Effects of Speech Supplementation Strategies on Intelligibility and Listener Attitudes for a Speaker with Mild Dysarthria

KATHERINE C. HUSTAD*

University of Wisconsin, Madison, WI, USA

Speech supplementation strategies (alphabet cues, topic cues, and combined topic and alphabet cues) have been shown to have a positive effect on speech intelligibility for many individuals with dysarthria, particularly those with severe and profound intelligibility deficits (Hustad, Auker, Natale, & Carlson, 2003a; Hustad, Jones, & Dailey, 2003b). However, less attention has been given to speakers with moderate and mild intelligibility problems; therefore, the effects of speech supplementation strategies are largely unknown for these individuals. The present study examined the effects of speech supplementation strategies on intelligibility scores and listener attitudes for one speaker with mild spastic dysarthria secondary to cerebral palsy. Results showed different findings from previous studies. In the present study, the only speech supplementation strategy that significantly increased intelligibility was alphabet cues. Attitude ratings for each strategy followed a different pattern than intelligibility scores. Results suggest that severity of dysarthria may play an important role in deciding which speech supplementation strategy to use.

Keywords: Augmentative and alternative communication; Speech intelligibility; Speech supplementation; Cerebral palsy; Dysarthria

A growing body of literature has reported that speech supplementation strategies can have an important effect on speech intelligibility (Beukelman & Yorkston, 1977; Beukelman, Fager, Ullman, Hanson, & Logemann, 2002; Crow & Enderby, 1989; Hustad et al., 2003a, 2003b) and on listener attitudes toward speakers who use the strategies (Hustad, 2001; Hustad & Gearhart, 2004). Alphabet cues, in which the speaker points to the first letter of each word as he or she produces the word, have been shown to increase intelligibility by 15 – 44% (Beukelman & Yorkston, 1977; Beukelman et al., 2002; Crow & Enderby, 1989; Hustad et al., 2003a, 2003b). Similarly, combined cues, in which the speaker points first to the topic of the message and then to the first letter of each word while simultaneously speaking, can lead to intelligibility increases of up to 35 – 40% (Hustad et al., 2003a, 2003b). Finally, topic cues, which involve the speaker pointing only to the topic of the message prior to producing the message, have been shown to increase intelligibility by 3 – 16% (Beukelman et al., 2002; Hustad et al., 2003a, 2003b). See Table 1 for a summary of the different speech supplementation strategies. Although each supplementation strategy has resulted in increases in intelligibility relative to habitual speech, studies suggest that the severity of the dysarthria of individual speakers may have an influence on which strategy is most effective (Hustad et al., 2003a, 2003b).

Hustad et al. (2003a) examined the effects of speech supplementation strategies on intelligibility scores for three speakers with profound dysarthria (i.e., below 10% intelligible). Results showed that, for each speaker, combined cues resulted in higher intelligibility scores than alphabet cues, topic cues, and no cues. Furthermore, alphabet cues resulted in higher
intelligibility scores than topic cues and no cues. Finally, for two of the three speakers, topic cues resulted in higher intelligibility scores than no cues.

In a study examining five speakers with severe dysarthria (i.e., intelligibility between 15 and 30%), a different pattern of results emerged from that observed for speakers with profound dysarthria. Hustad et al. (2003b) found that, for each speaker, combined cues resulted in higher intelligibility scores than topic cues and no cues; and alphabet cues resulted in higher intelligibility scores than topic cues and no cues. Interestingly, the difference between topic cues and no cues was not significant, nor was the difference between alphabet cues and combined cues. Findings of this study suggest that topic cues did little to enhance intelligibility alone or in combination with alphabet cues for speakers with severe dysarthria.

Beukelman et al. (2002) examined the influence of alphabet and topic supplementation on intelligibility scores for speakers with dyarthria of varying severity secondary to traumatic brain injury. Descriptive results for individual speakers with moderate dysarthria (between 35 and 65% intelligible) suggested that both alphabet cues and topic cues resulted in large increases in intelligibility. However, alphabet cues seemed to result in greater intelligibility increases than topic cues. The influence of combined cues on the intelligibility of speakers with moderate dysarthria has not been studied.

Little is known about the influence of speech supplementation strategies on the intelligibility of speakers with milder dysarthria (above 70% intelligible). Descriptive data from Beukelman and colleagues (2002) suggests that both alphabet cues and topic cues resulted in similar increases in intelligibility scores relative to habitual speech for one speaker with mild dysarthria. The influence of combined cues has not been examined for speakers with mild dysarthria. Speakers with mild dysarthria are an important population to study because they may not have access to or a consistent need for voice output AAC systems, which may prevent referrals for AAC services. However, these speakers may experience communication difficulties in certain situations, such as when they are communicating with unfamiliar communication partners or when they are in adverse environmental situations. For these individuals, simple low-tech speech supplementation strategies may be the only AAC options that are available.

Decisions regarding whether to adopt a particular speech supplementation strategy should be based upon its impact on communication, which includes not only intelligibility but also listener attitudes toward a speaker using the target strategy (Hustad & Gearhart, 2004). The Encyclopedia Brittanica (June, 15, 2004) defines attitude as “a predisposition to classify objects and events and to react to them with some degree of evaluative consistency.” It is widely accepted that the construct of attitude has three constituent components: affective, cognitive, and behavioral (Antonak & Livneh, 1988; Eiser, 1986; Greenwald, Brock, & Ostrom, 1968; Triandis, Adamopoulos, & Brinberg, 1984). Generally, the affective dimension refers to feelings, preferences,
and emotions of a person toward an attitude object. The cognitive dimension refers to beliefs, opinions, or thoughts toward an attitude object. The behavioral dimension refers to the course of action an individual would take with regard to the attitude object (Greenwald et al., 1968; Triandis et al., 1984). Hustad and Gearhart (2004) examined the attitudes of listeners toward seven speakers who had severe or profound dysarthria and who used speech supplementation strategies. Results showed that attitudes tended to be highest for the behavioral domain and that ratings for cognitive and affective domains were similar. This finding is encouraging because it suggests that listeners are willing to interact with speakers who have dysarthria, even if they think and feel less positively about the speakers. Hustad and Gearhart also found that ratings for all domains followed a similar pattern of results to intelligibility scores and were, in fact, highly correlated with intelligibility scores for speakers with severe and profound dysarthria who used speech supplementation strategies. In general, global attitude results showed that attitudes toward the different speech supplementation strategies varied with the severity of each speaker’s dysarthria. However, across all speakers, attitudes were more positive towards those who used alphabet cues and combined cues than for those who communicated only by habitual speech. Attitudes also were more positive when speakers used combined cues compared with topic cues. The influence of topic cues on attitudes varied. In the case of speakers with more severe dysarthria for topic cues, attitudes were more positive than for no cues; topic cues did not have an affect on attitudes towards speakers with less severe dysarthria.

Attitudes toward speakers with mild dysarthria have not been studied, but may have an important effect on whether listeners will interact with speakers who are using AAC to supplement their speech and, ultimately, on the number of communication opportunities that are available to a speaker. The present research compared the effects of three speech supplementation strategies (topic cues, alphabet cues, and combined topic and alphabet cues) and a control condition on intelligibility and listener attitudes for one speaker with mild dysarthria secondary to cerebral palsy. The following research questions were addressed: How do speech supplementation strategies (topic cues, alphabet cues, and combined cues) affect intelligibility scores relative to a habitual speech control condition; and How do speech supplementation strategies (topic cues, alphabet cues, and combined cues) affect attitude ratings relative to a habitual speech control condition.

### METHOD

#### Participants

One speaker with dysarthria and 24 listeners without disabilities participated in this experiment. The speaker produced a standard corpus of speech stimuli using three speech supplementation strategies (topic cues, alphabet cues, and combined topic and alphabet cues) and habitual speech. Listeners viewed video tapes of the speaker using each supplementation strategy, transcribed what they heard, and rated their attitudes toward the speaker following completion of each experimental task.

The speaker was a 52-year-old male who had a medical diagnosis of spastic cerebral palsy. His speech was characterized by spastic dysarthria with salient perceptual features that included strained-strangled vocal quality, imprecise consonants, short phrases, reduced rate of speech, and mild hypernasality. The speaker used speech as his primary mode of communication with familiar partners and a voice output augmentative communication device for communication with unfamiliar partners and for repair of communication breakdown. While his dysarthria was quite prominent, his intelligibility impairment was judged to be mild by a certified speech-language pathologist. Results of the Sentence Intelligibility Test (Yorkston, Beukelman, & Tice, 1996) revealed that intelligibility was 75% to unfamiliar listeners. The speaker also met the following criteria: (a) able to produce connected speech consisting of at least eight consecutive words; (b) American English as a primary language; (c) functional literacy skills at or above the sixth grade level; (d) corrected or uncorrected vision within normal limits per self-report; (e) hearing within normal limits per self report; and (f) able to accurately direct select letters and orthographically represented phrases from a communication board.

Twenty-four different listeners, 12 males and 12 females, were randomly assigned to view tapes of the speaker. Each listener viewed the speaker, producing a different narrative passage, in each of the four experimental conditions (habitual speech, topic cues, alphabet cues, and combined cues). Participation took approximately 1 hour. All listeners were currently attending college or graduate school. Half of the listeners were male and half were female. Inclusion criteria required that each listener: (a) pass a pure tone hearing screening test at 25 dB SPL for 250 and 500 Hz, and 1, 4, 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.
per self-report; (d) be native speakers of American English; and (e) have no identified language, learning, or cognitive disabilities per self-report. The mean age of listeners was 20.7 years (SD = 1.6 years).

**Materials**

The speaker produced four narrative passages that have been described elsewhere (Hustad & Beukelman, 2001, 2002; Hustad et al., 2003b). Each passage consisted of 10 sentences that ranged between five and eight words in length and had a sixth grade reading level. Passages pertained to a sporting event; a natural disaster; purchasing a vehicle; and independence day.

**Procedures**

The speaker with dysarthria produced each of the four narrative passages in each of four strategy conditions (using habitual speech, while using alphabet cues, while using topic cues, and while using combined cues). Prior to using each strategy, the speaker was provided with a verbal description of the strategy, an explanation of the purpose of the strategy, and a model for using the strategy. Before the experimental tapes were recorded, the speaker practiced using each of the three strategies on a set of rehearsal sentences. Learning time prior to recording was less than 15 min. per strategy.

The speaker was audio and video recorded in a quiet environment within his home. During recording, the speaker was required to use alphabet cues, topic cues, and combined cues with 100% accuracy. In addition, he was required to produce all words within each sentence, as per a written script shown on a laptop computer and an auditory model produced by an experimenter. The speaker was asked to repeat any sentence in which these criteria were not met. Repetition was necessary for fewer than 5% of the sentences. The full protocol took approximately 5 hours, including periodic breaks.

Following procedures detailed elsewhere (Hustad & Cahill, 2003; Hustad et al., 2003b), digital audio and video recordings were edited to create stimulus tapes for playback to listeners. Recordings were transferred to computer via digital-to-digital interface (IEEE 1394 for digital video and S/PDIF for audio). Because the video camera was positioned directly in front of the speaker, it was difficult to see clearly the topics and letters to which the speaker pointed during the topic, alphabet, and combined cues conditions. Consequently, presentation of target letters and topics was digitally enhanced so that listeners could easily see referents to which the speaker pointed. For alphabet cues, 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. For topic cues, 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. For combined cues, the topic was displayed first and then individual graphemes that corresponded with the production of each word were displayed following the same conventions as the topic and alphabet cues conditions. Instructions for each task were presented on the stimulus tapes for playback to listeners.

Individual listeners viewed the broadcast quality videotapes in a quiet, sound attenuating room. Each listener was seated directly in front of a 25-inch television monitor, which was approximately 3 feet away. A digital video cassette player and external speaker were attached to the monitor; and peak audio output levels were calibrated to be approximately 65 dB SPL from where listeners were seated.

Listeners were told that they would complete four different tasks in which the speaker used three different strategies (and a control condition where no strategies were used). A brief explanation of each condition was provided. Listeners were instructed to watch and listen and to follow all instructions presented on the videotape. They were told that all productions were grammatically correct and meaningful and that the sentences within each task constituted a short story. Listeners were encouraged to write down everything they thought they understood, taking their best guess if they were unsure. They were instructed to take as much time as they needed to complete their transcriptions and attitude ratings. They were not given the opportunity to replay target sentences.

The order of presentation of the four experimental conditions (habitual speech, topic cues, alphabet cues, and combined cues) was counterbalanced so that each of the 24 listeners viewed the cue conditions in a different sequence. Across all four tasks, each narrative passage was presented in only one task so that listeners heard different speech stimuli for each task. Assignment of individual narratives to each of the four experimental conditions was evenly distributed across listeners and cue conditions, so that data for each cue condition represented all narratives.
Dependent Measures
The two dependent variables for this study were intelligibility scores and listener attitude ratings. Intelligibility scores were determined by calculating the percentage of words transcribed correctly for each experimental task and listener. Transcriptions from listeners were scored by tallying the number of words that were an exact phonemic match to the target word. This number was then divided by the total number of words possible and multiplied by 100 to yield percent intelligibility scores for each listener in each cue condition.

Listener attitude ratings were obtained by having participants respond to three questions, targeting cognitive, affective, and behavioral domains of attitude on a seven-point Likert scale ranging from 1 (disagree strongly) to 7 (agree strongly). The questions were: I would be willing to communicate with this person in a class or at work if he/she used this strategy; I think this person is an effective communicator using this strategy; and I would feel comfortable communicating with this person in a class or at work if he/she used this strategy. Responses to each question were then averaged to yield a summary “attitude” rating that was subjected to statistical analysis.

Experimental Design and Analysis
This study employed a $2 \times 4$ split plot design for each of the two dependent variables (intelligibility scores and attitude ratings) (Kirk, 1995). The between subjects measure was gender and it had two groups, male and female. The within subjects measure was cue condition and its four categories were habitual speech, topic cues, alphabet cues, and combined cues. To control the Type I error rate, the Bonferroni procedure was used. For the parametric analysis of intelligibility scores, three omnibus tests (two main effects and an interaction) and six planned follow-up contrasts were performed. For the nonparametric analysis of attitude ratings, one omnibus test and six planned follow-up contrasts were performed. Each set of analyses was allotted a total z level of 0.01, which was evenly divided among the statistical tests. A probability value less than or equal to 0.001 (9/0.01) was necessary for an intelligibility test to be considered significant, and a probability value less than or equal to 0.0014 (7/0.01) was necessary for an attitude test to be considered significant.

RESULTS
Intelligibility
Mean intelligibility scores are shown in Figure 1. ANOVA results revealed a significant main effect of cue condition (see Table 2 for statistics). Six pairwise contrasts were examined to characterize this main effect (see Table 3 for statistics). Results showed that use of alphabet cues resulted in significantly higher intelligibility scores than use of habitual speech (11.33%), topic cues (19.25%), and combined cues (9.04%). Intelligibility scores associated with combined cues did not differ from intelligibility scores associated with topic cues. Furthermore, intelligibility scores for habitual speech did not differ from those for combined cues or topic cues. Neither the main effect of gender nor the interaction between gender and cue conditions was significant.

Attitude Ratings
Mean attitude ratings are shown in Figure 2. Nonparametric ANOVA results using the Friedman test revealed a significant main effect of cue condition ($p < 0.001$). Six pairwise contrasts were examined using the Wilcoxon signed rank test to characterize this main effect (see Table 3 for statistics). Results showed that alphabet cues resulted in significantly higher attitude ratings than habitual speech (1.08 Likert points), and topic cues (1.29 Likert points). In addition, combined cues resulted in significantly higher attitude ratings than habitual speech (0.85 Likert points), and topic cues (1.05 Likert points). Attitude ratings for habitual speech did not differ from those associated with topic cues; and attitude ratings for combined cues did not differ from those associated with alphabet cues. The main effect of gender was evaluated with the nonparametric Mann–Whitney $U$-test and was not significant.
DISCUSSION

Speech supplementation strategies such as alphabet cues, topic cues, and combined topic and alphabet cues have been shown to have positive effects on speech intelligibility for many individuals with dysarthria, particularly those with severe and profound intelligibility deficits (Hustad et al., 2003a, 2003b). However, less attention has been given to speakers with moderate and mild intelligibility problems. The present study examined the effects of speech supplementation strategies on intelligibility scores and listener attitudes for one speaker with mild dysarthria secondary to cerebral palsy. The findings were different from previous studies. In the present study, the only speech supplementation strategy that significantly increased intelligibility was alphabet cues. Interestingly, intelligibility scores for topic cues and combined cues did not differ from one another or from habitual speech. Attitude ratings for each strategy followed a different pattern than intelligibility scores. Findings are discussed in the sections that follow.

Effects of Supplementation Strategies on Intelligibility

Results of the present study are both consistent and inconsistent with previous research that has compared the influence of various supplementation strategies on intelligibility scores. A common finding across previous studies is that the use of combined cues resulted in higher intelligibility scores than did the use habitual speech and topic cues (Hustad et al., 2003a, 2003b). The obvious explanation for the result in the present study, that combined cues did not differ from one another or from habitual speech. Attitude ratings for each strategy followed a different pattern than intelligibility scores. Findings are discussed in the sections that follow.

TABLE 2 Split plot ANOVA Results for Intelligibility Scores

<table>
<thead>
<tr>
<th>Source</th>
<th>Sums of squares</th>
<th>df</th>
<th>F</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cues</td>
<td>4517.36</td>
<td>2.65</td>
<td>14.95*</td>
<td>0.41</td>
</tr>
<tr>
<td>Gender × cues</td>
<td>384.62</td>
<td>2.65</td>
<td>1.27</td>
<td>0.06</td>
</tr>
<tr>
<td>Error (effectiveness)</td>
<td>6649.77</td>
<td>58.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>1.26</td>
<td>1.00</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>Error (gender)</td>
<td>3452.48</td>
<td>22.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.001.

TABLE 3 Statistical contrasts for Intelligibility Scores and Attitude Ratings

<table>
<thead>
<tr>
<th>Contrast</th>
<th>Mean difference</th>
<th>df</th>
<th>Intell</th>
<th>Att</th>
<th>Intell</th>
<th>Att</th>
<th>Intell</th>
<th>Att</th>
<th>Intell (t)</th>
<th>Att (z)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC – NC</td>
<td>2.29</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>2.84</td>
<td>–</td>
<td>0.81</td>
<td>–</td>
<td>–3.04*</td>
<td>–</td>
</tr>
<tr>
<td>CC – TC</td>
<td>10.21</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>3.21</td>
<td>–</td>
<td>3.18</td>
<td>–</td>
<td>–3.42*</td>
<td>–</td>
</tr>
<tr>
<td>CC – AC</td>
<td>–9.04</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>2.45</td>
<td>–</td>
<td>–3.69*</td>
<td>–</td>
<td>–1.77</td>
<td></td>
</tr>
<tr>
<td>TC – NC</td>
<td>–7.92</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>3.43</td>
<td>–</td>
<td>–2.31</td>
<td>0.28</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < 0.001.

Note. Intell, intelligibility scores; Att, attitude ratings; NC, no cues; TC, topic cues; AC, alphabet cues; CC, combined cues. Test statistics for intelligibility scores are from parametric t-tests; test statistics for attitude ratings are from non-parametric Wilcoxon signed rank tests (large sample approximation).
cues may have actually misled or confused them because the topic cues did not provide unique content. That is, listeners may have over-interpreted the information provided within the topic cues and, in effect, may have been led astray as a result. Thus, topic cues may have mitigated the benefit of the alphabet cues that were provided within the combined cues strategy.

In the present study, the finding that the use of alphabet cues resulted in higher intelligibility scores than the use of topic cues and habitual speech, is consistent with previous studies. However, in the present study, the use of alphabet cues increased intelligibility by approximately 10%; in previous studies focused on speakers with cerebral palsy, the magnitude of this benefit was larger, with most studies demonstrating benefits of 25% or more. One explanation may relate to the relatively mild intelligibility deficit that the speaker experienced (i.e., 80% intelligibility for habitual speech) and the strong information bearing capability of the habitual speech signal. That is, because most of the content information could be transmitted successfully via the speech signal, there was less room for improvement and, consequently, less need for the alphabet cues.

Effects of Supplementation Strategies on Attitude Ratings

Results of attitude measures differed from previous studies in that attitude ratings did not directly parallel intelligibility scores. Three important differences were present between intelligibility findings and attitude findings: When the speaker used combined cues, attitude ratings were higher than when he used topic cues and no cues; and attitude ratings were the same for alphabet cues and combined cues. Hustad (2001) suggested that listeners may have more favorable attitudes when they believe a speaker to be trying harder. One way that a speaker may “try harder” is to provide listeners with specific and detailed information, as with alphabet and combined cues. Although combined cues did not enhance intelligibility, listeners may have perceived the effort made by the speaker positively, as reflected in attitude ratings. Similarly, listeners’ attitudes toward the speaker using alphabet cues and combined cues did not differ, perhaps because both strategies employed first letter information that was specific and frequent. Thus, listeners believed the speaker to be making every effort to enhance his intelligibility.

Another noteworthy observation regarding attitudes relates to the positive orientation of all ratings. For alphabet cues, average attitude ratings were approximately 6 on a seven-point Likert scale; while ratings were just under 5 on the scale for habitual speech and topic cues. In previous research examining listener attitudes toward speakers with severe and profound dysarthria, maximal attitude ratings were below 6. Overall, attitude ratings in the present study were very strong and indicated that listeners thought, felt, and behaved positively toward this speaker. One variable that may impact ratings is the homogeneity of listeners. In the present study and in previous studies examining attitudes of listeners toward speakers who implemented supplementation strategies, listeners tended to be similar with regard to age, education, and socioeconomic status. Results may have been different for listeners who varied on these and other variables.

Limitations

Clearly the results of the present study must be regarded cautiously for several important reasons. First, only one speaker with mild dysarthria participated. Second, the presentation of all of the cue conditions was optimized via digital enhancements so that the target topic and/or letter were clearly visible and 100% accurate. In addition, the digitally imposed cues made it so that the listener never actually had to look away from the speaker. In real life, listeners typically look back and forth between a speaker’s face and the visual display, potentially missing visual oral–facial information. Thus, digital optimization may serve to inflate the benefit of each of the cue conditions and also may compromise the extent to which results can be generalized to real speakers in real communicative environments. Also, the speaker produced a scripted set of utterances that formed cohesive narratives which were delivered to listeners in a monologue fashion, within a pristine listening environment. The effects of speech supplementation strategies on speech intelligibility and listener attitudes may be different if any one or several of these variables are altered.

Clinical Implications and Future Directions

Results of the present study suggest that alphabet cues can be a valuable tool for increasing intelligibility of speakers with mild dysarthria. This strategy may be most useful in communication breakdown situations for speakers with reasonable habitual intelligibility, like the individual in the present study. Alphabet cues increased intelligibility in the present study by approximately 10% (up to 90%). The finding that intelligibility peaked at 90% suggests that there may be a limit to the extent that alphabet cues can
increase intelligibility, at least within a controlled experimental context. However, this limit is likely sufficient to enable successful communication. Future research should examine additional speakers with mild and moderate dysarthria to increase the generalizability of the findings of the present study. Also, different communication situations that are more ecologically valid should be studied.

Acknowledgements

The author thanks the speaker with dysarthria for participating in this research and also Julie Auker for assisting with digital videotape post-production. Data collection from the speaker and listeners was supported by a New Investigator grant from the American Speech Language and Hearing Foundation.

References


