Unfamiliar Listeners’ Evaluation of Speech Supplementation Strategies for Improving the Effectiveness of Severely Dysarthric Speech

Katherine C. Hustad
Department of Communication Disorders, The Pennsylvania State University, University Park, Pennsylvania, USA

This study examined listener attitudes toward three speech supplementation strategies (topic cues, alphabet cues, and combined topic and alphabet cues) associated with the speech of four individuals with severe dysarthria. Listeners saw experimentally imposed visual images of each strategy in conjunction with auditory presentation of the habitual speech of four individuals with dysarthria. Using a 7-point Likert scale, listeners rated how effective they thought the speakers were, how willing they would be to communicate with the speakers, and how persistent they were in trying to understand the speakers in each strategy condition and a control condition in which no cues were provided. The results revealed that ratings of communication effectiveness, willingness to communicate with the speakers, and listener persistence were each more favorable in the combined cues condition than in any other cue condition. The results suggest that augmentative and alternative communication strategies providing frequent and specific cues regarding the content and constituent words of a message may enhance the attitudes of listeners.

KEY WORDS: attitudes, dysarthria, speech supplementation

Many individuals who use augmentative and alternative communication (AAC) systems employ natural speech as one of several modes of communication. In spite of reduced intelligibility, these individuals are often able to communicate successfully using speech in specific situations with certain communication partners. For those AAC users who are able to produce natural speech, it is sometimes the most convenient way to communicate even though listeners may have to work harder to understand the message (Yorkston, Beukelman, Strand, & Bell, 1999).

Low-technology compensatory strategies to enhance or augment the intelligibility of natural speech have received increasing attention in recent years (Beliveau, Hodge, & Hagler, 1995; Dongilli, 1994; Garcia & Cannito 1996; Hustad & Beukelman, 2000). Strategies such as alphabet supplementation (Beukelman & Yorkston, 1977; Crow & Enderby, 1989; Hustad & Beukelman, 2000), topic supplementation (Carter, Yorkston, Strand, & Hammen, 1996; Dongilli, 1994; Garcia & Cannito, 1996; Hustad & Beukelman, 2001), and combined (alphabet and topic) supplementation (Beliveau et al., 1995; Hustad & Beukelman, 2001), used in conjunction with natural speech, have been shown to increase the intelligibility of speech that is moderately to severely dysarthric.

Alphabet supplementation involves an alphabet board that speakers use to indicate the first letter of each word while simultaneously speaking it (Beukelman & Yorkston, 1977; Crow & Enderby, 1989; Hustad & Beukelman, 2000, 2001; Yorkston et al., 1999). This strategy appears to improve intelligibility by enhancing listener knowledge regarding the phonetic content of the message produced by the speaker with dysarthria. When this strategy is employed, listeners receive first-letter-of-word orthographic information that may serve to constrain the number of possible word choices, thus increasing the probability of correct word identification (Hustad & Beukelman, 2000). Improvements in sentence intelligibility associated with the presentation of alphabet cues to listeners have ranged from 15% (habitual speech intelligibility without cues = 59%, speech intelligibility with alphabet cues = 74%) (Crow & Enderby, 1989) to 45% (habitual speech intelligibility without cues = 25%, speech intelligibility with alphabet cues = 70%) (Beukelman & Yorkston, 1977).

Topic supplementation involves the use of a communication board containing key words or phrases that represent topics or contexts in the speaker’s life. Topics can be represented orthographically or through the use of picture symbols. Speakers use this strategy...
by indicating the topic of a forthcoming message prior to producing it using speech (Dongilli, 1994; Hustad & Beukelman, 2000, 2001; Yorkston et al., 1999). Hustad and Beukelman (2000, 2001) suggest that this strategy works by helping listeners narrow expectations for the semantic content of messages. Research has demonstrated that topic supplementation enhances sentence intelligibility to a lesser extent than alphabet supplementation (Hustad & Beukelman, 2001), with the magnitude of benefit in the 9% range (Carter et al., 1996; Dongilli, 1994; Garcia & Cannito, 1996; Hustad & Beukelman, 2001).

Finally, combined (alphabet and topic) supplementation involves having the speaker present the listener with a topic cue followed by a first-letter cue for each word as it is spoken. Research concerning the effects of combined cues on word (Beliveau et al., 1995) and sentence intelligibility (Hustad & Beukelman, 2001) has shown that combined cues result in greater gains than either topic or alphabet cues alone. Hustad and Beukelman (2001) found that the average benefit for sentence intelligibility across four speakers with severe dysarthria was 35% (habitual speech intelligibility without cues = 19%, speech intelligibility with combined cues = 54%).

Clearly, evidence concerning the effects of speech supplementation strategies on intelligibility suggests that they can markedly enhance the usefulness of speech. However, because of the interactive nature of communication, it is important to determine how communication partners view the use of these strategies. Lasker and Bedrosian (2000) emphasized that partner attitudes toward AAC systems and strategies can greatly influence the acceptance of AAC systems and strategies by a user. These authors have developed an AAC Acceptance Model that defines the construct of attitudes (for both partner and AAC user) according to three components: affective, behavioral, and cognitive. The affective component of attitude involves emotional and physiologic reactions to AAC such as comfort level or frustration. The behavioral component involves the willingness to use AAC or to interact with someone using AAC. The cognitive component of attitude relates to speaker and partner perceptions of communication skills.

Research on the attitudes of communication partners toward individuals who use AAC has been diverse with respect to AAC user characteristics and variables of interest. However, most of the research has focused primarily on what Lasker and Bedrosian (2000) identified as the cognitive component of attitude. Although a limited number of studies have been conducted, existing research seems to show that, as the sophistication of AAC technique increases, listener attitudes tend to become more favorable (Gorenflo & Gorenflo, 1991; Lasker, 1997). Specifically, Gorenflo and Gorenflo (1991) found that listeners rated a speaker most favorably when he used a voice output AAC device than when he used low-technology strategies. However, they also found that listeners rated the use of low-technology alphabet boards, on which the AAC user spelled each word of his target message, significantly more favorably than the use of natural speech and gestures (without AAC strategies). Similarly, Lasker (1997) found that listeners rated a speaker more favorably when he used a voice output AAC device than when he used a low-technology communication notebook or his natural speech (without AAC strategies). Interestingly, Lasker also found that preference for communication mode varied by communication partner, with family members and speech-language pathologists selecting natural speech as their preferred mode in a forced-choice task and same-age peers selecting the voice output AAC device as their preferred mode.

The purpose of the present study was to determine how different low-technology speech supplementation strategies influence the attitudes of naive listeners toward speakers with severe dysarthria. The speech supplementation strategies of interest were topic cues, alphabet cues, and combined cues (as well as a no cues control condition) presented in conjunction with natural speech. Following the presentation of speech stimuli, listeners were asked to rate three different variables—communication effectiveness, willingness to interact with speakers, and persistence in trying to understand speakers—on a 7-point Likert-type scale. Communication effectiveness, a cognitive component of attitude, is described as “the success with which a speaker is perceived to interact, or exchange information, in various communication situations compared with nondisabled speakers of similar age, background, and experience” (Hustad, Beukelman, & Yorkston, 1998, p. 297). Willingness to interact with speakers and persistence in trying to understand speakers are both behavioral components of attitude in that they measure the hypothetical and actual behaviors of the listener with respect to the speaker. Accordingly, this study sought to answer three questions:

1. What is the effect of speech supplementation strategies on listener perceptions of communication effectiveness?
2. What is the effect of speech supplementation strategies on listener willingness to interact with speakers?
3. What is the effect of speech supplementation strategies on listener ratings of persistence in trying to understand speakers?

METHOD

Research Design

A repeated-measures design was employed for each of the three dependent variables of interest in this study. The within-subjects repeated measure was
cue condition and its constituent groups were no cues, topic cues, alphabet cues, and combined cues. The dependent variables were listener ratings of communication effectiveness, willingness to interact with speakers, and persistence in trying to understand the speakers.

Participants

This study involved two types of participants: speakers and listeners. The speakers were individuals with dysarthria who produced speech stimuli and listeners were nondisabled individuals who listened to the speakers in each of the experimental conditions.

Speakers

Four women with severe dysarthria secondary to cerebral palsy served as speakers. Each speaker indicated that she used speech as her primary mode of communication and AAC as a secondary mode of communication in adverse situations or during communication breakdowns. Speakers had intelligibility between 15 and 25% as measured by the Sentence Intelligibility Test (Yorkston, Beukelman, & Tice, 1996), were native speakers of American English, and were able to produce and repeat connected utterances that were at least eight words in length. Two of the speakers had mixed spastic-athetoid cerebral palsy and two of the speakers had spastic cerebral palsy. All speakers had significant motor impairments that required them to use wheelchairs for mobility. See Table 1 for detailed information regarding each speaker.

Listeners

Seventy-four individuals without disabilities served as listeners for this study. However, because of missing data for the dependent measures of interest, the results reported are for 68 listeners who contributed information to each cell of the design. Data reported in this study are part of a larger study that examined the effects of context on decoding and comprehension of dysarthric speech (Hustad, 1999). Listeners met the following criteria: (a) no known hearing loss, per self-report; (b) age between 18 and 31 years; (c) no more than incidental experience listening to or communicating with persons having communication disorders; (d) native speakers of American English; and (e) no identified language, learning, cognitive, or other disabilities, per self-report. University-level literacy skills were assumed because all listeners were high school graduates who were either currently attending or had graduated from university. Corrected or uncorrected visual acuity that was within normal limits was also assumed based on listeners’ responses to question (e) above. Listeners had a mean age of 21 years (SD = 2.46). Gender composition was 8 males and 60 females. Gender was not a variable of interest for the larger study (Hustad, 1999); thus, no effort was made to balance the number of male and female listeners.

Materials

Speech Stimuli

The four speakers with dysarthria produced a total of 160 sentences across 16 narrative passages of 10 sentences each. The content of the narratives and their constituent sentences represented situational information such as vacations, holidays, sporting events, and buying a car that is common to native adult speakers of American English. The length and content of stimulus materials were equalized across several different linguistic parameters, including number of words per sentence, number of words per passage, number of syllables per sentences, type-token ratio of each passage, and reading level of each passage. Summary statistics for stimulus materials are presented in Table 2.

Topic Cues

Topic cues were short, descriptive phrases that reflected the main idea of each narrative passage. Only one topic cue was provided for each narrative; thus, a total of 16 different topic cues were employed in the study. Topic cues were constructed to be representative of the entire narrative and of each sentence within the narrative. The appropriateness of each sentence-topic pair was rated by 10 independent judges. For the sentences employed in this study, at least 90% of judges rated the associated topic as appropriate. See Table 3 for sample topic cues and associated sentences.

Procedures

Recording Speech Samples

Audio recordings of each speaker were made in a double-walled sound-proofed room using a digital audiotape (DAT) recorder. Each speaker produced the full corpus of stimulus materials consisting of 160 sentences. A Crown CM-312 microphone was positioned via a headband so that it was 5 cm from the speaker’s mouth regardless of head movement. Speakers viewed an orthographic representation of each stimulus sentence on a laptop computer placed directly in front of them and were asked to produce each sentence following the experimenter’s model. They were instructed to speak “naturally,” as they would in typical communication situations. The rate and prosody for each speaker were not controlled.
Constructing Stimulus Tapes

Recordings were transferred from DAT to a personal computer via a digital-to-digital sound card (S/PDIF interface) (44.1-kHz sampling rate; 16-bit quantization). Speech samples were edited to remove extraneous comments, and each sentence was amplitude normalized to 69 dB. Signal-to-noise ratios were above 45 dB for all recordings.

To associate visual images of each cue condition with speech samples, sound files were imported into Adobe Premiere 5.1, a digital video (DV) software package. For counterbalancing purposes, the full corpus of stimuli produced by each speaker was associated with each cue condition.

For the topic cues condition, an orthographic representation of the target topic was presented on a television screen for the duration of each sentence. For the alphabet cues condition, the first letter of each word produced by the speaker was presented in real time with the auditory speech signal. Listeners saw an alphabet board on a television screen, and the first letter of each word produced was circled in red for the duration of that word. In the combined cues condition, alphabet and topic cues were presented simultaneously on the television screen. Finally, in the no cues condition, a blank screen was presented for the duration of the utterance. Additional details regarding construction of stimulus tapes are provided elsewhere (Hustad & Beukelman, 2001). The final DV tapes presented to listeners were broadcast quality (DV-NTSC [National Television Standards Commission]).

Randomization and Counterbalancing

A Latin Square counterbalancing scheme (Campbell & Stanley, 1963; Cook & Campbell, 1979; Kirk, 1995) was employed to prevent the possibility of an order effect for cue conditions. Using the Latin Square, all possible permutations of presentation order for cue conditions were represented across all listeners. This was accomplished through the development of 24 unique stimulus tapes that each contained a full set of stimuli. On each of the stimulus tapes, individual speakers appeared once in a given cue condition so that listeners saw four different speakers, each associated with a different cue condition. Across all tapes, each speaker was presented in each cue condition five to seven times, and each speaker appeared first, second, third, and fourth in each cue condition one to two times. On each of the 24 stimulus tapes, individual sentences occurred only once. The length, complexity, and predictability of stimulus material were equated through randomization across cue conditions and speakers.

Presentation of Stimuli to Listeners

Stimulus tapes containing speech samples of the four women with dysarthria and visual images associated with each cue condition were presented to 24 small groups of listeners, with each group viewing a different stimulus tape. Listeners were positioned approximately 4 to 6 feet away from a 25-inch television monitor with one external speaker and a digital audio output.
video cassette player attached to it. The average loudness of the speech stimuli was approximately 65 dB SPL, and the ambient noise level in the room where the experiment occurred was approximately 35 dB SPL, yielding a signal-to-noise ratio of approximately 30 dB SPL.

As part of the larger experiment (Hustad, 1999), listeners were instructed that they would hear four speakers with cerebral palsy and that a different type of supplemental information (no cues, topic cues, alphabet cues, combined cues) would be presented with each speaker. Listeners were told that they would complete four tasks for each speaker-cue condition pair, two intelligibility tasks (one that involved narratives, and one that involved random sentences), and two comprehension tasks (one that involved narratives and one that involved random sentences). Following completion of intelligibility and comprehension tasks for each speaker-cue condition pair, listeners were asked to respond to a series of statements, pertaining to the present study, regarding their perceptions of the speaker’s communication effectiveness, their own willingness to communicate with the speaker, and their persistence in trying to understand the speaker. The entire experiment required approximately 2 hours to complete, and listeners were offered breaks between each experimental task.

**Dependent Measures**

Each of the three dependent variables was measured using a 7-point ordinal Likert-type scale, on which 1 = strongly disagree and 7 = strongly agree. Listeners were asked to circle the number that most closely represented their perceptions of each speaker for the following statements: (a) In general, this person is an effective communicator; (b) I would be willing to communicate with this person if I encountered her in the community; and (c) I was often so fatigued that I stopped trying to understand.

**Data Analysis**

Because data gathered for this study were ordinal in nature (ranging from 1–7), nonparametric statistics were used to conduct analyses of variance (ANOVAs) for each dependent variable. The Friedman repeated-measures nonparametric ANOVA was employed for omnibus tests and Wilcoxon signed rank tests were employed to examine pairwise differences.

**RESULTS**

**Perceptions of Communication Effectiveness**

The omnibus Friedman’s ANOVA for effectiveness was significant (chi square = 58.62; p < .001).
Wilcoxon signed rank follow-up contrasts examining pairwise differences among cue conditions revealed that ratings of effectiveness were significantly better for combined cues than for no cues, topic cues, and alphabet cues. Conversely, ratings of effectiveness were significantly worse for no cues than for topic and alphabet cues. Finally, listener ratings of effectiveness were better for alphabet cues than for topic cues. Statistics for each Wilcoxon signed rank contrast are shown in Table 4. Mean ratings across speakers for each cue condition are shown graphically in Figure 1.

**Listener Willingness to Interact With Speakers**

Listener ratings of willingness to interact with speakers for each cue condition showed similar results to those observed for communication effectiveness. The Friedman’s ANOVA omnibus test was significant (chi square = 33.05; p < .001). Pairwise follow-up contrasts again showed that combined cues resulted in significantly better ratings than no cues, topic cues, and alphabet cues. In addition, no cues resulted in significantly worse ratings than for topic and alphabet cues. The difference between ratings associated with alphabet and topic cues was not significant. Statistics for Wilcoxon signed rank contrasts are shown in Table 4. Mean ratings of willingness to interact with speakers for each cue condition are shown graphically in Figure 1.

**Listener Persistence**

The Friedman’s ANOVA omnibus test for listener ratings of persistence were significant (chi square = 35.94; p < .001). Follow-up contrasts showed that combined cues had significantly better ratings of persistence than no cues, topic cues, and alphabet cues. No cues had worse ratings of persistence than alphabet and combined cues. Ratings of persistence did not differ for topic and no cues. Statistics are shown in Table 4 and graphic representation of ratings is shown in Figure 1. For ease of interpretation, data for persistence are inverted in Figure 1 so that higher numbers reflect better outcomes. In the original statement on which listeners made ratings, lower numbers reflected better outcomes for this variable.

**DISCUSSION**

Overall, the results of the present study show that combined alphabet and topic cues, used in conjunction with natural speech, resulted in significantly higher ratings of listeners’ perceptions of the communication of speakers with severe dysarthria. However, it is important to consider these findings in light of the mean ratings, which were generally on the middle to lower end of the rating scale for each variable and cue condition. One reason for these ratings may have been that the speakers were severely impaired and, as such, their intelligibility continued to be compromised, even in the combined cues condition. In addition, listeners did not have the opportunity to interact with speakers in a context that was representative of real-world communication. Perhaps attitudinal measures would have been different if the experimental task had been more similar to a natural communicative context.

The results of attitudinal data reported in the present study are consistent with previously reported intelligibility data demonstrating that combined cues resulted in significantly higher intelligibility scores than alphabet cues, topic cues, and no cues (see Table 5 for a summary of these results from Hustad and Beukelman, 2001). Taken together, these results seem to suggest a positive relationship between attitudes toward speech supplementation strategies and intelligibility; that is, as intelligibility increases, attitudes generally become more favorable. In addition, the results of the present study are consistent with those of previous studies, suggesting that AAC techniques that are more complex result in better attitudinal ratings (Gorenflo & Gorenflo, 1991; Lasker, 1997).

**Communication Effectiveness**

All statistical findings for ratings of communication effectiveness in the present study are consistent with intelligibility data published elsewhere (Hustad & Beukelman, 2000) and summarized in Table 5 for these same speakers and cue conditions. In addition to the finding that combined cues result in more favorable ratings of effectiveness and intelligibility than any other cue condition, the results show that the no cues condition yielded poorer ratings of effectiveness and intelligibility than any other cue condition. Finally,
alphabet cues result in more favorable ratings of effectiveness and intelligibility than topic cues.

One reason for the positive effect of combined cues on effectiveness ratings may be that the listeners recognized attempts to enhance speech via strategy presentation and perceived the use of multiple strategies as trying "harder" or attempting to compensate to a greater extent for communication difficulties. The significant difference between effectiveness ratings for topic and alphabet cues conditions can also be explained with similar logic. That is, for the alphabet cues condition, listeners were provided with a greater quantity of information (i.e., one cue per word) than in the topic cues condition (i.e., one cue per utterance). Perhaps listeners equated more information with greater effectiveness.

Listener Willingness to Interact with Speakers

The results of listener ratings of their willingness to interact with speakers under each of the speech supplementation conditions showed a somewhat different pattern of results than those noted for effectiveness ratings. As was the case for effectiveness, combined cues yielded more favorable willingness ratings than any other cue condition, and no cues resulted in poorer willingness ratings than any other cue condition. Again, this result suggests that more information (i.e., both topic and alphabet cues) may have made listeners more willing to interact with the speakers. In addition, ratings of willingness for combined cues may reflect the associated increase in intelligibility observed with this one condition (see Table 5), such that as listeners were able to understand the speakers with greater accuracy, they became more willing to interact with them. Interestingly, willingness ratings did not differ significantly for topic and alphabet cues. This difference in results relative to the analogous comparison for effectiveness suggests that listeners seemed to regard willingness to interact with speakers somewhat more liberally. That is, listeners seemed to be equally willing to interact with speakers when any single strategy was employed, regardless of which one it was.

Listener Persistence

Listener ratings of their persistence in trying to understand speakers showed a pattern of results that differed from both effectiveness and willingness ratings. Listener persistence ratings were best for the combined cues condition. In addition, the alphabet cues condition had higher persistence ratings than the topic cues condition and the no cues condition. However, the topic cues condition did not differ significantly from the no cues condition. Listener ratings of persistence suggest that perhaps topic cues did not provide enough additional information to make lis-
teners consciously increase their efforts to understand the speakers. Relative to alphabet and combined cues, the presence of topic cues does not seem to encourage listeners to apply more effort to the task of understanding speakers.

CONCLUSIONS AND FUTURE DIRECTIONS

Overall, the results of the present study suggest that speakers who choose to use speech that is severely reduced in intelligibility may benefit markedly from the simultaneous use of topic and alphabet cues to supplement their speech. In experimental paradigms, this combined cueing strategy results in both increased intelligibility and in attitudes that are generally more favorable.

The results of the present study are preliminary in nature and should be regarded as such. This study examined only four speakers who had severe dysarthria due to cerebral palsy. The listeners of these speakers were not able to see them, and the speakers were not participating in a real communicative interaction. Instead, the listeners saw AAC strategies that were experimentally imposed visual images presented in conjunction with the habitual speech of four individuals with dysarthria. Because this research was experimental in nature, generalization to ecologically valid communication situations in which (a) the speaker actually implements the strategy, (b) the listener can see the speaker, and (c) speaker and listener can interact dynamically is difficult. Research to extend the findings of the present study and address variables (a) and (b) is currently under way.

Future research should also address the attitudes of familiar and unfamiliar partners toward individuals with different speech characteristics and severity levels. In addition, gender variables such as differences in attitudes between males and females should be investigated. Finally, a qualitative approach to the study of how AAC strategies and systems affect listener attitudes is necessary to provide detailed information regarding what listeners think, feel, and believe when faced with individuals who use AAC.

Address reprint requests to: Katherine C. Hustad, Department of Communication Disorders, The Pennsylvania State University, 110 Moore Building, University Park, PA 16803, USA.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Mean Intelligibility (%)</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>No cues</td>
<td>19.45</td>
<td>11.60</td>
</tr>
<tr>
<td>Topic cues</td>
<td>30.08</td>
<td>14.28</td>
</tr>
<tr>
<td>Alphabet cues</td>
<td>40.70</td>
<td>16.13</td>
</tr>
<tr>
<td>Combined cues</td>
<td>53.90</td>
<td>17.60</td>
</tr>
</tbody>
</table>

TABLE 5: Mean Intelligibility across Four Speakers by Cue Condition as Reported in Hustad and Beukelman (2001)

REFERENCES


