Original Paper

Folia Phoniatr Logop 2006;58:217–228
DOI: 10.1159/000091735

Folia Phoniatrica et Logopaedica

Estimating the Intelligibility of Speakers with Dysarthria

Katherine C. Hustad

Department of Communicative Disorders and Waisman Center,
University of Wisconsin-Madison, Madison, Wisc., USA

Key Words
Cerebral palsy · Speech intelligibility · Dysarthria

Abstract

Many speakers with dysarthria have reduced intelligibility, and improving intelligibility is often a primary intervention objective. Consequently, measurement of intelligibility provides important information that is useful for clinical decision-making. The present study compared two different measures of intelligibility obtained in audio-only and audio-visual modalities for 4 different speakers with dysarthria (2 with mild-moderate dysarthria; 2 with severe dysarthria) secondary to cerebral palsy. A total of 80 college-aged listeners provided word-by-word transcriptions and made percent estimates of intelligibility which served as dependent variables. Group results showed that transcription measures were higher than percent estimates of intelligibility overall. There was also an interaction between speakers and measures of intelligibility, indicating that the difference between transcription scores and percent estimates varied among individual speakers. Results revealed a significant main effect for presentation modality, with the audio-visual modality having slightly higher scores than the audio-only modality; however, presentation modality did not interact with speakers or with measures of intelligibility. Results suggest that standard clinical measurement of intelligibility using orthographic transcription may be more consistent than the use of more subjective percent estimates.

Introduction

Intelligibility is an important clinical construct in the management of speakers with dysarthria. Yorkston et al. [1] describe speech intelligibility as a measure of functional limitation, providing global information regarding how well the speech

KARGER

Fax +41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1021–7762/06/0583–0217$23.50/0

Accessible online at:
www.karger.com/fpl

Katherine C. Hustad, PhD, Assistant Professor
Department of Communicative Disorders
University of Wisconsin-Madison, 1975 Willow Drive
Madison, WI 53706 (USA), Tel. +1 608 262 8079
Fax +1 608 262 6466, E-Mail Hustad@waisman.wisc.edu

subsystems – respiration, phonation, articulation, and resonation – work together to produce speech that is understandable to listeners under idealized circumstances. Indeed, most speakers who have dysarthria also experience reduced intelligibility to some extent [1, 2]; therefore, improving intelligibility is often viewed as a primary intervention goal [1, 3].

Although intelligibility measures do not provide comprehensive information regarding the functionality of speech across partners and contexts, they do provide useful clinical information that can be employed for several important purposes. For example, intelligibility is widely used as an index of severity of the speech disorder, and as a tool for documenting and monitoring change in functional speech performance [1]. Critical clinical decisions such as whether or not to discontinue intervention may be made on the basis of intelligibility measures. Consequently, the means by which intelligibility is measured and the accuracy of those measures is of the utmost importance. Several strategies for measuring intelligibility have been described in the literature. In general, there are two categories of measures, those that are objective and those that are subjective [4].

Objective Measure of Intelligibility

Objective measures generally involve listener transcription of words comprising target sentences produced by speakers with dysarthria, or alternately, by selection of single words from a closed set of response choices. Listener responses are then compared with speakers' target productions and are scored as either correct or incorrect. The proportion of words identified correctly relative to the total number of words possible is then computed and multiplied by 100 to yield a percent intelligibility. Standard clinical tools such as the Sentence Intelligibility Test (SIT) [5] and the Assessment of Intelligibility of Dysarthric Speech [6] employ this type of transcription measurement of intelligibility. Similarly, the Phoneme Intelligibility Test [7] employs a forced choice paradigm, whereby listeners must identify the target word from a field of other choices to characterize intelligibility. Clinically, one advantage of this type of measurement is that listeners are forced to commit to paper their perceptions of what a speaker is saying, resulting in a quantifiable measure of the integrity of the speech signal under idealized circumstances.

Subjective Measure of Intelligibility

There are several different subjective or qualitative measures that have been used to characterize speech intelligibility. These measures generally require listeners to quantify their qualitative judgments of a speaker's intelligibility by assigning a number to what they heard. Variations on this method include use of equal-appearing interval scales [2, 8–10], direct magnitude estimation (DME) [10–14], and percent estimates [4, 15]. One of these subjective measures, percent estimates, was of interest for the present study.

Percentage estimates require listener(s) to estimate the percent of words, ranging from 0 to 100, that they understood [4, 11]. While percent estimates of intelligibility have been used less frequently in the research literature, they may be used more frequently in clinical situations than DME and equal-appearing interval scales. Specifically, DME and transcription measures require clinicians to: (a) record speakers producing a corpus of words and/or sentences, (b) recruit at least 1 and preferably several individual(s) to listen to the tape and scale what s/he hears, and (c) score and/or average the results. With current caseload and productivity demands placed on clinicians, it may be difficult to devote the necessary time to make this type of measure, particularly if repeated intelligibility measures are made to document progress throughout the course of therapy. Percentage estimates may be an attractive alternative to clinicians because this type of measure is straightforward and economical, requiring no additional time, recording, or listeners outside of therapy.

Accuracy of Percentage Estimates

Research examining the accuracy of percentage estimates has shown inconsistent findings. For example, in a study examining 8 speakers with dysarthria of varying severity, Yorkston and Beukelman [4] asked listeners to estimate the percent of words that they understood following auditory presentation of speech samples and prior to transcribing sentences and words. Results showed a strong positive correlation between the two types of measures, indicating that as one measure increased, so did the other. Although the authors did not directly compare mean scores on the two measures, graphically displayed descriptive data suggest that for 5 of 8 speakers, percent estimates overestimated transcription intelligibility scores; and for 3 of 8 speakers percent estimates underestimated transcription scores. No clear pattern with regard to severity was evident in these results. In a later study examining 6 speakers with dysarthria of varying severity, Carter et al. [15] asked listeners to estimate the number of words that they understood following transcription of the sentences. Results comparing the two measures showed that listener judgments were consistent with transcription measures for 3 speakers who had relatively good intelligibility. However, results also showed that listeners overestimated intelligibility (by approximately 10%) for 3 speakers who had markedly reduced intelligibility. Taken together, the findings of these studies are equivocal, revealing no clear pattern of findings. There were, however, several important methodological differences between the two studies. In particular, Yorkston and Beukelman [4] required listeners to make subjective judgments about what they heard prior to transcription and Carter et al. [15] required listeners to make subjective judgments about what they heard following transcription.

Another explanation for the inconsistent findings within and between the two studies is individual differences among speakers. However, differences in listener behavior for individual speakers were not examined or discussed in either study. This type of fine-grained comparison has the potential to provide important insight into speaker characteristics that may be associated with certain patterns of listener performance. As part of a larger study [16] the present study sought to examine differences in listeners' ability to estimate intelligibility for individual speakers with mild-moderate and severe dysarthria following the first of a set of four listening tasks. In addition, the effects of visual information on listeners' ability to estimate intelligibility were also of interest, as clinical measures of intelligibility commonly involve isolated auditory input only (audiotape) and previous research has not examined the impact of bimodal audio-visual input on accuracy of percent estimates of intelligibility. Toward this end, the following specific research questions were addressed.

(1) Are listener estimates of intelligibility consistent with transcription intelligibility scores across speakers with dysarthria and presentation modalities? Is the pattern of results for individual speakers consistent with group results?

Table 1. Demographic characteristics of speakers

	Speaker 1	Speaker 2	Speaker 3	Speaker 4
Age, years	34	37	24	46
Medical diagnosis	athetoid diplegia	spastic diplegia	athetoid quadriplegia	spastic diplegia
Dysarthria type	hyperkinetic	spastic	mixed spastic-hyperkinetic	spastic
SIT score, %	80	75	16	30

(2) Are listeners able to estimate intelligibility scores across speakers with dysarthria more accurately when they can both see and hear speakers than when they can only hear speakers? Is the pattern of results for individual speakers consistent with group results?

Method

Subjects

Speakers with Dysarthria

Four individuals with chronic dysarthria contributed speech samples for this experiment. Speakers were evaluated by a certified speech-language pathologist who diagnosed the dysarthria type and administered the SIT [5]. Each speaker was required to: (a) have speech intelligibility between 15 and 85% as measured by the SIT [5]; (b) be a native speaker of American English; (c) be between 18 and 50 years of age; (d) be able to produce connected speech consisting of at least 8 consecutive words, and (e) be able to repeat sentences of up to 8 words in length following a verbal model. All 4 speakers had a medical diagnosis of cerebral palsy and had cognitive skills that were informally judged to be within normal limits based on educational achievement and interpersonal interactions. Speakers fell into two severity groups, mild-moderate and severe. See table 1 for demographic information for each speaker.

Listeners

Eighty listeners participated in this experiment, with 20 different listeners assigned to each of the 4 speakers. Each listener was required to: (a) pass a pure-tone hearing screening at 25 dB SPL for 250, 500, 1,000, 4,000, and 6,000 Hz bilaterally; (b) be between 18 and 30 years; (c) have no more than incidental experience listening to or communicating with persons having communication disorders; (d) be a native speaker of American English, and (e) have no identified language, learning, or cognitive disabilities per self-report. The mean age of listeners in each group ranged from 20 to 21 years. Gender was not a variable of interest; consequently, no effort was made to balance the number of male and female listeners.

Materials

Speakers produced stimuli from the Hearing in Noise Test (HINT) [17], which consists of 25 lists of 10 sentences that range in length between 4 and 7 words. Sentences are phonemically balanced, and equated for naturalness, difficulty, length, and reliability. In addition, each sentence is declarative in nature, following a predictable subject-verb-object syntactic structure. Sample sentences are as follows: 'The boy ran down the path'; 'Strawberry jam is sweet'; 'The boy fell from the window'; 'Her shoes were very dirty'. For the present study, HINT lists 1–4 were employed.

Procedures

Recording Speech Samples

Each speaker was recorded producing the full corpus of 40 HINT sentences. Digital audio and video recordings of each speaker were made in a quiet environment using a Sony TRV 900 digital camcorder, an HHb PDR 1000 Portadat recorder, and a Crown CM-312 head-mounted microphone. The microphone was positioned so that it was aligned with the jaw line of the speaker and did not obscure the speaker's face or mouth. During recording, the camcorder was positioned approximately 3 feet away from and directly in front of the speaker, focused inferiorly at mid-torso and superiorly just above the speaker's head. Speakers were seated directly in front of the same pale-colored neutral background with the same external lighting positioned to eliminate shadows during recording.

Productions of each sentence were modeled by the experimenter and speakers were asked to repeat these productions. In addition, orthographic representations of stimulus sentences were displayed on a laptop computer, which was positioned directly in front of the speakers and out of the camera view. This dual presentation was used to ensure that potential differences in visual acuity, literacy, and short-term memory did not influence the speakers' productions of the target sentences. Speakers were required to produce target sentences verbatim and were asked to repeat any sentences in which all target words were not produced or any sentences in which the order of words was incorrect relative to the target sentence. Speakers were instructed to speak 'naturally', as they would in habitual communication situations. Rate and prosody for each speaker were not controlled. All speakers were able to complete the production task with little difficulty.

Stimulus Tape Preparation

Recordings were transferred to a personal computer via a digital-to-digital sound card (S/PDIF interface) (16-bit quantization; 44.1 kHz sampling rate) for audio tape and via a Firewire (IEEE 1394) card (video = 29.97 frames per second, 720 × 480 frame size) for videotape. Audio and video files were edited to remove extraneous noises and experimenter productions. Audio samples were peak amplitude normalized to 69 dB. Adobe Premiere 5.1 (1999) (computer software) was used to associate the high-quality audio from DAT with digital video clips following standard procedures [16, 18, 19]. This was accomplished by aligning high-quality audio files from DAT with the lower-quality native audio signal from the video recording using visual inspection of the two waveforms as well as auditory-perceptual judgments obtained by listening to the two audio samples simultaneously. Following alignment of the two samples, the native audio sample was deleted, leaving only the high-quality DAT sample associated with the video of each stimulus sentence.

Listeners received written instructions via video tape regarding the experimental task. Following the instructions, they were presented with the sequence of 10 experimental sentences two times. During the first presentation of all 10 sentences, they simply viewed and/or listened to the tape. During the second presentation, there was a 14-second pause between each sentence and written instructions to transcribe the preceding sentence. For tapes containing the audio-only stimuli, video-images of the speaker were digitally removed and replaced by a plain blue screen. All other aspects of the task were identical to the video presentation modality. Videotapes for both conditions maintained first-generation quality because of the digital format used for all aspects of video processing. Audio and video signals followed the National Television System Committee broadcast quality standards.

Randomization and Counterbalancing

As part of a larger study [16], listeners completed four experimental tasks. However, data reported in the present study reflect a smaller subset of two experimental tasks, one involving audio-only presentation and one involving audio-visual presentation. Order of presentation for audio-only and audio-video tasks was counterbalanced so that half of the listeners in each speaker group completed the audio-only task first and half completed the audio-video task first. Listeners heard a different HINT list for each task. To ensure equivalency of lists, presentation order of lists within each speaker group was quasi-counterbalanced, so that data reported for each speaker represent intelligibility measures obtained across all four HINT lists.

Experimental Task

Individual listeners completed the experiment in a quiet, sound-treated environment. Listeners were seated at a desk and positioned approximately 3 feet away from a 25-inch television monitor (approximately 3 1/2 feet from the ground) with 1 external speaker (approximately 2 1/2 feet from the ground), and a digital video cassette player attached to it. The peak output level of stimulus material was approximately 65 dB SPL from where listeners were seated and was measured periodically to assure that all listeners heard stimuli at the same output level.

Listeners were told that during the experiment they would see and/or hear the same person, who has a speech impairment because of cerebral palsy. They were informed that speakers would be producing a series of grammatically correct and meaningful utterances and that their job was to write down exactly what they thought the speaker said after hearing each utterance, following the instructions provided on the video tape. After transcribing each series of 10 sentences, listeners were instructed to estimate the percent of words that they understood from the preceding list of 10 sentences. Listeners were given as much time as they needed to complete each task. Following delivery of instructions, the experimenter left the sound-treated room and entered an adjacent control room from which she controlled the digital video player and observed the listener.

Scoring and Reliability

Transcription intelligibility scores for each listener were obtained by counting the number of words identified correctly, then dividing that number by the number of words possible for each list of 10 sentences, and finally multiplying by 100 to compute percent intelligibility. Individual words were judged as incorrect or correct based on whether they matched the target word phonemically. Misspellings and homonyms were accepted as correct.

Interscorer reliability involved having a judge, who was not involved in initial scoring of intelligibility data, rescore all transcription data for 8 of the 80 listeners (2 listeners from each speaker group). The original transcription results (in percent intelligibility) for the same listener across each of the four lists were then correlated with the rescored transcription results, yielding a Pearson product-moment correlation coefficient of 0.99 across all 8 listeners, indicating a very high level of reliability for scoring accuracy.

Experimental Design

A 2 × 2 × 4 split-plot design [20] was employed for this study. The first repeated measure was modality and its two categories were audio-only and audio-visual. The second repeated measure was intelligibility measure and its two categories were transcription and percent estimate. The between-subjects factor was speaker, with a different group of listeners assigned to each of the 4 speakers.

Results

A fully factorial 2 × 2 × 4 ANOVA was performed to address the research questions of interest. The type I error rate for the experiment was controlled so that each omnibus test was allotted an alpha level of 0.01, and follow-up contrasts were allotted a combined alpha level of 0.01, which was partitioned using the Dunn-Bonferroni procedure [20, 21].

Group means for intelligibility data by measure and modality are presented in figure 1. In addition, individual speaker intelligibility data by measure and modality are presented in figure 2. Results of ANOVA across speakers showed that each of the three main effects was significant. The omnibus test for modality indicated that intelligibility values (across measures and speakers) were significantly higher (an av-

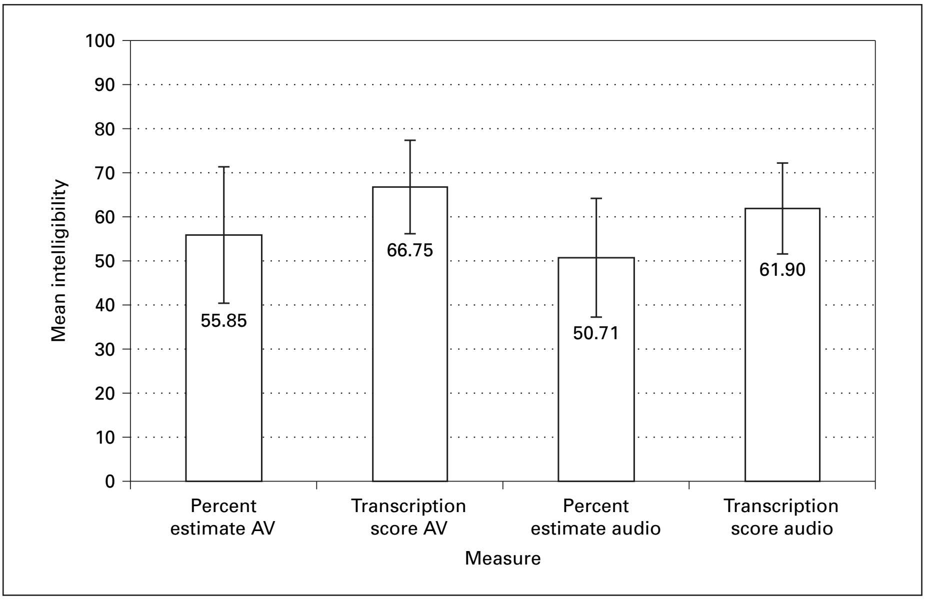

Fig. 1. Mean intelligibility scores (± 1 SD) by presentation modality (audio-only and audio-visual, AV) and intelligibility measure (transcription and percent estimate).

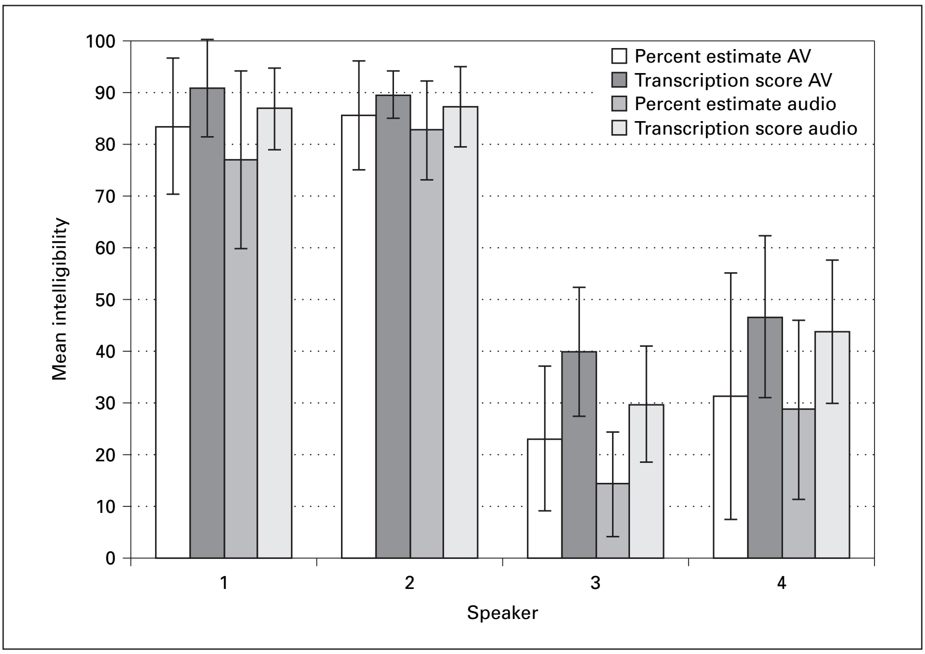

Fig. 2. Mean intelligibility scores (± 1 SD) by individual speaker, presentation modality, and intelligibility measure. AV = Audio-visual.

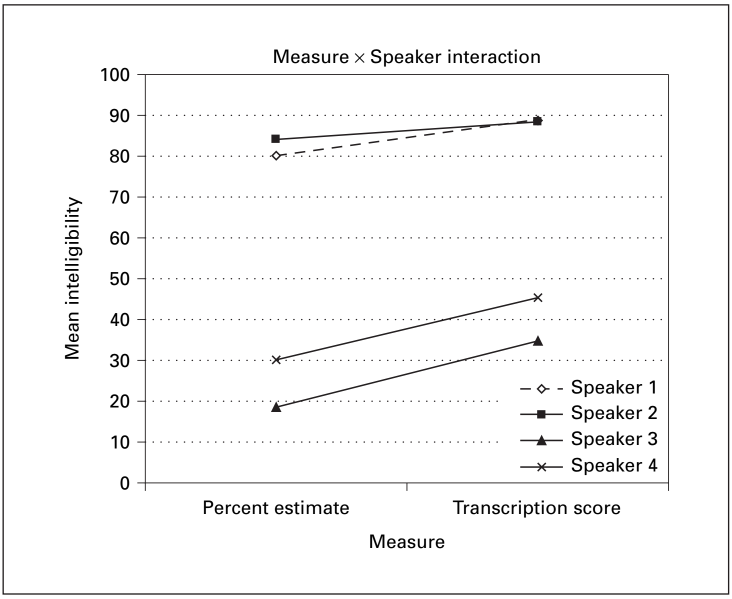

Fig. 3. Measure × speaker interaction.

erage of 5%) for the audio-visual condition than for the audio-only condition. The omnibus test for measure (across modalities and speakers) indicated that transcription intelligibility scores were significantly higher (an average of 11%) than percent estimates of intelligibility. Finally, the omnibus test for speaker (across modalities and measures) indicated that intelligibility scores varied among speakers.

Examination of two-way interactions revealed that the modality × speaker interaction was not significant, indicating that the difference between audio-only and audio-visual modalities was constant among speakers. Similarly, the modality × measure interaction was not significant, demonstrating that the difference between transcription scores and percent estimates was constant between audio-only and audio-visual modalities. Finally, the speaker × measure interaction was significant, indicating that the difference between transcription scores and percent estimates varied among speakers. Figures 3, 4, and 5 illustrate each of these two-way interactions, respectively. The three-way interaction, modality × measure × speaker, was not significant. Statistics for this ANOVA are shown in table 2.

To provide more detailed information regarding individual speakers and the speaker × measure interaction, the differences between transcription scores and percent estimates for each speaker were examined with follow-up contrasts. Results showed that for speakers 1, 3, and 4, percent estimates of intelligibility were significantly lower than transcription intelligibility scores. The magnitude of this difference was 8.6, 16.2 and 15.1%, respectively. The difference between percent estimates and transcription intelligibility scores was not significant for speaker 2. Statistics for these follow-up contrasts are shown in table 3.

Table 2. Fully factorial ANOVA for measures of intelligibility by modality and speaker

Source	d.f.	F
Between subjects		
Speaker	3	265.36*
Error	76	(290.79)
Within subjects		
Measure	1	74.00*
Measure × Speaker	3	4.78*
Error (Measure)	76	(131.69)
Modality	1	8.63*
Modality × Speaker	3	0.870
Error (Modality)	76	(231.81)
Modality × Measure	1	0.041
Modality × Measure × Speaker	3	0.320
Error (Modality × Measure)	76	(43.73)

Values in parentheses represent mean square errors.
* p < 0.01.

Table 3. Follow-up contrasts for measures of intelligibility by speaker

Contrast	Mean difference	d.f.	Standard error for contrast	t
Transcription – estimate				
Speaker 2	3.42	19	1.14	2.99
Speaker 1	8.60	19	2.22	3.88*
Speaker 3	16.15	19	2.73	5.93*
Speaker 4	15.12	19	3.09	4.90*

* p < 0.001.

Discussion

The present study compared the differences between intelligibility measures obtained via percent estimates and intelligibility scores obtained via word-by-word transcription in audio-only and audio-visual modalities for 4 different speakers with dysarthria secondary to cerebral palsy. Overall, results showed that transcription intelligibility scores were higher than percent estimates of intelligibility. However, there was an interaction between speakers and measures of intelligibility, indicating that the magnitude of the difference between transcription scores and percent estimates varied among individual speakers. See figure 3, for a graphic representation of this interaction. Results also revealed a significant main effect for presentation modality, with the audio-visual modality having higher scores than the audio-only modality; however, presentation modality did not interact with measures of intelligibility or with speakers. Findings indicated that the higher intelligibility scores associated with the visual presentation modality were consistent for both percent estimates and transcription scores. In addition, the higher intelligibility scores associated with the visual presentation modality were also consistent among all speakers.

Results of the present study provide unique information that differs from existing literature examining accuracy of percent estimates of intelligibility. Specifically, the present study found that, overall, percent estimates were consistently *lower* than transcription-based measures. Previous studies have been inconsistent, with percent estimates being higher, lower, or the same as transcription intelligibility scores, depending on the individual speaker [4, 15].

In attempting to interpret disparate findings, it is important to emphasize that speech intelligibility is a complex and multidimensional construct. Research has clearly demonstrated that intelligibility can be influenced by a number of factors, particularly linguistic-contextual variables [22] such as semantic and syntactic predictiveness of stimulus material [23–25]. In the present study, sentence stimuli were taken from the HINT [17]. Generally, these sentence stimuli have well-quantified psychometric and linguistic properties (e.g. phonemically balanced, controlled length, declarative in nature). Sentence stimuli used in Yorkston and Beukelman [4] were developed to incorporate target words from Tikofsky [26, p. 500] and 'were designed so that the target word was relatively difficult to predict without knowledge of the sentence. Carter et al. [15] employed sentences from the computerized assessment of intelligibility of dysarthric speech [27]. These sentences, taken from adult reading material, varied in length between 5 and 15 words and contained a variety of syntactic forms and syllabic structures. One potential explanation for the discrepant findings among existing studies may relate to the linguistic content of the speech stimuli produced by speakers. However, without access to the specific sentences used in the previous two studies, it is difficult to elaborate further upon this explanation.

In the present study, important individual differences among speakers were clearly present. For the 2 speakers who had severe dysarthria, percent estimates of intelligibility were approximately 15% lower than transcription scores. One reason for this underestimation may be related to the severity of the speakers' impairment. That is, perhaps listeners were able to understand so little of what the speakers said that they did not have confidence in the words that they actually did understand. This lack of confidence in the accuracy of their transcriptions may be reflected in listeners' gross underestimates of intelligibility.

For the 2 speakers whose dysarthria was less severe, the pattern of results is different. Similar to speakers 3 and 4, the percent estimates for speaker 1 were approximately 9% lower than transcription scores. However, for speaker 2, percent estimates and transcription scores were generally the same. One explanation for these differences may relate to the process of making a percent estimate. That is, listeners may have had difficulty separating an estimate of words identified correctly from an estimate of how much of the message was understood. For speakers with less severe dysarthria, it might be expected that these two indices would be similar. Another potential explanation for these differences may be related to the nature of the dysarthria affecting each speaker. Speaker 2 had spastic dysarthria and his productions were reasonably intelligible and consistent. Speaker 1, however, had hyperkinetic dysarthria which was characterized by inconsistent production of phonemes and words, and intermittent arrests of the voice, both associated with athetosis. Although listeners were able to understand most of what this speaker said, the inconsistent features of his speech may have been a variable that undermined listeners' confidence in their ability to understand him, ultimately resulting in lower percent estimates of his intelligibility. It is difficult to determine, based on published data, whether similar differences among speakers were present in the Carter et al. [15] study and the Yorkston and Beukelman [4] study.

Across the three studies comparing percent estimates with transcription intelligibility scores, several conclusions seem reasonable. First, the accuracy of percent estimates of intelligibility appears to be variable. In some cases percent estimates are consistent with transcription intelligibility scores, in other cases percent estimates are higher than transcription intelligibility scores, and in other cases percent estimates are lower than transcription intelligibility scores. The linguistic content and form of the speech stimuli and the severity and type of dysarthria of the speaker(s) are two variables that likely play a key role in the accuracy of percent estimates of intelligibility. Finally, results of the present study showed that although the audio-only modality generally yields lower intelligibility values than the audio-visual presentation modality, there was not a differential effect on either measure of intelligibility or on individual speakers. Thus, the current practice of using audiotapes to measure intelligibility continues to be valid.

Clinical Implications and Future Directions

The present study was experimental in nature and included only 4 participants with dysarthria associated with cerebral palsy. Accordingly, generalization to the clinical domain must be made with caution. However, results of the present research indicate that for 3 of the 4 speakers studied, percent estimates of intelligibility underestimated transcription scores. Considered together with existing research by Carter et al. [15] and Yorkston and Beukelman [4], results suggest that subjective percentage estimates may not be a reliable substitute for the more objective tally of words identified correctly provided by transcription intelligibility measures. Based on the relatively sparse and inconsistent available evidence (examining a total of 18 speakers across studies), it may be wise for clinicians who wish to quantify intelligibility to use transcription measures rather than percent estimates to ensure accurate and objective characterization of the functional limitation [1] experienced by speakers with dysarthria. Future research should seek to understand the impact of other variables such as listener experience, linguistic content and complexity of speech stimuli, and severity and type of dysarthria on measurement of intelligibility using both objective and subjective measures.

References

1 Yorkston KM, Beukelman DR, Strand EA, Bell KR: Management of Motor Speech Disorders in Children and Adults, ed 2. Austin, Pro-Ed, 1999.
2 Darley F, Aronson A, Brown J: Clusters of deviant speech dimensions in the dysarthrias. J Speech Hear Res 1969;12:462–496.
3 Ansel BM, Kent RD: Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. J Speech Hear Res 1992;35:296–308.
4 Yorkston K, Beukelman D: A comparison of techniques for measuring intelligibility of dysarthric speech. J Commun Disord 1978;11:499–512.
5 Yorkston K, Beukelman D, Tice R: Sentence Intelligibility Test for Macintosh. Lincoln, Communication Disorders Software. Distributed by Tice Technology Services, Lincoln, 1996.
6 Yorkston KM, Beukelman DR: Assessment of Intelligibility of Dysarthric Speech. Austin, Pro-Ed, 1981.
7 Yorkston K, Beukelman D, Tice R: Phoneme Intelligibility Test for Windows. Lincoln, Communication Disorders Software. Distributed by Tice Technology Services, Lincoln, NE, 1999.
8 Platt LJ, Andrews G, Young M, Quinn PT: Dysarthria of adult cerebral palsy. I. Intelligibility and articulatory impairment. J Speech Hear Res 1980;22:28–40.
9 Subtelney J: Assessment of speech with implications for training; in Bess F (ed): Childhood Deafness. New York, Grune & Stratton, 1977, pp 183–194.
10 Zraick R, Liss JM: A comparison of equal-appearing interval scaling and direct magnitude estimation of nasal voice quality. J Speech Lang Hear Res 2000;42:979–988.
11 Schiavetti N: Scaling procedures for the measurement of speech intelligibility; in Kent R (ed): Intelligibility in Speech Disorders: Theory, Measurement, and Management. Amsterdam, Benjamins, 1992, pp 11–34.
12 Schliesser HF: Psychological scaling of speech by students in training compared to that by experienced speech-language pathologists. Percept Mot Skills 1985;61:1299–1302.
13 Weismer G, Jeng J, Laures JS, Kent RD: Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders. Folia Phoniatr Logop 2001;53:1–18.
14 Weismer G, Laures J: Direct magnitude estimates of speech intelligibility in dysarthria: effects of a chosen standard. J Speech Lang Hear Res 2002;45:421–433.
15 Carter C, Yorkston K, Strand E, Hammen V: Effects of semantic and syntactic context on actual and estimated sentence intelligibility of dysarthric speakers; in Robin D, Yorkston K, Beukelman D (eds): Disorders of Motor Speech: Assessment, Treatment, and Clinical Characterization. Baltimore, Brookes, 1996, pp 67–87.
16 Hustad KC, Cahill MA: Effects of visual information and repeated familiarization on intelligibility of dysarthric speech. Am J Speech Lang Pathol 2003;12:198–208.
17 Nilsson M, Soli SD, Sullivan JA: Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and noise. J Acoust Soc Am 1994;95:1085–1099.
18 Hustad KC, Auker J, Natale N, Carlson R: Improving intelligibility of speakers with profound dysarthria and cerebral palsy. Augmentative Alternative Commun 2003;19:187–198.
19 Hustad KC, Jones T, Dailey S: Implementing speech supplementation strategies: effects on intelligibility and speech rate of individuals with chronic severe dysarthria. J Speech Lang Hear Res 2003;46:462–474.
20 Kirk R: Experimental Design: Procedures for the Behavioral Sciences, ed 3. Pacific Grove, Brooks/Cole Publishing, 1995.
21 Marascuilo LA, Serlin RC: Statistical Methods for the Social and Behavioral Sciences. New York, Freeman, 1988.
22 Connolly JH: Intelligibility: a linguistic view. Br J Disord Commun 1986;21:371–376.
23 Garcia J, Cannito M: Influence of verbal and nonverbal contexts on the sentence intelligibility of a speaker with dysarthria. J Speech Lang Hear Res 1996;39:750–760.
24 Hustad KC, Garcia JM: The influences of alphabet supplementation, iconic gestures, and predictive messages on intelligibility of a speaker with cerebral palsy. J Med Speech Lang Pathol 2002;10:279–285.
25 Kalikow DN, Stevens KN, Elliot LL: Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. J Acoust Soc Am 1977;61:1337–1351.
26 Tikofsky RS: A revised list for the estimation of dysarthric single word intelligibility. J Speech Hear Res 1970;13:59–64.
27 Yorkston K, Beukelman D, Traynor C: Computerized Assessment of Intelligibility of Dysarthric Speakers. Austin, Pro-Ed, 1984.