Aided and Unaided Speech Supplementation Strategies: Effect of Alphabet Cues and Iconic Hand Gestures on Dysarthric Speech

Katherine C. Hustad
Waisman Center,
University of Wisconsin—Madison

Jane Mertz Garcia
Kansas State University,
Manhattan

Purpose: This study compared the influence of speaker-implemented iconic hand gestures and alphabet cues on speech intelligibility scores and strategy helpfulness ratings for 3 adults with cerebral palsy and dysarthria who differed from one another in their overall motor abilities.
Method: A total of 144 listeners (48 per speaker) orthographically transcribed sentences spoken with alphabet cues (aided), iconic hand gestures (unaided), and a habitual speech control condition; scores were compared within audio–visual and audio-only listening formats.
Results: When listeners were presented with simultaneous audio and visual information, both alphabet cues and hand gestures resulted in higher intelligibility scores and higher helpfulness ratings than the no-cues control condition for each of the 3 speakers. When listeners were presented with only the audio signal, alphabet cues and gestures again resulted in higher intelligibility scores than no cues for 2 of the 3 speakers. Temporal acoustic analyses showed that alphabet cues had consistent effects on speech production, including reduced speech rate, reduced articulation rate, and increased frequency and duration of pauses. Findings for gestures were less consistent, with marked differences noted among speakers.
Conclusions: Results illustrate that individual differences play an important role in the value of supplemental augmentative and alternative communication strategies and that aided and unaided strategies can have similar positive effects on the communication of speakers with global motor impairment.

KEY WORDS: speech therapy, cerebral palsy, intelligibility, augmentative communication, alphabet supplementation

Cerebral palsy is a nonprogressive, neurologically based, motor impairment that is usually diagnosed within the first 18 months of life (Pellegrino & Dormans, 1998; Rosenbaum et al., 2000). Demographic research suggests that the prevalence of cerebral palsy is approximately 2 per 1,000 persons (Winter, Autry, Boyle, & Yeargin-Allsopp, 2000). Up to 88% of those with cerebral palsy have dysarthria (Wolfe, 1950). Although some individuals who have dysarthria associated with cerebral palsy do not experience communication problems, survey results indicate that approximately 40% of children with cerebral palsy are not able to meet their communication needs using speech in at least some situations of daily life (Kennes et al., 2002). Of this 40%, approximately 26% are able to use natural speech to meet some communication needs, whereas the remaining 14% are unable to

1092-4388/05/4805-0996

produce any functional speech (Kennes et al., 2002). Although analogous data do not exist for adults, it is probably reasonable to conclude, on the basis of the underlying static nature of cerebral palsy, that adults face communication problems in similar proportion. Clearly, for those individuals who cannot use speech for any functional purposes, augmentative and alternative communication (AAC) systems are essential for all communication. However, for those who are able to use speech to meet at least some communication needs, AAC systems and strategies can play an important role in supplementing speech or filling in the gaps where speech fails.

This paper focuses on individuals with cerebral palsy who are able to use natural speech to meet some communication needs. Although this may seem to be a relatively straightforward designation, in reality, the usefulness of natural speech can vary on the basis of a number of variables. Examples include the communication partner, the context, the complexity and predictability of messages, use of gestures and facial expression, use of any other compensatory strategies, and the integrity of the acoustic signal (Connolly, 1986; Hustad, Beukelman, & Yorkston, 1998; Yorkston, Strand, & Kennedy, 1996).

In conceptualizing variables that influence the usefulness of natural speech, often measured as intelligibility, it is important to consider communication as an interactive and dynamic process, occurring between at least two people, that is, bidirectional and multimodal (Kraat, 1987). Indeed, at a minimum, communication requires a sender, who encodes a message, a transmission channel through which the message is sent, and a receiver who is able to decode successfully the intended message. However, a number of other variables contribute to successful communication, or mutuality (Lindblom, 1990). Lindblom described a model that provides a simplified account of the contribution of signal-dependent factors (i.e., the integrity of the acoustic signal produced by the speaker) and signal-independent factors (i.e., linguistic and world knowledge possessed by the listener) to mutual understanding between speaker and listener. In this model, signal-dependent information and signal-independent information are conceptualized as inversely related so that when signal-dependent acoustic information is of high quality, signal-independent information is less important to achieving mutual understanding. Conversely, when signal-dependent acoustic information is of lower quality, signal-independent information plays a critical role in compensating for the degraded speech signal. Recently, Hustad, Jones, and Dailey (2003) proposed an extension to Lindblom's model. They suggested that supplementation strategies which present extrinsic speaker-implemented cues should be considered separately from intrinsic signal-independent knowledge possessed by the listener. Alphabet supplementation and iconic hand gestures as deliberate speaker-implemented strategies used in conjunction with natural speech were of particular interest for the present study.

Speech Supplementation Strategies

Alphabet Cues

When a speaker uses alphabet cues, he or she points to the first letter of each constituent word of a target message on a communication board while simultaneously speaking. Alphabet cues are considered an "aided" AAC strategy because use of this technique requires an alphabet board that is external to the body (Lloyd & Fuller, 1986). Several studies have shown marked increases in the intelligibility of dysarthric speech (between 15% and 44%) when alphabet cues are available to listeners (Beukelman & Yorkston, 1977; Crow & Enderby, 1989; Hustad, Auker, Natale, & Carlson, 2003; Hustad & Beukelman, 2001; Hustad, Jones, & Dailey, 2003). Research suggests that intelligibility changes associated with alphabet cues are related to three factors. First, when speakers implement alphabet cues, rate of speech is markedly reduced (Hustad, Jones, et al., 2003) and intelligibility is increased, even when listeners are unable to see the letters to which the speakers are pointing (Beukelman & Yorkston, 1977; Crow & Enderby, 1989). Second, provision of grapho-phonemic information for the first letter of each word in experimental contexts without concomitant changes in the speech signal associated with implementation of alphabet cues results in increased intelligibility (Hustad & Beukelman, 2001). Third, when speakers implement alphabet supplementation and listeners are able to see the alphabet cues, intelligibility is increased to a greater extent than when listeners are unable to see the alphabet cues (Beukelman & Yorkston, 1977).

Iconic Hand Gestures

When a speaker uses iconic hand gestures to supplement a message, he or she produces gestures concurrently with speech. Iconic hand gestures are one form of *gesticulation* (the natural hand movements made while speaking) that illustrate or add meaning to content words of the spoken message (Garcia, Cannito, & Dagenais, 2000). Hand gestures are considered an unaided type of AAC symbol or strategy because use of the technique does not require anything external to the body (Lloyd & Fuller, 1986). Research suggests that for most speakers, scripted iconic hand gestures increase intelligibility between 10% and 40% (Garcia & Cannito, 1996; Garcia & Dagenais, 1998). However, for some speakers, implementation of hand gestures may actually decrease speech intelligibility (Garcia & Dagenais, 1998). Preliminary research suggests that iconic hand gestures, like alphabet cues, may have an impact on the acoustic signal. Specifically, Garcia and Cobb (2000) found that rate of speech seemed to increase and interword intervals seemed to decrease when hand gestures were implemented while speaking.

Alphabet cues and hand gestures differ in several ways. First, hand gestures are unaided, offering the user the important feature of being “built-in” and thus are completely portable and available in nearly any communication situation. Alphabet cues are aided and therefore require an external alphabet board, which may be impractical or impossible to access in some situations. Another difference between alphabet cues and hand gestures relates to the nature of the information provided. Iconic hand gestures could be considered an *extralinguistic* type of cue because the symbol set is not clearly defined. That is, gesticulations are idiosyncratic and universal definitions that are accepted by all members of the language community do not exist (Garcia et al., 2000). Alphabet cues, on the other hand, could be considered a *linguistic* type of cue because the symbol set, orthographic letters, is clearly defined. Similarly, the grapheme-to-phoneme correspondence, necessary for a listener to benefit from a speaker’s use of alphabet cues, is governed by linguistic rules shared by the language community. Finally, alphabet cues and gestures differ in their generativeness, or whether the symbols can be used to create an infinite number of meanings. Alphabet cues would be considered generative because the first letter of any possible word could be represented by a speaker and understood by a listener who is literate using this strategy. Hand gestures would not be considered generative because the number of possible meanings that could be expressed by a speaker and understood by a listener who is a member of the same language community is limited to those words and concepts that are “gesturable.” Indeed, certain meanings would be difficult or impossible to express, particularly concepts with reduced iconicity such as *happy* or *love*. Consequently, the information provided by gestures may be less explicit, leaving more room for interpretation.

Although there are some key differences between hand gestures and alphabet cues used to supplement speech, there are also some important similarities—namely that both are overlaid on speech. In addition, both strategies have been shown to affect speech production. At present, it is unknown whether there is an intelligibility advantage for alphabet cues or for gestures; however, a preliminary report (Hustad & Garcia, 2002) comparing intelligibility scores between the two strategies for one speaker in an audio–visual presentation format showed no difference.

Effects of Presentation Modality

As described previously, implementation of hand gestures and implementation of alphabet cues seem to affect speech-production parameters for some speakers, independent of the visual cues provided by each strategy (Garcia & Cobb, 2000; Garcia, Dagenais, & Cannito, 1998; Hustad, Jones, et al., 2003). Preliminary research suggests that gestures may actually increase overall rate of speech for some speakers (Garcia & Cobb, 2000), whereas alphabet cues have been clearly shown to decrease overall rate of speech (Beukelman & Yorkston, 1977; Crow & Enderby, 1989). Speaker-implemented hand gestures and alphabet cues within an audio-only modality have not been compared previously; therefore, it is unknown if one strategy has a more positive effect on the speech signal than the other. This information is critical to understanding the means by which each intervention is effective.

Listener Perceptions of Strategies

The attitudes and perceptions of potential communication partners toward a particular communication strategy may have an important impact on the willingness of a speaker to adopt a strategy and on the willingness of a potential communication partner to interact with a speaker using that strategy (Hustad & Gearhart, 2004). Recent research has shown that attitude ratings tend to increase linearly with intelligibility scores; therefore, attitude ratings and intelligibility scores are strongly correlated (Hustad & Gearhart, 2004). Although studies have examined the attitudes of listeners toward alphabet cues, topic cues, and combined topic and alphabet cues (Hustad, 2001; Hustad & Gearhart, 2004), there have been no published studies that have examined listeners’ attitudes or perceptions of speakers using iconic hand gestures to supplement speech. This information could have an effect on which strategy a speaker chooses to adopt.

Purpose of the Present Study

In the present study, which is an extension of preliminary work (Hustad & Garcia, 2002), the effects of two supplemental AAC strategies (alphabet cues and hand gestures) and a control condition (no-cues habitual speech) were examined within both audio-only and audio–visual viewing conditions for 3 speakers. Specifically, this study addressed the following questions for each speaker: (a) How do the different cue conditions compare with regard to intelligibility scores and listener ratings of the helpfulness of each strategy across the audio–visual and audio-only presentation modalities; (b) how do the different cue conditions compare with regard to intelligibility scores and listener ratings of the helpfulness of each strategy within the audio–visual presentation modality; (c) how do the different cue conditions compare with regard to intelligibility scores and listener ratings of the helpfulness of each strategy within the audio-only presentation modality; and (d) how does the pattern of results differ for individual speakers?

Method

Participants

One hundred forty-four individuals without disabilities participated as listeners in this study. These individuals were presented with videotapes of the speakers using habitual speech (no cues), alphabet supplementation, and iconic hand gestures. Listeners transcribed productions in each condition and made ratings of the helpfulness of each strategy. Forty-eight listeners participated for each of the 3 speakers, with 24 assigned to the audio–visual group for each speaker and 24 assigned to the audio-only group for each speaker. All listeners were undergraduate students. The mean age of listeners who viewed Speaker A was 20.73 years (SD = 1.94). The mean age of listeners who viewed Speaker B was 19.83 years (SD = 2.33). The mean age of listeners who viewed Speaker C was 22.65 years (SD = 5.87). Gender was not a variable of interest, and previous work (Garcia & Cannito, 1996) has shown no significant difference between male and female listeners; therefore, no effort was made to balance the number of male and female participants.

All listeners met the following inclusion criteria: (a) pass a pure-tone hearing screening at 25 dB HL for 1 kHz, 2 kHz, and 4 kHz bilaterally; (b) between 18 and 45 years in age; (c) no more than incidental experience listening to or communicating with persons having communication disorders; (d) native speakers of American English; (e) self-report of normal or near-normal vision with correction; and (f) no identified language, learning, or cognitive disabilities per self-report.

Materials

Speakers with dysarthria. Three individuals with severe dysarthria secondary to cerebral palsy participated in this study. These individuals produced a standard corpus of speech stimuli using habitual speech, alphabet supplementation, and iconic hand gestures. Productions were audio and video recorded, edited, and presented to listeners without disability.

Speaker A was a 42-year-old woman who had a medical diagnosis of spastic diplegia. Her dysarthria was characterized by spastic perceptual features as judged by a certified speech–language pathologist. Perceptual characteristics included short phrases, imprecise articulation, strained–strangled vocal quality, and reduced loudness. Speaker A used speech as her primary mode of communication and did not use any formal AAC strategies to supplement or replace her speech. Her intelligibility on the Sentence Intelligibility Test (SIT; Yorkston, Beukelman, & Tice, 1996) was 27%, and her rate of speech while producing sentences from the SIT was 81 words per minute (wpm). A certified physical therapist evaluated upper extremity motor control for the two supplementation tasks along a continuum of severity ranging from *no or minimal involvement* to *severely impaired* motor functioning. When using alphabet supplementation, Speaker A's hand and arm movements were judged to be mildly impaired and thus adequate when using her less involved side (dominant hand). When using gestures, the physical therapist judged Speaker A's motor movements to be moderately impaired, with reduced shoulder function for larger hand movements. Hand movements were characterized by general postures, with minimal motion of the fingers. Overall, the quality of limb movements was judged to be better for alphabet cues than for gestures.

Speaker B was a 33-year-old woman who also had a medical diagnosis of spastic diplegia. Her dysarthria was characterized by mixed spastic–ataxic features as judged by a certified speech-language pathologist. Perceptual characteristics included excess and equal stress, short phrases, imprecise articulation, and strained–strangled vocal quality. Speaker B used speech as her primary mode of communication and used a voice-output AAC system to repair communication breakdowns and to communicate with strangers. Her intelligibility on the SIT (Yorkston, Beukelman, & Tice, 1996) was 20%, and her rate of speech while producing sentences from the SIT was 70 wpm. The physical therapist's evaluation of Speaker B indicated that overall motor impairment observed during use of alphabet supplementation and gestures was moderate for both strategies. When using alphabet supplementation, arm movements were judged to be mildly impaired, and hand movements were judged to be moderate–severely impaired. Movements were considered adequate for the task of pointing to alphabet cues. When using gestures, Speaker B's arm movements and hand movements were both judged by the physical therapist to be moderately impaired, with some dyscoordination noted. Overall, the quality of limb movements was judged to be better for gestures than for alphabet cues.

Speaker C was a 37-year-old man who had a medical diagnosis of spastic cerebral palsy (geographic distribution unspecified). His dysarthria was characterized by spastic features as judged by a certified speech-language pathologist. Perceptual characteristics included imprecise articulation, strained–strangled and wet vocal quality, short phrases, and hypernasality. Speaker C used his natural speech and a voice-output AAC device with equal frequency as his primary modes of communication. His intelligibility on the SIT (Yorkston, Beukelman, & Tice, 1996) was 7%, and his rate of speech while producing sentences from the SIT was 63 wpm. The physical therapist's rating of Speaker C indicated that overall motor impairment observed during use of both alphabet supplementation and gestures was minimal to mild for both strategies. When using alphabet supplementation, impairment in both arm and hand movement was judged to be minimal to none. Movements were considered adequate for the task of pointing to alphabet cues. When using gestures, the physical therapist judged Speaker C's arm movements to be mildly impaired and hand movements to be moderately impaired, particularly fine-tuning of hand movements. Overall, the quality of limb movements was judged to be better for alphabet cues than for gestures.

All 3 speakers had graduated from high school and had completed some postsecondary education but had not earned a degree. On the basis of informal interactions and educational history, all speakers demonstrated cognitive and language abilities that were within acceptable limits. In addition, speakers met the following inclusion criteria: (a) ability to produce connected speech consisting of at least eight consecutive words; (b) speech intelligibility between 5% and 30% as measured by the SIT (Yorkston, Beukelman, & Tice, 1996); (c) use of speech as a mode of communication; (d) native speaker of American English; (e) functional literacy skills at or above the 6th-grade level; (f) corrected or uncorrected vision within normal limits, per self-report; (g) hearing within normal limits, per self-report; (h) ability to accurately direct-select letters and orthographically represented phrases from a communication board; and (i) ability to produce scripted gestures using dominant hand.

Speech Stimuli

Speakers produced 24 different sentences that have been used in previous studies examining the impact of iconic gestures on intelligibility (Garcia & Cobb, 2000; Garcia & Dagenais, 1998; Garcia & Hayden, 1999; Hustad & Garcia, 2002). Sentences were each six to eight words in length and were composed of one- to two-syllable words. In addition, all sentences were imperative in nature. Each sentence incorporated two scripted iconic gestures that corresponded to key content words in the sentences and could be illustrated using one upper extremity. For example, while saying the target sentence, "Shut the door and lock it," the speaker simultaneously illustrated "shut" (palm facing forward and moving away from body) and "lock" (turning motion with hand). The nonverbal content of each test sentence was evaluated by a group of 12 undergraduate college students (Garcia & Dagenais, 1998). These participants viewed a videotaped recording of an unimpaired speaker producing the two scripted gestures for each sentence (without speech) and then wrote "a complete sentence corresponding to the message conveyed" (Garcia & Dagenais, 1998, p. 1284). The written responses of each participant were rated on a 5-point scale reflecting how accurately their transcribed sentence interpreted the two gestures as well as its semantic relatedness to the target sentence. The minimum score was 0 (0 rating $\times$ 12 participants) and the maximum score for each sentence was 48 (4 rating $\times$ 12 participants). The 24 sentences used in the present study were selected to reflect a broad range of nonverbal information content, with an overall average rating of 22.5 out of 48 possible (47% of nonverbal content could be interpreted without any auditory information). Additional details regarding development and validation of sentences can be found in Garcia and Dagenais (1998). The 24 test sentences with their gestural descriptions are provided in the Appendix.

Procedures

Data Collection From Speakers With Dysarthria

Speech samples used in the present study were collected at the same time as speech samples from two different published studies that focused on the effects of speech supplementation strategies on intelligibility (Hustad, Auker, et al., 2003; Hustad, Jones, et al., 2003). Collectively, 8 speakers participated in the two studies; however, only the 3 speakers presented in this article were able to use hand gestures with functional accuracy and consistency. Speech stimuli and listeners who participated in the present study are unique to this research and not a part of other published studies.

The entire experimental protocol for each speaker was implemented during a single 5-hr session. For the present study, speakers produced target sentences from Garcia and Dagenais (1998) in each of three speaking conditions: using habitual speech, using alphabet supplementation, and using iconic hand gestures. The 3 speakers each completed the experimental tasks in a different order to prevent the possibility of an order effect associated with learning the different strategies.

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

Speakers were instructed in the use of alphabet supplementation and hand gestures prior to recording. For each strategy, the experimenter provided speakers with a verbal description of the strategy, an explanation of the purpose of the strategy, and a model for using the strategy.

For alphabet supplementation, speakers were required to point to the first letter of each word on a 12 in. × 14 in. communication board while simultaneously speaking the word. Speakers were coached so that the timing of letter selection and production of each word coincided. Speakers practiced using alphabet supplementation on a set of rehearsal sentences until they were able to use the strategy comfortably and accurately. During recording, speakers were required to meet three criteria: (a) produce all words within each target sentence in the appropriate sequence; (b) select the appropriate first letter cue for each word; and (c) point to the first letter of each word while simultaneously producing it. These three criteria were monitored online during recording by the first author. Speakers were asked to repeat any sentence for which all three criteria were not met. Fewer than 5% of productions required repetitions.

For hand gestures, speakers were required to produce target sentences while simultaneously implementing two scripted gestures for each sentence. Speakers learned the scripted gestures for each sentence by watching a videotape of a non–neurologically impaired speaker produce target sentences and gestures. Two productions of the target sentence and corresponding gestures were presented on the tape. Immediately after this presentation, speakers were asked to produce the sentence with its corresponding hand gestures themselves. After attempting the target sentence and gestures, speakers were allowed to view the nonimpaired speaker again, if necessary. During recording, speakers were required to meet three criteria: (a) produce all words within each target sentence in the appropriate sequence; (b) produce each of the two target hand gestures; and (c) time the onset of each hand gesture with the content word(s) to which it referred. These three criteria were monitored online during recording, again by the first author. Speakers were asked to repeat any sentence for which all three criteria were not met. Fewer than 5% of productions required repetitions. Two iterations of each sentence and its corresponding hand gestures were recorded. The second of the speaker's productions for each sentence was used in this study.

For the habitual speech condition, speakers were required to produce each target sentence following a verbal model. An orthographic presentation of each target sentence was also presented on the computer located directly in front of the speaker. During recording, speakers were required to meet one criterion: produce all target words in each sentence in the appropriate sequence. This criterion was monitored online by the first author. None of the target sentences required repetition.

Constructing stimulus tapes. Procedures for the construction of stimulus tapes followed those detailed elsewhere (Hustad, Auker, et al., 2003; Hustad & Cahill, 2003; Hustad, Jones, et al., 2003). In brief, digital video and audio recordings were transferred to computer by digital-to digital interface (IEEE 1394 for digital video and S/PDIF for audio) so that there was no generational loss of quality. Video recordings of stimulus sentences produced in each of the three experimental conditions (habitual speech, alphabet supplementation, hand gestures) were edited using Adobe Premiere 6.0 (computer software) for Macintosh. High-quality audio recordings from DAT of each stimulus sentence were similarly edited and peak amplitude normalized using SoundForge 4.5 (computer software) to assure that maximum loudness levels of the recorded speech stimuli were consistent across speakers and sentences. The normalized audio files were then matched with the native audio associated with the video samples using auditory–perceptual judgments and visual matching of the native waveform and the higher quality amplitude-normalized waveform. Following alignment of the two audio samples, the native audio sample was deleted, leaving only the high-quality, amplitude-normalized sample from DAT associated with the video of each stimulus sentence. During this process, productions from all speakers were systematically evaluated by a research assistant to assure that each constituent word of each sentence was produced in each speaking condition.

Because the video camera was positioned directly in front of the speakers, it was nearly impossible to decipher the target letters to which the speakers were pointing when they implemented alphabet supplementation. Therefore, videotapes were digitally enhanced so that the alphabet cues were clearly visible. Specifically, 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.

In each of the three experimental tasks, there were eight different sentences. Listeners were presented with different sentences in each task. To guard against an order effect and to assure that the results of the experiment would not be biased by the assignment of particular sentences to tasks, two sets of experimental tapes were made with different sentences quasi-randomly assigned to each of the three tasks. Quasi-random assignment involved assuring that each sentence was assigned to a different experimental task than on the first set of tapes.

Tapes for the audio–visual and audio-only conditions differed in only one way: Listeners were presented with a blue screen in place of the visual image of the speaker during production of the stimulus sentences. All stimulus tapes contained the same sequence of events, task instructions, familiarization sentences, individual sentence numbers, stimulus sentences, and a visual prompt to transcribe each sentence. Digital videotapes for each task were transferred, following broadcast-quality standards, to VHS tape for playback to listeners.

Data Collection From Listeners

Presentation of stimuli. Listeners viewed the stimulus tapes individually or in groups of 2 in a quiet listening environment. While viewing the stimulus tapes, listeners were seated approximately six feet away from a 27-in., high-resolution television monitor with a video cassette recorder attached to it. To approximate conversational speech, the peak output level of stimulus material was recalibrated to approximately 65 dB SPL from where listeners were seated. The sound output level was calibrated prior to each experimental session to assure that all listeners heard stimuli at the same output level.

Administration instructions. The experimenter explained to listeners that they would complete three different listening tasks that would last for a total of about 30 min. For those listeners in the audio–visual groups, the experimenter explained that they would see the same person who has speech problems producing a series of 24 different sentences (8 sentences in each of three tasks). In one task, listeners would see the speaker pointing to the first letter of each word while simultaneously talking; in another task, they would see the speaker producing hand gestures while simultaneously talking; and in another task they would see the speaker talking without any strategies. For those listeners in the audio-only groups, the experimenter provided a similar explanation; however, listeners were told that they would only be able to hear the speaker. Listeners were informed of the strategy that the speaker was using for each condition, even though they were unable to see the speakers using the strategy.

All listeners were told that they would see/hear each sentence one time. After each sentence, listeners were to follow the instructions on the videotape, which directed them to write down exactly what they thought the speaker said, taking their best guess if they were unsure. At this point, listeners were presented with three familiarization sentences from the SIT to orient them to the experimental paradigm and the speaker whom they would be hearing/watching. After each sentence, listeners were prompted to orthographically transcribe what they thought the speaker had said. They were not given feedback regarding the accuracy of their transcriptions. Prior to beginning the experimental tasks, the experimenter explained that she would be controlling the videotape from the back of the room and listeners could take as much time as necessary to write their responses. Following their transcriptions of the eight sentences for each experimental condition, listeners were instructed to rate the overall helpfulness of the communication strategy using a scale of 1 (*not helpful*) to 7 (*very helpful*).

Randomization and counterbalancing. To prevent an order effect and/or a learning effect, the order of presentation of the three experimental conditions (habitual speech, alphabet supplementation, and hand gestures) was counterbalanced. Specifically, the six possible presentation sequences were each viewed by 4 different listeners (for a total of 24 listeners per group).

Scoring and Reliability

The dependent variable, intelligibility, was measured by calculating the percentage of words transcribed correctly for each experimental task and listener. Transcriptions from each listener were scored by the same experimenter, who tallied the number of words identified correctly on the basis of whether each was an exact word-for-word match to the corresponding word in the sentence. The dependent variable, listener helpfulness ratings, was obtained from questionnaires that listeners completed following each experimental task.

To determine intrascorer reliability, the same experimenter who completed the initial scoring of intelligibility data rescored all transcription data for 36 of the 144 listeners (6 listeners from each speaker group and presentation condition; 25% of the sample). The original transcription results (in percentage intelligibility) were then compared with the rescored transcription results. Point-by-point agreement in scoring across all 36 listeners was calculated using the following formula: percentage agreement = [(agreements)/(agreements + disagreements) × 100]; the resulting agreement was 99.72%.

To determine interscorer reliability, a second individual who was unfamiliar with the experiment scored all transcription data for 36 randomly selected listeners from the pool of 144 listeners. Again, 6 listeners from each speaker group and presentation condition were represented. The original transcription results (in percentage intelligibility) were then compared with those of the unfamiliar scorer. Point-by-point agreement across the two judges for all 36 listeners was 99.3%. Thus, there was a very high level of reliability for intra- and interjudge scoring.

Experimental Design and Analysis

This study employed a 3 × 2 × 3 split-plot design (Kirk, 1995) for each of the two dependent variables, intelligibility and listener ratings of the helpfulness of each strategy. The research design incorporated two between-subjects factors (speaker group and presentation mode) and one within-subject factor (cue condition). For the between-subjects factors, 48 listeners were first randomly assigned to each of the three speaker groups. Then, within each speaker-group, 24 listeners were randomly assigned to each of the two presentation mode groups, one that received all speech stimuli in the audio–visual presentation modality and one that received all stimuli in the audio-only presentation modality. The within-subject factor of cue condition had three categories: alphabet cues, iconic hand gestures, and no cues.

Because a small number of speakers participated in this study and because results for individual speakers were of particular interest, separate statistical analyses were performed for each speaker. Furthermore, because speakers were heterogeneous, there was concern that important differences would be lost if data were analyzed in a single analysis of variance (ANOVA) for each of the two dependent measures. Statistical analyses used a nested-model ANOVA (Marascuilo & Serlin, 1988) because the questions of interest were localized within the two presentation modes for each speaker group. Nonparametric ANOVAs were used to analyze Likert-scale helpfulness ratings; parametric ANOVAs were used for intelligibility data. The research questions and specific contrasts of interest were identified a priori; therefore, only the omnibus tests and follow-up analyses that specifically addressed the targeted questions were performed. This is generally considered to be a conservative approach because fewer statistical tests are performed, thus reducing the probability of a Type I error. The specific tests focusing on the questions of interest were as follows: (a) the main effect of cue conditions on intelligibility scores and helpfulness ratings across all other variables; (b) the nested effect of cue conditions on intelligibility scores and helpfulness ratings within the audio–visual presentation mode; and (c) the nested effect of cue conditions on intelligibility scores and helpfulness ratings within the audio-only presentation mode. Each omnibus test was allotted an alpha level of .01. Follow-up tests for significant omnibus tests were allotted an alpha of .01 per family, which was partitioned using the Bonferroni procedure. Follow-up tests were considered significant if the observed probability was less than .0033.

Results

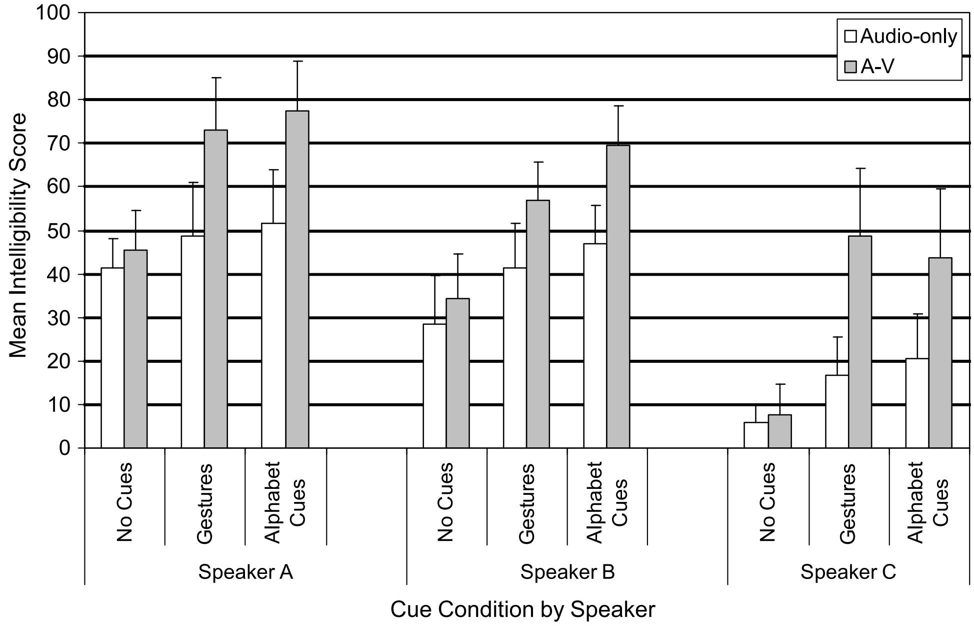
Intelligibility data are displayed graphically by speaker, cue condition, and presentation mode in Figure 1. Listener ratings of the helpfulness of each strategy are displayed graphically by speaker, cue condition, and presentation mode in Figure 2.

Main Effect of Cue Conditions

Statistical analyses revealed a significant main effect of cue conditions for both intelligibility scores and for helpfulness ratings for each of the 3 speakers. Follow-up tests for each of the 3 speakers showed that alphabet cues resulted in significantly higher intelligibility scores and higher helpfulness ratings than the no-cues condition. The magnitude of this difference for intelligibility scores was 21.16% for Speaker A, 26.71% for Speaker B, and 25.28% for Speaker C. Similarly, the magnitude of the difference between alphabet cues and no cues for ratings of helpfulness was 1.44 Likert points for Speaker A, 1.75 Likert points for Speaker B, and 1.60 Likert points for Speaker C.

Follow-up tests also showed that the use of hand gestures while speaking resulted in significantly higher intelligibility scores and higher helpfulness ratings than the no-cues conditions for each speaker. The magnitude of the difference between gestures and no cues for intelligibility scores was 17.48% for Speaker A, 17.53% for Speaker B, and 25.87% for Speaker C. The difference in ratings of helpfulness was 1.27 Likert points for Speaker A, 1.02 Likert points for Speaker B, and 1.71 Likert points for Speaker C.

Figure 1. Mean intelligibility scores (+*SD*) by speaker, cue condition, and presentation mode.



For Speakers A and C, there was no statistical difference between alphabet cues and hand gestures for intelligibility scores and helpfulness ratings. However, for Speaker B, alphabet cues resulted in significantly higher intelligibility scores and helpfulness ratings than the use of hand gestures. The magnitude of this difference for intelligibility scores and helpfulness ratings, respectively, was 9.18% and 0.73 Likert points. See Tables 1, 2, and 3 for statistics pertaining to the ANOVAs and follow-up contrasts.

Nested Main Effect of Cue Conditions Within the Audio–Visual Presentation Mode

The nested main effect of cue conditions within the audio–visual presentation mode was statistically

Figure 2. Mean helpfulness rating (+*SD*) by speaker, cue condition, and presentation mode.

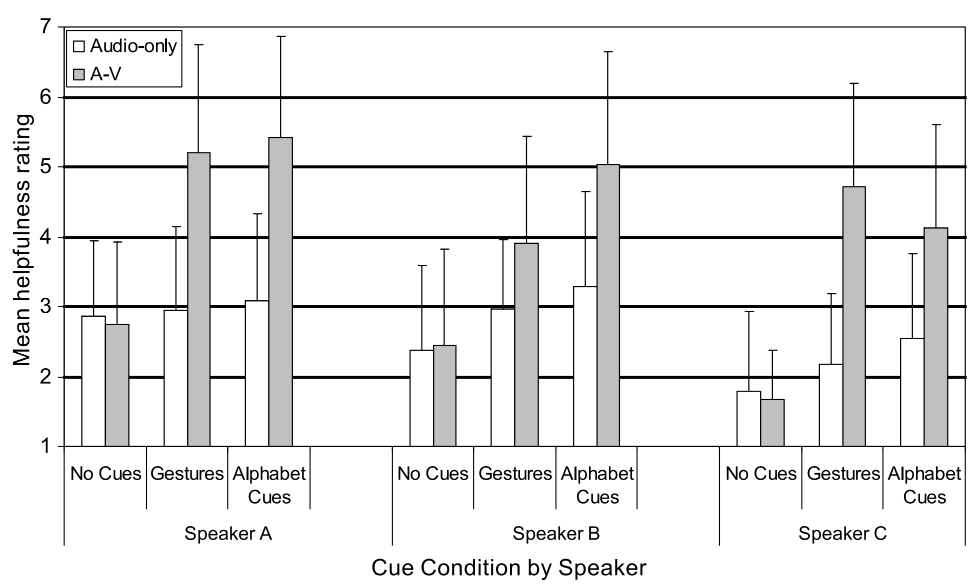


Table 1. Parametric ANOVA results of omnibus tests of nested main effects model for intelligibility data.

Source	Sum of squares	*df*	*F*	η^2
Speaker A				
Cues	12,264.93	1.91	40.86*	.47
Error (cues)	14,107.68	89.52		
Cues in mode				
Cues in AV	14,381.51	1.99	55.84*	.71
Error	5,924.22	45.79		
Cues in A only	1,372.24	1.99	6.72*	.23
Error	4,694.64	45.94		
Speaker B				
Cues	17,685.60	1.77	104.91*	.69
Error (cues)	7,923.54	83.44		
Cues in mode				
Cues in AV	15,279.79	1.93	113.01*	.83
Error	3,109.69	44.27		
Cues in A only	4,165.81	1.59	31.38*	.58
Error	3,053.85	36.47		
Speaker C				
Cues	20,942.02	1.87	67.11*	.59
Error (cues)	14,666.34	87.89		
Cues in mode				
Cues in AV	23,949.98	1.82	83.44*	.78
Error	6,601.52	41.77		
Cues in A only	2,740.49	1.97	27.21*	.54
Error	2,316.37	45.30		

*$p < .01$.

significant for intelligibility scores and helpfulness ratings for each of the 3 speakers. Follow-up contrasts for each speaker showed that alphabet cues and hand gestures resulted in higher intelligibility scores and helpfulness ratings than no cues. The mean benefit from alphabet cues was 31.93% and 2.67 Likert points for Speaker A, 35.24% and 2.58 Likert points for Speaker B, and 36.03% and 2.46 Likert points for Speaker C. The mean benefit from the use of hand gestures was 27.56% and 2.46 Likert points for Speaker A, 22.46% and 1.46 Likert points for Speaker B, and 40.89% and 3.04 Likert points for Speaker C. For Speakers A and C, there was no statistical difference between alphabet cues and hand gestures with regard to intelligibility scores and Likert ratings. For Speaker B, alphabet cues resulted in intelligibility scores and Likert ratings of helpfulness that were significantly higher for alphabet cues than for hand gestures (12.8% and 1.13 Likert points).[1]

[1]Data from Speaker A pertaining only to intelligibility in the audio–visual condition have been published previously (see Hustad & Garcia, 2002). Previously published results represent approximately 8% of the dataset presented in the current article. Furthermore, in Hustad and Garcia's study, different analyses were conducted to answer different experimental questions.

Table 2. Nonparametric Friedman omnibus test results of nested main effects model for helpfulness data.

Source	*n*	*df*	χ^2
Speaker A			
Cues	24	2	18.16*
Cues in mode			
Cues in AV	24	2	26.89*
Cues in A only	24	2	0.12
Speaker B			
Cues	24	2	34.85*
Cues in mode			
Cues in AV	24	2	29.74*
Cues in A only	24	2	7.76
Speaker C			
Cues	24	2	35.63*
Cues in mode			
Cues in AV	24	2	35.88*
Cues in A only	24	2	11.38*

*$p < .01$.

Nested Main Effect of Cue Conditions Within the Audio-Only Presentation Mode

The nested main effect of cue conditions within the audio-only presentation mode was statistically significant for intelligibility scores for all 3 speakers. The nested main effect for Likert ratings of helpfulness was statistically significant only for Speaker C. Follow-up contrasts showed that alphabet cues resulted in higher intelligibility scores than the no-cues condition for all 3 speakers. The mean benefit from alphabet cues was 10.38% for Speaker A, 18.18% for Speaker B, and 14.53% for Speaker C. For Speaker C, ratings of helpfulness were also significantly higher (0.92 Likert points) for alphabet cues than for no cues; helpfulness ratings did not differ for any other contrast. For Speakers B and C, the use of hand gestures while speaking resulted in significantly higher intelligibility scores than the no-cues conditions. The mean benefit was 12.61% for Speaker B and 10.85% for Speaker C. The difference between alphabet cues and gestures was not significant for intelligibility for any of the speakers.

Discussion

Alphabet cues and iconic hand gestures are two speech supplementation strategies that can be used to enhance intelligibility in speakers with dysarthria. These strategies differ in several important ways, including the linguistic nature of the cues provided (linguistic vs. extralinguistic) and the means necessary

Table 3. Follow-up contrasts for the main effect of cues for intelligibility data (Intell) and listener ratings of the helpfulness of strategies (Helpful).

	Mean difference		*df*		Test statistic	
Contrast	Intell	Helpful	Intell	Helpful	Intell (*t*)	Helpful (*z*)
Speaker A						
Main effect of cues						
AC vs. NC	21.16	1.44	47	47	7.91*	−3.91*
Gestures vs. NC	17.48	1.27	47	47	6.73*	−3.62*
AC vs. gestures	3.67	.17	47	47	1.660	−0.26
Cues within AV						
AC vs. NC	31.93	2.67	23	23	9.66*	−3.93*
Gestures vs. NC	27.56	2.46	23	23	8.70*	−4.04*
AC vs. gestures	4.37	0.21	23	23	1.30	−0.29
Cues within A only						
AC vs. NC	10.38	—	23	—	3.63*	—
Gestures vs. NC	7.39	—	23	—	2.51	—
AC vs. gestures	2.99	—	23	—	1.02	—
Speaker B						
Main effect of cues						
AC vs. NC	26.71	1.75	47	47	13.03*	−3.13*
Gestures vs. NC	17.53	1.02	47	47	8.69*	−4.35*
AC vs. gestures	9.18	0.73	47	47	6.09*	−3.27*
Cues within AV						
AC vs. NC	35.24	2.58	23	23	14.21*	−4.13*
Gestures vs. NC	22.46	1.46	23	23	9.00*	−3.57*
AC vs. gestures	12.78	1.13	23	23	6.04*	−2.96*
Cues within A only						
AC vs. NC	18.18	—	23	—	8.39*	—
Gestures vs. NC	12.61	—	23	—	4.38*	—
AC vs. gestures	5.58	—	23	—	2.94	—
Speaker C						
Main Effect of cues						
AC vs. NC	25.28	1.60	47	47	10.39*	−4.12*
Gestures vs. NC	25.87	1.71	47	47	9.04*	−4.13*
AC vs. gestures	−0.59	−0.10	47	47	−0.25	−0.45
Cues within AV						
AC vs. NC	36.03	2.46	23	23	11.36*	−4.23*
Gestures vs. NC	40.89	3.04	23	23	12.90*	−4.13*
AC vs. gestures	−4.86	−0.58	23	23	−1.22	−1.57
Cues within A-only						
AC vs. NC	14.53	0.75	23	23	7.18*	−2.98*
Gestures vs. NC	10.85	0.37	23	23	5.57*	−1.44
AC vs. gestures	3.68	0.38	23	23	1.69	−1.63

*$p < .0033$.

to use the strategy (aided vs. unaided). Although previous studies have shown that both strategies are effective, the two strategies have not been directly compared with one another in regard to their effects on intelligibility and listener perceptions of each strategy. The present study investigated the effects of iconic hand gestures and alphabet cues, along with a habitual speech control condition, on intelligibility and helpfulness ratings. Results were examined across two presentation modalities: audio–visual and audio-only, and within each presentation modality. Results across the two presentation modalities and results within the audio–visual presentation modality were the same except that the magnitude of the effects of the cue conditions was larger in the audio–visual presentation modality. There were some general differences and some speaker-specific differences in the findings between the audio–visual modality and the audio-only modality.

These findings and their relation to Lindblom's (1990) model of mutuality are discussed below.

Effects of Cue Conditions When Listeners Could See and Hear Speakers

Findings showed that use of alphabet cues and use of hand gestures while speaking resulted in significantly higher intelligibility scores and helpfulness ratings than the no-cues control condition for each of the 3 speakers. These findings are consistent with previous studies examining speaker-implemented alphabet cues (Beukelman & Yorkston, 1977; Crow & Enderby, 1989; Hustad, Auker, et al., 2003; Hustad, Jones, et al., 2003) and hand gestures (Garcia & Cannito, 1996; Garcia & Dagenais, 1998) and illustrate that both extralinguistic and linguistic information can have a powerful impact on intelligibility, relative to habitual speech. Indeed, listeners seemed able and willing to make use of either type of signal-independent information to fill in the gaps where signal-dependent acoustic information fails. As shown in Figure 1, the magnitude of the benefit for alphabet cues was relatively consistent among the three speakers (32%–36%). One explanation is that the linguistic content of the cues was not dependent on the motor skills of the speakers because digital enhancements were provided so that listeners could see the letter cues. This may have served to reduce the variability in performance among speakers with regard to clarity of their pointing gestures. The magnitude of benefit for gestures showed greater variability among the 3 speakers (22%–41%). One obvious explanation for this finding relates to differences in upper extremity motor control for the 3 speakers. Some of the speakers were able to produce better hand gestures than others, which may have resulted in a greater benefit from gestures for some speakers.

The pattern of results differed for the 3 speakers with regard to the differences between intelligibility scores and helpfulness ratings for alphabet cues compared with hand gestures. Two speakers showed no statistically significant difference in intelligibility scores between alphabet cues and gestures (Speakers A and C), indicating that both strategies were equally effective in enhancing intelligibility. This finding was further corroborated by the helpfulness ratings for the two strategies, which also did not differ. Thus, listeners performed similarly when speakers used each strategy; they also perceived the helpfulness of the strategies similarly. However, for one individual, Speaker B, alphabet cues resulted in higher intelligibility scores and helpfulness ratings than hand gestures. The magnitude of the benefit in intelligibility scores was 13%. This finding is especially puzzling in light of motor assessment results from the physical therapist, which indicated that the quality of upper extremity movement for gestures was better than the quality of upper extremity movement noted for alphabet cues. For this particular speaker, the information provided by alphabet cues may have mapped onto the speech signal in a more meaningful way, leaving less room for ambiguity, than the hand gestures. Consequently, the linguistic information provided by alphabet cues was perhaps easier for listeners to interpret than the extralinguistic hand movements of gestures, especially because the alphabet cues were digitally optimized and clearly visible. Another explanation for this difference, elaborated below, relates to the impact of each strategy on the acoustic signal produced by the speaker.

Consistent with other research examining listener perceptions of speech supplementation strategies (Hustad, 2001; Hustad & Gearhart, 2004), listeners' ratings of the helpfulness of each strategy followed the same pattern of results as intelligibility scores for each speaker. This seems to suggest that listeners may make their ratings of helpfulness on the basis of intelligibility. In the present study, listeners did not receive any quantitative or qualitative feedback regarding their orthographic transcriptions, yet their helpfulness ratings imply that they may have had reasonable insight into their performance on intelligibility tasks. It is also noteworthy that for the listeners who were in the audio extremity visual group, ratings of helpfulness for each speaker were generally positive (between 4 and 5.5 on a 7-point scale) for both the alphabet cues and gestures conditions.

Effects of Cue Conditions When Listeners Could Only Hear Speakers

For the audio-only presentation mode, findings differed from those observed for the audio–visual condition in the pattern of results and the magnitude of the differences. Results showed that alphabet cues were again superior to no cues with regard to intelligibility scores for all 3 speakers, and that gestures resulted in higher intelligibility scores than no cues for 2 of the 3 speakers (Speakers B and C). The magnitude of benefit for both alphabet cues and gestures was similar in the audio-only condition, ranging from approximately 8%–18% across speakers. Alphabet cues and hand gestures did not differ from one another in the audio-only condition for any of the 3 speakers. Collectively, these findings suggest that for some speakers, both alphabet cues and gestures may alter the acoustic signal in a way that enhances intelligibility, independent of visual information. That is, implementation of both strategies changes the signal-dependent acoustic information provided to listeners, resulting in a positive effect on intelligibility scores. However, the way in which the

Table 4. Summary of temporal acoustic measures completed for each speaker.

	Speaker A			Speaker B			Speaker C		
	Hab	AC	Ges	Hab	AC	Ges	Hab	AC	Ges
Speech rate in words per minute (including all pauses)	159	33	86	83	18	84	104	34	86
Articulation rate in words per minute (excluding all pauses greater than 200 ms)	165	103	117	95	50	98	125	97	128
% of total time spent articulating	95	33	73	87	41	86	84	36	67
% of total time spent pausing	5	67	27	13	59	14	16	64	33
Mean number of pauses greater than 200 ms	0	6	3	2	6	2	2	6	3
Average length of pauses (ms)	0	1,466	531	376	2,570	450	418	1,396	628

Note. Hab = habitual speech (no cues); AC = alphabet cues; Ges = iconic hand gestures.

strategies alter the acoustic signal and whether both strategies have similar effects on speech production cannot be determined from intelligibility data alone.

To characterize the nature of acoustic changes associated with strategy implementation, a descriptive secondary analysis was undertaken. In this analysis, simple, temporal acoustic measures of total utterance duration and duration of pauses greater than 200 ms (following Turner & Weismer, 1993) were obtained from spectrographic display. Measurements were made on the same six sentences for each speaker in each of the three speaking conditions (25% of the sample). Descriptive results, provided in Table 4, suggest that there were several differences in temporal aspects of speech among the three speaking conditions and among the 3 different speakers.

When implementing alphabet cues, temporal data suggest that overall speech rate was substantially slower (reduced by 67%–80%) for all 3 speakers relative to the no-cues habitual speech condition. Furthermore, articulation rate (speech rate excluding pauses greater than 200 ms) was also substantially slower (reduced by 22%–50%) when speakers implemented alphabet cues relative to the no-cues habitual speech condition. In addition, speakers spent the majority of their talking time pausing (between 59% and 67%) when they implemented alphabet cues. Examination of pause location indicated that speakers clearly paused between each word, with pauses ranging between 1.4 s and 2.6 s in duration. When speakers used habitual speech, very little of their talking time was spent pausing (5%–16%), and they paused less frequently (0–2 times per utterance) and for shorter durations (0.5 s on average). These clear temporal changes in speech associated with implementation of alphabet cues are likely to be one primary reason that intelligibility increased, even when listeners could not see the speakers or the alphabet cues. In particular, word boundaries were clearer for listeners, there was more time to process individual words, and speakers articulated more slowly, perhaps making better approximations of each word, when alphabet cues were implemented.

When implementing hand gestures, temporal characteristics of speech appeared to be less consistent than those associated with alphabet cues. Overall speech rate was reduced by 45% for Speaker A, 0% for Speaker B, and 28% for Speaker C, relative to the no-cues habitual speech condition for each speaker. It is also noteworthy that although habitual speech rate varied for these 3 individuals, speech rate associated with implementation of gestures was the same for all 3 speakers (approximately 85 wpm). Articulation rate was approximately 30% slower for Speaker A when gestures were produced simultaneously with speech; however, for the other 2 speakers, articulation rate was not affected by the implementation of gestures. The 3 speakers spent a variable percentage of their talking time pausing when they implemented gestures (14%–33%). For Speakers A and C, gestures resulted in at least twice as much pause time relative to habitual speech; for Speaker B, the percentage of time spent pausing was the same for habitual speech and gestures. With regard to pause frequency, each speaker paused two or three times while using gestures and zero or two times during habitual speech. However, pauses were generally longer in duration during the gestures condition. This variability in temporal characteristics of speech associated with production of gestures makes generalization across speakers difficult. Although the results from this study confirm previous reports that the use of hand gestures while speaking can cause speakers to modify aspects of production (Garcia & Cobb, 2000; Garcia et al., 1998), temporal differences in speech associated with production of gestures did not seem to map onto intelligibility data in a clear way (e.g., temporal data for Speaker A might suggest there would be intelligibility gains associated with implementation of gestures relative to habitual speech in the audio-only condition; however, this was not the case). Clearly, other variables not examined in the present article, such as formant-frequency–related measures, may provide additional explanatory data for the findings related to gestures.

Implications for the Model of Mutuality

Results of the present study have demonstrated that two speech supplementation strategies, alphabet cues and hand gestures—both considered signal-independent information—can have an important effect on speech intelligibility for individuals with severe dysarthria. Particularly notable is that although hand gestures and alphabet cues differ in important ways with regard to the nature of the information provided to listeners, both have similar effects on overall intelligibility. However, it would be inappropriate to conclude that these strategies operate simply by providing signal-independent information to listeners. Indeed, in the present study results showed that important signal-dependent changes occur when both strategies are implemented and that these changes alone can have an independent effect on intelligibility.

Using a simple additive model, the relative impact of signal-dependent information and signal-independent information on intelligibility can be parsed through calculations involving subtraction among mean intelligibility scores for each speaker. Descriptive analyses of data for alphabet cues suggest that changes in signal-dependent information alone (alphabet cues minus no cues within the audio-only condition) accounted for 10% of the overall intelligibility gains for Speaker A, 18% for Speaker B, and 14% for Speaker C. Similarly, changes in signal-independent information alone (alphabet cues minus no cues within the audio–visual condition) accounted for 32% of the overall intelligibility gains for Speaker A, 35% for Speaker B, and 36% for Speaker C. Thus, for alphabet cues, intelligibility gains appear to have been primarily associated with the signal-independent information provided by the cues themselves for each of the 3 speakers.

For hand gestures, the relative impact of signal-dependent and signal-independent information for each speaker was different than observed for alphabet cues. Signal-dependent changes in the acoustic signal alone that occurred with implementation of hand gestures (gestures minus no cues within the audio-only condition) accounted for 8% of the gains in intelligibility for Speaker A, 13% for Speaker B, and 10% for Speaker C. Similarly, changes in signal-independent information alone (gestures minus no cues within the audio–visual condition) accounted for 28% of the overall intelligibility gains for Speaker A, 22% for Speaker B, and 41% for Speaker C. Again, intelligibility gains appear to have been primarily associated with the signal-independent information provided by the gestures themselves for each of the 3 speakers. It is important to note, however, that this type of additive reasoning may oversimplify the complexity of the interactions among variable that contribute to speech intelligibility. However, it does lend some insight into the relative contributions of signal-dependent and signal-independent information for different speakers. Overall, the findings of the present study suggest that listeners have a strong propensity to make use of any and all sources of information available to them to attain mutuality with a speaker.

Limitations

The present study was experimental in nature and highly controlled with regard to a number of variables. Speakers followed a careful script that dictated their expression of gestures, alphabet cues, and message content, which likely resulted in findings that may be different than in real communication situations. Furthermore, digital enhancements of the alphabet cues may have resulted in elevated intelligibility scores for the speakers relative to what might actually be the case when listeners do not have the benefit of a pop-up box containing the target letter. The scripted and rehearsed nature of the hand gestures and the idealized presentation of the alphabet cues in the present study may have important consequences for the ecological validity of the findings of the present study. Listeners transcribed speech stimuli under optimal experimental listening conditions, but they did not have the opportunity to interact with the speakers, which may have an important influence on what they were able to understand. Results of the present study may reflect elevated intelligibility relative to what would occur in a real communication situation because of certain idealized experimental conditions. Conversely, results of the present study may actually be worse than what would occur in a real communication situation where there is opportunity for give and take between speaker and listener, there is more redundancy within the interaction, and both speaker and listener have a real communicative goal. Only research that takes place in the milieu can inform these speculations.

Clinical Implications and Conclusions

Clearly and not surprisingly, the most important benefit from both alphabet cues and gestures seems to occur when listeners can both see and hear speakers. In this situation, both intelligibility scores and listener ratings of the helpfulness of each strategy were enhanced for both alphabet cues and hand gestures implemented while speaking. If a similar magnitude of benefit (22%–41% gain in intelligibility) were observed in spontaneous speech with the implementation of alphabet cues or gestures, it would be difficult to deny the success of the intervention from a clinical perspective.

In considering the overall use of alphabet cues versus gestures, there appear to be important speaker-specific differences that impact the effectiveness of each strategy. For some speakers, gestures and alphabet cues may be equally effective. However, alphabet cues may be more effective in maximizing intelligibility for other speakers. One important clinical consideration is the speaker's motor abilities, particularly upper extremity control, for implementing hand gestures in a way that can be easily interpreted by others. The results from this study further substantiate the clinical value of speaker-implemented supplementation strategies, even for individuals with dysarthria who have moderately impaired upper extremity movement.

Intelligibility scores from the audio-only condition along with temporal acoustic data also illustrate that speakers alter production features of their speech when they implement speech supplementation strategies. Although both strategies had similar effects on intelligibility scores when visual information was eliminated, the temporal features of speech were very different for the two strategies. Findings were homogeneous and predictable among speakers with regard to alphabet cues and heterogeneous with regard to gestures. These somewhat unexpected temporal observations highlight the importance of the detailed study of explanatory factors associated with intelligibility. Additional research examining spectral acoustic differences may provide further information that aids in the understanding of how these strategies impact speech and, ultimately, intelligibility.

In summary, the present study represents the first report comparing the speech supplementation strategies of alphabet cues and hand gestures. Additional research is necessary to replicate the findings of the current study with a larger number of speakers who have different types of motor impairments to increase the validity and generalizability of the results. In addition, detailed profiling of upper extremity function of speakers may provide important information regarding the basic motor skills necessary for producing functional gestures and ideal motor skills for producing optimal gestures.

Acknowledgments

Portions of this article were presented at the 2002 Conference on Motor Speech Disorders in Williamsburg, VA. We thank the speakers with dysarthria for participating in this research and Julie Auker for assisting with digital videotape postproduction.

This research was supported by a New Investigator grant from the American Speech-Language- Hearing Foundation and a grant from the National Institute on Deafness and Other Communication Disorders (R03 DC005536).

References

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.

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.

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., Cannito, M. P., & Dagenais, P. A. (2000). Hand gestures: Perspectives and preliminary implications for adults with acquired dysarthria. *American Journal of Speech-Language Pathology, 9,* 107–115.

Garcia, J. M., & Cobb, D. S. (2000). The effects of gesturing on speech intelligibility and rate in ALS dysarthria: A case study. *Journal of Medical Speech-Language Pathology, 8,* 353–357.

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.

Garcia, J. M., Dagenais, P. A., & Cannito, M. P. (1998). Intelligibility and acoustic differences in dysarthric speech related to use of natural gestures. In M. P. Cannito, K. Yorkston, & D. Beukelman (Eds.), *Neuromotor speech disorders: Nature, assessment, and management* (pp. 213–227). Baltimore: Brookes.

Garcia, J. M., & Hayden, M. (1999). Young and older listener understanding of a person with severe dysarthria. *Journal of Medical Speech-Language Pathology, 7,* 109–112.

Hustad, K. C. (2001). Unfamiliar listeners' evaluation of the effectiveness of speech supplementation strategies for improving the effectiveness of severely dysarthric speech. *Augmentative and Alternative Communication, 17,* 213–220.

Hustad, K. C., Auker, J., Natale, N., & Carlson, R. (2003). Improving intelligibility of speakers with profound dysarthria and cerebral palsy. *Augmentative and Alternative Communication, 19,* 187–198.

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

Hustad, K. C., & Cahill, M. A. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology, 12,* 198–208.

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. *Journal of Medical Speech-Language Pathology, 10,* 279–285.

Hustad, K. C., & Gearhart, K. J. (2004). Listener attitudes toward individuals with cerebral palsy who use speech supplementation strategies. *American Journal of Speech-Language Pathology, 13,* 168–181.

Hustad, K. C., Jones, T., & Dailey, S. (2003). Implementing speech supplementation strategies: Effects on intelligibility and speech rate of individuals with chronic severe dysarthria. *Journal of Speech, Language, and Hearing Research, 46,* 462–474.

Kennes, J., Rosenbaum, P., Hanna, S., Walter, S., Russell, D., Raina, P., Bartlett, D., & Galuppi, B. (2002). Health status of school-aged children with cerebral palsy: Information from a population-based sample. *Developmental Medicine & Child Neurology, 44,* 240–247.

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

Kraat, A. (1987). *Communication interaction between aided and natural speakers: A state of the art report.* Madison: Trace R&D Center, University of Wisconsin—Madison.

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

Lloyd, L. L., & Fuller, D. R. (1986). Toward an augmentative and alternative communication symbol taxomony: A proposed super-ordinate classification. *Augmentative and Alternative Communication, 2,* 165–171.

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

Pellegrino, L., & Dormans, J. (1998). Definitions, etiology, and epidemiology of cerebral palsy. In J. Dormans & L. Pellegrino (Eds.), *Caring for children with cerebral palsy* (pp. 3–30). Baltimore: Brookes.

Rosenbaum, P., Walter, S., Hanna, S., Palisano, R., Russell, D., Raina, P., Wood, E., Bartlett, D., & Galuppi, B. (2000). Prognosis for gross motor function in cerebral palsy. *Journal of the American Medical Association, 288,* 1357–1363.

Turner, G. S., & Weismer, G. (1993). Characteristics of speaking rate in the dysarthria associated with amyotrophic lateral sclerosis. *Journal of Speech, Language, and Hearing Research, 36,* 1134–1144.

Winter, S., Autry, A., Boyle, C., & Yeargin-Allsopp, M. (2000). Trends in the prevalence of cerebral palsy in a population-based study. *Pediatrics, 110,* 1220–1225.

Wolfe, W. (1950). A comprehensive evaluation of fifty cases of cerebral palsy. *Journal of Speech and Hearing Disorders, 15,* 234–251.

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

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

Received June 18, 2004

Revision received December 10, 2004

Accepted March 8, 2005

DOI: 10.1044/1092-4388(2005/068)

Contact author: Katherine C. Hustad, 1500 Highland Avenue, Madison, WI 53705. E-mail: hustad@waisman.wisc.edu

Appendix. Stimulus sentences and accompanying hand gestures.

Sentence	Hand gestures
Pick up the phone and dial the number.	Grasping motion with hand, then circular movement with index finger
Slice the bread and butter it.	Downward movement with closed hand, then back and forth movement with closed hand
Start the car and put it in gear.	Turning motion with hand, then arm movement down
Stand up and then sit down.	Raising hand with palm up, then lowering hand with palm down
Take a big step or you will trip.	Forward movement of hand with palm facing self, then tilting hand over
Find a large bottle with a small top.	Hand held out with fingers extended from thumb, then fingers moved closer to thumb
Hit the ball over the infield.	Swinging motion with hand, then index finger pointing away toward the left
Take the plate and pass it on.	Taking from the left, then moving hand toward right
Crawl through the pipe and down the shaft.	Forward movement with hand perpendicular to body, then hand movement downward with pointed fingers
Back the car next to the curb.	Backward movement of hand perpendicular to body, then movement of hand to the right
Stop and turn around where you are.	Palm extended in halting motion, then circular movement of index finger in horizontal plane
Sharpen the knife before you slice it.	Back and forth movement with closed hand, then downward movement with closed hand
Take off your glasses and set them down.	Movement of hand away from eye, then placing hand down to the side
Shut the door and lock it.	Palm facing forward and moving away from body, then turning motion with hand
Pick up the trash and throw it away.	Grabbing motion to the left, then tossing to the right
Throw it up and watch it fall.	Opening hand in upward movement, then palm forward with waving motion down
Turn it up so I can hear it.	Turning motion with hand, then cupping hand to ear
Cover your eyes to block the sunlight.	Moving palm of hand toward eyes, then turning palm to face out
Roll it up and put it down.	Vertical, circular motion of closed hand, then placing hand down to the side
Be quiet so I can hear the news.	Index finger toward mouth, then cupped hand to ear
Read the list from the top to bottom.	Pointing away from body, then moving finger to a lower position
Catch the ball and throw it back.	Catching motion with back of hand moving toward self, then throwing motion
Gather all his things and toss them out.	Circular sweeping motion with hand, then motioning away from body
Take it out and put it on top.	Pulling movement toward body, then moving hand to higher position with palm down