AJSLP

Research Article

Differentiating Typical From Atypical Speech Production in 5-Year-Old Children With Cerebral Palsy: A Comparative Analysis

Katherine C. Hustad,[a,b] Ashley Sakash,[b] Aimee Teo Broman,[c] and Paul J. Rathouz[b,c]

Objective: Early diagnosis of speech disorders in children with cerebral palsy (CP) is of critical importance. A key problem is differentiating those with borderline or mild speech motor deficits from those who are within an age-appropriate range of variability. We sought to quantify how well functional speech measures differentiated typically developing (TD) children from children with CP.
Method: We studied speech production in 45 children with CP (26 with clinical speech motor impairment [SMI] and 19 with no evidence of speech motor impairment [NSMI]) and in 29 TD children of the same age. Speech elicitation tasks were used. Intelligibility, speech rate, and intelligible words per minute were examined.
Results: All measures differentiated between all 3 groups of children with considerable precision based on area under the receiver operating characteristic curve (AUC) data. AUC was highest for overall intelligibility, which ranged from .88 to .99. Intelligible words per minute also yielded very strong AUCs, ranging from .81 to .99. In each of the receiver operating characteristic models, discrimination between groups was highest for children with speech motor impairment versus TD children. Data indicated that 90% of TD children had overall intelligibility above 87% at 5 years of age, but that no child was 100% intelligible. Furthermore, 90% children with SMI had intelligibility below 72%.
Conclusion: Findings suggest that functional speech measures differentiate very clearly between children with and without CP and that even children who do not show evidence of speech motor impairment have functional differences in their speech production ability relative to TD peers.

Cerebral palsy (CP) is a complex neuromotor disorder that has, as its defining feature, very early onset motor impairment, secondary to an underlying neuropathology (Bax, Goldstein, Rosenbaum, & Levinton, 2005). CP has a prevalence of 3.1 per 1,000 children, as identified at 8 years of age in the United States (Christensen et al., 2014). Children with CP are at a considerable risk for a range of speech, language, and communication problems. These problems may include language delay, dysarthria, limitations in using gestures, or a combination of these. Studies suggest that at least 60% of children with CP have some type of communication problem (Bax, Tydeman, & Flodmark, 2006). About half of children with CP have dysarthria (Nordberg, Miniscalco, Lohmander, & Himmelmann, 2013). Research in our laboratory has used a classification model for characterizing constellations of speech and language deficits seen in children with CP (Hustad, Gorton, & Lee, 2010; Hustad, Oakes, McFadd, & Allison, 2016). In its simplest form, this model has four categories or levels. These are children who have no speech motor impairment (NSMI) and appear otherwise typical in terms of their communication ability, children with speech motor impairment (SMI) who have typical language, children with SMI who have language impairment, and children who are unable to speak due to anarthria. Our work suggests that these profiles may not be clinically discernable until children are around 4 years of age, but that earlier speech and language characteristics are highly predictive of later profile group membership

[a]Department of Communication Sciences and Disorders, University of Wisconsin–Madison
[b]Waisman Center, University of Wisconsin–Madison
[c]Department of Biostatistics and Medical Informatics, School of Medicine and Public Health, University of Wisconsin–Madison
Correspondence to Katherine C. Hustad: kchustad@wisc.edu

Editor-in-Chief: Kristie Spencer
Editor: Karen Forrest
Received May 18, 2018
Revision received August 31, 2018
Accepted September 19, 2018
https://doi.org/10.1044/2018_AJSLP-MSC18-18-0108
Publisher Note: This article is part of the Special Issue: Selected Papers From the 2018 Conference on Motor Speech—Clinical Science and Innovations.

Disclosure: The authors have declared that no competing interests existed at the time of publication.

(Hustad, Allison, McFadd, & Riehle, 2014; Hustad et al., 2017).

A key challenge in assessing young children with CP is determining whether speech motor impairment, or dysarthria, is present. Many children with CP are delayed in the onset of speech production (Hustad et al., 2017), which limits clinical assessment of speech motor abilities at early ages. In children with CP who are able to produce speech, frank dysarthria is often very clear, manifesting perceptually as some combination of articulatory imprecision, hypernasality, short breath groups, weak or breathy voice, variations in loudness, reduced intelligibility, and reduced speech rate (Allison & Hustad, 2018; Workinger & Kent, 1991). Recent work suggests that articulatory features of pediatric dysarthria tend to be the most pronounced of the speech subsystem deficits (Allison & Hustad, 2018; Lee, Hustad, & Weismer, 2014). However, diagnosis of dysarthria in children is confounded by the fact that dysarthric speech features are similar to and overlap with typical developmental speech features (e.g., reduced rate; reduced intelligibility; articulatory omissions, distortions, substitutions). In typical children, these features resolve with development; however, in children with dysarthria, reduced rate, reduced intelligibility, and persistent articulatory errors continue through life (Kim, Martin, Hasegawa-Johnson, & Perlman, 2010; Nordberg, Miniscalco, & Lohmander, 2014; Platt, Andrews, & Howie, 1980). In addition, children, by their nature, even those who are typically developing (TD), tend to exhibit considerable individual developmental variability in their speech. This variability is seen in speech sound acquisition (Smit, Hand, Freilinger, Bernthal, & Bird, 1990) and speech motor control development (Smith & Zelaznik, 2004; Vick et al., 2012). Together, this developmental variability, along with the overlap between developmental features and dysarthric features of speech, can make clinical identification of dysarthria in young children difficult, particularly for those children with CP who may have a mild or borderline presentation.

In recent work, we have found that children with CP who present with no clinical speech motor impairment (NSMI) may have subtle subclinical differences in speech performance. Our work suggests that, as a group, these children tend to have excellent receptive language skills that fall at or above age level (Hustad, Sakash, Broman, & Rathouz, 2018). However, they tend to have difficulty with executive function. Specifically, children with CP and NSMI showed reduced executive function, on average, relative to typical children, and the proportion of children with CP and NSMI with executive function deficits was significantly higher than the expected proportion in a normal population of children (Sakash, Broman, Rathouz, & Hustad, 2018). In addition, intelligibility of children with NSMI seems to lag behind their TD peers, suggesting that they may have subtle deficits that make them different from their typical counterparts (Hustad, Schueler, Schultz, & DuHadway, 2012). Studies examining speech rate for children with NSMI have not been conducted; however, speech rate data could provide important information regarding functional speech characteristics for this group of children.

One promising approach to early differentiation between children with CP who have speech motor impairment and those who do not involves measurement of functional speech indices such as intelligibility, speech rate, and a hybrid measure of speech efficiency (intelligible words per minute [IWPM]). Because reductions in intelligibility and speech rate are hallmark features of dysarthria, we expect that, by comparing performance of children with CP to performance of their TD age mates, we may be able to begin to identify age-specific performance cut-points or benchmarks. These benchmarks could serve as useful clinical indicators of both severity and the need for intervention. The ability to differentiate children with CP who have speech motor impairment from those with no speech motor impairment may have important consequences for intervention and later outcomes. In particular, identification of speech motor impairment at the earliest possible age and the ability to predict longer term outcomes for speech may lead to earlier speech intervention and/or treatment focused on early augmentative and alternative communication systems and strategies.

The purpose of this study was to quantify how well measures of speech intelligibility, speech rate, and a hybrid measure of speech efficiency (IWPM) differentiate TD children from children with CP at 5 years of age. We chose to study children at 5 years of age because we could reliably classify children into profile groups (SMI, NSMI), which allowed us to have clinically valid characterizations of speech motor ability against which to consider how well the functional speech measures of interest differentiated children. We place particular and comparative focus on children with CP who have clinically identifiable SMI, children with CP who have NSMI (clinically identifiable), and TD children in this study. We also sought to determine which measures best separated children in order to begin to advance age-based guidelines for differential diagnosis of functional speech motor deficits in CP. The following specific research questions were addressed:

1. How well does each measure (overall intelligibility, speech rate for sentences, and IWPM for sentences) separate groups of children at the age of 5 years (i.e., TD vs. SMI, TD vs. NSMI, NSMI vs. SMI)?
2. What is the 10th percentile threshold of “normal” for the targeted measures for TD children? How well does this threshold differentiate among children in the three groups for each measure? We chose the 10th percentile (the point above which 90% of children are performing) owing to (a) the potential future value of each measure for screening children who could benefit from further evaluation and intervention and (b) the relatively low “cost,” within the CP population, of a false positive test versus that of a false negative test.

We hypothesized that measures of functional speech, in particular, intelligibility (Hustad, Oakes, & Allison,

2015) speech rate, and IWPM would each provide robust differentiation between children with CP who have SMI versus TD children and between children with SMI versus children with NSMI at 5 years of age. We further hypothesized that our measures would differentiate children with NSMI versus TD children, but that this differentiation would not be as strong as for comparisons involving children with SMI. Among the measures, we hypothesized that speech intelligibility would be the strongest differentiator, but that measures of speech rate and IWPM would also differentiate among children. We expected that 10th percentile thresholds from typical children on all three measures would provide strong differentiation among all groups.

Method

Participants

Children With CP

Forty-five children with CP participated as speakers. All children were participants in a longitudinal study on communication development in children with CP. Inclusion criteria for the larger study required that children (a) have a medical diagnosis of CP, (b) have hearing abilities within normal limits as documented by either formal audiologic evaluation or distortion product otoacoustic emission screening, and (c) be able to produce utterances of at least four words in length in an elicitation task. The final criterion was necessary because we were interested in examining variables involving speech rate and speech intelligibility in connected speech. It is important to note, however, that our inclusion criteria shaped the participant pool in important ways, with a particular bias toward children who had speech motor control that was sufficient to produce connected speech. As will be shown in the Results section, a full range of dysarthria severity, as reflected in intelligibility scores, was represented in this study.

Of the 45 children with CP, 19 had NSMI and 26 had evidence of SMI, as determined by a research speech-language pathologist. SMI was identified through clinical observation of the presence or absence of common features of dysarthria, including facial asymmetry; drooling; hypernaslity; short breath groups; breathy, harsh, or wet vocal quality; imprecise articulation; and consonant or vowel substitutions, distortions, or omissions that were not age appropriate. Perceptual judgments were made during a delayed imitation task in which the child produced a standard set of sentences ranging from two to seven words in length following an adult model (Hodge & Daniels, 2007) and during a spontaneous speech sample between the child and a parent or the child and a clinician. Note that neither intelligibility nor speech rate data were used in the classification of children. We have documented reliability of our clinical classification of children with SMI and NSMI in previous studies (Hustad et al., 2010, 2016). In our research paradigm, classification is made by two speech-language pathologists independently. For this study, classification agreement was 100%.

The mean age across children with CP was 60.8 months (*SD* = 4.9). The sample comprised 24 boys and 21 girls. All children were from homes where American English was the primary language. The children were born in the United States between 2001 and 2008.

Tables 1 and 2 present demographic characteristics of children with CP, including medical diagnoses and Gross Motor Function Classification System (Palisano et al., 1997) rating, which is a five-level index of functional self-initiated gross motor abilities (with an emphasis on sitting and walking) designed specifically for children with CP. The Gross Motor Function Classification System is a standard measure of gross motor function throughout North America and Europe.

TD Children

Twenty-nine TD children also participated as speakers. Children were recruited from the local community, including through a university preschool, through word of mouth, and through public postings. Inclusion criteria required that children (a) have TD speech, (b) have TD language, (c) have no history of developmental delay per parent report, and (d) have hearing abilities within normal limits as documented by either formal audiologic evaluation or distortion product otoacoustic emission screening. Standardized speech and language screening measures and audiologic screening were administered to ensure that children met inclusion criteria. Speech was screened using the Arizona Articulatory Proficiency Scale–Third Edition (Fudala, 2001). Language was screened using the Preschool Language Scale–Fourth Edition Screening Test (Zimmerman, Steiner, & Pond, 2012). The mean age across children who were included in the study was 61.0 months (*SD* = 5.5). The sample comprised 11 boys and 18 girls. All children were from homes where American English was the primary language. The children were born in the United States between 2001 and 2008.

Nondisabled Adult Listeners

Nondisabled listeners were 370 adults (five per child). Listeners were recruited from the university community via public postings and were primarily undergraduate students. Listeners were compensated for their participation. Five different listeners were randomly assigned to each child; each listener heard only one child producing all stimulus material. Inclusion criteria required that listeners (a) pass pure-tone hearing screening at 25 dB HL for 250 Hz, 500 Hz, 1 kHz, 4 kHz, and 6 kHz bilaterally; (b) be between 18 and 45 years of age; (c) have no more than incidental experience listening to or communicating with persons having communication disorders; (d) be a native speaker of American English; and (e) have no identified language, learning, or cognitive disabilities per self-report. The mean age of listeners was 21.5 years (*SD* = 3.2). There were 105 male and 265 female participants.

Table 1. Demographics of all children.

Children	Male:female	Total children	M_{age} (months)	*SD* age (months)
CP	24:21	45	60.8	4.9
SMI	11:15	26	62.1	5.4
NSMI	13:6	19	58.7	3.6
TD	11:18	29	61.0	5.5
All children	35:39	74	60.1	5.2

Note. SMI and NSMI designation was made by clinical assessment by a research speech-language pathologist. CP = cerebral palsy; SMI = children with speech motor impairment; NSMI = children with no speech motor impairment; TD = typically developing.

Acquisition of Speech Samples: Materials and Procedures

All speech stimuli produced by children in this study were taken from the Test of Children's Speech (Hodge & Daniels, 2007). The Test of Children's Speech is a set of single words and sentences that systematically vary in length and are developmentally appropriate (lexically, phonetically, syntactically, and morphologically) for young children. Children produced stimuli using delayed imitation following an adult model. We used this approach because it allowed us to compare listener transcription data with known target responses, thus ensuring that intelligibility scores were an accurate reflection of whether listeners perceived the target words correctly. Our research group has employed this as a standard approach to intelligibility measurement for many years.

Table 2. Demographic characteristic of children with cerebral palsy (CP).

Characteristic	SMI	NSMI	Total participants with CP
Sex			
Male	11	13	24
Female	15	6	21
GMFCS			
I	9	14	23
II	6	3	9
III	3	1	4
IV	7	1	8
V	1	0	1
Type of CP			
Spastic	21	17	38
Hemiplegia	11	10	21
Diplegia	4	5	9
Triplegia	1	0	1
Quadriplegia	4	1	4
Unknown	1	1	2
Dystonic	0	0	0
Choreoathetotic	1	0	1
Ataxic	2	1	3
Mixed	0	0	0
Unknown	2	1	3

Note. GMFCS ratings were made by clinician observation, and type of CP was obtained from medical research records. SMI = children with speech motor impairment; NSMI = children with no speech motor impairment; GMFCS = Gross Motor Function Classification System.

For this research, we used the same stimuli for each child to ensure equivalence among utterances and children. Children produced 38 different single-word utterances and up to 60 different multiword utterances. Of the multiword utterances, 10 were two words in length, 10 were three words in length, 10 were four words in length, 10 were five words in length, 10 were six words in length, and 10 were seven words in length.

To ensure consistency across modeled productions, adult recordings of each target utterance were presented to each child, along with an image depicting the utterance via a laptop computer situated in front of the child. Children were asked to repeat what they heard following the recorded model. Children's productions were monitored online by a research assistant to ensure that clean samples free of overlap with examiner speech and free of extraneous noises were obtained. Children were asked to repeat utterances when these criteria were not met. Among the 19 children with NSMI, all children were able to produce utterances up to five words in length, 17 children were able to produce six-word utterances, and 17 children were able to produce seven-word utterances. Among the 26 children with SMI, all children were able to produce utterances up to four words in length, 22 children were able to produce five-word utterances, 20 children were able to produce six-word utterances, and 16 children were able to produce seven-word utterances. Among the 29 TD children, all children were able to produce utterances up to seven words in length.

The research protocol was administered by a speech-language pathologist in a sound-attenuating room. Speech samples from children were recorded using a digital audio recorder (Marantz PMD 570) at a 44.1-kHz sampling rate (16-bit quantization). A condenser studio microphone (Audio-Technica AT4040) was positioned next to each child using a floor stand and was located approximately 18 in. from the child's mouth. The level of the signal was monitored and adjusted on a mixer (Mackie 1202 VLZ) to obtain optimized recordings and to avoid peak clipping.

Acquisition of Intelligibility Data: Materials and Procedures

Preparing Speech Samples for Playback to Listeners

Digital audio recordings were transferred to personal computer. Recordings were edited to remove extraneous noises and the examiner's voice; individual files were created for each stimulus utterance produced by each child. Audio samples were peak amplitude normalized to assure that maximum loudness levels of the recorded speech samples were the same across children and utterances, while preserving the amplitude contours of the original productions. This also enabled calibration of peak output levels for playback to listeners.

Data Collection From Listeners

Listeners completed two listening tasks, one in which they heard a single child producing all single-word stimulus utterances and one in which they heard the same child producing all multiword stimulus utterances. The order of presentation for the single-word listening task and the multiword listening task was counterbalanced among the listeners of each child to prevent a potential order effect. The individual stimulus items within each task were randomized for each listener so that no two listeners heard the stimulus items in the same sequence. Note that, for the multiword utterance stimuli, utterances of varying length were randomly intermingled to prevent a potential learning effect.

During the experiment, listeners were seated individually in a sound-attenuating suite in front of a 19-in. flat panel computer screen with a keyboard placed directly in front of them. An external speaker was connected to a computer and situated adjacent to the computer screen. The peak audio output level was calibrated to approximately 75 dB SPL from where listeners were seated and was checked periodically to ensure that all listeners heard stimuli at the same output level.

Speech stimuli were delivered via an in-house computer program that presented stimulus utterances and stored typed orthographic transcriptions. Listeners were allowed to hear each utterance one time. Listeners were told that the purpose of the study was to determine how understandable children were to unfamiliar listeners like themselves. They were instructed that children would be producing real words and to take their best guess if they were unsure as to what the child said. Listeners were provided with instructions on how to use the experimental software to advance through the experiment. In addition, they heard two sample utterances to familiarize themselves with the experimental task. Data from the sample utterances were excluded from analyses.

Analysis of Data: Speech Measures

Intelligibility

This measure was selected because it is a measure of functional speech performance (Dykstra, Hakel, & Adams, 2007). Improved intelligibility is often a primary goal of intervention, and individuals with CP frequently have reduced intelligibility due to dysarthria. Objective measurement of intelligibility involving orthographic transcription of target words by listeners has been considered the "gold standard" for clinical quantification and is often used as an outcome measure in dysarthria (Hustad & Weismer, 2007). Orthographic transcriptions of children's speech obtained from listeners were scored using our in-house computer program. The program automatically tallied the number of transcribed words that were an exact phonemic match (including bound and free grammatical morphemes) to the stimulus words in the sentences produced by the children. Misspellings and homonyms were accepted as correct, as long as all phonemes in the spoken version of the transcribed words matched the target words. The number of words identified correctly for each child across the five listeners was summed and divided by the number of words possible for single-word stimuli, yielding a percentage for single-word intelligibility. Similarly, the number of words identified correctly for two- to seven-word utterances across the two listeners per child was summed and divided by the number of words possible, yielding a percentage for overall intelligibility.

Speech Rate

This measure was of interest because reduced rate is a well-documented characteristic of dysarthria, 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 (Allison, 2015; Hustad et al., 2010). Acoustic waveforms for each sentence of two or more words in length were analyzed for duration in Praat (Boersma & Weenink, 2015). Sentence duration was defined as the time between sentence initiation and termination. Using the spectrographic signal, sentence initiation was identified by the onset of audible or visible acoustic energy associated with production of the first phoneme of the sentence, and sentence termination was identified by the offset of acoustic energy associated with production of the final phoneme of the sentence. Speech rate was calculated for each child by summing the total number of words produced across all utterances that were two or more words in length, divided by the sum of the utterance durations, then multiplied by 60 to yield words per minute (wpm) values for each child.

IWPM

This measure was examined because it is a hybrid measure incorporating both intelligibility and temporal acoustic features of speech. IWPM provides an index of speech efficiency (Yorkston & Beukelman, 1981) and simultaneously provides information about two areas of deficits that are common in dysarthria and interact with one another. Using both speech rate data and intelligibility data, IWPM was calculated for overall intelligibility (across two- to seven-word utterances) by taking the sum of all words identified correctly by listeners and dividing by the sum of the utterance durations, then multiplied by 60 to yield IWPM.

Statistical Analyses

The first research question addressed how well each measure (overall intelligibility, speech rate for sentences, and IWPM for sentences) discriminated among groups of children at 5 years of age. Our analysis focused on three key comparisons for each of the measures: SMI versus TD, SMI versus NSMI, and NSMI versus TD. For each comparison, we estimated sensitivity and specificity for varying thresholds of the three speech measures and plotted these quantities in receiver operating characteristic (ROC) curves, along with sensitivity and specificity confidence intervals at various thresholds (Robin et al., 2011). ROC analyses were complemented with estimation and confidence intervals for the area under the ROC curve (AUC) for each comparison (DeLong, DeLong, & Clarke-Pearson, 1988).

The second research question sought to determine the 10th percentile threshold of performance for the targeted measures for TD children and then to determine how well this threshold differentiated among children in the three groups for each measure. We were interested in the 10th percentile values because they would allow us to compare the lowest values for typical development and differentiate them from the highest levels of atypical development. Sensitivity and specificity values at the 10th percentile of TD thresholds were examined for pairwise group results (TD vs. SMI, SMI vs. NSMI, TD vs. NSMI). We interpreted sensitivity and specificity findings descriptively within the context of the literature.

Results

How Well Do Measures Discriminate Among Groups of Children?

The distributions of the three speech measures for each group are shown in Figure 1. ROC curves along with sensitivity and specificity confidence intervals at various thresholds for each of the comparisons of interest (SMI vs. TD, SMI vs. NSMI, TD vs. NSMI) for overall intelligibility, speech rate, and IWPM are shown in Figures 2–4. The AUC, along with lower and upper confidence limits, for each comparison of interest and each measure of interest is provided in Table 3.

Results for overall intelligibility show that the AUC was very high for all comparisons (see Figure 2), suggesting that overall intelligibility differentiated between the three groups of children with a very high degree of accuracy. AUC was highest for children with SMI versus TD children (.99), followed by children with SMI versus children with NSMI (.93), and finally children with NSMI versus TD children (.88).

Results for speech rate yielded the same pattern of results as those for overall intelligibility (see Figure 3). However, generally, the AUC was smaller for all comparisons than it was for overall intelligibility. The AUC was highest for children with SMI versus TD children (.93), followed by children with SMI versus children with NSMI (.81), and finally children with NSMI versus TD children (.68).

Finally, results for IWPM again yielded the same pattern as those for overall intelligibility and speech rate (see Figure 4). AUCs were generally lower than those observed for overall intelligibility, but higher than those observed for speech rate. AUC was highest for children with SMI versus TD children (.99), followed by children with SMI versus children with NSMI (.91), and finally children with NSMI versus TD children (.81). For each of the three measures, the higher AUCs for children with SMI versus TD children and children with SMI versus children with NSMI suggest that children with SMI are readily differentiated from the other groups.

Figure 1. Plots of individual data for children with typical development (TD), no speech motor impairment (NSMI), and speech motor impairment (SMI) for overall intelligibility, words per minute, and intelligible words per minute. Note that dashed lines represent 10th percentile values for typical children on each measure. Outliers are identified with triangles. There were a total of three children who had outlier values on one or more measures for a total of six outlier values across children. One child was in the NSMI group and was a low outlier for all three measures; the second child was in the NSMI group and was a high outlier for words per minute and intelligible words per minute; the third child was TD and was a low outlier for intelligibility. IQR = interquartile range.

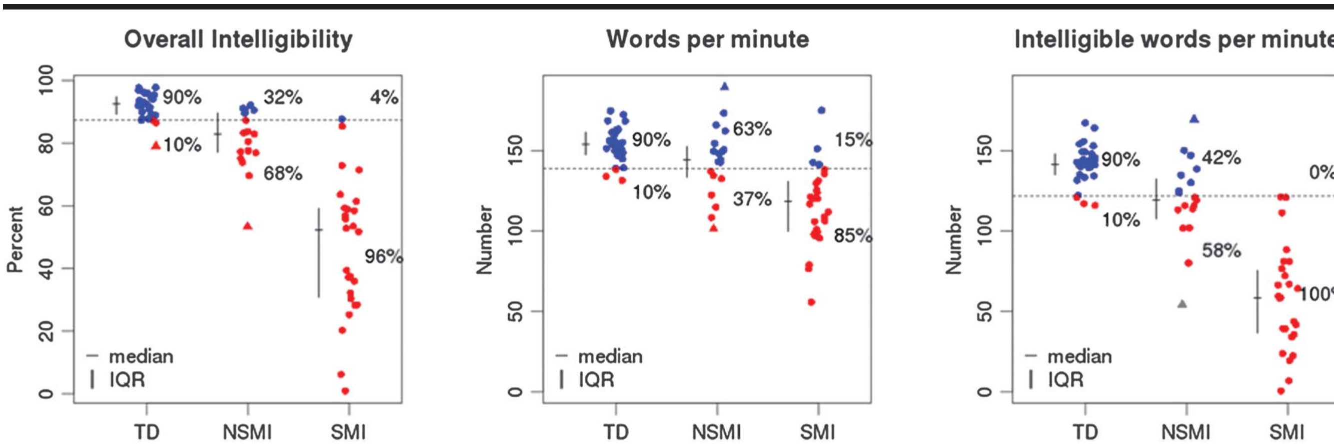
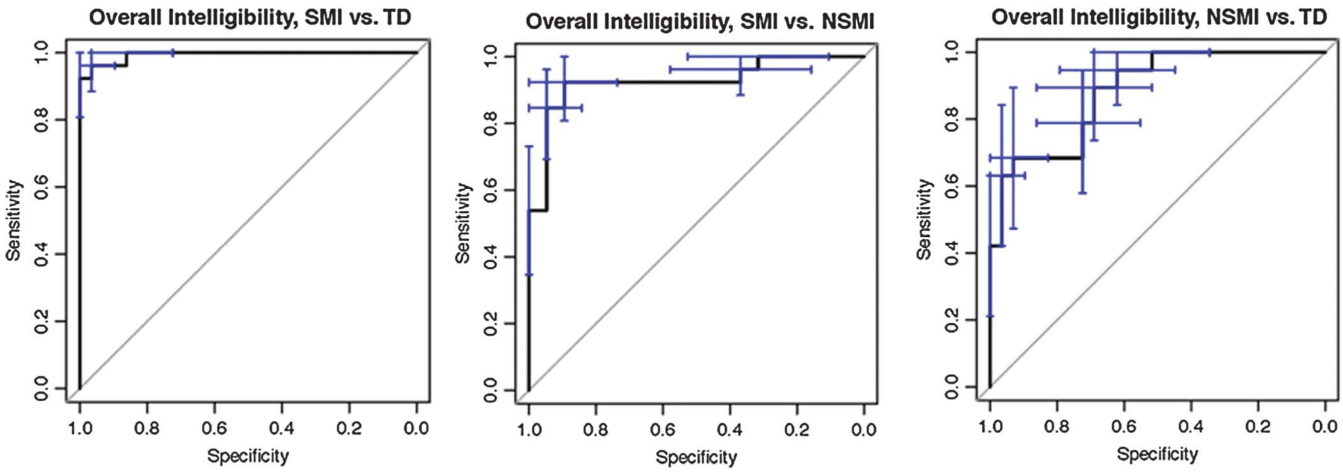


Figure 2. Receiver operating characteristic curves for overall intelligibility by contrast group. Horizontal bars represent confidence intervals for specificity; vertical bars represent confidence intervals for sensitivity. SMI = children with speech motor impairment; TD = typically developing children; NSMI = children with no speech motor impairment.



How Well Do 10th Percentile Benchmarks for TD Children Separate Among Groups of Children?

Examination of Figure 1 suggests that 90% of TD children had overall intelligibility at or above 87%, thus establishing 87% as the 10th percentile threshold for TD children for the purposes of this study. For children with NSMI, 32% had intelligibility at or above the 87% TD threshold, suggesting that most children with NSMI fell below the 10th percentile TD performance range. For children with SMI, 4% of children had intelligibility at or above the 87% TD threshold, suggesting that almost all children with SMI had intelligibility below the 10th percentile TD performance range.

Speech rate data, also shown in Figure 1, indicate that 90% of TD children had a speech rate of 139 wpm or faster, thus establishing 139 wpm as the 10th percentile threshold for TD children for this study. For children with NSMI, 63% of children had speech rate at or above the 139 wpm TD threshold, suggesting considerable overlap with TD speech rate findings. For children with SMI, 15% of children had speech rate at or above the 139 wpm TD threshold, suggesting that most children with SMI had speech rate below the 10th percentile TD threshold.

IWPM data in Figure 1 show that 90% of TD children had IWPM above 122, thus establishing 122 as the 10th percentile threshold for TD children for this study.

Figure 3. Receiver operating characteristic curves for speech rate by contrast group. Horizontal bars represent confidence intervals for specificity; vertical bars represent confidence intervals for sensitivity. SMI = children with speech motor impairment; TD = typically developing children; NSMI = children with no speech motor impairment.

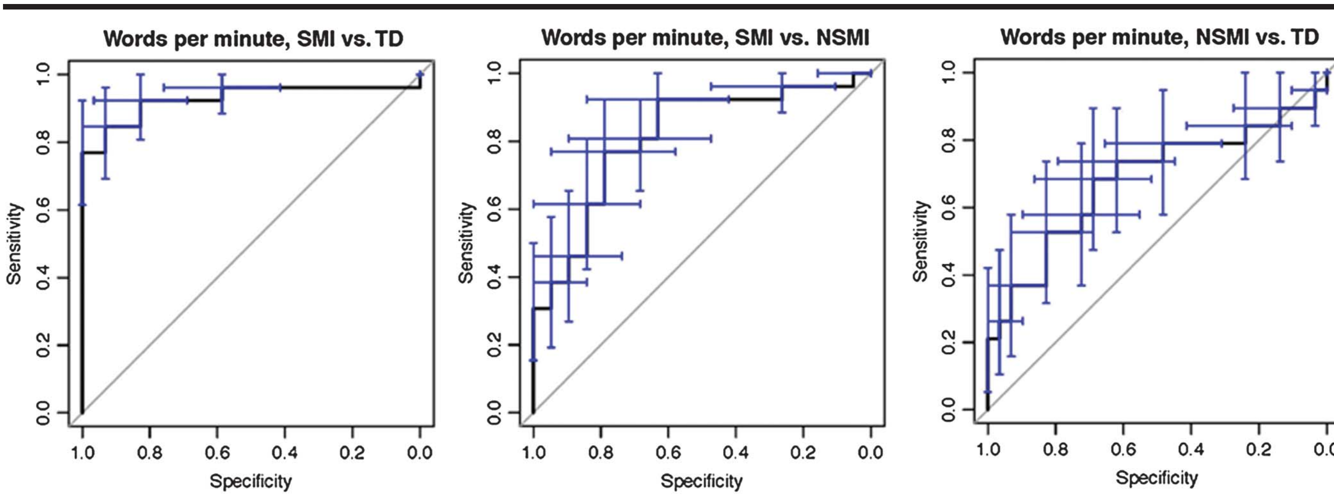
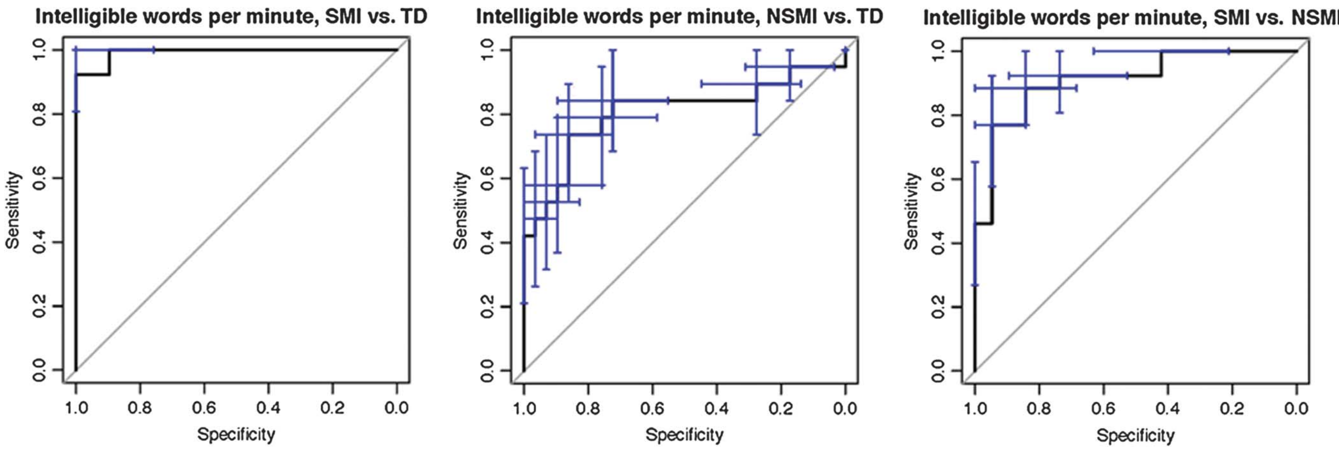


Figure 4. Receiver operating characteristic curves for intelligible words per minute by contrast group. Horizontal bars represent confidence intervals for specificity; vertical bars represent confidence intervals for sensitivity. SMI = children with speech motor impairment; TD = typically developing children; NSMI = children with no speech motor impairment.



For children with NSMI, 42% of children had IWPM at or above the 122 TD threshold, suggesting nearly half of children with NSMI had IWPM that overlapped with TD children. For children with SMI, 0% of children had IWPM at or above the 122 TD threshold, suggesting that all children with SMI had IWPM below the 10th percentile TD threshold.

Sensitivity and specificity values at 10th percentile TD thresholds for each of the three measures are shown in Table 4 for pairwise group results (TD vs. SMI, SMI vs. NSMI, TD vs. NSMI). Results show that 10th percentile TD values for all measures have very high sensitivity and specificity for differentiating between TD and SMI groups, suggesting that the thresholds of 87% overall intelligibility, 139 wpm and 122 iwpm, each do well at revealing both true positive cases (SMI) and true negative cases (TD). Results also showed that these same 10th percentile TD values for all measures had very high sensitivity (and lower specificity) for differentiating between TD and NSMI, suggesting that all measures do well at revealing true positive cases (children with NSMI) but do not do as well at revealing true negative cases (TD children). Finally, 10th percentile TD values for all measures had high specificity for differentiating between SMI and NSMI groups, suggesting that all measures identify true negative cases (children with NSMI) well but are not as strong with identifying true positives (children with NSMI).

Table 3. Area under the ROC curve (AUC) is estimated along with 95% confidence intervals (lower and upper limits).

Comparison	Speech measure	LCL	UCL	AUC
Overall intelligibility	SMI vs. TD	.98	1.00	.99
	SMI vs. NSMI	.84	1.00	.93
	NSMI vs. TD	.79	.98	.88
Speech rate	SMI vs. TD	.84	1.00	.93
	SMI vs. NSMI	.68	.94	.81
	NSMI vs. TD	.51	.85	.68
Intelligible words per minute	SMI vs. TD	.98	1.00	.99
	SMI vs. NSMI	.83	1.00	.91
	NSMI vs. TD	.66	.95	.81

Note. SMI = children with speech motor impairment; TD = typically developing children; NSMI = children with no speech motor impairment.

Discussion

In this study, there were two key findings. First, each group of children, those with SMI, NSMI, and TD, could be readily differentiated from the other groups using each of the three measures of functional speech ability (intelligibility, speech rate, IWPM). However, overall speech intelligibility provided the highest level of accuracy in differentiating among groups, and speech rate provided the lowest. Second, 10th percentile thresholds of typical development on all three measures showed high specificity in separating among SMI versus TD and SMI versus NSMI groups; conversely, however, all three measures showed high sensitivity in separating among NSMI and TD groups. Results are discussed below.

Differentiating Among Groups Using Functional Speech Measures

In this study, children with CP were separated into profile groups based on whether or not there was clinically identifiable evidence of SMI as determined by a research speech-language pathologist with extensive expertise in CP and in pediatric speech development. Children who did not have NSMI comprised one group, whereas children with clear evidence of clinical SMI were another group. Previous studies in our laboratory have suggested that children with SMI have unique characteristics that differentiate them from children with SMI—notably, their language growth (Hustad et al., 2018), their speech

Table 4. Sensitivity and specificity at 10th percentile thresholds for typically developing (TD) children.

	10th Percentile threshold for TD	Comparison	Sensitivity	Specificity
Overall intelligibility	87%	SMI vs. TD	.90	.96
		SMI vs. NSMI	.32	.96
		TD vs. NSMI	.90	.68
Speech rate	139 wpm	SMI vs. TD	.90	.85
		SMI vs. NSMI	.63	.85
		TD vs. NSMI	.90	.37
Intelligible words per minute	122 wpm	SMI vs. TD	.90	1.00
		SMI vs. NSMI	.42	1.00
		TD vs. NSMI	.90	.58

Note. Specificity for TD comparison is not exactly 90% due to discrete data. SMI = children with speech motor impairment; NSMI = children with no speech motor impairment.

rate (Hustad et al., 2010), and their intelligibility (Hustad et al., 2012). Overall, this study provides further validation for the presence of unique clinical profile groups among children with CP. Specifically, we found that each of the three groups of children (TD, SMI, NSMI) was clearly different from the others on all three functional speech measures.

We expected that intelligibility would be the best differentiator among groups, and this hypothesis was supported by our data. IWPM also provided strong differentiation among groups, which is not surprising given that it is dependent on both intelligibility and speech rate measures. Speech rate was the weakest differentiator, particularly for separating TD from children with NSMI. The finding regarding speech rate was somewhat surprising, given that previous studies have found speech rate to be a key differentiator among groups of children with CP (Hustad et al., 2010) and between typical children and children with dysarthria (Allison & Hustad, 2018). Three points are noteworthy here. First, although speech rate was the weakest differentiator of the variables examined, it was still strong in the absolute sense of differentiating among groups. Second, children were required to be able to produce utterances of at least four words in length to be included in this study. This may have biased the sample toward children with better speech motor coordination, which, in turn, may have led to less variability among children on speech rate. Finally, we considered speech rate inclusive of pauses in this study. If we had examined articulation rate exclusive of pauses, findings may have been different.

Results of this study were consistent with our earlier findings regarding children with NSMI, demonstrating that children with CP and NSMI were unique relative to typical children and children with CP who had SMI. However, the finding that the AUC was smallest for each measure in the contrasts examining TD children versus children with NSMI suggests that children with NSMI may be more similar to their TD peers than they are to their peers with SMI. This makes sense given that the clinical presentation of children in the NSMI groups is not suggestive of speech motor impairment. Examination of speech rate findings showed that 63% of children with NSMI had speech rates above the 10th percentile for TD children; thus, there was considerable overlap between NSMI and TD groups. Interestingly, there was far less overlap between TD children and children with NSMI for intelligibility data. One explanation for reduced intelligibility in the presence of relatively typical speech rate is that these children may have speech sound disorders that are more phonological than motor in nature. In addition, some children with NSMI may have had expressive language or memory delays that were not considered here but may contribute to intelligibility. Further study is necessary to examine these possibilities and to quantify segmental integrity relative to developmental expectations for these children.

Specificity and Sensitivity of 10th Percentile TD Benchmarks

When we deconstructed continuous data plotted in the ROC curves for each measure to examine how well only the 10th percentile thresholds of typical development differentiated between groups, results were consistent with overall findings using AUC results and supported our hypotheses that these thresholds would provide strong differentiation among all groups. Sensitivity and specificity results indicated that 10th percentile TD thresholds had excellent specificity and sensitivity for differentiating TD children from children with SMI. In essence, this means that each threshold value for each measure was very good at identifying both true positive cases (SMI) and true negative cases (TD) in 5-year-old children. In evaluating TD versus NSMI, 10th percentile thresholds for TD had very good sensitivity, but specificity was weaker. That is, the threshold value of each measure did a very good job of identifying true positive cases (in this context, the children with NSMI). Interestingly, when 10th percentile TD thresholds for differentiating between SMI versus NSMI were examined, specificity was very high, indicating that thresholds, again, did a good job of revealing the children with NSMI, who in this case were the true negatives. Overall, these findings suggest that the 10th percentile

TD values for intelligibility, speech rate, and IWPM have the potential to provide a powerful set of age-based benchmarks for identifying children who have CP with SMI versus NSMI and for differentiating them from TD children. Notably, findings highlight that children with NSMI are highly differentiable from children with SMI and from TD children and suggest that these children are not as typical in their speech development as they otherwise might appear.

Clinical Implications

Results of this study suggest that functional measures of speech performance can readily differentiate between typical children and children with CP, regardless of whether they have clinical speech motor impairment. Intelligibility differentiated among groups of children most readily and typical performance at the 10th percentile (with 90% of children falling above this threshold) provided a good threshold for 5-year-old children. In this study, the threshold for transcription intelligibility of TD children by naïve listeners was 87%. Although IWPM was also a strong differentiator among groups, it is unclear whether it added anything beyond the findings from intelligibility.

It is notable that 87% intelligibility as a benchmark for 5-year-old children to be considered "typical" is generally consistent with other recent, though small-scale, intelligibility studies using similar methodologies (Hodge & Gotzke, 2014). Larger scale studies have not been conducted.

Another implication of this study is that, as a group, even the children with CP who did not have clinical evidence of speech motor impairment (NSMI group) could be differentiated from typical children on measures of functional speech performance. That is, children with CP and NSMI did not look typical, particularly on their intelligibility and on their rate of IWPM. However, their speech rate was less readily differentiable from typical children. This finding has important clinical implications because children with NSMI have not been the target of assessment or intervention for speech. Results of this study suggest that it may be useful for all children with CP, at 5 years of age, to be evaluated for potential functional speech deficits, with an eye toward milder problems. There are currently a range of potential interventions available (see Pennington, Lombardo, Steen, & Miller, 2018; Pennington, Miller, Robson, & Steen, 2010; Pennington et al., 2013) that could improve speech performance and reduce the difference between these children and their TD counterparts.

Limitations and Future Directions

In this study, we examined only children who were 5 years of age. Age stratification is critical in the study of children because developmental change with chronological age adds variability and is a potential confound to any findings. Studies of both younger and older children are necessary to determine how well functional speech measures differentiate between typical and disordered children at different ages and to identify age-specific cut-points beginning at the onset of word production.

We focused on children who could produce connected speech consisting of four or more words in this study. Although our sample represented children across the full severity continuum in terms of intelligibility, it did exclude a subset of children with very severe speech motor control problems. Future studies should examine differences among groups of children on speech samples that vary in length (i.e., utterances under four words long). Studies should also examine how well single-word intelligibility data differentiates between groups of children.

Although we had a relatively large sample of 5-year-old children with CP (n = 45), our sample of TD children was smaller (n = 26). Given the considerable variation across TD children that has been documented in speech development and given the well-known heterogeneity among children with speech-language disorders, especially CP, studies including larger samples of children are necessary to ensure that a full range of performance levels is captured. This will help foster a more complete understanding of both the lower and upper quantiles of performance on measures that are functionally meaningful for clinical assessment of those at risk for speech problems.

This study did not examine relationships between segmental speech performance and intelligibility. An important additional direction for research is to examine how segmental integrity predicts or explains speech intelligibility. Such information could have important implications for speech interventions, particularly for children with CP and NSMI.

Acknowledgments

This study was funded by National Institute on Deafness and Other Communication Disorders Grants R01DC009411 and R01DC015653, awarded to the first author. Support was also provided by a core grant to the Waisman Center (U54 HD090256) from the National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors thank the children and their families who participated in this research and the graduate and undergraduate students at the University of Wisconsin–Madison who assisted with data collection and data reduction.

References

Allison, K. M. (2015). *Classification of pediatric dysarthria in cerebral palsy*. Madison: University of Wisconsin–Madison, Waisman Center, Language Analysis Laboratory.

Allison, K. M., & Hustad, K. C. (2018). Acoustic predictors of pediatric dysarthria in cerebral palsy. *Journal of Speech, Language, and Hearing Research, 61*(3), 462–478.

Bax, M., Goldstein, M., Rosenbaum, P., & Levinton, A. (2005). Proposed definition and classification of cerebral palsy. *Developmental Medicine & Child Neurology, 47*(8), 571–576.

Bax, M., Tydeman, C., & Flodmark, O. (2006). Clinical and MRI correlates of cerebral palsy: The European cerebral palsy study. *Journal of American Medical Association, 296*(13), 1602–1608.

Boersma, P., & Weenink, D. (2015). Praat: Doing phonetics by computer [Computer program]. Retrieved from https://www.praat.org

Christensen, D., Van **Naarden Braun, K., Doernberg, N. S., Maenner, M. J., Arneson, C. L., Durkin, M. S., . . . Yeargin-Allsopp, M.** (2014). Prevalence of cerebral palsy, co-occurring autism spectrum disorders, and motor functioning—Autism and developmental disabilities monitoring network, USA, 2008. *Developmental Medicine & Child Neurology, 56*(1), 59–65.

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics, 44*(3), 837–845.

Dykstra, A., Hakel, M., & Adams, S. (2007). Application of the ICF in reduced speech intelligibility in dysarthria. *Seminars in Speech and Language, 28,* 301–311.

Fudala, J. B. (2001). *Arizona Articulatory Proficiency Scale–Third Edition (AAPS-3)*. Los Angeles, CA: Western Psychological Services.

Hodge, M., & Daniels, J. (2007). *TOCS+ intelligibility measures*. Edmonton: University of Alberta.

Hodge, M., & Gotzke, C. L. (2014). Construct-related validity of the TOCS measures: Comparison of intelligibility and speaking rate scores in children with and without speech disorders. *Journal of Communication Disorders, 51,* 51–63.

Hustad, K. C., Allison, K., McFadd, E., & Riehle, K. (2014). Speech and language development in 2-year-old children with cerebral palsy. *Developmental Neurorehabilitation, 17*(3), 167–175.

Hustad, K. C., Allison, K. M., Sakash, A., McFadd, E., Broman, A. T., & Rathouz, P. J. (2017). Longitudinal development of communication in children with cerebral palsy between 24 and 53 months: Predicting speech outcomes. *Developmental Neurorehabilitation, 20*(6), 323–330.

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, and Hearing Research, 53,* 1496–1513.

Hustad, K. C., Oakes, A., & Allison, K. (2015). Variability and diagnostic accuracy of speech intelligibility scores in children. *Journal of Speech, Language, and Hearing Research, 58*(6), 1695–1707.

Hustad, K. C., Oakes, A., McFadd, E., & Allison, K. M. (2016). Alignment of classification paradigms for communication abilities in children with cerebral palsy. *Developmental Medicine & Child Neurology, 58*(6), 597–604.

Hustad, K. C., Sakash, A., Broman, A. T., & Rathouz, P. J. (2018). Longitudinal growth of receptive language in children with cerebral palsy between 18 months and 54 months of age. *Developmental Medicine & Child Neurology, 60,* 1156–1164.

Hustad, K. C., Schueler, B., Schultz, L., & DuHadway, C. (2012). Intelligibility of 4-year-old children with and without cerebral palsy. *Journal of Speech, Language, and Hearing Research, 55*(4), 1177–1189.

Hustad, K. C., & Weismer, G. (2007). A continuum of interventions for individuals with dysarthria: Compensatory and rehabilitative approaches. In G. Weismer (Ed.), *Motor speech disorders* (pp. 261–303). San Diego, CA: Plural.

Kim, H., Martin, K., Hasegawa-Johnson, M., & Perlman, A. (2010). Frequency of consonant articulation errors in dysarthric speech. *Clinical Linguistics & Phonetics, 24*(10), 759–770.

Lee, J., Hustad, K. C., & Weismer, G. (2014). Predicting speech intelligibility with a multiple speech subsystems approach in children with cerebral palsy. *Journal of Speech, Language, and Hearing Research, 57*(5), 1666–1678.

Nordberg, A., Miniscalco, C., & Lohmander, A. (2014). Consonant production and overall speech characteristics in school-aged children with cerebral palsy and speech impairment. *International Journal of Speech-Language Pathology, 16*(4), 386–395.

Nordberg, A., Miniscalco, C., Lohmander, A., & Himmelmann, K. (2013). Speech problems affect more than one in two children with cerebral palsy: Swedish population-based study. *Acta Paediatrica, 102*(2), 161–166.

Palisano, R., Rosenbaum, P., Walter, S., Russell, D., Wood, E., & Galuppi, B. (1997). Development of the Gross Motor Function Classification System. *Developmental Medicine & Child Neurology, 39,* 214–223.

Pennington, L., Lombardo, E., Steen, N., & Miller, N. (2018). Acoustic changes in the speech of children with cerebral palsy following an intensive program of dysarthria therapy. *International Journal of Language & Communication Disorders, 53*(1), 182–195.

Pennington, L., Miller, N., Robson, S., & Steen, N. (2010). Intensive speech and language therapy for older children with cerebral palsy: A systems approach. *Developmental Medicine & Child Neurology, 52*(4), 337–344.

Pennington, L., Roelant, E., Thompson, V., Robson, S., Steen, N., & Miller, N. (2013). Intensive dysarthria therapy for younger children with cerebral palsy. *Developmental Medicine & Child Neurology, 55*(5), 464–471.

Platt, L. J., Andrews, G., & Howie, P. (1980). Dysarthria of adult cerebral palsy: II. *Phonemic analysis of articulation errors. Journal of Speech and Hearing Disorders, 23,* 41–55.

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J. C., & Müller, M. (2011). pROC: An open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics, 12,* 77.

Sakash, A., Broman, A. T., Rathouz, P. J., & Hustad, K. C. (2018). Executive function in school-aged children with cerebral palsy: Relationship with speech and language. *Research in Developmental Disabilities, 78,* 136–144.

Smit, A. B., Hand, L., Freilinger, J. J., Bernthal, J. E., & Bird, A. (1990). The Iowa articulation norms project and its Nebraska replication. *Journal of Speech and Hearing Disorders, 55,* 779–798.

Smith, A., & Zelaznik, H. N. (2004). Development of functional synergies for speech motor coordination in childhood and adolescence. *Developmental Psychobiology, 45*(1), 22–33.

Vick, J. C., Campbell, T. F., Shriberg, L. D., Green, J. R., Abdi, H., Rusiewicz, H. L., . . . Moore, C. A. (2012). Distinct developmental profiles in typical speech acquisition. *Journal of Neurophysiology, 107*(10), 2885–2900.

Workinger, M., & Kent, R. (1991). Perceptual analysis of the dysarthria in children with athetoid and spastic cerebral palsy. In C. Moore, K. Yorkston, & D. Beukelman (Eds.), *Dysarthria and apraxia of speech: Perspectives on management* (pp. 109–126). Baltimore, MD: Brookes.

Yorkston, K., & Beukelman, D. (1981). Communication efficiency of dysarthirc speakers as measured by sentence intelligibility and speaking rate. *Journal of Speech and Hearing Disorders, 46,* 296–301.

Zimmerman, I., Steiner, V., & Pond, R. (2012). *Preschool Language Scale–Fifth Edition Screening Test*. San Antonio, TX: The Psychological Corporation.