 of the children had evidence of speech motor impairment with language comprehension skills that reflected typical developmental expectations (referred to here as DYS) (average intelligibility = 31.95%). The sample comprised 7 boys (mean age 54.2 months (SD .84)) and 7 girls (mean age 53.71 months (SD 2.08)). Among the children, 6 had spastic diplegia, 3 had hemiplegia, 1 had spastic quadriplegia, 2 had dyskinesia, and specific CP diagnoses were unknown for 2 children.

In a related study examining the same children (Hustad, Schueler, Schulz, & DuHadway, in press), intelligibility judgments for this dataset were obtained from seventy adult listeners between the ages of 18-40 years with normal hearing. Five different listeners were randomly assigned to each child; each listener heard only one child producing all stimulus material. Listeners were allowed to hear each utterance only once and were instructed to orthographically transcribe each utterance.

Analysis procedure

Speech samples and listener judgments from the Hustad et al., (2010; in press) studies were further evaluated for the current study. First, speech recordings were analyzed to extract fundamental frequency, relative intensity, and duration features for each word. A total of 246 utterances produced by children without dysarthria (NSMI) and 157 utterances produced by children with dysarthria (DYS) were manually annotated using Praat (Boersma & Wernick, 2011) to demarcate the beginning and end of each word ($r = 0.98$ inter-rater reliability for 10% of the dataset). Praat was also used to generate time-stamped F0 and intensity values for each utterance. Custom software then operated on the annotations and the time-stamped pitch and intensity values to calculate the following features per word: mean F0, peak F0, slope of F0, relative change in peak F0 above the utterance mean F0 (delta peak F0), time to peak F0, mean intensity, peak intensity, time to peak intensity, and word duration. Where appropriate, each feature was normalized by utterance measures of that feature. For example, word duration was normalized by the average word duration within an utterance to examine relative changes in word lengthening. Pitch tracking errors, when found, were manually corrected. In all, the acoustic dataset consisted of a matrix of acoustic data for 1068 words produced by children in group NSMI and 565 words produced by children in group DYS.

Word intelligibility data were extracted from the transcriptions of all 5 listeners who heard each child. All word transcriptions were categorized into two groups: those that were intelligible (operationally defined as words that at least 3 of 5 listeners transcribed correctly) or as unintelligible (operationally defined as words that 1 or fewer listeners transcribed correctly). Note that words that were correctly identified by 2 listeners were excluded from this analysis since they reflected neither intelligible nor unintelligible productions ($N = 73$ for NSMI and $N = 54$ for DYS). For each CP group, one-way analyses of variance were conducted on each acoustic feature to identify which prosodic cues separated high and low intelligibility words. To account for multiple comparisons, the Bonferroni correction factor was used to adjust the alpha level.

Table 1 provides descriptive statistics for a representative subset of acoustic features by speaker group. Overall, words produced by children without dysarthria were more intelligible than those produced by children with dysarthria (percentage of words identified correctly by 3 or more listeners; NSMI = 82.8%; DYS = 32.9%). Peak F0 was the only prosodic feature that differed across high versus low intelligibility words produced by children without dysarthria. In contrast, several prosodic features separated high and low intelligibility words in the group with dysarthria. Specifically, when peak F0 was in range with that used by children without dysarthria, intelligibility increased. If instead speakers

produced an excessive change in peak F0 within a word, it was more likely to be misjudged. The rise time of the F0 also impacted intelligibility. Words were more intelligible when the peak F0 occurred within the first third of its duration rather than later in the word. However, words with elongated duration were less intelligible. Unlike F0 and duration, intensity features did not separate high and low intelligibility words in either group.

Discussion & Conclusions

The current investigation examined whether prosodic cue use would differ across children with and without dysarthria and whether prosodic features would differ across intelligible versus unintelligible words. Consistent with findings in adults with CP, children with CP and dysarthria in the current dataset modulated prosody to a greater extent than those without dysarthria (cf. Ciocca et al., 2002; Patel 2003; Patel & Campellone, 2009). Secondly, several prosodic features separated high and low intelligibility words produced by children with dysarthria. Specifically, when words were produced with heightened F0 they tended to be better understood. There was, however, a trade-off between heightening F0 and intelligibility. Words were likely to be intelligible if the peak in F0 occurred within the first third of the word and if the F0 excursion was not excessive. These findings are consistent with reports of the adverse effects of excessive F0 variation and range in speakers with severe dysarthria (Falk et al., 2012; Patel & Campellone, 2009; Schlenck, Bettrich, & Wilmes, 1993). In contrast, the only feature that reached statistical significance for the group without dysarthria was Peak F0. This finding is also consistent with previous work in that prosodic modulation may not be necessary when segmental integrity is already high (De Bodt et al., 2002; Schlenck et al., 1993).

Another finding was that intelligibility was lower for words that were elongated in duration for children with dysarthria. One explanation for this finding is the effect of severity. Individuals with more severe dysarthria have reduced intelligibility and also tend to use a slower rate of speech. Without controlling for the confounding variable of severity (typically measured via intelligibility), it is not possible to understand fully the relationship between rate and intelligibility.

Given individual differences in prosodic cue use, further analyses are planned to examine the cue combinations used by individual speakers and to assess the impact on intelligibility. A larger dataset comprising systematically controlled levels of dysarthria severity would be highly informative for shedding light on the interaction between prosody, intelligibility and severity.

Although traditional dysarthria interventions focus on global strategies such as rate reduction and loudness enhancement across an entire utterance, results of the present study suggest that local strategies that enhance word-level prosodic features may boost intelligibility and comprehensibility. Children with CP and dysarthria may benefit from targeted intervention aimed at prosody in the early phases of treatment to ensure modulation that is both appropriate phonologically and temporally aligned to enhance intelligibility. Rather than “icing on the cake”, prosodic modulation may in fact be the scaffolding that helps bootstrap improved intelligibility.

Acknowledgements

This research was funded by grant R01DC009411 from the National Institutes of Health. The authors would like to thank Diana Franco for help with acoustic analysis, Karl Wiegand for statistical analysis and Ryan Ma for producing the schematic visualization.

References

- Ansel BM, Kent RD. Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *Journal of Speech and Hearing Research*. 1992; 35:296–308. [PubMed: 1573870]
- Arneson C, Durkin M, Benedict R, Kirby R, Yeargin-Allsopp M, Braun K, Doernber B. Prevalence of cerebral palsy: Autism and Developmental Disabilities Monitoring Network, three sites, United States, 2004. *Disability and Health Journal*. 2009; 2:45–48. [PubMed: 21122742]
- Bax M, Tydeman C, Flodmark O. Clinical and MRI correlates of cerebral palsy: The European cerebral palsy study. *Journal of American Medical Association*. 2006; 296(13):1602–1608.
- Boersma, P.; Weenink, D. *PRAAT: doing phonetics by computer* (Version 5.2.25). [Computer software]. Institute of Phonetic Sciences; Amsterdam: 2011. from <http://www.praat.org/> [March 1, 2011]
- Ciocca V, Whitehill TL, Ng SS. Contour tone production by Cantonese speakers with cerebral palsy. *Journal of Medical Speech-Language Pathology*. 2002; 10:243–248.
- De Bodt M, Hernandez-Diaz Huici M, van de Heyning P. Intelligibility as a linear combination of dimensions in dysarthric speech. *Journal of Communication Disorders*. 2002; 35(3):283–292. [PubMed: 12064788]
- Falk TH, Chan W-Y, Shein F. Characterization of atypical vocal source excitation, temporal dynamics and prosody for objective measurement of dysarthric word intelligibility. *Speech Communication*. 2012; 54:622–631.
- Higgins CM, Hodge MM. Vowel area and intelligibility in children with and without dysarthria. *Journal of Medical Speech-Language Pathology*. 2002; 10:271–277.
- Hodge, M.; Daniels, J. *TOCS+ Intelligibility Measures*. University of Alberta; Edmonton, AB: 2007.
- Hustad KC, Gorton K, Lee J. Classification of speech and language profiles in 4-year-old children with cerebral palsy: A prospective preliminary study. *Journal of Speech, Language and Hearing Research*. 2010; 53:1496–1513. doi: 1092-4388_2010_09-0176 [pii] 10.1044/1092-4388(2010/09-0176).
- Hustad KC, Sassano K. Effects of rate reduction on severe spastic dysarthria in cerebral palsy. *Journal of Medical Speech-Language Pathology*. 2002; 10:287–292.
- Hustad KC, Schueler B, Schulz L, DuHadway C. Intelligibility of 4 year old children with and without cerebral palsy. *Journal of Speech, Language and Hearing Research*. In press.
- Kent R, Kent J, Weismer G, Martin R, Sufit R, Brooks B, Rosenbek J. Relationships between speech intelligibility and the slope of second-formant transitions in dysarthric subjects. *Clinical Linguistics and Phonetics*. 1989; 3(4):347–358.
- Kim Y, Kent RD, Weismer G. An acoustic study of the relationships among neurologic disease, dysarthria type, and severity of dysarthria. *Journal of Speech, Language and Hearing Research*. 2011; 54:417–429. doi: 10.1044/1092-4388(2010/10-0020).
- Liu H-M, Tsao F-M, Kuhl PK. The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy. *Journal of Acoustical Society of America*. 2005; 117(6):3879–3889.
- Patel R. Acoustic characteristics of the question-statement contrast in severe dysarthria due to cerebral palsy. *Journal of Speech, Language & Hearing Research*. 2003; 46(6):1401–1415.
- Patel R. The acoustics of contrastive prosody in adults with cerebral palsy. *Journal of Medical Speech-Language Pathology*. 2004; 12:189–193.
- Patel R, Campellone P. Production and Identification of Contrastive Stress in Dysarthria. *Journal of Speech Language and Hearing Research*. 2009; 52:206–222.

- Schlenck K, Bettrich R, Wilmes K. Aspects of disturbed prosody in dysarthria. *Clinical Linguistics and Phonetics*. 1993; 7(2):119–128.
- van Doorn J, Sheard C. Fundamental frequency patterns in cerebral palsied speech. *Clinical Linguistics and Phonetics*. 2001; 15(7):585–601.
- Whitehill T, Ciocca V. Perceptual-phonetic predictors of single-word intelligibility: A study of Cantonese dysarthria. *Journal of Speech, Language, and Hearing Research*. 2000a; 43:1451–1465.
- Whitehill T, Patel R, Lai J. The use of prosody by children with severe dysarthria: A Cantonese extension study. *Journal of Medical Speech Language Pathology*. 2008; 16(4):293–301.

Table 1

Descriptive statistics by prosodic feature. Note significant differences are marked by *.

Intelligibility	Peak F ₀ (Hz)	Normalized Peak F ₀ (proportion above utterance peak F ₀)	PF ₀ /mean utterance F ₀ (proportion)	Normalized Time to Peak F ₀ (proportion of word length)	Delta Peak F ₀ (Hz)	Mean Peak intensity (dB)	Word duration (s)
High 3-5 (N=882)	*311.54	0.13	1.01	0.4	35.19	74.73	0.36
Low 0-1 (N=114)	324.88	0.11	0.98	0.37	37.58	75.17	0.34
High 3-5 (N= 186)	*325.43	*0.11	0.99	*0.34	*34.32	76.19	0.42
Low 0-1 (N=325)	341.67	0.15	1.01	0.41	43.12	76.03	0.48

NMSI

DYS