Implementing Speech Supplementation Strategies: Effects on Intelligibility and Speech Rate of Individuals With Chronic Severe Dysarthria

Katherine C. Hustad
Tabitha Jones
Suzanne Dailey
The Pennsylvania State University
University Park

A growing body of experimental research suggests that speech supplementation strategies can markedly increase speech intelligibility for individuals with dysarthria (D. Beukelman & K. Yorkston, 1977; E. Crow & P. Enderby, 1989; L. Hunter, T. Pring, & S. Martin, 1991; K. C. Hustad & D. R. Beukelman, 2001). However, studies in which speech supplementation strategies were actually implemented by speakers with dysarthria are limited, leaving their clinical efficacy unknown. The present study compared intelligibility and speech rate differences following speaker implementation of 3 strategies (topic, alphabet, and combined topic and alphabet supplementation) and a habitual (noncued) speech control condition for 5 speakers with severe dysarthria. Results revealed that combined cues and alphabet cues yielded significantly higher intelligibility scores and slower speech rates than topic cues and noncued speech. Overall, topic cues and noncued speech did not differ from one another with regard to intelligibility or speech rate. Combined cues and alphabet cues did not differ from one another with regard to intelligibility; however, speech rate was significantly different between the 2 strategies. Individual differences among speakers were generally consistent with group findings. Results were somewhat different from previous research in which strategies were experimentally superimposed on the habitual speech signal. However, findings provide evidence that alphabet cues and combined cues can have an important effect on intelligibility for speakers with severe dysarthria.

KEY WORDS: alphabet supplementation, dysarthria intervention, augmentative and alternative communication, cerebral palsy, motor speech disorders

Intelligibility is a central construct in speech-language pathology, particularly for speakers with dysarthria, who may live with chronically compromised intelligibility that impedes their ability to communicate effectively in daily life. In fact, Ansel and Kent (1992) suggested that "reduced intelligibility is the most clinically and socially important aspect of [the dysarthrias] and improved intelligibility is a fundamental goal of management" (p. 297).

Clinically, intelligibility is often regarded as an attribute of the speaker or a measure of the integrity of the speech signal (Kent, 1993; Yorkston, Beukelman, Strand, & Bell, 1999). However, from a theoretical perspective, researchers recognize intelligibility as a multifaceted

1092-4388/03/4602-0462

and interactive construct that is influenced by a myriad of variables, including the acoustic signal produced by the speaker and the linguistic–contextual knowledge of the listener (Connolly, 1986; Hustad, Beukelman, & Yorkston, 1998; Yorkston, Strand, & Kennedy, 1996).

In his model of mutuality, Lindblom (1990) discussed two sources of information, signal-dependent and signal-independent information, that affect mutual understanding between speaker and listener. Signal-dependent information refers to the acoustic signal and the extent to which the signal alone is able to convey the speaker's intended message. By definition, dysarthria involves neurologically based impairments of the speech subsystems that nearly always compromise the integrity of signal-dependent acoustic information (Yorkston et al., 1999). In contrast, signal-independent information refers to "what is in the listener's brain" (Lindblom, 1990, p. 225), or implicit knowledge. This includes such things as linguistic knowledge, paralinguistic knowledge, and experiential knowledge. Linguistic knowledge, perhaps the most important source of signal-independent information, encompasses understanding of and expectations for semantics, syntax, morphology, and phonology. Paralinguistic knowledge encompasses the ability of listeners to derive meaning from nonverbal cues including gestures, facial expression, speech-related facial cues, and eye contact. Finally, experiential knowledge encompasses life events, experiences, and world knowledge shared between speaker and listener and/or members of a cultural group. Lindblom's model of mutuality conceptualizes signal-dependent and signal-independent information on continua ranging from poor to rich. Further, this model characterizes these two sources of information as negatively correlated, so that as the acoustic signal becomes increasingly degraded, listeners must rely to a greater extent on signal-independent information to achieve mutual understanding with the speaker. Conversely, when signal-dependent information is "rich" listeners do not need to invoke signal-independent information to the same extent, because all information necessary to achieve mutual understanding with the speaker can be obtained from the acoustic signal.

Not surprisingly, research examining the contribution of intrinsic signal-independent variables to intelligibility suggests that different types of signal-independent information affect intelligibility differentially. For example, semantic predictability has been shown to have an important impact on intelligibility (Duffy & Giolas, 1974; Garcia & Cannito, 1996; Hustad & Garcia, 2002; Kalikow, Stevens, & Elliot, 1977). In addition, paralinguistic information such as iconic hand gesture, used in conjunction with speech, can dramatically improve intelligibility for speakers with dysarthria (Garcia & Cannito, 1996; Garcia & Dagenais, 1998). The effects of speech-related visual–facial cues on intelligibility are less clear. For example, research examining individuals with hearing impairment has consistently shown that intelligibility is higher when listeners can both see and hear speakers, relative to when listeners can only hear speakers (Erber, 1975; Monsen, 1983). However, research examining speakers with dysarthria has shown that visual–facial cues, exclusive of other gestures, have variable effects on intelligibility, with some speakers experiencing no benefit when listeners can both see and hear them (Garcia & Dagenais, 1998; Hunter, Pring, & Martin, 1991; Hustad & Cahill, in press).

In the present study, an extension to Lindblom's (1990) model is proposed whereby an additional type of signal-independent information is incorporated into the conceptual framework—explicit linguistic cues that are deliberately provided to enhance listener knowledge and thus improve intelligibility. The explicit linguistic cues or speech supplementation strategies (Hustad et al., 1998; Yorkston et al., 1999) of interest include alphabet cues, topic cues, and combined topic and alphabet cues that are used in conjunction with speech. These strategies and current research findings have been described in detail elsewhere (see Hustad & Beukelman, 2001, 2002). Consequently, only a brief summary is provided here.

In the alphabet supplementation strategy, listeners are provided with first letter cues for each word of a message while the speaker simultaneously produces each word (Beukelman & Yorkston, 1977). This strategy provides word-specific information, thus helping to make the message more intelligible by constraining the number of possible word choices available to the listener (Hustad & Beukelman, 2001). Only three published studies have examined the effects of alphabet cues on intelligibility of sentences. However, these studies showed that this strategy can improve sentence intelligibility between 25% and 44% (Beukelman & Yorkston, 1977; Crow & Enderby, 1989; Hustad & Beukelman, 2001) for speakers with dysarthria.

In the topic supplementation strategy, listeners are provided with a topic cue before the speaker produces the message. Like alphabet cues, topic cues also serve to constrain listener expectations for forthcoming messages. However, topic cues provide listeners with information that is broader in nature. A small body of research suggests that nonspecific orthographic topic cues can increase sentence intelligibility between 5% and 10% (Carter, Yorkston, Strand, & Hammen, 1996; Dongilli, 1994; Garcia & Cannito, 1996; Hustad & Beukelman, 2001) for individuals with dysarthria.

Finally, in the combined supplementation strategy (topic and alphabet cues presented simultaneously), listeners are provided first with a topic and then with first letter cues for each word of the message as the speaker produces it. Thus, the listener receives both broad contextual information and narrow word-specific information pertaining to the content of the speaker's message. The one study that examined the effects of combined cues on sentence intelligibility showed that intelligibility improved by an average of 33%, relative to a no-cues control condition, when listeners were presented with combined cues (Hustad & Beukelman, 2001).

Although several studies have compared an individual speech supplementation strategy with a control condition, only one study has directly compared the effects of each of these three strategies on sentence intelligibility. Hustad and Beukelman (2001) examined the effects of topic, alphabet, and combined cues on the intelligibility of 4 speakers with severe dysarthria. In this study, speakers produced target sentences using habitual speech. Linguistic contextual information for each of three cue conditions was then experimentally superimposed onto the habitual speech signal. Of particular importance, listeners were only able to see target linguistic–contextual cues, not the faces of the speakers in this study. Results showed that the combined cues condition yielded higher intelligibility scores than in any other cue condition, no cues yielded lower intelligibility scores than in any other cue condition, and alphabet cues yielded higher intelligibility scores than topic cues. Relative to the same speech signal without any associated cues, intelligibility increased 36% for combined cues, 24% for alphabet cues, and 12% for topic cues. With regard to the Lindblom (1990) model, the results of Hustad and Beukelman (2001) demonstrated that signal-independent information, in the form of explicit cues, can have an independent and significant effect on sentence intelligibility. However, other research suggests that when speakers actually implement alphabet cues, speech rate is markedly reduced and this signal-dependent reduction in rate alone makes a contribution to improvements in intelligibility (Beukelman & Yorkston, 1977; Crow & Enderby, 1989). Indeed, the findings of Beukelman and Yorkston (1977) suggest that the joint effects of rate reduction and alphabet cues may be additive, resulting in greater intelligibility improvements than either in isolation (an average gain of 44% relative to habitual speech). Further research is necessary to verify this, as Beukelman and Yorkston (1977) studied only 2 speakers and 5 listeners.

From a clinical perspective, understanding how speech supplementation strategies affect intelligibility when they are actually implemented by speakers with dysarthria is critical for clinical decision making. The aim of the present study was to replicate the findings of Beukelman and Yorkston (1977) and to extend the findings of Hustad and Beukelman (2001) by incorporating into the methodology signal-dependent changes that occur when speakers actually implement alphabet, topic, and combined cues. Further, the present study sought to address the Lindblom (1990) model of mutuality by characterizing the joint effects of alterations in signal-dependent information and signal-independent explicit information relative to noncued (habitual) speech. Specifically, the research questions of interest pertained to whether differences exist within the dependent variables of (a) intelligibility and (b) speech rate when speakers actually implement alphabet cues, topic cues, and combined cues, relative to a noncued speech control condition. Also of interest for the present study were individual differences among speakers with respect to the pattern and magnitude of change in intelligibility and rate associated with each condition.

Following the results of Hustad and Beukelman (2001), we hypothesized that alphabet cues, topic cues, and combined cues would all significantly increase intelligibility relative to noncued speech. More specifically, we expected that combined cues would result in higher intelligibility scores than any other cue condition, that noncued speech would result in lower intelligibility scores than any other cue condition, and that alphabet cues would yield higher intelligibility scores than topic cues. Further, we hypothesized that the magnitude of the effects for combined cues and alphabet cues would be greater for the present study than those observed by Hustad and Beukelman (2001), because overall speech rate would be slower due to speaker implementation of strategies. We expected the magnitude of the effect of topic cues to be similar to that in previous research. We anticipated that individual speakers would show the same general pattern of results as that shown by the group data.

With regard to speech rate changes associated with implementation of strategies, we hypothesized that rate would be markedly slower for combined cues and alphabet cues than for topic cues and noncued speech because of letter identification and pointing requirements for the former two strategies. We expected the difference between combined and alphabet cues and the difference between noncued speech and topic cues to be nonsignificant. We expected the magnitude of rate change for alphabet and combined cues among individual speakers to be diverse because of the different motor skills of each speaker. However, the pattern of results was expected to be similar among speakers.

Method

Participants

Two types of participants were involved in this study: speakers with dysarthria and listeners. Speakers were recorded producing a standard corpus of speech stimuli using three speech supplementation strategies (topic cues, alphabet cues, and combined cues) and noncued speech. Listeners viewed videotapes of speakers and transcribed what they heard.

Speakers With Dysarthria

Five individuals with severe dysarthria served as speakers. Four had dysarthria secondary to cerebral palsy, and one had dysarthria secondary to a traumatic brain injury. Speaker demographics, including type of dysarthria and distribution of motor impairment, are presented in Table 1. Inclusion criteria required that each speaker (a) be able to produce connected speech consisting of at least eight consecutive words; (b) have speech intelligibility between 15% and 30%, as measured by the Sentence Intelligibility Test (SIT; Yorkston, Beukelman, & Tice, 1996); (c) use speech as a mode of communication; (d) be a native speaker of American English; (e) have functional literacy skills at or above the 6th grade level; (f) have corrected or uncorrected vision within normal limits per self-report; (g) have hearing within normal limits per self-report; and (h) be able to accurately direct select letters and orthographically represented phrases from a communication board.

Listeners

One hundred twenty individuals served as listeners. Twenty-four different listeners were randomly assigned to each of the five speakers and viewed the same speaker in each of the four experimental conditions (no cues, topic cues, alphabet cues, and combined cues). Inclusion criteria required that each listener (a) pass a pure tone hearing screening at 25 dB SPL for 250 Hz, 500 Hz, 1 kHz, 4 kHz, and 6 kHz bilaterally; (b) be between 18 and 35 years of age; (c) have no more than incidental experience listening to or communicating with persons having communication disorders; (d) be native speakers of American English; and (e) have no identified language, learning, or cognitive disabilities per self-report.

All listeners were either currently attending college or graduate school or had completed college or graduate school. As such, college-level literacy skills were assumed. The mean age of listeners in each group ranged from 20.46 to 21.43 years. Gender composition was 30 men and 90 women. Gender was not a variable of interest; consequently, no effort was made to balance the number of male and female listeners.

Materials

Four narrative passages and their associated topic cues, described in Hustad and Beukelman (2001, 2002), were used as speech stimuli. Briefly, passages consisted of 10 syntactically and semantically predictable sentences that followed standard American English conventions for content, form, and use. Sentences systematically varied in length between 5 and 8 words, with each passage containing a total of 65 words and 90–92 syllables. In addition, each passage had a 6th grade reading level. Passages used in the present study pertained to a sporting event, a natural disaster, purchasing a vehicle, or Independence Day.

Procedures

Data Collection From Speakers With Dysarthria

Speakers with dysarthria completed four experimental tasks, which took approximately 5 hr (including

Table 1. Perceptual speech characteristics of participants with dysarthria.

Speaker	1	2	3	4	5
Age (years)	33	58	33	42	42
Gender	M	F	F	F	M
Speech diagnosis	Mixed spastic-hyperkinetic dysarthria	Spastic dysarthria	Mixed spastic-ataxic dysarthria	Spastic dysarthria	Spastic dysarthria
Medical diagnosis	CP Mixed spastic-athetoid quadriplegia	CP Spastic quadriplegia	CP Mixed spastic-ataxic diplegia	CP Spastic diplegia	TBI Spastic quadriplegia
Selection method	Direct selection via head pointer	Direct selection via hand	Direct selection via hand	Direct selection via hands	Direct selection via hands
Comm. strategies	Speech as sole mode	Speech as sole mode	Speech and voice output comm. device	Speech as sole mode	Speech as sole mode
Education	Completed 1 year of college	Completed HS equivalent	Completed 2 years of college	Completed HS	Completed HS
SIT score (%)	20	26	20	27	22

Note. CP = cerebral palsy; TBI = traumatic brain injury; comm. = communication; HS = high school; SIT = Sentence Intelligibility Test.

breaks), during one day. In the experimental tasks, speakers produced all four passages using alphabet cues, topic cues, combined cues, and noncued (habitual) speech. Each speaker completed the four experimental tasks in a different order to prevent the possibility of an order effect associated with learning the different strategies. In addition, the four passages were presented in a different order within each experimental task and among speakers to prevent a familiarization effect with the narrative passages.

Strategy Learning. Prior to recording the experimental narrative passages, speakers were instructed in the use of each target strategy. Instruction involved a verbal description of the strategy and its purpose, and modeling of the strategy. Speakers practiced using the strategy on a set of rehearsal sentences, which were similar to the experimental passages, until they were able to use the strategy comfortably and accurately.

For the topic cues condition, strategy use involved correctly pointing to the predetermined topic of each utterance on a premade communication board prior to speaking the utterance. For the alphabet cues condition, strategy use involved correctly pointing to the first letter of each word while speaking the word. The timing of letter selection and speech production was controlled in the alphabet cues condition so that speakers selected the letter and then either simultaneously produced the target word or subsequently produced the target word immediately afterward. For the combined cues condition, speakers pointed to the topic of each sentence and then pointed to the first letter of each constituent word, following the same requirements as those for alphabet cues and topic cues. All 5 speakers had rapid success with learning each of the strategies and all were able to implement strategies without difficulty while producing the experimental passages.[1] Learning time prior to recording experimental passages was less than 15 min per strategy for each speaker.

Recording the Speakers. Recordings of speakers were made in a quiet environment in each speaker's home. Recordings were made using a Canon XL-1 digital camcorder, a Tascam DA-P1 digital audiotape (DAT) recorder, and a Sony ECM-77B lapel microphone. Speakers were seated in front of a chroma blue background for all recordings, and a Cool Lux U30001 Tri-Light, mounted above the videocamera, was used for lighting. A laptop computer, used for presenting speech stimuli, was positioned directly in front of each speaker and out of the camera's view. Video recordings focused on the speakers' upper body so that a lap-mounted communication board as well as facial features were clearly visible.

[1]Speaker 5 was unable to complete the combined cues condition due to fatigue—this condition was presented as the last experimental task for him.

During recording, speakers were required to use each strategy with 100% accuracy for topic and letter selection as well as verbatim production of all words within each sentence per the written script shown on the laptop computer and an auditory model produced by an experimenter. Speakers were asked to repeat any sentence in which they selected an inappropriate first letter or topic, any sentence in which they spoke the word before indicating the first letter, and any sentence in which they did not produce the constituent words exactly as written. Repetition was necessary for fewer than 10% of the sentences comprising the narrative passages across speakers.

Constructing Stimulus Tapes. Digital video recordings were transferred to a personal computer via a Firewire (IEEE 1394) card, maintaining the sampling rate and frame size of the original recording (video = 29.97 frames per second, 640 × 480 frame size). Video recordings were edited using Adobe Premiere 6.0 computer software for Macintosh. Editing involved separating digital recordings of the four narrative passages produced by each speaker in each speaking condition into constituent sentences (40 per speaker per cue condition) for sequencing into experimental tapes.

DAT recordings (16-bit quantization; 48000 Hz sampling rate) were transferred to a personal computer via a digital-to-digital sound card. Audio samples for each narrative passage produced by each speaker in each speaking condition were then segmented into individual sentences. To ensure that maximum loudness levels of the recorded speech stimuli were consistent across speakers and sentences, waveforms for individual sentences were normalized using SoundForge 4.5 (Sonic Foundry, 1998). Normalization involved adjustment of the peak amplitude of each sentence waveform to a constant level. All other amplitude values in each sentence waveform were then adjusted by the same amount as the peak, thus preserving the dynamic range of the speech signal for each sentence.

Although an audio track was recorded simultaneously with the video track on the digital videotape, precise control of audio recording levels was difficult, particularly with dysarthric speakers who often had marked variability in their loudness. Consequently, it was preferable to use audio samples collected via DAT, which would allow for greater experimental control of the quality of speech samples. Use of DAT samples necessitated manually matching a separate audio track with the digital video via Adobe Premiere. Matching was accomplished through visual inspection of the two waveforms (original from the digital video camera and higher quality from DAT) along with auditory–perceptual judgments obtained by listening to the two audio samples simultaneously to ensure perfect alignment. Following alignment of the two samples, the original audio sample was deleted, leaving only the high quality DAT sample associated with the video of each stimulus sentence.

Because the videotapes were filmed from directly in front of the speaker, it was difficult to see clearly the topic and/or letter cue to which the speaker was pointing on the videotape. To compensate for this, the videotapes were digitally enhanced so that the cues were clearly visible, as they might be if a listener was sitting next to the speaker. For the topic cues condition, the topic of each sentence was represented orthographically in a box to the right of the speaker's face on the videotape and was shown for a duration of 3 s immediately prior to the onset of speech and corresponding approximately with the pointing gesture of the speaker. Similarly, for the alphabet cues condition, the first letter of each word was represented in a box to the right of the speaker's face on the videotape. The onset of each grapheme corresponded to the physical pointing gesture of the speaker and was displayed for the duration of the target word, as indicated by visual inspection of the speech waveform. For combined cues, the topic was presented for a duration of 3 s prior to the onset of speech and corresponding with the pointing gesture of the speaker; then the first letter of each word was presented following the speaker's pointing gesture and lasting for the duration of the word.[2]

Data Collection From Listeners

Presentation of Stimuli. Listeners viewed the broadcast-quality digital videotapes individually in a quiet, sound-treated room. During the experiment, individual listeners were seated at a desk and positioned approximately 3 feet away from a 25-in. television monitor with one external speaker 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 ensure that all listeners heard stimuli at the same output level.

Administration Instructions. The experimenter explained to listeners that they would complete four different listening tasks lasting a total of about 60 min. All tasks would involve the same individual, who has a speech impairment. In each task, the speaker would be producing a different set of grammatically correct and meaningful sentences that formed a 10-sentence short story. In one task listeners would see the speaker pointing to the topic of the story prior to producing each sentence, in another task they would see the speaker pointing to the first letter of each word while simultaneously talking, in another task they would see the speaker pointing to the topic and the first letter of each word while talking, and in another task they would see the speaker talking without any strategies. The experimenter told the listeners that the purpose of the study was to determine whether particular kinds of information helped people, like themselves, understand the speaker better. In addition, the experimenter informed the listeners that there would be two consecutive presentations of each story. During the first presentation, they were to watch and listen without writing anything down. During the second presentation of the same 10-sentence story, listeners were to follow the instructions on the videotape, which directed them to write down exactly what they thought the speaker said and take their best guess if they were unsure. Finally, the experimenter explained that she would be controlling the videotape from an adjacent control room and that listeners could take as much time as necessary to write their responses.

Randomization and Counterbalancing. To prevent an order effect and/or a learning effect, the order of presentation of the four experimental conditions (no cues, topic cues, alphabet cues, and combined cues) was counterbalanced so that individual listeners in each speaker group viewed the cue conditions in a different sequence. Each of the four narrative passages was presented in only one experimental condition, so that all listeners heard four unique narratives across the four experimental conditions. Furthermore, assignment of individual narratives to the four experimental conditions was evenly distributed across speaker groups and cue conditions, so that averages for each condition reflected listener performance across all narratives.

Scoring and Reliability

Transcriptions from each listener were scored by one of the experimenters, who tallied the number of words identified correctly on the basis of whether they matched the target word phonemically (misspellings and homonyms were accepted as correct). This number was then divided by the number of words possible and multiplied by 100 to yield a percent intelligibility score for each task. Because scoring of transcription data required some subjective judgment, the reliability of experimenter scoring was verified by an individual who was not involved in initial scoring. This person rescored all transcription data for 10 of the 120 listeners (2 listeners from each speaker group). The original transcription results (in percent intelligibility) were then correlated with the rescored transcription results, yielding a Pearson product–moment correlation coefficient of .95 across all 10 listeners.

Speech rate data were obtained using a wideband spectrographic display to view all utterances produced by each of the 5 speakers. Measures of the total utterance duration, defined as the time between the onset of speech-related acoustic energy and the offset of speech-related acoustic energy, were made for each utterance. Duration values were converted to words per minute by dividing

[2]For Speaker 5, the combined cues condition was created digitally using videotape from the alphabet cues condition.

the number of words in each utterance by the total duration of the utterance and multiplying by 60. Interjudge reliability involved the same procedures as those described for intelligibility. A second judge independently measured the durations of two utterances from each speaking condition for each speaker (40 utterances total). These measures were correlated with the original measures, yielding a Pearson product–moment correlation of .93 across all 40 utterances.

Experimental Design

This study used a 4 × 5 split plot design (Kirk, 1995) for each of the two dependent measures, intelligibility and speech rate. The within-subjects measure for both dependent variables was cue condition and its four categories were noncued speech, topic supplementation, alphabet supplementation, and combined supplementation. For intelligibility data, the between-subjects measure was speaker group, with a different group of 24 listeners assigned to each of the 5 speakers. For rate data, the between-subjects measure was the individual speakers' productions (40 sentences per speaker per condition).

Results

Research questions of interest were specified a priori; therefore, a planned contrast approach to analysis of variance (ANOVA) was used in which only the contrasts of interest were tested (Hertzog & Rovine, 1985; Kirk, 1995; Marascuilo & Levin, 1983; Marascuilo & Serlin, 1988; Seaman, Levin, & Serlin, 1991). This approach is considered more conservative than the traditional ANOVA, as fewer tests are performed, thus reducing the probability of Type I error. For intelligibility data, an alpha of .05 was partitioned among 36 contrasts of interest (the same six contrasts were tested for group data and for individual speaker data) via the Dunn Bonferroni procedure (Marascuilo & Serlin, 1988). This yielded a per contrast alpha level of .001. The same procedures were followed for tests involving speech rate data.

Intelligibility

Descriptive statistics for intelligibility data are presented in Table 2, and group means are displayed graphically in Figure 1. Mean intelligibility results across speakers showed that combined cues resulted in significantly higher intelligibility scores than noncued speech and topic cues. Alphabet cues also resulted in significantly higher intelligibility scores than topic cues and noncued speech. However, the differences between noncued speech versus topic cues and alphabet cues versus combined cues were not significant. Statistics for contrasts are shown in Table 3.

The same six pairwise contrasts were examined for each individual speaker to determine whether differences in the pattern of results were present. Results for each

Table 2. Intelligibility (Intell.) data (% of words identified correctly) and speech rate data (in words per minute) for individual speakers by cue condition.

Cue condition	Speaker 1		Speaker 2		Speaker 3		Speaker 4		Speaker 5		*M*	
	Intell.	Rate	Intell.	Rate	Intell.	Rate	Intell.	Rate	Intell.	Rate	Intell.	Rate
NC												
M	20.92	52.91	40.17	102.29	36.56	58.71	60.67	93.23	40.67	31.92	39.79	67.82
SD	10.56	11.31	12.21	31.05	13.25	12.78	16.77	23.49	23.36	14.82	20.15	32.97
Range	6–48	29–76	17–65	51–167	11–57	34–87	18–83	53–161	8–89	17–96	6–89	17–167
TC												
M	21.21	63.44	44.79	97.64	37.34	57.51	60.17	88.60	48.21	18.57	42.34	67.84
SD	12.48	10.45	11.99	22.83	14.74	12.87	14.99	23.74	22.61	32.04	20.24	29.70
Range	50–49	44–92	25–72	56–137	11–66	33–87	20–85	47–169	22–94	18–117	5–94	18–169
AC												
M	60.96	18.28	66.42	18.42	69.44	17.72	87.08	27.07	76.63	11.09	71.90	18.31
SD	20.36	2.91	19.64	2.21	17.31	3.16	8.24	3.93	17.51	2.78	19.23	5.96
Range	22–95	15–27	31–89	15–26	40–91	12–26	66–98	20–40	37–94	7–19	22–89	7–40
CC												
M	70.79	17.74	67.12	18.69	76.57	19.86	89.50	31.86	76.96	—	76.19	22.04
SD	18.86	3.05	18.58	3.18	12.32	4.31	10.89	5.05	16.11	—	17.22	6.96
Range	28–97	13–24	28–91	13–28	54–97	11–29	57–98	23–47	48–100	—	28–100	11–47

Note. NC = no cues; TC = topic cues; AC = alphabet cues; CC = combined cues. Dashes represent missing data (see footnote 1).

Figure 1. Mean percent intelligibility (±*SD*) across speakers by cue condition.

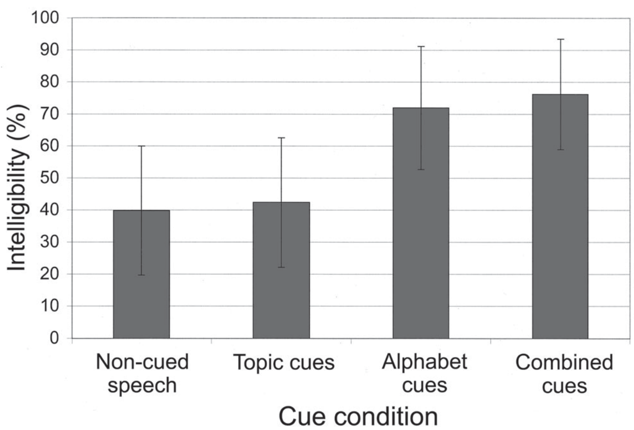
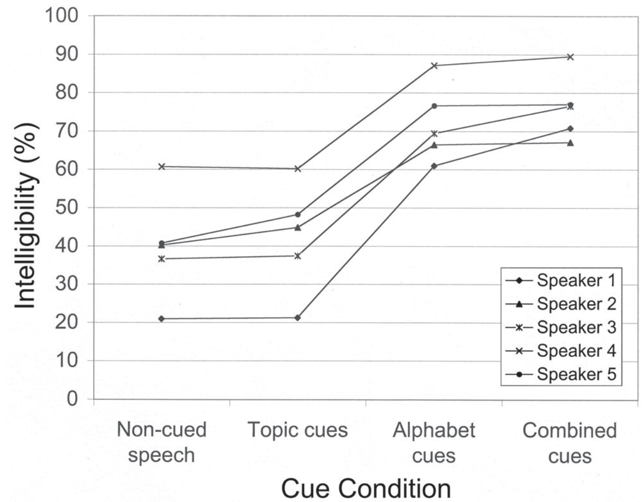


Figure 2. Mean percent intelligibility for individual speakers by cue condition.



speaker were the same as the group findings. Combined cues resulted in higher intelligibility scores than noncued speech and topic cues, alphabet cues resulted in higher intelligibility scores than topic cues and noncued speech, and the differences between noncued speech versus topic cues and alphabet cues versus combined cues were not significant. A graphical display of intelligibility scores for each speaker and cue condition is provided in Figure 2.

Speech Rate

Descriptive statistics for rate data are presented in Table 2, and group means are displayed graphically in Figure 3. Mean rate results across speakers showed that combined cues resulted in a slower speech rate than noncued speech and topic cues; however, combined cues resulted in a slightly faster speech rate than alphabet cues.[3] In addition, alphabet cues resulted in a slower speech rate than noncued speech and topic cues. The difference between topic cues and noncued speech was not significant. Statistics for these contrasts are shown in Table 3.

The same six pairwise contrasts were examined for individual speakers, and findings for some speakers differed from the group results. For each speaker, combined cues resulted in a slower rate than noncued speech and topic cues, and alphabet cues resulted in a slower rate than noncued speech and topic cues. The difference between combined cues and alphabet cues was significant for Speakers 3 and 4, with combined cues resulting in a slightly faster speech rate than alphabet cues for both speakers. Finally, the difference between topic cues and noncued speech was significant only for Speaker 1, whose rate declined when he implemented topic cues. A graphical display of speech rate data for each speaker and cue condition is provided in Figure 4.

[3]Speech rate data for Speaker 5 in the combined cues condition were the same as speech rate data from the alphabet cues condition because combined cues were digitally created for this speaker. Therefore, speech rate data associated with combined cues for Speaker 5 were not included in these analyses.

Discussion

The present study examined the effects of speaker-implemented speech supplementation strategies on intelligibility and speech rate for 5 speakers with severe dysarthria. This study is the first of its kind to examine the effects of actual implementation of all three strategies (topic cues, alphabet cues, and combined cues) and thus provides preliminary clinical efficacy data.

Intelligibility

Results of the present study were consistent with the findings of Hustad and Beukelman (2001) in that combined cues and alphabet cues resulted in higher intelligibility scores than topic cues and noncued (habitual) speech. However, the present study differed from previous research in that intelligibility scores associated with implementation of topic cues did not differ from those associated with noncued speech, and intelligibility scores associated with combined cues did not differ from those associated with alphabet cues. This same pattern of results was present for each individual speaker as well as in group data across speakers, suggesting that this was a robust effect.

Perhaps the most compelling explanation for the discrepant findings between the current study and previous research is the impact of speaker implementation of the

Table 3. Statistical contrasts for individual speaker intelligibility (Intell.) data (% correct) and rate data (in words per minute).

	Mean difference		*df*		*SE* for contrast		*t*	
Contrast	Intell.	Rate	Intell.	Rate	Intell.	Rate	Intell.	Rate
All speakers								
CC–NC	36.39	−54.76	119	159	2.00	2.27	18.18*	−24.17*
CC–TC	33.84	−54.76	119	159	1.88	1.89	18.05*	−28.96*
CC–AC	4.28	1.92	119	159	1.83	0.38	2.33	5.07*
AC–TC	29.57	−49.53	119	199	1.89	1.89	15.59*	−26.29*
AC–NC	32.11	−49.51	119	199	2.06	2.13	15.58*	−23.23*
TC–NC	2.55	0.02	119	199	1.69	1.28	1.50*	0.02
Speaker 1								
CC–NC	49.88	−35.16	23	39	4.38	1.89	11.39*	−18.52*
CC–TC	49.58	−45.69	23	39	3.49	1.77	14.21*	−25.77*
CC–AC	9.83	0.54	23	39	4.31	0.44	2.28	1.22
AC–TC	39.75	−45.15	23	39	4.39	1.79	9.05*	−25.23*
AC–NC	40.04	−34.62	23	39	4.69	1.88	8.54*	−18.35*
TC–NC	0.29	10.53	23	39	3.02	1.43	0.09	7.36*
Speaker 2								
CC–NC	29.96	−83.59	23	39	4.23	4.81	6.37*	−17.39*
CC–TC	22.33	−78.94	23	39	4.75	3.48	4.71*	−22.66*
CC–AC	1.71	0.28	23	39	5.97	0.56	0.27	0.50
AC–TC	20.62	−79.22	23	39	4.76	3.66	6.37*	−21.65*
AC–NC	25.2	−83.87	23	39	5.35	5.00	4.72*	−16.76*
TC–NC	4.63	−4.65	23	39	3.46	4.36	1.34	−1.07
Speaker 3								
CC–NC	40.01	−38.84	23	39	3.82	1.96	10.46*	−19.82*
CC–TC	39.23	−37.65	23	39	3.02	2.05	12.98*	−18.32*
CC–AC	7.13	3.15	23	39	3.41	0.78	2.09	4.05*
AC–TC	32.09	−40.79	23	39	4.01	2.06	8.00*	−19.79*
AC–NC	32.88	−41.99	23	39	4.51	2.08	7.29*	−20.19*
TC–NC	0.78	−1.19	23	39	3.84	1.76	0.20	−0.68
Speaker 4								
CC–NC	28.83	−61.43	23	39	4.09	3.79	7.05*	−16.21*
CC–TC	29.33	−56.75	23	39	3.29	3.75	8.90*	−15.12*
CC–AC	2.42	4.79	23	39	2.36	0.86	1.03	5.53*
AC–TC	26.92	−61.53	23	39	3.06	3.48	8.79*	−17.66*
AC–NC	26.42	−66.22	23	39	3.49	3.54	7.59*	−18.73*
TC–NC	0.50	−4.68	23	39	3.96	3.39	0.13	−1.38
Speaker 5								
CC–NC	36.29	—	23	—	4.51	—	8.05*	—
CC–TC	28.75	—	23	—	4.06	—	7.09*	—
CC–AC	0.33	—	23	—	3.57	—	0.09	—
AC–TC	28.42	−20.94	23	39	4.12	2.80	6.89*	−7.48*
AC–NC	28.42	−20.83	23	39	4.12	2.18	6.89*	−9.54*
TC–NC	7.54	0.11	23	39	4.59	1.61	1.64	0.07

Note. NC = no cues; TC = topic cues; AC = alphabet cues; CC = combined cues. Dashes represent missing data (see footnote 1).
*$p < .001$.

strategies. It is unlikely, though, that actual implementation was responsible for the different effects of strategies involving topic cues (alone and in combination with alphabet cues) between the two studies. Findings from the present study showed that speaker implementation of combined and alphabet cues resulted in marked global speech rate reductions relative to topic cues and noncued speech. However, differences in rate between alphabet cues

Figure 3. Mean speech rate in words per minute (±*SD*) across speakers by cue condition.

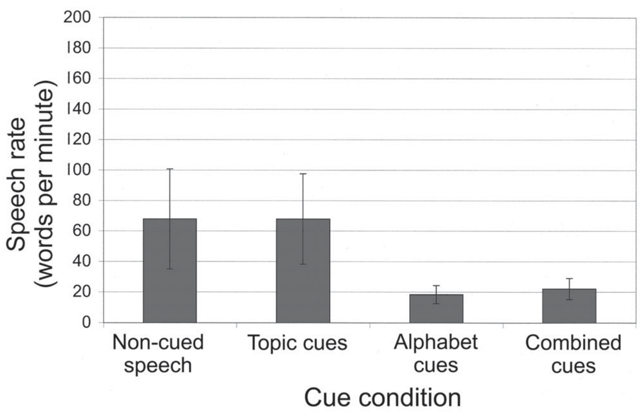
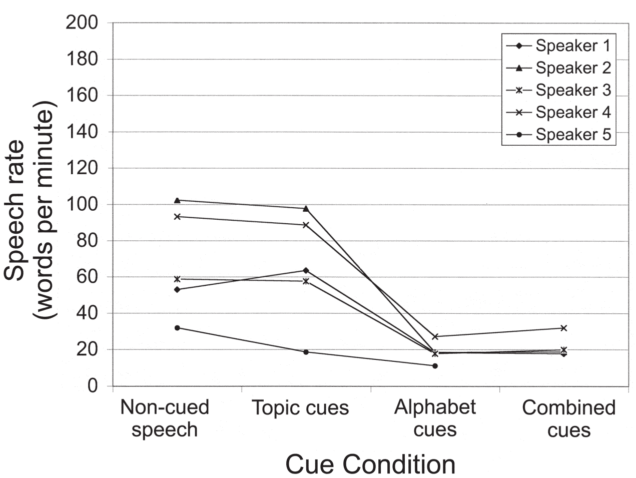


Figure 4. Mean speech rate in words per minute for individual speakers by cue condition.



versus combined cues and topic cues versus noncued speech did not appear to be clinically meaningful (less than two words per minute). Consequently, rate changes would not likely be a reasonable explanation for the pattern of results in the present study.

In addition to speaker implementation of strategies, there was another important difference between the present study and Hustad and Beukelman (2001). In the present study, listeners were able to see speakers implementing strategies, whereas in the previous study listeners saw only the linguistic information associated with the cue conditions. Lindblom's (1990) model of mutuality would predict that intelligibility would generally be higher when listeners can see a speaker with dysarthria because additional signal-independent information (i.e., visual information associated with speech production) is present. The literature regarding the impact of visual information on speech intelligibility in dysarthria, however, is equivocal (Garcia & Cannito, 1996; Garcia & Dagenais, 1998; Hustad & Cahill, in press). In a recent study, using methodology and speakers similar to those in the present study, Hustad and Cahill (in press) found that intelligibility did not differ for 4 of 5 speakers with dysarthria when listeners were presented with audiovisual information relative to audio-only information. It is, therefore, difficult to speculate with any degree of certainty about the contribution of visual information to the findings of the present study.

Another variable that may have influenced the pattern of results is individual differences among speakers between the present study and its predecessor, Hustad and Beukelman (2001). In particular, the speakers in Hustad and Beukelman were somewhat more severely involved than those in the present study. Although SIT (Yorkston, Beukelman, & Tice, 1996) scores were similar for speakers in the two studies (between 20% and 30% for the present study; between 15% and 25% for the previous study), group data for the present study indicated that intelligibility scores for the noncued speech condition were nearly double those for the analogous condition in the Hustad and Beukelman (2001) study. The effect of severity may be partially responsible for the failure of topic cues (alone or in combination with alphabet cues) to enhance intelligibility. Again, following Lindblom (1990), because speakers in the present study were generally more intelligible than those in the previous study, listeners may have been able to understand speech well enough to infer the topic and, therefore, did not gain any advantage from explicit presentation of topic cues. In the previous study, reduced intelligibility along with reduced signal-independent information may have increased the impact of topic cues, ultimately resulting in significant increases in intelligibility for the two conditions in which topic cues were provided.

Regarding magnitude of benefit from speech supplementation strategies, the present study generally showed greater intelligibility increases than Hustad and Beukelman (2001) for conditions involving alphabet and combined cues. In addition, the magnitude of benefit from alphabet cues in the present study was generally similar to the findings of Beukelman and Yorkston (1977), particularly for the 2 speakers in the present study who had similar habitual intelligibility to those in the Beukelman and Yorkston (1977) study. Presumably, the greater magnitude of benefit in studies where speakers implemented strategies was due to the joint effects of reduced speech rate and explicit signal-independent information. Additional research is necessary to examine the independent contribution of speech rate changes to supplemented intelligibility.

Two other related and noteworthy observations pertaining to intelligibility results from the present study warrant some discussion. First, combined cues increased intelligibility to an average of 76% across speakers. Two of the speakers had intelligibility increases to approximately 75% and one had an increase to almost 90%. To date there has been little discussion of the issue of a minimal intelligibility score that would be considered functional for spoken communication. Perhaps this is the case because *functional* is likely to vary on the basis of many signal-independent factors discussed previously, particularly those associated with experiential knowledge. However, Beukelman and Yorkston (1979) suggested that information transfer, or the ability of listeners to derive meaning from a spoken message, may be optimized when intelligibility is 80% or above. Following this suggestion, combined or alphabet cues may be a viable and comprehensive intervention for several of the speakers in the present study. For speakers whose intelligibility increased to approximately 70%, other augmentative means of communication would likely be necessary, at least in some situations, to enable the partners of these individuals to derive meaning from their spoken messages.

Another important observation from the data of the present study is the variability among listeners. Across speakers, intelligibility scores for combined cues varied by over 70 percentage points and for alphabet cues by over 65 percentage points, as shown in Table 2. Clearly, some listeners were able to make valuable use of these strategies, whereas others were not. However, it is noteworthy that when combined or alphabet cues were used, at least 1 listener for each speaker achieved an intelligibility score between approximately 90% and 100%. For all speakers, then, data from the present study suggest that combined or alphabet cues would likely be a viable communication strategy with at least some unfamiliar communication partners. The tremendous variability among listeners highlights the importance of access to multiple communication modes and strategies so that speakers have contingency plans if a partner is unable to understand a message.

With regard to Lindblom's (1990) model of mutuality, intelligibility results suggest that some types of signal-independent information, in the form of explicit linguistic cues, have a marked effect on intelligibility. In particular, alphabet cues, which provide unique information that cannot be obtained through implicit linguistic knowledge, seem to be the most useful of the solitary explicit linguistic cues examined. Topic cues, however, did not serve to enhance intelligibility, seemingly because they provided information that was already available via listeners' implicit linguistic knowledge. Additional studies are necessary to separate the impact of signal-dependent variables from signal-independent ones as well as to examine the effects of the different signal-independent variables that seem to play a role in speech intelligibility.

Speech Rate

Overall, findings from the present study demonstrate that strategies requiring speakers to point to the first letter of each word while simultaneously producing speech result in marked reductions in overall speech rate relative to speakers' own noncued speech and relative to normative speech rate data for neurologically intact speakers.[4] The average magnitude of rate decrease relative to the speakers' noncued productions was approximately 51 words per minute for combined and alphabet cues—a reduction of approximately 70%. As expected, the average rate difference across speakers between topic cues and noncued speech was not significant. Finally, combined cues resulted in a small, but significant, increase in speech rate compared with alphabet cues. This finding is somewhat difficult to explain given the similarity of the two strategies with regard to speech requirements. Again, the magnitude of this difference (approximately 2 words per minute) is probably not clinically meaningful, and this difference was not consistent across all speakers.

Changes in speech rate for individual speakers did not consistently follow the same pattern as group results for strategies involving topic cues (alone and combined with alphabet cues). For example, Speaker 1 produced speech significantly more slowly for topic cues than for noncued speech. In addition, two speakers (1 and 2) did not show a significant difference in rate between alphabet and combined cues. Possible explanations include fatigue effects associated with participation in this experiment and individual differences in motor control capabilities.

Results of the present study demonstrate that seemingly important signal-dependent changes in rate occur when speakers implement alphabet and combined cues. However, the present study provides only gross information about changes in speech rate that occur when speakers implement these strategies. Additional research is necessary to better understand the exact nature of production changes that likely accompany implementation of alphabet and combined cues.

Limitations and Future Directions

The present research was experimental in nature and involved a small number of speakers with dysarthria. As such, there are a number of variables that limit the generalizability of these findings.

[4]The average rate at which nondysarthric individuals read sentences is approximately 190 words per minute (Yorkston & Beukelman, 1981).

Stimulus material produced by the speakers was carefully controlled with regard to semantic and syntactic linguistic features. In addition, speakers were provided with an auditory model and visual prompts prior to and during production of speech stimuli. Consequently, speaker productions likely reflect optimized performance with regard to speech production and expressive language. Indeed, the impact of speech supplementation strategies may be different when speech motor demands, language formulation demands, and strategy implementation demands co-vary during real communication situations.

Data from the present study reflect the effects of each speech supplementation strategy following brief instruction and minimum exposure to the strategy for both speakers and listeners. Longer term experience with each strategy for both speakers and listeners may yield different findings than those of the present study. For example, listeners' performance may improve to a greater extent following practice and explicit instruction regarding how to make optimal use of the information provided by each strategy. In addition, speech production characteristics may change in unpredictable ways as speakers become more automatic in their ability to use strategies.

Another important limitation to the ecological validity of the present study is that listeners heard each narrative two consecutive times. This was done to maintain consistency with the methods of Hustad and Beukelman (2001) so that data would be more directly comparable. However, in real communication situations, listeners generally hear a narrative only one time, unless they request repetition from the speaker. Consequently, intelligibility results of the present study may be idealized.

All listeners were considered unfamiliar partners in this study, meaning that they did not have any prior experience listening to or interacting with individuals having dysarthria. Results would likely be different if familiar communication partners were studied. It is possible that strategies would result in even higher average intelligibility scores for partners who are familiar with the speech patterns of the individual speakers.

Presentation of cues associated with each strategy was digitally optimized because it was difficult to collect data from speakers in which the communication board containing letters and topics was clearly visible on videotape. In real communication situations where supplementation strategies are used, the effects of strategy use may differ when the ability of listeners to see the cues themselves is less than perfect.

Finally, listeners in the present study did not have the opportunity to interact with speakers, as would be the case in real communication situations. It is likely that when interaction between speaker and listener is possible, strategies may have a different, and perhaps greater, effect on intelligibility.

In spite of these limitations, results of this study provide solid evidence that the use of speech supplementation strategies, particularly alphabet cues and combined cues, can markedly improve the intelligibility of speakers with severe dysarthria. However, findings of the present study indicate that intelligibility improvements associated with combined cues were due, primarily, to the joint influence of alphabet cues and reduced speech rate, with topic cues making no significant contribution. Consequently, results suggest that alphabet supplementation should be considered before combined supplementation because it is easier to implement and is likely to have a similar impact on intelligibility. Future research should seek to evaluate the impact of the aforementioned ecologically related variables as well as speaker and listener attitudes toward use of these strategies with a larger number of speakers who vary with respect to severity and type of dysarthria.

Acknowledgments

This research was supported, in part, by a New Investigator grant from the American Speech-Language-Hearing Foundation. We wish to express special thanks to the speakers who participated in this study and to Dr. Robert Prosek for the use of his laboratory space for collecting data from listeners.

References

Ansel, B. M., & Kent, R. D. (1992). Acoustic-phonetic contrasts and intelligibility in the dysarthria associated with mixed cerebral palsy. *Journal of Speech and Hearing Research, 35*, 296–308.

Beukelman, D., & Yorkston, K. (1977). A communication system for the severely dysarthric speaker with an intact language system. *Journal of Speech and Hearing Disorders, 42*, 265–270.

Beukelman, D. R., & Yorkston, K. (1979). The relationship between information transfer and speech intelligibility of dysarthric speakers. *Journal of Communication Disorders, 12*, 189–196.

Carter, C., Yorkston, K., Strand, E., & Hammen, V. (1996). Effects of semantic and syntactic context on actual and estimated sentence intelligibility of dysarthric speakers. In D. Robin, K. Yorkston, & D. Beukelman (Eds.), *Disorders of motor speech: Assessment, treatment, and clinical characterization* (pp. 67–87). Baltimore: Paul H. Brookes.

Connolly, J. H. (1986). Intelligibility: A linguistic view. *British Journal of Disorders of Communication, 21*, 371–376.

Crow, E., & Enderby, P. (1989). The effects of an alphabet chart on the speaking rate and intelligibility of speakers with dysarthria. In K. Yorkston & D. Beukelman (Eds.), *Recent advances in clinical dysarthria* (pp. 100–108). Boston: College Hill.

Dongilli, P. (1994). Semantic context and speech intelligibility. In J. Till, K. Yorkston, & D. Beukelman (Eds.), *Motor* *speech disorders: Advances in assessment and treatment* (pp. 175–191). Baltimore: Paul H. Brookes.

Duffy, J. R., & Giolas, T. G. (1974). Sentence intelligibility as a function of key word selection. *Journal of Speech and Hearing Research, 17*, 631–637.

Erber, N. P. (1975). Auditory-visual perception of speech. *Journal of Speech and Hearing Disorders, 35*, 481–492.

Garcia, J., & Cannito, M. (1996). Influence of verbal and nonverbal contexts on the sentence intelligibility of a speaker with dysarthria. *Journal of Speech and Hearing Research, 39,* 750–760.

Garcia, J. M., & Dagenais, P. A. (1998). Dysarthric sentence intelligibility: Contribution of iconic gestures and message predictiveness. *Journal of Speech, Language, and Hearing Research, 41*, 1282–1293.

Hertzog, C., & Rovine, M. (1985). Repeated-measures analysis of variance in developmental research: Selected issues. *Child Development, 56*, 787–809.

Hunter, L., Pring, T., & Martin, S. (1991). The use of strategies to increase speech intelligibility in cerebral palsy: An experimental evaluation. *British Journal of Disorders of Communication, 26*, 163–174.

Hustad, K. C., & Beukelman, D. R. (2001). Effects of linguistic cues and stimulus cohesion on intelligibility of severely dysarthric speech. *Journal of Speech, Language, and Hearing Research, 44*, 497–510.

Hustad, K. C., & Beukelman, D. R. (2002). Listener comprehension of severely dysarthric speech: Effects of linguistic cues and stimulus cohesion. *Journal of Speech, Language, and Hearing Research, 45,* 545–558.

Hustad, K. C., Beukelman, D. R., & Yorkston, K. M. (1998). Functional outcome assessment in dysarthria. *Seminars in Speech and Language, 19*, 291–302.

Hustad, K. C., & Cahill, M. A. (in press). Effects of visual information and repeated sequential listening experience on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology.*

Hustad, K. C., & Garcia, J. M. (2002). The influences of alphabet supplementation, iconic gestures, and predictive messages on intelligibility of a speaker with cerebral palsy. *The Journal of Medical Speech Language Pathology, 10,* 279–285.

Kalikow, D. N., Stevens, K. N., & Elliot, L. L. (1977). Development of a test of speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America, 61*, 1337–1351.

Kent, R. (1993). Speech intelligibility and communicative competence in children. In A. P. Kaiser & D. B. Gray (Eds.), *Enhancing children's communication: Foundations for intervention* (Vol. 2, pp. 223–239). Baltimore: Paul H. Brookes.

Kirk, R. (1995). *Experimental design: Procedures for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

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

Marascuilo, L. A., & Levin, J. R. (1983). *Multivariate statistics in the social sciences: A researcher's guide*. Monterey, CA: Brooks/Cole.

Marascuilo, L. A., & Serlin, R. C. (1988). *Statistical methods for the social and behavioral sciences*. New York: W.H. Freeman.

Monsen, R. B. (1983). The oral speech intelligibility of hearing-impaired talkers. *Journal of Speech and Hearing Disorders, 43*, 286–296.

Seaman, M. A., Levin, J. R., & Serlin, R. C. (1991). New developments in pairwise multiple comparisons: Some powerful and practicable procedures. *Psychological Bulletin, 110,* 577–586.

Sonic Foundry. (1998). Sound Forge 4.5 [Computer software]. Madison, WI: Author.

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

Yorkston, K. M., Beukelman, D. R., Strand, E. A., & Bell, K. R. (1999). *Management of motor speech disorders in children and adults* (2nd ed.). Austin, TX: Pro-Ed.

Yorkston, K., Beukelman, D., & Tice, R. (1996). Sentence Intelligibility Test for Windows. [Computer software]. Lincoln, NE: Tice Technology.

Yorkston, K., Strand, E., & Kennedy, M. (1996). Comprehensibility of dysarthric speech: Implications for assessment and treatment planning. *American Journal of Speech-Language Pathology, 5*(1), 55–66.

Received July 19, 2002

Accepted November 7, 2002

DOI: 10.1044/1092-4388(2003/038)

Contact author: Katherine C. Hustad, PhD, Department of Communication Sciences and Disorders, The Pennsylvania State University, 110 Moore Building, University Park, PA 16803. E-mail: kch2@psu.edu