Intelligibility of 4 year old children with and without cerebral palsy

Katherine C. Hustad, Ph.D.[*,+], **Brynn Schueler**[*,+], **Laurel Schultz**[*,+], and **Caitlin DuHadway**[*,+]

[*]Department of Communicative Disorders, University of Wisconsin – Madison

[+]Waisman Center, University of Wisconsin – Madison

Abstract

Purpose—We examined speech intelligibility in typically developing (TD) children and three groups of children with cerebral palsy (CP) who were classified into speech / language profile groups following Hustad et al. (2010). Questions addressed differences in transcription intelligibility scores among groups, the effects of utterance length on intelligibility, the relationship between ordinal ratings of intelligibility and orthographic transcription intelligibility scores, and the difference between parent and naïve listener ordinal ratings.

Method—Speech samples varying in length from 1–7 words were elicited from 23 children with CP (mean age 54.3 months) and 20 typically developing children (mean age 55.1 months). 215 naïve listeners made orthographic transcriptions and ordinal ratings of intelligibility. Parent ordinal ratings of intelligibility were obtained from a previous study (Hustad et al., 2010).

Results—Intelligibility varied with speech / language profile group and utterance length, with different patterns observed by profile group. Ratings of intelligibility by parents and naïve listeners did not differ and were both highly correlated with transcription intelligibility scores.

Conclusions—Intelligibility was reduced for all groups of children with CP relative to TD children, suggesting the importance of speech-language intervention and the need for research investigating variables associated with changes in intelligibility in children.

Keywords

cerebral palsy; dysarthria; speech development; speech intelligibility; typical speech

---

The fundamental purpose of spoken communication is to be understood (Kent, Miolo, & Bloedel, 1994). Speech is the primary means through which most humans express themselves. However, speech must be sufficiently intelligible in order to enable successful communication. Intelligibility has been defined as the extent to which an acoustic signal, generated by a speaker, can be correctly recovered by a listener (Kent et al., 1989; Yorkston & Beukelman, 1980). Intelligibility is a dyadic construct and its measurement reflects the joint efforts of the speaker (who produces the signal) and the listener (who interprets the signal) (Lindblom, 1990). To be "intelligible", speech does not need to be perfect or even "normal". In fact, productions may be characterized by a variety of errors involving one or more speech subsystems (perhaps most classically articulatory omissions, substitutions, and distortions) and still be intelligible. The key issue in intelligibility is whether listeners are able to map the acoustic signal onto the intended *lexical* units in spite of segmental- or suprasegmental- level problems. Many variables have been shown to influence

---

Contact author: Katherine C. Hustad, 1500 Highland Ave., Madison, WI 53705, kchustad@wisc.edu, 608 265-9977 (voice).

intelligibility, for example length of the message (K.C. Hustad, 2007b; Miller, Heise, & Lichten, 1951; Yorkston & Beukelman, 1981), and semantic predictability of the message (Boothroyd & Nittrouer, 1988; Garcia & Cannito, 1996; K.C. Hustad, 2007a). Intelligibility is a multidimensional construct that is dynamic in nature. For children and adults with dysarthria, reductions in intelligibility are pervasive, and can lead to important limitations in functional communication. Enhancing intelligibility is often a primary goal of communication intervention (Ansel & Kent, 1992).

Speech intelligibility and speech motor development in children

In children, acquisition of intelligible speech is a developmental process, beginning early in the first year of life with vocal play, babbling, and word approximations, and continuing through childhood, culminating in the ability to produce fully intelligible, adult-like connected speech. At the same time, children are also acquiring language abilities and there are complex interactions between speech and language development that are not well understood, but are of considerable theoretical and clinical interest (Nip, Green, & Marx, 2011; Smith, 2006; Smith & Goffman, 2004). The particular age at which typically developing children become fully intelligible to everyday listeners and the range of acceptable variability among children at different ages have not been unequivocally established. Among the few extant studies examining intelligibility in typical children, several different methodologies have been employed which make comparison among studies and subsequent interpretation difficult. Existing studies examining children have primarily employed listeners who are "experts" (commonly speech-language pathologists, phoneticians, or graduate students in speech-language pathology). Research has demonstrated that there are important differences between what experienced listeners understand, and what naïve listeners understand (Baudonck, Buekers, Gillebert, & VanLierde, 2009; K.C. Hustad, Kent, & Beukelman, 1998). It is therefore possible that intelligibility data generated by experts may not represent the everyday listener that a child might encounter. Information regarding age-level intelligibility expectations based on listeners who are not experienced with phonetic analyses is necessary to provide a broader-based perspective on the functionality of children's speech at different ages.

Studies of children have examined intelligibility at different linguistic levels (i.e. single words, elicited sentences, spontaneous speech). Findings for four year old children and expert listeners have revealed scores that vary by as much as 20% (ranging from 64% to 95% intelligibility), depending on the study. Importantly, participant inclusion criteria have also varied, so it is not surprising that one study in which children who represented a range of speech sound production abilities (including some with phonological disorders) showed mean intelligibility of connected speech produced by four year old children to be 76% (Gordon-Brannan & Hodson, 2000), while two other studies showed intelligibility to be well above 90% (Austin & Shriberg, 1997; Rice et al., 2010). Similarly, intelligibility of elicited sentences produced by four year old children was found to be 74% in one study (again reflecting children with a range of speech sound production abilities) (Gordon-Brannan & Hodson, 2000) and 95% in another study (Chin, Tsai, & Gao, 2003). Such variability has also been noted among studies of intelligibility of single word productions, with one study finding intelligibility of four year old children to be 64% (Gordon-Brannan & Hodson, 2000), another finding intelligibility to be 76% (S. R. Morris, Wilcox, & Schooling, 1995). It is unknown whether differences between studies reflect normal variability among children, methodological differences related to speech stimuli, the presence/absence of speech sound delay or disorder, or other such variables.

As with other literature examining the influence of linguistic level (words, sentences, conversation / narrative) on intelligibility [see (Miller, et al., 1951; O'Neill, 1957; Sitler,

Schiavetti, & Metz, 1983)], literature examining intelligibility of four-year old children seems to suggest that linguistic context (via sentences or spontaneous conversation / narrative) generally leads to better intelligibility than single word productions. Specific findings regarding the extent to which length and / or complexity of utterances may influence intelligibility are limited to general designations of single words, sentences, and conversations. However, it has been hypothesized that the reason for intelligibility differences at various linguistic levels relates to listeners' ability to capitalize on their linguistic knowledge and thus utilize bootstrapping and inferencing abilities to aid in the decoding of the acoustic signal (K.C. Hustad & Beukelman, 2001).

In the speech motor literature, a growing body of research has focused on understanding interactions between speech motor control and length and complexity of utterance in typically developing children, adults, and in individuals with various disorders (Kleinow & Smith, 2000, 2006; Maner, Smith, & Grayson, 2000; Sadagopan & Smith, 2008). Such studies have provided evidence of increased variability in speech motor control for longer utterances produced by typically developing children (Maner, et al., 2000; Sadagopan & Smith, 2008), increased phonological errors for sentences that were more syntactically complex (Prelock & Panagos, 1989) and increased dysfluencies for sentences that were more syntactically complex (Bernstein Ratner & Sih, 1987). Studies have not examined the impact of length and complexity of utterances on speech intelligibility. Such research would be useful toward establishing age-specific and length/complexity-specific intelligibility benchmarks which could inform assessment and treatment of children with motor speech disorders.

Children with cerebral palsy

Children with cerebral palsy (CP) face a special set of challenges during the process of speech development. First, like all children, they face the challenge of development itself and the associated gradual emergence of intelligible speech. Second, they face the impact of a neuromotor disorder, which, for many children, may bring with it intelligibility reductions above and beyond those associated with typical development. Third, they must contend with the poorly understood, but likely pervasive, interactions between development and a superimposed neuromotor disorder that may affect cognitive and language development as well as speech development. The process of acquiring speech is clearly more complex and likely more protracted for children with CP than for children who are typically developing.

Although considerable efforts have been directed toward understanding gross motor development and gross motor function in children with CP, far less attention has been paid to speech, language and communication abilities in children with CP. Recently, we proposed a speech and language classification system for children with CP (K. C. Hustad, Gorton, & Lee, 2010) based on the presence of speech-motor involvement and language / cognitive involvement. Our model describes eight different speech and language profile groups, and we presented preliminary data on four-year-old children with CP which validated a four-group collapsed version of the model. The collapsed speech and language profile groups among children with CP were as follows: 1.) *Children with no speech motor impairment* (NSMI) based on clinical assessment; language abilities were age-appropriate (typically developing) based on clinical assessment (24% of the sample). 2.) *Children with speech motor impairment, and language comprehension within typical expectations* (SMI-LCT) based on clinical assessment (26% of the sample). 3.) *Children with speech motor impairment, and language comprehension impairment* (SMI-LCI) based on clinical assessment (18% of the sample). 4.) *Children with anarthria* (ANAR) were unable to produce functional speech; language abilities were age appropriate (typically developing), impaired, or unknown based on clinical assessment (32% of the sample).

A key finding from Hustad and colleagues (2010) was that the discriminant function that best differentiated among children in the different groups consisted of two variables: speech rate and vowel space. Speech intelligibility was also included as a variable but it did not make a substantial contribution to differentiation among groups. However, intelligibility was measured using parent estimates, indicated by ratings on a 7-point ordinal scale of how intelligible parents thought their child's speech was to unfamiliar communication partners. Parent report of speech intelligibility has received very limited research attention, but one study examining parent estimates of their child's intelligibility to unfamiliar partners showed findings that were consistent with speech-language pathologists' assessment results (Coplan & Gleason, 1988). In our previous study, parent ratings of intelligibility differed among groups of children with CP in the expected directions; however, the extent to which parents may have had a biased view of their child's performance is unclear.

In the present study, we sought to examine intelligibility as measured by orthographic transcription via naïve listeners. We used speech samples from the same children as those reported in Hustad et al., 2010, along with a control group of typically developing, age-matched 4-year old children. Our specific questions were as follows: 1.) Are there overall differences in transcription intelligibility among groups of children? 2.) Does intelligibility change across utterances of different lengths within groups of children? 3.) What is the relationship between ordinal ratings of intelligibility for children with CP and orthographic transcription-based intelligibility scores from naïve listeners? 4.) Are there differences between parents' ratings of intelligibility and unfamiliar listeners' ratings of intelligibility?

Method

Participants

Children with cerebral palsy—Twenty-three children with CP participated as speakers. All children were participants in a longitudinal study on communication development in children with CP. Children included in the present study are a subset of those reported on in Hustad et al., (2010). Children in group ANAR from Hustad et al. (2010) were excluded from the present study as children in group ANAR, by definition, were unable to produce speech. Inclusion criteria for the larger study required that children: 1.) have a medical diagnosis of CP; and 2.) have hearing abilities within normal limits as documented by either formal audiological evaluation or distortion product otoacoustic emission screening.

As described in Hustad et al. (2010), 8 of the children had no evidence of speech motor impairment (Group NSMI); 9 of the children had evidence of speech motor impairment with language comprehension skills that reflected typical developmental expectations (SMI-LCT); and 6 of the children had evidence of speech motor impairment with language comprehension skills that reflected language impairment (SMI-LCI). The mean age across the children was 54.3 months (SD 1.6). The sample comprised 10 boys (mean age 54.2 months (SD 1.7)) and 13 girls (mean age 54.5months (SD 1.4)). Table 1 present demographic characteristics of children in each profile group, including medical diagnoses and Gross Motor Function Classification System (GMFCS)(Palisano et al., 1997) rating, which is a 5-level index of functional self-initiated gross motor abilities (with an emphasis on sitting and walking) designed specifically for children with CP. The GMFCS is a standard measure of gross motor function throughout North America and Europe.

Typically developing children—Twenty typically developing 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: 1.) have typically developing speech; 2.) have typically developing language; 3.) have no history of developmental delay per parent report; and 4.) have hearing

abilities within normal limits as documented by either formal audiological 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-3 (AAPS) (Fudala, 2001). Per the AAPS examiner's manual, children were required to achieve standard scores at or above 86 to be considered "within normal limits" and thus be included in the present study. Language was screened using the Preschool Language Scale-4 Screening Test (PLS-4 Screener); children were required to meet "pass" criteria for all sections of the PLS-4 Screener to be included in the present study. The mean age across children who were included in the study was 55.1 months (SD 3.2). The sample comprised 8 boys (mean age 55.0 months (SD 2.6)) and 12 girls (mean age 55.1 months (SD 3.6)).

Non-disabled adult listeners—Two-hundred fifteen non-disabled individuals participated as listeners. Listeners were recruited from the university community via public postings and were primarily undergraduate students. Listeners were compensated monetarily 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. Participants were 78 males and 137 females. The mean age of listeners was 21.6 years (SD = 3.3).

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 (TOCS+) (Hodge & Daniels, 2007). The TOCS+ 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. This is a standard approach to intelligibility measurement in motor speech disorders.

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 70 different multiword utterances. Of the multi-word utterances, 10 were 2-words in length, 10 were 3-words in length, 10 were 4-words in length, 10 were 5-words in length, 10 were 6-words in length, and 10 were 7-words in length. Morphological complexity of TOCS sentences varied within the 10 sentence groups; however, linguistic characteristics of all utterances fell below a 48 month developmental level, and thus were considered to be within the language ability profile for all children except for those in group SMI-LCI, who, by definition, had language impairment. Table 2 provides summary data regarding the linguistic features of the stimulus sentences, including the expected age range for production by utterance 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 on-line 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. All typically developing children were able to produce all stimuli; however, due to speech motor control issues, some of the children with CP who had dysarthria were unable to produce longer utterances.

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 inches 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 to 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 multi-word stimulus utterances. The order of presentation for the single-word listening task and the multi-word 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 multi-word utterance stimuli, utterances of varying length were randomly intermingled to prevent a potential learning effect. Listeners made ordinal ratings regarding their perception of what they heard after completing the listening tasks.

During the experiment, listeners were seated individually in a sound attenuating suite in front of a 19 inch 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 intelligibility

Orthographic transcription of intelligibility—Orthographic transcriptions of children were scored using our in-house computer program. The program automatically tallied the number of transcribed words that were an exact phonemic match to the stimulus words in the sentences produced by the children. Mis-spellings 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 individual listener of each child was summed and divided by the number of words possible for each stimulus length condition. In situations where children did not produce utterances of longer lengths, and therefore intelligibility data were not obtained for particular utterances, missing data were omitted casewise. The resulting intelligibility scores for each listener of each child were used for intelligibility analyses.

Ordinal ratings of intelligibility—After completing orthographic transcription tasks, listeners were asked to rate how well they thought other adults would be able to understand the child. Ratings were made on a 7-point equal appearing scale, where 1 = difficult or impossible to understand, and 7 = very easy to understand. To enable comparison with our previous study (see Hustad et al., 2010) in which we asked *parents* to rate how well they thought other adults understood their child, we averaged ratings across the 5 listeners for each child so that we would have only one rating per child. We also utilized parent ratings of intelligibility reported in our previous study to compare differences between parent and listener ratings and to examine the relationship between ratings and intelligibility scores obtained through orthographic transcription.

Experimental design and statistical procedures

A 4 x 7 repeated measures nested design was employed for this study. The within subjects variable was stimulus length and its levels were utterances of seven different lengths. The between subjects variable was group, with children separated into typically developing (TD), NSMI, SMI-LCT, and SMI-LCI.

Research questions of interest focused on 1.) overall intelligibility differences between groups of children; 2.) differences across stimulus lengths within each of the groups of children; 3.) relationships between ordinal ratings of intelligibility and intelligibility scores obtained via orthographic transcription; and 4.) differences between parent ordinal ratings and naïve listener ordinal ratings. All questions were addressed using non-parametric analysis procedures because statistical assumptions were not met for data from children with CP. To address the first question, the Kruskal-Wallis test was performed on intelligibility data averaged across utterance lengths to determine if there were group differences. Pairwise follow-up contrasts were performed using the Mann-Whitney U to determine which groups differed. To control the type I error rate, an alpha level of .01 was assigned to this set of analyses and was partitioned evenly among the 7 tests (1 omnibus test; 6 pairwise comparisons). To be considered significant, a probability value less than or equal to .0014 was necessary.

To address the second research question, the Friedman test was performed within each group of children to determine if there were overall differences in intelligibility across utterances of different lengths. To control the type I error rate, an alpha level of .01 was partitioned among the four tests. To be considered significant, a probability value less than or equal to .0025 was necessary. Trends were examined descriptively.

For the third question, three correlations, assigned a family-wise alpha of .01, were examined using Spearman's rho correlation coefficients. To be considered significant, a probability value less than or equal to .003 was necessary.

For the fourth question, differences between parent ordinal ratings and naïve listener ordinal ratings of intelligibility were examined across all children with CP using the Wilcoxon Signed Rank test. An alpha of .01 was allotted to this analysis.

Results

Intelligibility differences among groups

Descriptive data showing overall group differences are presented in Figure 1. Descriptive results suggest that there were differences between all four groups of children, with TD children having higher intelligibility scores than all groups of children with CP. Kruskal Wallis ANOVA results revealed that the overall difference among groups was significant ($TS_3 = 161.13.94$; p<.001). Follow-up Mann-Whitney U contrasts are presented in Table 3. Results indicated that all between group differences were significant except for the difference between children in group SMI-LCT and SMI-LCI.

Intelligibility changes across utterance lengths within groups

Descriptive data showing intelligibility trends across utterance lengths for the different groups are shown in Figure 2. Descriptive results suggest that there may be two patterns of findings for the effects of utterance length on intelligibility. The first pattern, observed for children in the TD and NSMI groups suggests that intelligibility peaks at 4-word utterances and then decreasing after that point. The second pattern, observed for children in group SMI-LCT and SMI-LCI, suggests that intelligibility fluctuates, perhaps decreasing with longer utterances. Figure 3 shows data for individual children comprising the different groups. The variability in intelligibility scores among children in all groups is noteworthy, but particularly within the dysarthria groups.

Nonparametric inferential statistics, presented in Table 4, showed that there were significant differences in intelligibility across utterance lengths for children in the TD, NSMI, and SMI-LCT groups. However, there was not a significant difference across utterances lengths for children in the SMI-LCI group.

Relationship between ordinal ratings and transcription intelligibility scores

Scatterplots showing the relationship between transcription intelligibility scores of naïve listeners, and ordinal ratings are shown in Figures 4–6. In each of these figures, a positive linear relationship is shown between ordinal ratings of intelligibility and transcription intelligibility scores. Spearman's rho correlation coefficients examining the relationship between naïve listener intelligibility scores and naïve listener ratings of how well they thought other adults would be able to understand each child revealed a correlation of rho = .83 (p = .001) for children in the TD group (See Figure 4); for children in the CP group, rho = .79 (p<.001) (See Figure 5). Similarly, the correlation between parent ratings of how well they thought other adults would be able to understand their child and naïve listener intelligibility scores revealed rho = .78 (p<.001) for children with CP (See Figure 6). Note, however, that there were three children with intelligibility scores less than 20%, but parent ratings of 4 or greater.

Differences between ordinal ratings of intelligibility made by parents and naïve listeners for children with CP

For children with CP, the difference between parent ordinal ratings (mean = 4.28; SD = 1.89) and mean listener ordinal ratings of intelligibility (mean = 3.44; SD = 1.29) was .84 scale points (on a 7-point ordinal scale). Using the Wilcoxon Signed Rank test, this difference was not statistically significant ($Z = -2.34$; $p = .02$), at the pre-specified alpha level of .01. Therefore differences between parent and listeners ratings within the three groups of children with CP were not explored.

Discussion

There were four key findings from the present study. First, overall intelligibility differed among most groups of children based on profile group membership. Second, there seemed to be two descriptive patterns for the effects of utterance length on intelligibility, with children who did not have dysarthria showing intelligibility increasing to a point and then decreasing after that point, and children who had dysarthria showing a general fluctuating or decreasing pattern of intelligibility as utterance length increased. The third finding was that ratings of intelligibility by both parents and naïve listeners were highly correlated with transcription intelligibility scores, suggesting that both groups of listeners make reasonable subjective estimates of intelligibility. Finally, mean ordinal intelligibility ratings for parents and naïve listener did not differ, lending further validity to use of parent ratings of intelligibility. Each of these findings is discussed in detail, below.

Intelligibility differences among groups and impact of utterance length on intelligibility

Children in the TD group had higher mean intelligibility scores than children in each of the three CP groups. Further, mean intelligibility varied with length of utterance for children in the TD group. Consistent with other studies, mean intelligibility of single words for children in the TD group was lower than for sentences of any length. Single word intelligibility was approximately 80%, a finding that was slightly higher than reported in previous studies (Gordon-Brannan & Hodson, 2000; S. R. Morris, et al., 1995). Intelligibility scores peaked at 92% for four-word utterances for children in the TD group, and then decreased to 85% for longer utterances. Rice and colleagues (2010) demonstrated that the mean length of utterance in words (MLU-W) for children between the ages of four and five years was slightly over 4, and mean length of utterance in morphemes (MLU-M) was slightly over 4.5. Rice and colleagues reported intelligibility to be 92%, and our findings were consistent with this. Although data from Rice and colleagues reflect intelligibility of spontaneous language samples, presumably children were using language at or around their MLU-W, and MLU-M. The finding that intelligibility of typical children in the present study peaked within the range of their expected MLU (-M and –W) suggests that there may be important synergies between speech and language performance for typical children. In the context of this interpretation, it is worth noting that the language features of the target utterances produced by children were well within developmental expectations based on chronological age (see Table 2). Thus, one hypothesis regarding this finding may be that longer utterances (beyond the child's habitual speaking length) tax speech motor performance in ways that result in reduced intelligibility. Although previous studies documenting a detrimental effect of length and complexity of utterance on speech performance in children have not examined intelligibility (Bernstein Ratner & Sih, 1987; Maner, et al., 2000; Prelock & Panagos, 1989), reduced intelligibility might be one such effect.

The finding that children in the NSMI group had lower intelligibility scores than children in the TD group was surprising given that these children had no clinical evidence of speech or language impairment. A key factor associated with this intelligibility difference appeared to

be length of utterance. Examination of Figure 2 reveals that children in the TD and NSMI groups had very similar mean intelligibility scores for 1- and 2-word utterances, but that intelligibility scores differed by about 10% for utterances of longer lengths. Although children in the NSMI group showed the same general pattern of results as TD children, their peak intelligibility scores occurred for 2-, 3-, and 4- word utterances, reaching approximately 82%, and then decreasing to about 72% for longer utterances. One explanation for the gap of 9–13% between findings for TD children and children in the NSMI group for utterances of 3-words and longer is that although speech may appear to be within age-appropriate limits, these children may have had subtle speech production difficulties that only emerged when the motor system was sufficiently taxed (as may be the case with longer utterances). Again, it is important to highlight that language assessment for children in the NSMI group encompassed only characterization of receptive skills, and all children were within an age appropriate range. Although we assume that the language features of the stimulus sentences used in this study were within the language capability profiles of the children in group NSMI, it is possible that some of these children may have had expressive language deficits, independent of their receptive skills, which impacted their ability to produce the target utterances. Examination of Figure 3 suggests that there was greater variability within and across children in the NSMI group than observed in the TD group, particularly as length of utterance increased, again perhaps indicating a less stable speech production system in children with CP or potential expressive language deficits among particular children. Additional study involving use of physiologic and acoustic tools as well as measures of expressive language is necessary to fully understand the underlying bases for these intelligibility differences in longer utterances.

Children in group SMI-LCT had mean intelligibility scores that were approximately 20% higher than children in group SMI-LCI in the present study; however because of the considerable variability among children, this difference was not statistically significant. This finding is consistent with our previous study where we found that intelligibility based on parent report did not differ for these two groups of children (K. C. Hustad, et al., 2010). However, the rather sizeable descriptive difference in intelligibility scores between groups leads to a question regarding the potential relationships between intelligibility, language impairment and severity of overall impairment. Specifically, it is possible that language impairment leads to reductions in intelligibility because of complex interactions between speech and language abilities. Conversely, it may be that children who have more severe speech motor involvement generally have more extensive neurological involvement and thus are more likely to experience co-occurring deficits such as language impairment. Several studies have suggested that children with more severe gross motor impairment tend to have a greater number of concomitant problems (Bax, Tydeman, & Flodmark, 2006), including cognitive involvement (Sigurdardottir et al., 2008). Examination of gross motor classification data (GMFCS) for children in groups SMI-LCT and SMI-LCI (Table 1) suggests that a range of GMFCS levels were represented in both groups of children. However, children in group SMI-LCT were not represented in GMFCS level V, the most severe level, and children in group SMI-LCI were not represented in in GMFCS level I, the mildest level. The relationship between speech motor involvement and gross motor function has not been systematically studied in children with CP, but there is evidence to suggest that gross motor function and manual ability are often discrepant and are poorly correlated in CP (Carnahan, Arner, & Hagglund, 2007). Further study is clearly necessary to begin to understand whether language problems have an independent detrimental effect on speech production in children with CP who have dysarthria, or whether children with more severe dysarthria (and therefore greater neurological involvement) are simply more likely to have associated language problems.

For the TD, NSMI, and SMI-LCT groups of children, intelligibility showed significant changes with utterance length. In particular, intelligibility seemed to be highest for 4-word utterances (and lower for longer and shorter utterances) for children in the TD and NSMI groups. Fluctuation was generally erratic for children in the SMI-LCT group. These observations are difficult to explain in the context of other intelligibility studies, including studies examining intelligibility of adults with dysarthria, which have demonstrated that intelligibility is usually higher for longer utterances (Yorkston & Beukelman, 1981). One key factor that differentiates the present study from other studies on intelligibility in dysarthria is that participants were young children with developing speech motor and language systems. As alluded to previously, one important and poorly understood question is how neurological involvement of the speech musculature (as in dysarthria), when superimposed on development and in some cases on other co-occurring impairments (i.e. language impairment), may impact speech production. The measures examined in the present study are gross in the sense that they provide little, if any, explanatory information regarding underlying bases for observed intelligibility deficits. However, the lack of benefit to intelligibility from linguistic context that has been well documented in the adult dysarthria literature, suggests that a developmental factor may be responsible. An obvious contender is speech motor control. Again, studies that seek to explain the underlying bases for intelligibility deficits are necessary to understand how the neurologically compromised developing speech motor system may be impacted by a host of communication-related variables.

A final observation regarding length of utterance and intelligibility was that children with dysarthria and language impairment (SMI-LCI) had more difficulty producing longer utterances than those with only dysarthria (SMI-LCT). In fact, none of the children with co-occurring language impairment were able to produce utterances longer than 4-words in length. Again, the question regarding whether language problems confound speech production abilities, or whether increased global neurological involvement results in co-occurring motor and language limitations arises. Nonetheless, data showing reduced intelligibility relative to age expectations for children with dysarthria in this study point to the potential benefit of augmentative and alternative communication (AAC) strategies to support speech at all utterance lengths.

Ordinal ratings of intelligibility and transcription intelligibility scores

There is an extensive body of literature documenting the validity of parent report for measures such as gross motor function (C. Morris, Kurinczuk, Fitzpatrick, & Rosenbaum, 2006) cognition (Johnson et al., 2004), and language (Fenson et al., 1993; Pan, Rowe, Spier, & Tamis-LeMonda, 2004). However, few studies have examined the validity of parent report for measuring intelligibility. The limited research that does exist provides support for the validity of parent ratings relative to clinical assessment findings (Coplan & Gleason, 1988). Several key findings in the present study provide further support for the use of parent ratings of intelligibility for children with CP. Specifically, in the present study, intelligibility ratings by parents were highly correlated with transcription intelligibility scores of naïve listeners and this relationship was almost identical to the correlation between ordinal ratings of intelligibility made by naïve listeners and transcription intelligibility scores from those same listeners. In both sets of analyses, however, there were situations where ordinal ratings were considerably higher than would be expected based on transcription intelligibility scores (see Figures 5 and 6). These situations tended to occur for children in the SMI-LCT group more so than the other groups. However, for most children, intelligibility ratings increased in a strong linear fashion with transcription intelligibility scores. Furthermore the magnitude of ratings made by parents and ratings made by naïve listeners did not differ, suggesting that both groups rate the intelligibility of children with CP similarly. Parent ratings of

intelligibility were not obtained for typically developing children in the present study. A parallel set of analyses focused on typically developing children would be useful to provide a more complete picture of the validity of parent ratings for speech intelligibility.

Clinical implications

This study revealed several clinically-relevant findings. Perhaps most importantly, children with CP who had dysarthria (groups SMI-LCT and SMI-LCI) had markedly reduced speech intelligibility relative to their typically developing peers. This was not surprising, given that dysarthria is almost always accompanied by reductions in speech intelligibility. However, the observed magnitude of the reductions in intelligibility in the present study suggests that most if not all children with dysarthria likely had significant problems meeting all of their communication needs with speech alone. Even dysarthric children with relatively higher speech intelligibility (e.g. 75%) could benefit from intervention to support their speech development and to enhance intelligibility. In addition to traditional motor speech interventions, AAC systems and strategies can be important tools to enhance the usefulness of natural speech and to enhance intelligibility, and therefore should be considered for young children with CP who have dysarthria to support speech, language, and communication development.

Children with CP and dysarthria who had the added burden of language impairment (group SMI-LCI) had similar speech intelligibility to those who had dysarthria without co-occurring language impairment, but generally had more expressive communication limitations. Our previous work has demonstrated that these children are often and disproportionately overlooked for AAC interventions (K.C. Hustad & Miles, 2010), perhaps because of the complexity of their communication challenges. Findings from the present study demonstrating the severe intelligibility problems that these children face provide further evidence for the importance of AAC to enhance both speech and language development and to advance social participation for these children.

Children who had CP and did not have clinical speech problems (Group NSMI), may in fact have subtle speech production deficits or expressive language deficits that resulted in reduced speech intelligibility for longer utterances. By definition, children in this group are not routinely identified for speech and language services unless they have other language or learning difficulties that are not associated with speech motor difficulties. Findings of the present study suggest that perhaps speech language pathologists should look more closely at these children. Given the mild nature of problems that they may experience, therapy may be very beneficial for enhancing intelligibility to age-appropriate levels.

Four year old children who were typically developing had intelligibility that varied with length of utterance. Findings of the present study were generally consistent with studies in which expert transcribers were employed and suggest that intelligibility of 4 year old children varies between about 80% and 92%, depending on the length of utterance and the individual child.

Finally, results of the present study suggest that parents were good judges of intelligibility and their ratings were generally consistent with the ratings of unfamiliar listeners. Although valid, parent ratings are subjective and should not be used as a substitute for quantitative and objective measures of intelligibility which have many important clinical uses. Instead, parent ratings can be used as a complement to traditional clinical measures to provide another source of information regarding the functionality of children's speech.

Acknowledgments

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]

Austin, D.; Shriberg, LD. Phonology Project Technical Report No. 3. Madison, WI: Waisman Center, University of Wisconsin; 1997. Lifespan reference data for ten measures of articulation completence using the speech disorders classificaiton system (SDCS).

Baudonck NLH, Buekers R, Gillebert S, VanLierde KM. Speech intelligibility of Flemish children as judged by their parents. Folia Phoniatrica et Logopaedica. 2009; 61:288–295. [PubMed: 19738387]

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.

Bernstein Ratner N, Sih C. Effects of gradual increases in sentence length and complexity on children's dysfluency. Journal of Speech and Hearing Disorders. 1987; 52:278–287. [PubMed: 3455450]

Boothroyd A, Nittrouer S. Mathematical treatment of context effects in phoneme and word recognition. Journal of the Acoustical Society of America. 1988; 84(1):101–114. [PubMed: 3411038]

Carnahan K, Arner M, Hagglund G. Association between gross motor function (GMFCS) and manual ability (MACS) in children with cerebral palsy: A population-based study of 359 children. BMC Musculoskeletal Disorders. 2007; 8:50–57. [PubMed: 17584944]

Chin SB, Tsai PL, Gao S. Connected speech intelligibility of children with cochlear implants and children with normal hearing. American Journal of Speech Language Pathology. 2003; 12:440–451. [PubMed: 14658996]

Coplan J, Gleason JR. Unclear speech: Recognition and significance of unintelligible speech in preschool children. Pediatrics. 1988; 82:447–452. [PubMed: 3405680]

Fenson, L.; Resznick, S.; Thal, D.; Bates, E.; Hartung, J.; Reilly, J. The MacArthur Communicative Development Inventories. San Antonio, TX: Psychological Corporation; 1993.

Fudala, JB. Arizona Articulatory Proficiency Scale-3. 3. Los Angeles: Western Psychological Services; 2001.

Garcia J, Cannito M. Influence of verbal and nonverbal contexts on the sentence intelligibility of a speaker with dysarthria. Journal of Speech and Hearing Research. 1996; 39:750–760. [PubMed: 8844555]

Gordon-Brannan M, Hodson BW. Intelligibility/severity measurements of prekindergarten children's speech. American Journal of Speech-Language Pathology. 2000; 9:141–150.

Hodge, M.; Daniels, J. TOCS+ Intelligibility Measures. Edmonton, AB: University of Alberta; 2007.

Hustad KC. Contribution of two sources of listener knowledge to intelligibility of speakers with cerebral palsy. Journal of Speech, Language, and Hearing Research. 2007a; 50:1228–1240.

Hustad KC. Effects of speech stimuli and dysarthria severity on intelligibility and listener confidence ratings for speakers with cerebral palsy. Folia Phoniatrica. 2007b; 59:306–317.

Hustad KC, Beukelman DR. Effects of linguistic cues and stimulus cohesion on intelligibility of severely dysarthric speech. Journal of Speech, Language, and Hearing Research. 2001; 44:497–510.

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.

Hustad KC, Kent RD, Beukelman DR. Dectalk and Macintalk speech synthesizers: Intelligibility differences for three listener groups. Journal of Speech, Language, and Hearing Research. 1998; 41:744–752.

Hustad KC, Miles LK. Alignment between augmentative and alternative communication needs and school-based speech-language services provided to young children with cerebral palsy Early Childhood Services. 2010; 3:129–140.

Johnson S, Marlow N, Wolke D, Davidson L, Marston L, O'Hare A, et al. Validation of a parent report measure of cognitive development in very preterm infants. Developmental Medicine & Child Neurology. 2004; 46:389–397. [PubMed: 15174530]

Kent R, Kent J, Weismer G, Martin R, Sufit R, Brooks B, et al. Relationships between speech intelligibility and the slope of second-formant transitions in dysarthric subjects. Clinical Linguistics and Phonetics. 1989; 3(4):347–358.

Kent R, Miolo G, Bloedel S. The intelligibility of children's speech: A review of evaluation procedures. American Journal of Speech-Language Pathology. 1994; 3:81–95.

Kleinow J, Smith A. Influences of length and syntactic complexity on the speech motor stability of the fluent speech of adults who stutter. Journal of Speech, Language and Hearing Research. 2000; 43(2):548–559.

Kleinow J, Smith A. Potential interactions among linguistic, autonomic, and motor factors in speech. Dev Psychobiol. 2006; 48(4):275–287. [PubMed: 16617462]

Lindblom B. On the communication process: Speaker-listener interaction and the development of speech. Augmentative and Alternative Communication. 1990; 6:220–230.

Maner KJ, Smith A, Grayson L. Influences of utterance length and complexity on speech motor performance in children and adults. Journal of Speech, Language, and Hearing Research. 2000; 43:560–573.

Miller GA, Heise GA, Lichten W. The intelligibility of speech as a function of the context of the test materials. Journal of Experimental Psychology. 1951; 41:329–335. [PubMed: 14861384]

Morris C, Kurinczuk JJ, Fitzpatrick R, Rosenbaum PL. Who best to make the assessment? Professionals' and families' classifications of gross motor function in cerebral palsy are highly consistent. Archives of Disease in Childhood. 2006; 91:675–679. [PubMed: 16638783]

Morris SR, Wilcox KA, Schooling TL. The preschool speech intelligibility measure. American Journal of Speech Language Pathology. 1995; 4:22–28.

Nip IS, Green JR, Marx DB. The co-emergence of cognition, language, and speech motor control in early development: a longitudinal correlation study. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. Journal of Communication Disorders. 2011; 44(2):149–160. [PubMed: 21035125]

O'Neill JJ. Recognition of intelligibility test materials in context and isolation. Journal of Speech and Hearing Disorders. 1957; 22:87–90. [PubMed: 13406828]

Palisano R, Rosenbaum P, Walter S, Russell D, Wood E, Galuppi B. Development of the gross motor function classification system. Developmental Medicine & Child Neurology. 1997; 39:214–223. [PubMed: 9183258]

Pan BA, Rowe M, Spier E, Tamis-LeMonda CS. Measuring productive vocabulary of toddlers in low-income families: Concurrent and predictive validity of three sources of data. Journal of Child Language. 2004; 31:587–608. [PubMed: 15612391]

Prelock P, Panagos J. The influence of processing mode on the sentence producitons of language-disordered and normal children. Clinical Linguistics & Phonetics. 1989; 3(3):251–263.

Rice ML, Smolik F, Perpich D, Thompson T, Rytting N, Blossom M. Mean length of utterance levels in 6-month intervals for children 3 to 9 years with and without language impairments. Journal of Speech, Language and Hearing Research. 2010; 53(2):333–349.

Sadagopan N, Smith A. Developmental changes in the effects of utterance length and complexity on speech movement variability. Journal of Speech, Language and Hearing Research. 2008; 51:1138–1151.

Sigurdardottir S, Eriksdottir A, Gunnarsdottir E, Meintema M, Arnadottir U, Vik T. Cognitive profile in young Icelandic children with cerebral palsy. Developmental Medicine and Child Neurology. 2008:50.

Sitler RW, Schiavetti N, Metz DE. Contextual effects in the measurement of hearing-impaired speakers' intelligibility. Journal of Communication Disorders. 1983; 11:22–30.

Smith A. Speech motor development: Integrating muscles, movements, and linguistic units. Journal of Communication Disorders. 2006; 39:331–349. [PubMed: 16934286]

Smith, A.; Goffman, L. Interaction of motor and language factors in the development of speech production. In: Maassen, B.; Kent, RD.; Peters, HFM.; van Lieshout, PHHM.; Hulstijn, W., editors. Speech Motor Control in Normal and Disordered Speech. Oxford: Oxford University Press; 2004. p. 225-252.

Yorkston K, Beukelman D. A clinician-judged technique for quantifying dysarthric speech based on single-word intelligibility. Journal of Communication Disorders. 1980; 13:15–31. [PubMed: 7354139]

Yorkston K, Beukelman D. Communication efficiency of dysarthirc speakers as measured by sentence intelligibility and speaking rate. Journal of Speech and Hearing Disorders. 1981; 46:296–301. [PubMed: 7278175]

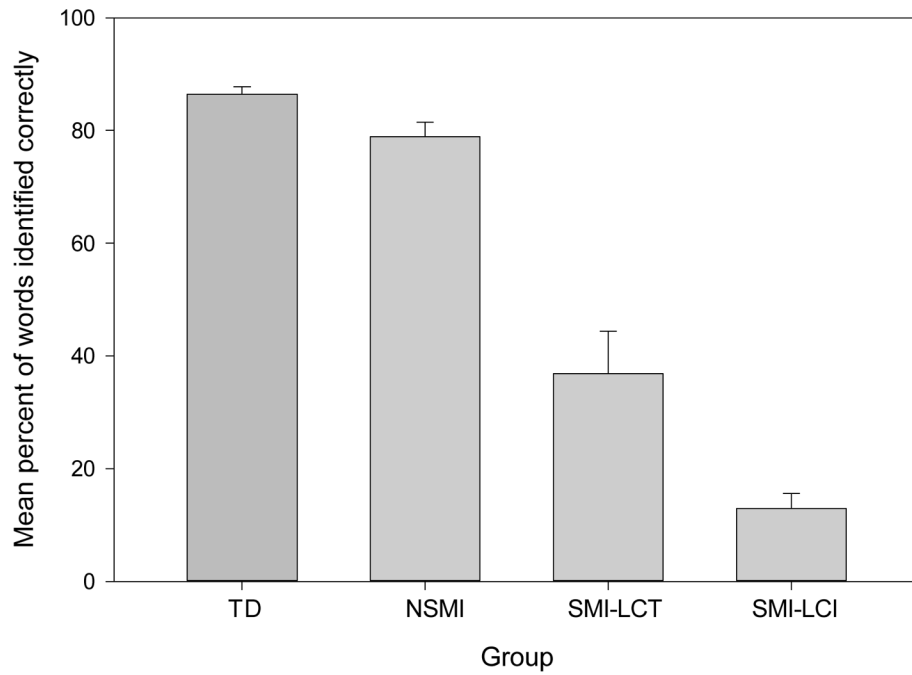Overall intelligibility (across stimulus lengths) by group

Figure 1.
Mean orthographic transcription intelligibility by group. Error bars represent standard error.
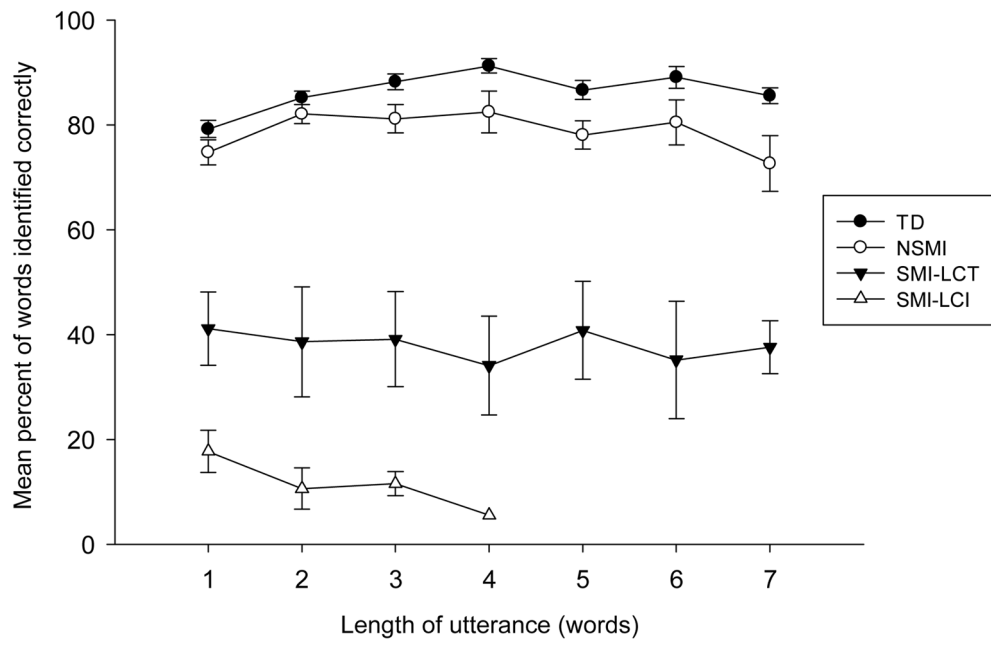
Mean intelligibility by stimulus length and severity group

Figure 2.
Mean orthographic transcription intelligibility by group and length of utterances. Error bars represent standard error.
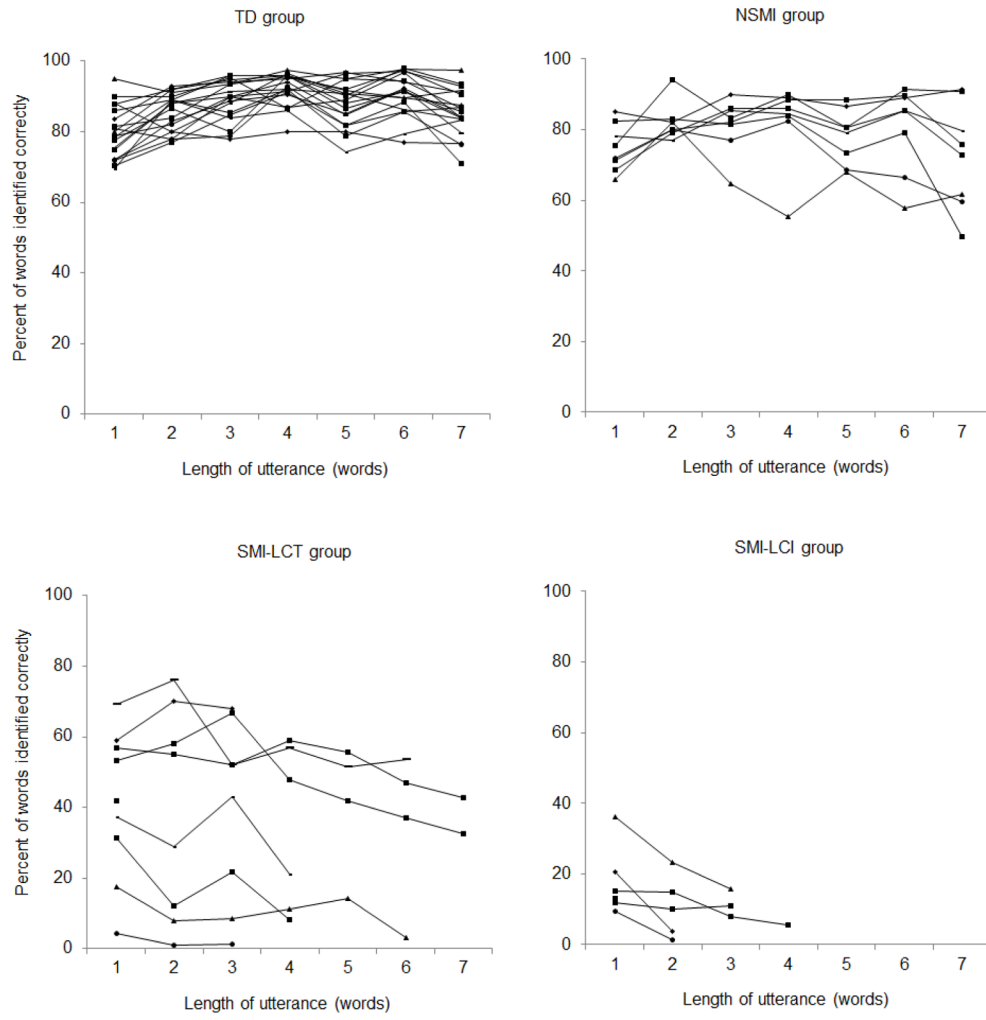
Figure 3.
Mean orthographic transcription intelligibility by group, individual child, and length of utterances.

Naive listener transcription intelligibility scores
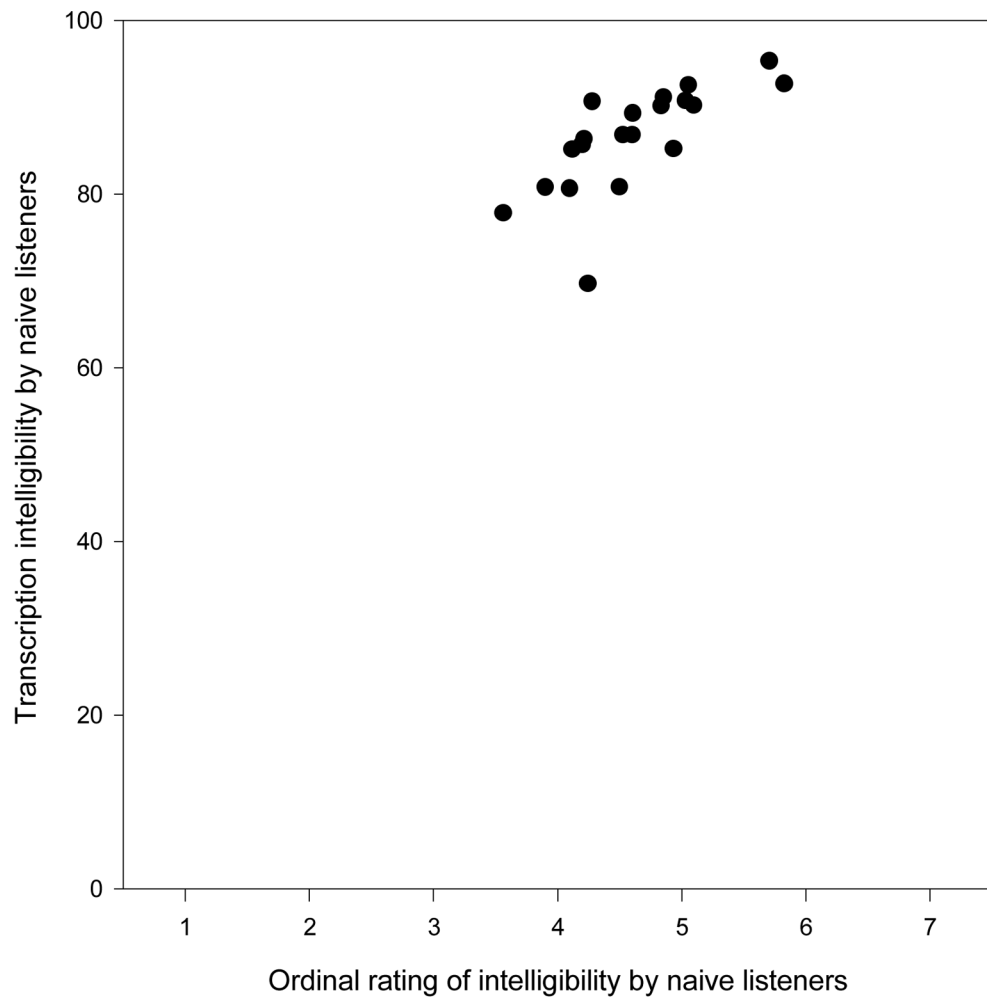by naive listener ordinal ratings of intelligibilty
for children in the TD group

Figure 4.
Naïve listener transcription intelligibility scores by naïve listener ordinal ratings of intelligibility for children in the TD group.

Naive listener transcription intelligibility scores by naive listener ordinal ratings of intelligibility for children with CP by group
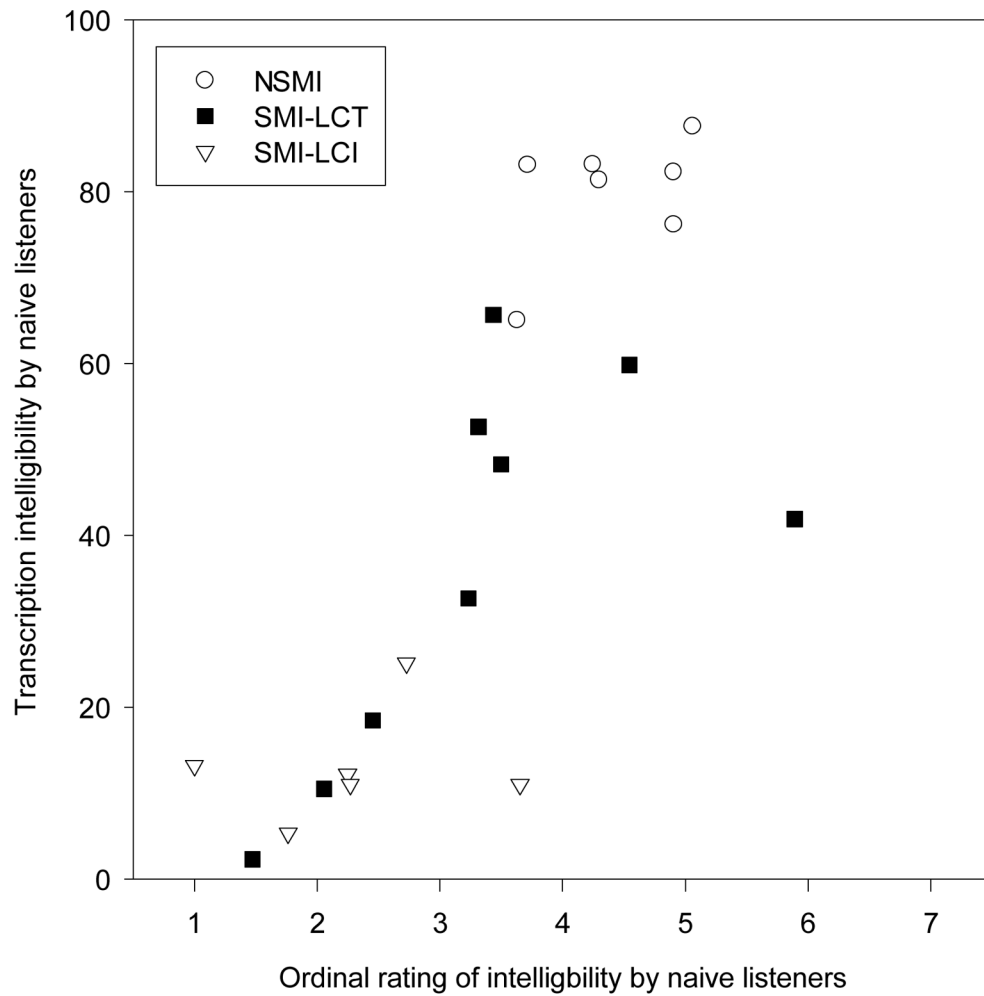
Figure 5.
Naïve listener transcription intelligibility scores by naïve listener ordinal ratings of intelligibility for children with CP by group.

Naive listener transcription intelligibility scores by parent ordinal ratings of intelligibility for children with CP by group
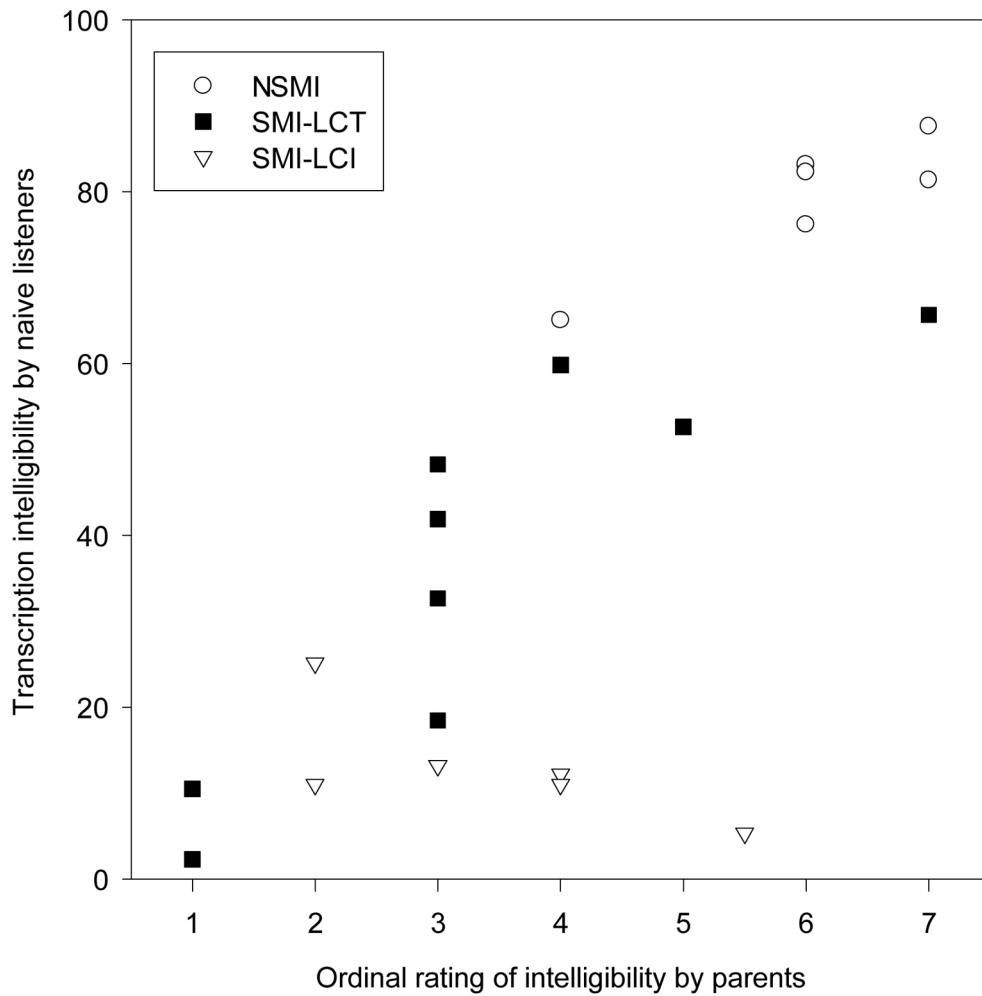
Figure 6.
Naïve listener transcription intelligibility scores by parent ordinal ratings of intelligibility for children with CP by group.

Table 1

Demographics of children with CP.

	Group NSMI (n = 8)	Group SMI-LCT (n = 9)	Group SMI-LCI (n = 6)
Mean Age (SD)	54.7 (1.1)	53.5 (1.7)	54.9 (1.8)
Male: female ratio	5:3	2:7	3:3
Type of CP			
Diplegia	4	2	1
Hemiplegia (left)	3	1	0
Hemiplegia (right)	0	1	1
Quadriplegia	0	2	2
Dyskinesia	0	2	0
Ataxia	0	0	1
Mixed	0	0	1
Unknown	1	1	0
GMFCS level			
Level I	2	1	0
Level II	6	4	2
Level III	0	2	2
Level IV	0	2	0
Level V	0	0	2

Table 2

Linguistic characteristics of TOCS sentences.

	MLU-W	MLU-M	Number of bound morphemes	Total number of words	Number of different words	Type token ratio	Brown's stage	Expected age range
2-word sentences	2.0	2.20	2	20	20	1.0	II	21–35 mos
3-word sentences	3.0	3.45	5	30	29	.97	Early IV	28–45 mos
4-word sentences	4.0	4.40	4	40	37	.93	Late V	37–52 mos
5-word sentences	5.0	6.0	10	50	42	.84	Post V	41+ mos
6-word sentences	6.0	6.7	7	60	46	.77	Post V	41+ mos
7-word sentences	7.0	7.6	6	70	59	.84	Post V	41+ mos

Table 3

Mann-Whitney U contrasts examining differences in intelligibility scores across all utterance lengths among groups of children.

Contrast	Mean difference	TS	Std. Error	Observed p-value
TD vs. NSMI	8.41	45.70	11.45	<.0010[*]
TD vs. SMI-LCT	50.37	108.79	10.97	<.0010[*]
TD vs. SMI-LCI	74.31	131.93	12.73	<.0010[*]
NSMI vs. SMI-LCT	41.97	63.09	13.21	<.0010[*]
NSMI vs. SMI-LCI	65.91	86.23	14.68	<.0010[*]
SMI-LCT vs. SMI-LCI	23.94	23.144	14.32	.1060

[*] statistical significance at p<.0014

Table 4

Friedman tests examining within group differences across utterance lengths.

Groups	df	X^2	Observed p-value
TD			
Length of utterance	6	168.51	<.0010 [*]
NSMI			
Length of utterance	6	34.54	<.0010 [*]
SMI-LCT			
Length of utterance	6	19.89	.003 [*]
SMI-LCI			
Length of utterance	3	8.28	.041

[*] statistical significance at p<.005