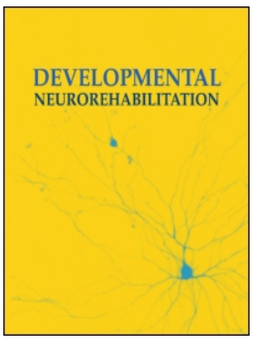
Taylor & Francis
Taylor & Francis Group



Developmental Neurorehabilitation

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/ipdr20

Validity of Parent Ratings of Speech Intelligibility for Children with Cerebral Palsy

Ashley Sakash , Tristan Mahr & Katherine C. Hustad

To cite this article: Ashley Sakash , Tristan Mahr & Katherine C. Hustad (2020): Validity of Parent Ratings of Speech Intelligibility for Children with Cerebral Palsy, Developmental Neurorehabilitation, DOI: 10.1080/17518423.2020.1830447

To link to this article: https://doi.org/10.1080/17518423.2020.1830447

Published online: 25 Oct 2020.

DEVELOPMENTAL NEUROREHABILITATION
https://doi.org/10.1080/17518423.2020.1830447

Taylor & Francis
Taylor & Francis Group

ARTICLE

Validity of Parent Ratings of Speech Intelligibility for Children with Cerebral Palsy

Ashley Sakash[a], Tristan Mahr[b], and Katherine C. Hustad[a]

[a]Waisman Center, University of Wisconsin, Madison, WIS, USA; [b]Waisman Center,Department of Communication Sciences and Disorders, University of Wisconsin, Madison

ABSTRACT

Aim: To examine the relationship between subjective parent ratings of intelligibility and objectively measured intelligibility scores for children with cerebral palsy (CP) with differing levels of speech severity.
Method: Fifty children (84–96 months) with CP were classified into groups based on intelligibility scores during a speech elicitation task – high intelligibility (90% or higher), mild-moderate intelligibility reduction (61–89%), and severe intelligibility reduction (60% or lower). Parent ratings of understandability (on a 7-point scale) were compared to intelligibility scores gathered from 100 naïve listeners.
Results: For children with mild-moderate and severe intelligibility reduction, there was a large range of variability in parent ratings. For children with high intelligibility, ratings were consistent with intelligibility scores. There was a range of intelligibility scores within each rating, especially in the middle of the scale.
Conclusions: For children with mild-moderate intelligibility deficits, parent ratings may best be used in conjunction with objective measurement of intelligibility.

ARTICLE HISTORY
Received July 08, 2020
Revised September 23, 2020
Accepted September 27, 2020

Keywords
Intelligibility; cerebral palsy; parent rating; severity; measurement

Throughout development, health care providers and educators rely on parents to be accurate reporters of their child's abilities. From early infancy, parents are asked to report any concerns they have regarding their child's hearing and vision abilities. It is a common practice for parents to complete the Ages & Stages Questionnaire from one month through 65 months in order to help pediatricians monitor gross motor, fine motor, and communication skill development.[1,2] For children at risk of delay, developmental screening based on parent report is increasingly relied upon.[3] Using parent report in assessment is efficient both in time and cost and is particularly useful for populations where formal standardized testing is not possible due to age or level of cognitive or motor functioning. Ultimately, many critical decisions, including referrals to specialty care, are based on parent report.

Parents are generally considered to be reliable raters when it comes to making judgments about their child's overall development.[4–7] However, parent report of certain skills, such as speech abilities, has not always been validated in the context of quantitative data. Recently, concerns have been identified using parent report to measure child speech/language skills. In particular, studies have noted inconsistencies between parent report and objective measures, with parents over- or underestimating their child's abilities.[4,8] Despite these findings, parent ratings remain one of the most commonly used methods of assessing speech and language skill development in children, especially for the measurement of speech intelligibility.

Speech intelligibility can be defined as how well a speaker's acoustic signal can be accurately recovered by a listener.[9,10] Speech intelligibility is a multidimensional construct, influenced by many variables, and can be measured in several ways.[11–14] Direct measurement of intelligibility can involve language sample analysis, or transcription methods involving naïve listeners, which is considered by some to be a gold standard.[12] With transcription methods, unfamiliar listeners provide orthographic transcriptions of speech samples which are scored against a reference key of target sentences. The percent of words identified correctly relative to the key is typically determined and this value is a child's intelligibility score. This approach to measurement is objective; however, it is time consuming and impractical for use in clinical settings. Having a quick, efficient, reliable measure of speech intelligibility is important in clinical practice for evaluating and monitoring changes in speech performance.[15] Parent estimates of speech intelligibility offer an indirect method of measurement and although there are concerns about accuracy of parent perception, parent report measures are quick, efficient, cost-effective, and have been found to provide valuable information.[16–18]

Parent estimates of intelligibility were used to develop guidelines and cut points for typical speech intelligibility development that is widely used by health care professionals including speech-language pathologists and pediatricians.[18] More recently, parent estimates of intelligibility have been used to develop the Intelligibility in Context Scale (ICS), which involves having parents make subjective ratings in response to seven different intelligibility-related questions.[16,17] In one recent study, Lagerberg and colleagues compared scores from the ICS to intelligibility scores from a single-word assessment procedure in children with speech and/or language disorders ranging in age from 4 to 10 years old.[19] Results indicated a weak but statistically significant relationship between parent ratings on the ICS and intelligibility scores from the single-word assessment.[19] Findings suggest that there is a relationship

CONTACT Ashley Sakash asakash@wisc.edu Waisman Center, University of Wisconsin Madison, WIS 53705

between subjective parent ratings and objective measures of intelligibility; however, children in the study by Lagerberg and colleagues varied greatly in their speech severity and the authors did not control for severity of speech impairment. Although measures such as the ICS are widely used, no studies have validated parent estimate data against objective measures of intelligibility of connected speech. Thus, it is unclear how subjective parent ratings map onto objectively measured multi-word speech intelligibility.

Children with cerebral palsy (CP) are a population at risk for developmental issues and present with a wide range of speech production abilities. Previous work has shown that about 50% of children with CP have been clinically diagnosed with dysarthria, a neurologically based speech disorder[20] that often leads to reductions in speech intelligibility. Speech differences and reductions in speech intelligibility have also been found even for children with CP without a clinical diagnosis of dysarthria relative to typically developing peers.[21,22] Past studies have shown that reduced intelligibility has significant adverse consequences for social participation and overall quality of life.[23] Therefore, it is critical to use a reliable measure of speech intelligibility to properly evaluate the need for intervention and guide treatment decision making. In a recent study, Natzke and colleagues examined the extent to which parent ratings of intelligibility revealed developmental change in speech over two years for children with CP who varied in severity.[24] Results indicated that parent ratings of intelligibility were not sensitive to growth, while at the same time, measures of transcription intelligibility showed increases in intelligibility over time for children with mild and moderate intelligibility deficits.[24] Further, this same study found that parent ratings of intelligibility were not correlated with transcription intelligibility within severity groups, but examination of parent ratings was only a small component of the study, and the predictive ability of parent ratings was not examined.

In the current study, we examine the parent ratings employed by Natzke and colleagues in greater detail, focusing on questions of the variability in ratings by severity group, how well empirically measured intelligibility scores predicted parent ratings within severity groups, and the variability in intelligibility scores within rating levels. A key goal was to examine the validity of parent ratings relative to clinical orthographic transcription measures of speech intelligibility in order to determine whether parent ratings could serve as a proxy for direct clinical measures.

We asked the following specific research questions:

(1) Within empirically defined severity groups:
   a. What is the range of variability in parent ratings for children within the same severity groups?
   b. Do parent ratings differ, on average, between severity groups?
   c. To what extent do intelligibility scores predict parent ratings within severity group?
(2) Within ordinal ratings of intelligibility, regardless of severity group membership, what is the range of variability in intelligibility scores for each parent rating level (1–7)?

We hypothesized that the range of variability in parent ratings within the same severity group would be small if parents were sensitive to their child's level of speech severity. We expected that ratings would differ between severity groups again, if parents were sensitive to severity. We expected that transcription intelligibility scores would be predictive of parent ratings based on the findings from Natzke and colleagues.[24] Within each ordinal rating of intelligibility, we expected there to be a relatively constrained range of transcription intelligibility scores.

Method

Participants

Approval for this study was given by the University of Wisconsin-Madison Institutional Review Board. Informed consent was obtained for all participants.

Children with CP

Participants were 50 children with CP drawn from a larger ongoing longitudinal study on communication development in children with CP. Criteria for inclusion in 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. For the current study, all participants also met the following criteria: 3) age between 84 and 96 months, 4) use speech as their primary method of communication, 5) be able to produce sentences at least 3 words in length in a sentence repetition task, and 6) parents completed an informal communication questionnaire. A total of 50 children (26 males) met these inclusion criteria and were included in the present study. Note that data for 42 of these children were also utilized in the study by Natzke and colleagues.[24] The eight additional children in the present study were not included in the study by Natzke and colleagues due to an incomplete data collection session at one of the three age points in the earlier study. Table 1 reports demographic information for child participants. Children included in the current study were receiving their usual therapy services as provided in their community. Parents were asked to indicate on a questionnaire whether their child was currently receiving speech and language therapy. This information is presented in Table 1.

Previous research has found that speech severity acts as a potentially confounding variable which may yield an artificial correlation with intelligibility.[27] The current study was designed to investigate the relationship between parent ratings of intelligibility and listener-derived measures of transcription intelligibility while controlling for severity of speech impairment. Therefore, children were classified into severity groups based on measured multi-word intelligibility scores. In the dysarthria literature, intelligibility scores are widely used as

Table 1. Demographic characteristics of children with CP.

	high intelligibility $n = 21$	mild-moderate intelligbility reduction $n = 15$	severe intelligbility reduction $n = 14$
Male:female ratio	13:8	8:7	5:9
Age in months: Mean (SD)	90.57 (4.57)	87.67 (3.94)	91 (4.80)
Language SS: Mean (SD)	101.1 (15.62)	93.8 (14.33)	73.57 (26.31)
Intellectual Disability	2	1	9
Dysarthria	6	12	14
Multi-word Intelligibility: Mean (SD)	94.57 (2.68)	79.59 (8.12)	26.44 (18.50)
Cerebral palsy type			
Spastic	19	11	9
Diplegia	6	2	2
Hemiplegia (left)	5	4	2
Hemiplegia (right)	5	3	2
Triplegia	0	1	0
Quadriplegia	3	1	1
Dyskinetic	0	0	0
Ataxic	0	3	2
Mixed	0	0	1
Hypotonic	0	0	1
Unknown	2	1	1
GMFCS			
I	12	7	0
II	8	4	8
III	1	0	3
IV	0	4	2
V	0	0	1
MACS			
I	7	6	3
II	10	8	7
III	2	1	4
IV	1	0	0
V	1	0	0
Currently receiving speech and language therapy	12	8	6

GMFCS = Gross Motor Classification System[25]; MACS = Manual Abilities Classification Scale[26]

a marker of speech severity.[27] The groupings we used in the current study were determined based on findings and operational definitions used in our previous work with children with CP.[24,28] Currently, there is no widely accepted standard for severity designations so we chose to use groupings that we've used previously for consistency purposes. Children with intelligibility scores of 90% or greater made up the high intelligibility group (n = 21), children with scores between 61% and 89% made up the mild-moderate intelligibility reduction group (n = 15), and children with scores of 60% or below made up the severe intelligibility reduction group (n = 14).

Nondisabled adult listeners

One hundred healthy adults participated as listeners in the current study. Listeners transcribed speech samples produced by children, which in turn yielded intelligibility scores. Listeners primarily consisted of undergraduate students and were recruited from a university setting through public postings and social media. All listeners met the following inclusion criteria: 1) pass a pure tone hearing screening administered via headphones at 25 dB HL at 250, 500, 1000, 4000, and 6000 Hz in both ears; 2) be between 18 and 45 years of age; 3) have no more than incidental experience listening to or communicating with persons having communication disorders; 4) be a native speaker of American English; and 5) have no identified language, learning or cognitive disabilities per self-report. Listeners comprised 76 females and 24 males. The mean age of listeners was 21.5 (SD = 3.8) years.

Materials and procedures

Each child completed a standard research protocol consisting of speech, language, cognitive, and oral-motor assessments at each visit. A complete description of assessments in the standard protocol is provided in ref. 25. Assessments were administered by a licensed and ASHA certified research speech-language pathologist (SLP). Parents of participants completed a set of standardized and informal questionnaires. Of interest to the current study were 1.) parent responses to specific questions on an informal communication questionnaire; and 2.) multi-word intelligibility scores reflecting the percent of words produced by child in a speech elicitation task that naïve listeners transcribed correctly.

Acquisition of parent ratings of speech intelligibility

Prior to visiting the research laboratory, parents of children with CP completed an informal communication questionnaire. Parents were asked to answer yes/no and to provide scaled ratings for questions regarding their child's use of speech, use of other modes of communication, and child's speech/language

therapy. Of interest in the present study, parents were asked to make ratings on a 7-point scale regarding how understandable their child was to themselves and to others. Results from parent responses to the following question were analyzed: Overall, how understandable is your child to others (even if he/she doesn't sound 'normal')? Parents were instructed to circle a number from 1 (very easy to understand) to 7 (very hard to understand).

Acquisition of transcription intelligibility scores from children

Each child with CP completed a structured imitative speaking task in the laboratory. During this task, children were audio-recorded while repeating a list of sentences from the Test of Children's Speech (TOCS+), a developmentally appropriate set of speech stimuli that systematically vary in length.[29] Stimuli consisted of sentences that ranged from 2 to 7 words with 10 sentences of each stimulus length for a total of 60 sentences. By eliciting the same set of stimuli from children, we were able to ensure that intelligibility scores reflected listeners' perception of target words relative to a known set of items. The speaking task took place with the child seated at a table in a sound-attenuating suite next to a research SLP. Speech samples from children were recorded using a digital audio recorder (Marantz PMD 570, D & M Holdings Inc., Tokyo, Japan) at a 44.1-kHz sampling rate (16-bit quantization). A condenser studio microphone (Audio-Technica AT4040, Audio-Technica U.S., Inc., Stow, OH) 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, Mackie Designs Inc., Woodinville, WA) to obtain optimized recordings and to avoid peak clipping.

Adult recordings of each target stimulus sentence along with the written words and images depicting the sentence were presented to children via a 12.9 in. Apple iPad Pro. Children were asked to repeat what they heard upon completion of the recorded adult model. All child productions were monitored in real time by a research assistant to ensure that speech samples were free from overlap with the model and free from extraneous noises.

Digital audio recordings were transferred to a desktop computer and edited to remove extraneous noises and the prerecorded adult model. Individual files were then created for each stimulus item produced by each child. Audio samples were peak amplitude normalized to ensure that maximum loudness levels of the recorded speech samples were the same across children and stimulus items while preserving the amplitude contours of the original productions. Speech stimuli were presented via in-house software to listeners seated in a sound-attenuating suite. The external speaker was calibrated on a regular basis by a research assistant to ensure the peak output level was 75 dB SPL from where listeners were seated.

Each listener was presented with all speech stimuli spoken by a single child. The in-house software randomized the presentation order of stimulus items for each listener. Listeners were instructed to provide orthographic transcriptions of each utterance – that is, to type what they thought the child had said. Two listeners provided transcriptions for each utterance and for each child. Each listener heard only one child producing all stimulus items. This assignment between listener and child was used to mitigate learning effects that might occur with the same listener hearing the same child or same speech stimuli several times.[30] Presentation of stimulus items was randomized. No two listeners heard the stimulus items presented in the same order.

In-house software scored each typed word generated by listeners as either correct or incorrect based on whether the listener transcription matched the target transcription phonemically. Misspellings and homonyms were accepted as correct, provided that all phonemes in the transcription matched the target. The total number of words transcribed correctly by each of the two listeners per child were added together and then divided by the total number of words possible (across the two listeners) and multiplied by 100 to yield a percent intelligibility score for each child.

Reliability

For each child, we computed the difference in average intelligibility between the two listeners. The average difference in intelligibility scores between the two listeners was 3.4 percentage points (SD = 3.2). We used an average-score, consistency-based, one-way random effects model, and we found strong agreement among average ratings, ICC(2) =.994, 95% CI = [.989, .996].

Analysis

Parent ratings of children's intelligibility are on an ordinal scale, while transcription intelligibility scores from listeners are on a ratio scale. For question 1a examining the range of variability among parent ordinal ratings within severity groups, our analyses are descriptive in nature. For question 1b, we used ordinal logistic regression to examine whether parent ratings differ between severity groups and to estimate group differences. For question 1 c, we regressed the parent ratings onto severity group, intelligibility, and a group by intelligibility interaction. Intelligibility was mean-centered within each group so that we could interpret the effect of intelligibility within each group. To address question 2 examining the range of variability in intelligibility scores at each level of the 7-point ordinal rating scale, we used descriptive statistics.

Analysis was carried out in R [vers. 4.0.0].[31] We fit the ordered logistic regression models with clm function in the ordinal package [vers. 2019.12.10].[32] We report marginal means and p-value-adjusted contrasts estimated using the emmeans package [vers. 1.4.7].[33]

Results

Question 1a: What is the range of variability in parent ratings for children within the same severity groups?

Table 2 shows the frequency of each ordinal parent rating by severity group. As can be seen in the table, the frequency of ratings differed somewhat between severity groups; however, there was a considerable range of variability with overlapping

Table 2. Frequency of parent rating responses by severity group. The cumulative proportion indicates the proportion of ratings less than or equal to the given row. For example, for the high intelligibility group, all ratings were 3 or less.

	High Intelligbility			Mild-moderate Intelligibility Reduction			Severe Intelligbility Reduction		
Intelligibility to others (rating)	n	Proportion	Cumulative	n	Proportion	Cumulative	n	Proportion	Cumulative
(most intelligible) 1	13	0.62	0.62	3	0.20	0.20	0	0.00	0.00
2	3	0.14	0.76	1	0.07	0.27	1	0.07	0.07
3	5	0.24	1.00	2	0.13	0.40	4	0.29	0.36
4	0	0.00	1.00	4	0.27	0.67	2	0.14	0.50
5	0	0.00	1.00	5	0.33	1.00	3	0.21	0.71
6	0	0.00	1.00	0	0.00	1.00	3	0.21	0.93
(least intelligible) 7	0	0.00	1.00	0	0.00	1.00	1	0.07	1.00

parental ratings among groups. For children in the high intelligibility group, the range of ratings was small. Ratings from parents in the high intelligibility group ranged from 1 to 3, and 62% of parents rated their child a 1, indicating they believe their child was highly intelligible to others. For children in the mild-moderate intelligibility reduction group, there was a descriptively larger range of ratings that parents assigned to their children, reflecting greater variability. Ratings from parents of children in the mild-moderate intelligibility reduction group ranged from 1 to 5. In this group, 67% of parents rated their child less than or equal to 4. Put differently, in the mild-moderate intelligibility reduction group, there was a greater range of parent rating scores associated with the same proportion of children. For children in the severe intelligibility reduction group, the range of parent ratings was the greatest. Ratings from parents of children in the severe intelligibility reduction group ranged from 2 to 7. None of the other groups used ratings of 6 or 7, but in the severe intelligibility reduction group, 28% of children received these ratings.

Question 1b: Do Parent Ratings Differ, on Average, between Severity Groups?

Figure 1 shows the average parent rating for children in each severity group. The estimated average parent rating was 1.7, 95% CI [1.3, 2.1], for the high group; 3.6, [2.8, 4.3], for the mild-moderate intelligibility reduction group; and 4.4, [3.6, 5.1] for the severe intelligibility reduction group. There was a significant difference in average ratings between the high intelligibility and severe intelligibility reduction groups, Severe − High = 2.7, SE = 0.44, $z = 6.05$, $p < .001$, and the mild-moderate intelligibility reduction and the high intelligibility groups, Mild-Mod − High = 1.9, SE = 0.44, $z = 4.27$, $p < .001$. There was not a statistically significant difference between the mild-moderate and severe intelligibility reduction groups, Severe − Mild-Mod = 0.8, SE = 0.54, $z = 1.49$, $p = .29$

Question 1c: To What Extent Do Intelligibility Scores Predict Parent Ratings within Severity Group?

Figure 2 shows transcription intelligibility scores from listeners by parent rating for children in each severity group. It is important to note that in the ordinal parent rating scale, lower numbers indicate higher understandability ratings and higher numbers indicate lower understandability ratings. The opposite is true for transcription intelligibility scores where lower numbers indicate lower intelligibility scores and higher numbers indicate higher intelligibility scores. For all three groups, the effect of transcription intelligibility on parent ratings of intelligibility went in the expected direction: increases in transcription intelligibility predicted lower parent ratings (higher understandability ratings). This effect, however, was

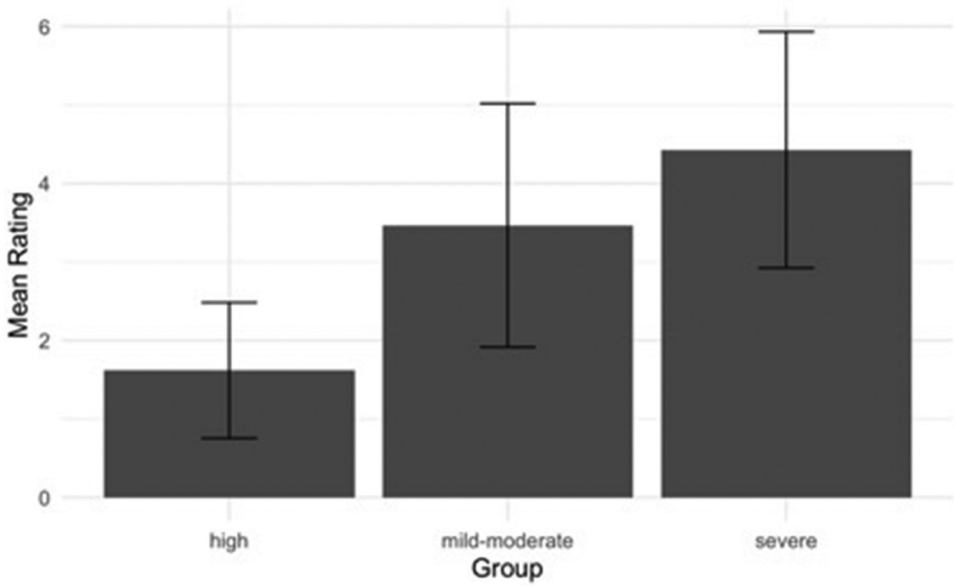


Figure 1. Average parent rating by severity group. Error bars represent ± 1 SD.

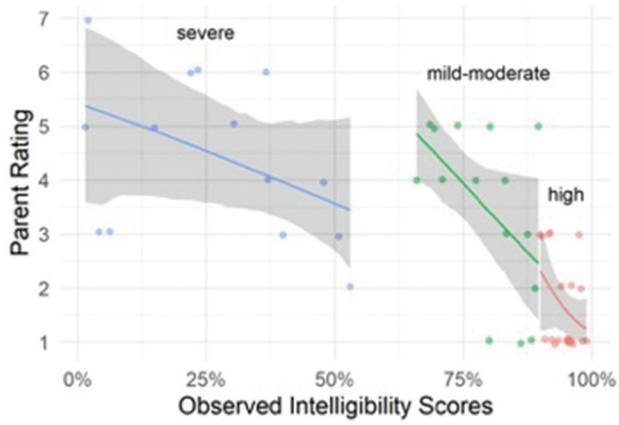


Figure 2. Intelligibility scores by parent rating for each severity group. Lines represent regression estimates for expected rating and ribbons represent 95% bootstrap interval.

statistically significant only for the mild-moderate intelligibility reduction group. For the mild-moderate intelligibility reduction group, the mean transcription intelligibility score was 80% (SD = 8%). For a 1-SD change in transcription intelligibility in the mild-moderate intelligibility reduction group (from 80% to 88%), the expected change in parent ratings was −0.84 ordinal scale points, 95% CI [−1.48, −0.19]. For the high intelligibility group, the mean transcription intelligibility score was 95% (SD = 2%). A 1-SD change for the high intelligibility group (from 95% to 97%) predicted a change in rating of −0.32 ordinal scale points, [−0.73, 0.09]. For the severe intelligibility reduction group, the mean transcription intelligibility score was 26% (SD = 19%). A 1-SD change for the severe intelligibility reduction group (from 26% to 45%) predicted a change of −0.70 ordinal scale points, [−1.42, 0.02]. For the high and severe intelligibility reduction groups, the 95% confidence interval did not exclude 0, therefore the intelligibility effects did not attain statistical significance.

Question 2: Within ordinal ratings of intelligibility, regardless of severity group membership, what is the range of variability in intelligibility scores for each parent rating level?

Table 3 shows transcription intelligibility scores associated with each parent rating (1–7), regardless of severity group. The average transcription intelligibility score decreased as parent ratings increased (indicating greater severity), but there was substantial variability in the transcription intelligibility scores within each parent rating. Children who received a rating of 1 from their parent (n = 16) had an average intelligibility of 93% with a range of 80–99%. Children who received a rating of 2 from their parent (n = 5) had an average intelligibility of 86% with a range of 53–98%. Children who received a rating of 3 from their parent (n = 11) had an average intelligibility of 67% with a range of 4–98%. Children who received a rating of 4 from their parent (n = 6) had an average intelligibility of 64% with a range of 37–83%. Children who received a rating of 5 from their parent (n = 8) had an average intelligibility of 54% with a range of 2–90%. Children who received a rating of 6 from their parent (n = 3) had an average intelligibility of 27% with a range of 22–37%. The one child who received a rating of 7 from their parent had an intelligibility score of 2%. These findings suggest that a rating of 1 identified children with high intelligibility or mildly reduced intelligibility and ratings of 6 and 7 identified children with severely reduced intelligibility. Parent ratings in the middle of the scale were associated with a wide range of intelligibility scores.

Table 3. Intelligibility scores by each parental rating.

	Intelligibility scores from listeners			
Intelligibility to others (rating)	n	Mean Intelligibility	SD	Range
(most) 1	16	93%	4.9	80–99%
2	5	86%	18.8	53–98%
3	11	67%	35.4	4–98%
4	6	64%	17.8	37–83%
5	8	54%	33.0	2–90%
6	3	27%	8.1	22–37%
(least) 7	1	2%		

Discussion

The purpose of this study was to examine the validity of parent ratings of intelligibility by quantifying the relationship between subjective ordinal parent ratings of intelligibility and objective transcription intelligibility scores for children with CP who had differing levels of speech severity. We classified children into three severity groups based on objectively measured multi-word intelligibility scores, following gold standard practices in motor speech disorders. We analyzed parent responses to the question of how understandable their child is to others (using a 7-point ordinal scale). We examined parent response data in the context of multi-word objective intelligibility scores obtained from a speech elicitation task where adult listeners provided orthographic transcriptions of child utterances. Key findings are discussed below

Parent ratings within severity groups

Children in the high intelligibility group were operationally defined as those with objective transcription intelligibility scores between 90% and 99%. Results of this study showed that for the high intelligibility group, the range of parent ratings was relatively narrow, ranging from 1 to 3 with an average rating of 1.7, suggesting that parents of children with high intelligibility had accurate insight into their child's speech. The average parent rating for the high intelligibility group was significantly different from both other groups suggesting that the 7-point ordinal scale was sensitive enough to differentiate between children with high intelligibility and lower intelligibility. Contrary to what we hypothesized, transcription intelligibility scores did not make a significant contribution to predicting parent ratings for the high intelligibility group. One potential explanation for this is likely a ceiling effect. That is, children, by definition, had high intelligibility scores and high parent ratings of intelligibility and thus did not vary reliably.

Children in the mild-moderate intelligibility reduction group were operationally defined as those with objective transcription intelligibility scores between 61% and 89%. The range of parent ratings was larger for the mild-moderate intelligibility reduction group compared to the high intelligibility group.

This observation is not surprising given that the mild-moderate intelligibility reduction group as defined comprised a wider range of objective intelligibility scores. However, it is noteworthy that parents used 5 of 7 possible ratings for children in this group, employing most of the range of the scale. Parent ratings for this group ranged from 1 to 5 with an average rating of 3.6. We also found that within this mild-moderate intelligibility reduction severity group, transcription intelligibility scores made a significant contribution to predicting parent ratings, indicating that as intelligibility increased, parent ratings reliably decreased. However, the magnitude of this relationship was relatively small, such that a change in objective intelligibility of 1-SD was associated with an expected change in parent ratings of −0.84 scale points. In this case 1-SD was 8%, thus for a parent to change their intelligibility rating by approximately 1 scale point, intelligibility would need to change by slightly over 8%.

Children in the severe intelligibility reduction group were operationally defined as those with objective transcription intelligibility scores below 60%. There was notable variability in the range of parent ratings for children with severely reduced intelligibility. Ratings ranged from 2 to 7 with an average rating of 4.4. Transcription intelligibility scores for this group did not make a significant contribution to predicting parent ratings. As with the mild-moderate intelligibility reduction group, this finding suggests that parents of children with severely reduced intelligibility may be less accurate at rating their child's speech intelligibility. Five parents of children in the severe intelligibility reduction group used a rating of 3 or less, indicating that they thought their child had a relatively high level of understandability to others. However, transcription intelligibility scores from unfamiliar listeners indicated that each of these children received intelligibility scores below 55%, indicating severely reduced intelligibility. Interestingly, there was no significant difference in average parent ratings between the mild-moderate and severe intelligibility reduction group, suggesting that parents of children with intelligibility impairments rated their children similarly, or with a similar range of variability across the scale, regardless of the severity of their impairment. Findings reveal that parent perceptions of intelligibility may not be well calibrated with objective measures of intelligibility. One possibility is that parents ratings reflect a more holistic view of communication, rather than intelligibility specifically. Since we did not have other measures of communicative success or communicative participation in the present study, we are not able to examine how parent ratings may relate to such measures, however, this would be an interesting avenue for future investigation.

Variability in intelligibility scores for each parent rating level

Results of this study showed the average transcription intelligibility score increased as parent ratings decreased (indicating that generally parent ratings tracked with transcription intelligibility). However, there was marked variability in the intelligibility scores within each ordinal rating, except at the high and low ends of the scale. A parent rating of 1 was associated with a 20% range of intelligibility scores of (80–99%) and a rating of 7 was associated with a 35% range of intelligibility scores of (2–37%). This finding indicates that parents who gave high ratings and parents who gave low ratings generally had an impression of their child's intelligibility that was roughly consistent with their objectively measured intelligibility scores, and thus parent ratings may be valid for these children. However, for parents that gave ratings between 2 and 5 on the ordinal scale, the difference between high and low transcription intelligibility scores within each ordinal level ranged from 45% (for an ordinal rating of 4) to 94% (for an ordinal rating of 3). For this middle range there was little consistency between transcription intelligibility scores and parent-assigned ordinal ratings, suggesting that use of parent ratings may not be a valid method for characterizing intelligibility for children who are not on the poles of the scale. Similar to findings from previous studies, it may be that parents are over- or underestimating their child's understandability.[5,8] Another possible explanation is that parents have different internal calibration to the rating scale. One parent's perception of a 3 vs. 5 on the scale may be different than another parent's.[34] Similar findings suggesting ambiguity in relationships among ratings within the middle range of subjective scales have been found in the CP literature, particularly in studies including communication ratings.[35] One important feature of the present study is that we examined severity groups separately so as not to misrepresent relationships among variables which would show a natural high correlation owing solely to the effects of severity. Results of this study highlight that children in the middle of a severity continuum often look very different than those on the poles of the continuum. If we do not consider severity groups separately, we risk making critical erroneous assumptions that can have a detrimental impact on assessment and treatment of children with neurodevelopmental disabilities.

Limitations and future directions

We had a relatively small number of children in this study. The group sizes were also relatively small with 21 in the high intelligibility group, 15 in the mild-moderate intelligibility reduction group, and 14 in the severe intelligibility reduction group. Future research should be conducted with a larger sample of children with CP.

The range of transcription intelligibility scores within the high intelligibility group was small (90–99%). As previously mentioned, the groupings we used in the current study were determined based on our previous work with children with CP.[24,28] It is possible that this limited range of scores resulted in a ceiling effect and lack of a finding of a significant correlation between transcription intelligibility scores and parent ratings for the high intelligibility group. In addition, the range of transcription intelligibility scores for the mild-moderate intelligibility reduction group was quite large (61–89%). Since there is not a standard in the field for severity designations, future studies should explore using different criteria to create severity groups based on intelligibility scores.

In the current study, we did not examine the relationship between parent ratings of intelligibility and child current speech and language services. It's possible that parents of children who receive speech and language services are more

skilled at understanding their child due to strategies they observe or learn during their child's therapy sessions. Therefore, future studies should examine the impact of therapy on parent ratings of intelligibility.

Only one holistic parent rating measure was used in the current study. A measure that breaks down parent ratings by context or partner, such as the ICS may provide a different, and perhaps more sensitive, view of parent-reported intelligibility.[15,16] Previous research has found differences between parent ratings of intelligibility for different communication partners, with parents rating themselves as better able to understand their child compared to strangers.[16] It may be that the question we used in the current study ("How understandable is your child to others?") was too broad and different parents considered different communication partners (e.g., siblings, teachers, strangers) when answering the question. It may also be the case that our rating scale was too fine-grained: Parents might not be able to reliably differentiate 7 degrees of intelligibility, but a simpler scale with 5 or 4 points would show a stronger relationship with observed intelligibility scores.

We did not have other measures of communicative or social participation in the current study. By including measures of participation, future research should examine the relationship between parent ratings of intelligibility and parent ratings of participation.

Conclusions

There are aspects of objectively measured speech intelligibility that provide an advantage to the listener; however, there are also aspects that provide a disadvantage. Objective measurement of speech intelligibility involves child productions of controlled speech material presented to unfamiliar listeners in an idealized listening environment with high-quality audio equipment. However, the detriment of this method of measurement is the lack of interaction among partners with no context, no nonverbal cues, and no opportunities to repair communication breakdowns – variables which can help improve intelligibility and variables that parents may be considering when rating their child's understandability. Parents are often their child's best communication partner. When parents make ratings of their child's intelligibility, they may be thinking more comprehensively, considering their child's environment, interests and preferences, their child's use of nonverbal cues (e.g., gestures, facial expressions, use of alternative and augmentation communication (AAC) while speaking, frequent communication partners) – all of which can impact their estimate of their child's understandability. Parent perspective may provide important insight into functional participation level communication abilities that are not captured by direct measurement of intelligibly.

In summary, findings from the current study suggest that for children with CP, an impressionistic parent rating scale may not provide a valid index of a child's speech intelligibility, especially for those children who are not on the extreme ends of the severity continuum. For children in the mild-moderate intelligibility deficit range, a subjective parent report measure does not provide a clear picture alone – parent report may be best used in conjunction with an objective measurement of speech intelligibility as these two methods of measurement can provide a more complete picture when administered together.

Acknowledgments

The authors thank the children and families who participated in this research as well as the graduate and undergraduate students in the Wisconsin Intelligibility, Speech, and Communication Lab at the University of Wisconsin-Madison who assisted in data collection and data reduction.

Disclosure Statement

The authors have no financial relationships relevant to this article to disclose. The authors have no conflicts of interest to disclose.

Funding

This study was funded by grant R01 DC009411 from the National Institute on Deafness and Other Communication Disorders, National Institutes of Health. Support was also provided by a core grant to the Waisman Center, U54 HD090256, from the National Institute of Child Health and Human Development, National Institutes of Health.

References

1. Squires J, Bricker D, Potter L. Revision of a parent-completed developmental screening tool: ages and stages questionnaires. J Pediatr Psychol. 1997;22(3):313–28. doi:10.1093/jpepsy/22.3.313.
2. Squires J, Potter L, Bricker D. The ASQ user's guide for the ages & stages questionnaires:. A parent-completed, child-monitoring system. Baltimore, MD, USA: Paul H Brookes Publishing; 1995.
3. Emerson ND, Morrell HE, Neece C. Predictors of age of diagnosis for children with autism spectrum disorder: the role of a consistent source of medical care, race, and condition severity. J Autism Dev Disord. 2016;46(1):127–38. doi:10.1007/s10803-015-2555-x.
4. Nordahl-Hansen A, Kaale A, Ulvund SE. Language assessment in children with autism spectrum disorder: concurrent validity between report-based assessments and direct tests. Res Autism Spectr Disord. 2014;8(9):1100–06. doi:10.1016/j.rasd.2014.05.017.
5. Nordahl-Hansen A, Kaale A, Ulvund SE. Inter-rater reliability of parent and preschool teacher ratings of language in children with autism. Res Autism Spectr Disord. 2013;7(11):1391–96. doi:10.1016/j.rasd.2013.08.006.
6. Feldman HM, Dale PS, Campbell TF, Colborn DK, Kurs-Lasky M, Rockette HE, Paradise JL. Concurrent and predictive validity of parent reports of child language at ages 2 and 3 years. Child Dev. 2005;76(4):856–68. doi:10.1111/j.1467-8624.2005.00882.x.
7. Coplan J. Parental estimate of child's developmental level in a high-risk population. American J Dis Children. 1982;136:101–04.
8. Venker CE, Haebig E, Edwards J, Saffran JR, Weismer SE. Brief Report: early lexical comprehension in young children with ASD: comparing Eye-Gaze methodology and parent report. J Autism Dev Disord. 2016;46(6):2260–66. doi:10.1007/s10803-016-2747-z.
9. Kent R, Weismer G, Kent J, Rosenbek J. Toward phonetic intelligibility testing in dysarthria. J Speech Hearing Disord. 1989;54:482–99. doi:10.1044/jshd.5404.482.
10. Yorkston K, Beukelman D. A clinician-judged technique for quantifying dysarthric speech based on single-word intelligibility. J Commun Disord. 1980;13:15–31. doi:10.1016/0021-9924(80)90018-0.

11. Connolly JH. Intelligibility: A linguistic view. British J Disord Commun. 1986;21:371–76. doi:10.3109/13682828609019848.
12. Lagerberg TB, Åsberg J, Hartelius L, Persson C. Assessment of intelligibility using children's spontaneous speech: methodological aspects. Int J Lang Commun Disord. 2014;49:228–39.
13. Kent R. Speech intelligibility and communicative competence in children. In: Kaiser AP, Gray DB, editors. Enhancing children's communication: foundations for intervention. Vol. 2. Baltimore, MD, USA: Paul H. Brookes; 1993. p. 223–39.
14. Kent R, Miolo G, Bloedel S. The intelligibility of children's speech: A review of evaluation procedures. American J Speech Lang Pathol. 1994;3:81–95. doi:10.1044/1058-0360.0302.81.
15. Yorkston K, Beukelman D, Strand E, Hakel M. Management of motor speech disorders in children and adults. 3rd ed. Austin (TX): Pro Ed; 2010.
16. McLeod S, Crowe K, Shahaeian A. Intelligibility in context scale: normative and validation data for English-speaking preschoolers. Lang Speech Hear Serv Sch. 2015;46(3):266–76. doi:10.1044/2015_LSHSS-14-0120.
17. McLeod S, Harrison LJ, McCormack J. The intelligibility in context scale: validity and reliability of a subjective rating measure. J Speech Lang Hearing Res. 2012;55(2):648–56. doi:10.1044/1092-4388(2011/10-0130).
18. Coplan J, Gleason JR. Unclear speech: recognition and significance of unintelligible speech in preschool children. Pediatrics. 1988;82:447–52.
19. Lagerberg TB, Hellström A, Lundberg E, Hartelius L. An investigation of the clinical use of a single-word procedure to assess intelligibility (Swedish test of intelligibility for children) and an evaluation of the validity and reliability of the intelligibility in context scale. J Speech Lang Hearing Res. 2019;62(3):668–81. doi:10.1044/2018_JSLHR-S-18-0018.
20. Nordberg A, Miniscalco C, Lohmander A, Himmelmann K. Speech problems affect more than one in two children with cerebral palsy: swedish population-based study. Acta Paediatr. 2013;102(2):161–66. doi:10.1111/apa.12076.
21. Hustad KC, Sakash A, Broman AT, Rathouz PJ. Differentiating typical from atypical speech production in 5-year-old children with cerebral palsy: a comparative analysis. American J Speech Lang Pathol. 2019;28(2S):807–17. doi:10.1044/2018_AJSLP-MSC18-18-0108.
22. Hustad KC, Schueler B, Schultz L, DuHadway C. Intelligibility of 4-year-old children with and without cerebral palsy. J Speech Lang Hear Res. 2012;55:1177–89.
23. Dickinson H, Parkinson K, Ravens-Sieberer G, Schirripa G, Thyen U, Arnaud C, Beckung E, Fauconnier J, McManus V, Michelsen SI, et al. Self-reported quality of life of 8–12 year old children with cerebral palsy: A cross-sectional European study. The Lancet. 2007;369:2171–78. doi:10.1016/S0140-6736(07)61013-7.
24. Natzke P, Sakash A, Mahr T, Hustad KC. Measuring speech production development in children with cerebral palsy between 6 and 8 years of age: relationships among measures. Lang Speech Hear Serv Sch. 2020;1–15:882–896.
25. Palisano R, Rosenbaum PL, Walter S, Russell D, Wood E, Galuppi B. Development of the gross motor function classification system. Dev Med Child Neurol. 1997;39:214–23. doi:10.1111/j.1469-8749.1997.tb07414.x.
26. Eliasson AC, Krumlinde-Sundholm L, Rosblad B, Beckung E, Arner M, Öhrvall A-M, Rosenbaum P. The manual ability classification system (MACS) for children with cerebral palsy: scale development and evidence of validity and reliability. Dev Med Child Neurol. 2006;48:549–54. doi:10.1017/S0012162206001162.
27. Weismer G, Martin R. Acoustic and perceptual approaches to the study of intelligibility. In: Kent R, editor. Intelligibility in Speech Disorders. Philadelphia: John Benjamins Publishing Co.; 1992. p. 67–118.
28. Hustad KC, Sakash A, Natzke PE, Broman AT, Rathouz PJ. Longitudinal growth in single word intelligibility among children with cerebral palsy from 24 to 96 months of age: predicting later outcomes from early speech production. J Speech Lang Hearing Res. 2019;62(6):1599–613. doi:10.1044/2018_JSLHR-S-18-0319.
29. *TOCS+ Intelligibility Measures*. [computer program]. Edmonton: University of Alberta; 2007.
30. Hustad KC, Oakes A, Variability AK. Diagnostic accuracy of speech intelligibility scores in children. J Speech Lang Hear Res. 2015;58(6):1695–707. doi:10.1044/2015_JSLHR-S-14-0365.
31. R: A language and environment for statistical computing [computer program]. Vienna (Austria): R Foundation for Statistical Computing; 2019.
32. Christensen RHB. Analysis of ordinal data with cumulative link models—estimation with the R-package ordinal. In: R-package version; 2015. p. 28.
33. Lenth R, Singmann H, Love J, Buerkner P, Herve M emmeans: estimated marginal means, aka least-squares means (Version 1.4.7) [Computer software]. In:2020.
34. Van Doornik A, Gerrits E, McLeod S, Terband H. Impact of communication partner familiarity and speech accuracy on parents' ratings of their child for the Intelligibility in Context Scale: dutch. Int J Speech Lang Pathol. 2018;20(3):350–60. doi:10.1080/17549507.2018.1472808.
35. Choi JY, Park J, Choi YS, Y-R G, ES P. Functional communication profiles in children with cerebral palsy in relation to gross motor function and manual and intellectual ability. Yonsei Med J. 2018;59(5):677–85. doi:10.3349/ymj.2018.59.5.677.