Listener Comprehension of Severely Dysarthric Speech: Effects of Linguistic Cues and Stimulus Cohesion

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

David R. Beukelman
Department of
Special Education and
Communication Disorders
University of Nebraska
Lincoln
and
Munroe-Meyer Institute
University of Nebraska
Medical Center
Omaha

This study is the second in a two-part series examining the effects of linguistic variables on listener processing of dysarthric speech. The first study (see K. C. Hustad & D. R. Beukelman, 2001) examined the effects of experimentally imposed topic cues, alphabet cues, and combined cues along with a control condition in which no cues were provided on *intelligibility* of unrelated and related sentences produced by 4 women with severe dysarthria secondary to cerebral palsy. The present study examined the effects of these same variables on listener *comprehension* of severely dysarthric speech produced by the same speakers. In addition, the relationship between intelligibility and comprehension was examined for each cue and stimulus cohesion condition. Consistent with intelligibility results, the present study found that combined cues resulted in higher comprehension scores than any other cue condition and that no cues resulted in lower comprehension scores than any other cue condition for both related and unrelated sentences. In addition, comprehension scores were higher for alphabet cues than for topic cues in the related-sentences condition. Findings dissimilar from intelligibility results were as follows: (a) comprehension scores associated with alphabet and topic cues did not differ for unrelated sentences, and (b) comprehension scores were higher for related than for unrelated sentences in each cue condition. Finally, significant positive relationships between intelligibility and comprehension data were found only for topic cues in the unrelated-sentences condition and for alphabet cues in the related-sentences condition.

KEY WORDS: speech supplementation strategies, dysarthria, augmentative and alternative communication, comprehension, intervention

Communication, by nature, is dyadic, necessitating the involvement of at least two communication partners who take turns playing the roles of speaker and listener throughout the course of an interaction. When the speech of one communication partner is compromised because of dysarthria, his or her ability to convey meaning and the listener's subsequent ability to interpret that meaning appropriately may be greatly reduced. The effect of dysarthria on speech production is often quantified using intelligibility measures (Weismer & Martin, 1992; Yorkston & Beukelman, 1978; Yorkston, Beukelman, Strand, & Bell, 1999). Clinically, intelligibility generally refers to how well a listener is able to separate the acoustic signal into surface-level constituent components. That is, speech intelligibility measures assess how well the acoustic signal can be parsed into the phonetic and lexical units that constitute it. Intelligibility measures are based upon the listener's word-by-word transcription, or *decoding*, of the speech signal and are quantified by dividing the number of correctly identified words by the number of possible words (Yorkston et al., 1999; Yorkston, Beukelman, & Tice, 1996). Although intelligibility measures provide important information regarding the integrity of the speech signal, they may not provide an entirely accurate characterization of how well the intended meaning of a message is effectively transmitted. An alternative measure that provides a different perspective on the capability of the speech signal to convey meaning is the listener's ability to *comprehend* the message.

Marslen-Wilson (1989) describes the process of speech comprehension as involving two separate, yet parallel, processes. First, listeners must engage in acoustic-phonetic decoding of the speech signal. During this process, listeners attend to form-based functions whereby acoustic-phonetic input is matched with representations in the mental lexicon. Next, listeners must engage in context-based functions in which they integrate form-based information into higher level representations in order to comprehend the meaning of the message. In terms of clinical measurement, comprehension differs from intelligibility in that comprehension tasks require listeners not only to decode the acoustic signal, but to process the linguistic information carried within that signal at a higher level. Research examining dysarthric speech has focused primarily on the listener's ability to decode the speech signal as indicated by intelligibility measures, with little attention paid to higher level comprehension of the message. However, measures of listener comprehension may be a more ecologically valid way to characterize the adequacy of dysarthric speech because listeners are required to respond to the content of messages in a manner similar to what occurs in real communication situations (i.e., conversation). Measures that assess comprehension might include the listener's ability to answer questions about the content of a message or narrative (Drager & Reichle, 2001) and the listener's ability to summarize the content of a narrative passage (Higginbotham, Drazek, Kowarsky, Scally, & Segal, 1994).

Comprehension of Degraded Speech

Several studies have examined the ability of listeners to comprehend disordered or distorted speech. However, most of this literature comes from the area of speech synthesis. Synthesized speech differs from natural speech in several important ways. In natural speech, there are a number of sources of variability that are not present in synthesized speech, such as within-speaker variability, cross-speaker variability, segment realization variability, and word-environment variability (Klatt, 1979). Some guarded generalizations from the literature pertaining to processing and comprehension of synthesized speech are warranted because both dysarthric and synthesized speech consist of a non-normal or degraded acoustic signal, which in turn results in intelligibility that is compromised to varying extents.

Duffy and Pisoni (1992) suggest that speech processing may differ when the human language comprehension system is presented with synthesized speech. They hypothesize that when initial input is degraded, additional cognitive resources are necessary to resolve ambiguous, missing, or misleading acoustic-phonetic cues and thus decode the speech signal. Because humans are assumed to have finite cognitive processing resources, the increased processing demands imposed by a degraded speech signal may leave fewer resources available for higher level comprehension processes, thus resulting in difficulty comprehending the message.

Research examining comprehension of speech produced by lower quality speech synthesizers seems to support Duffy and Pisoni's (1992) hypothesis. Various real-time (Ralston, Pisoni, Lively, Greene, & Mullenix, 1991; Paris, Gilson, Thomas, & Silver, 1995) and postperceptual (Higginbotham et al., 1994; Paris et al., 1995) measures indicate that processing time is increased and comprehension is compromised when the integrity of the speech signal is markedly reduced. In particular, Paris and colleagues found that listeners had more difficulty responding to higher level postperceptual inferential questions than to recognition questions with lower quality VOTRAX synthesized speech. Similarly, Higginbotham and colleagues found that listeners made more comprehension errors, as measured by their ability to summarize information, when lower quality Echo II speech synthesis was compared with higher quality DECTalk synthesis.

In general, Duffy and Pisoni's (1992) theory would predict that comprehension should be poorer than intelligibility (decoding) for degraded speech signals. However, direct comparisons between these two constructs are difficult to make, as they are generally captured using different tasks and different measurement scales. Consequently, the most statistically valid means of comparing these two measures may be to examine differences in patterns of results between comprehension and decoding data and the relationship between these two measures.

Relationship Between Comprehension and Intelligibility

Research exploring the relationship between comprehension and intelligibility is sparse and difficult to interpret. For example, Giolas and Epstein (1963) examined the relationship between isolated word transcription and paragraph comprehension, assessed via responses to comprehension questions, using low pass filtering to degrade the speech signal. Results within each of 7 filtering conditions showed no relationship between word intelligibility and discourse comprehension with average correlation coefficients ranging from –.225 to .1683.

In direct contradiction to these results, Beukelman and Yorkston (1979) found a correlation of .90 between word intelligibility scores and paragraph comprehension scores and a correlation of .95 between paragraph intelligibility scores and paragraph comprehension scores across 9 speakers with dysarthria of varying severity levels. As Weismer and Martin (1992) suggest, these high correlation coefficients reflect not only the relationship between intelligibility and comprehension, but also a confounding variable, severity of dysarthria, which is highly correlated with intelligibility. The presence of this confounding variable virtually assures high correlation coefficients. A more appropriate test of the relationship between comprehension and intelligibility would involve the elimination of severity as a variable.

Conclusions from existing literature regarding the relationship between intelligibility and comprehension are difficult to draw. These discrepant findings may be explained by methodological differences as neither study examined the same variables or measured variables in the same way. Clearly, additional research is necessary to determine the extent to which comprehension and intelligibility measures relate to one another.

Effects of Linguistic Cues on Intelligibility and Comprehension of Dysarthric Speech

Several augmentative and alternative communication (AAC) techniques, known as speech supplementation strategies (Hustad & Beukelman, 2000; Hustad, Beukelman, & Yorkston, 1998; Yorkston et al., 1999), have been shown to dramatically increase speech intelligibility in individuals who have dysarthria (Beliveau, Hodge, & Hagler, 1995; Beukelman & Yorkston, 1977; Crow & Enderby, 1989; Hunter, Pring, & Martin, 1991; Hustad & Beukelman, 2001). In general, these strategies are designed to enhance different types of listener knowledge, such as the semantic context of the message and/or the initial phoneme of each word produced by the speaker (Hustad & Beukelman, 2000, 2001). Specifically, alphabet supplementation, a strategy in which speakers use an alphabet board to indicate the first letter of each word while simultaneously speaking, has been shown to improve sentence intelligibility between 15% and 44% (Beukelman & Yorkston, 1977; Crow & Enderby, 1989; Hustad & Beukelman, 2001). The difference in magnitude of this effect appears to be attributable to methodological differences among studies (Hustad & Beukelman, 2001). Topic supplementation, a strategy in which speakers indicate the topic of the message on a communication board before speaking, has been shown to increase sentence intelligibility between 5% and 10% (Carter, Yorkston, Strand, & Hammen, 1996; Dongilli, 1994; Garcia & Cannito, 1996; Hustad & Beukelman, 2001). Finally, combined supplementation, which involves the use of both topic and alphabet cues in conjunction with speech, has been shown to increase sentence intelligibility by approximately 24% (Hustad & Beukelman, 2001) when cues are experimentally superimposed on a habitual speech signal (i.e., speakers did not actually implement the cues).

Although there is a growing body of literature examining the effects of speech supplementation strategies on intelligibility of dysarthric speech, research examining the effects of these same strategies on comprehension of dysarthric speech does not presently exist. Because comprehension of a speaker's messages is a desired outcome of communication intervention for individuals with chronic dysarthria, effects of strategies such as alphabet supplementation, topic supplementation, and combined supplementation on listener comprehension of the speaker's messages have important implications for effective treatment. In addition, differences in findings for intelligibility and comprehension measures have the potential to inform our understanding regarding the relationship between these two processes.

The present study examined the effects of speech supplementation strategies (topic cues, alphabet cues, and combined cues) on listener comprehension of severely dysarthric speech and the relationship between measures of intelligibility and measures of comprehension. A companion paper (Hustad & Beukelman, 2001) examined the effects of these strategies on intelligibility of dysarthric speech, and the present study addresses parallel research questions. Results will be discussed relative to intelligibility findings from Hustad and Beukelman (2001). The following research questions were addressed in the present study:

1. Are there differences in listener comprehension scores among cue conditions (no cues, topic cues, alphabet cues, and combined cues) for unrelated sentence stimuli and for related sentence stimuli constituting a narrative?
2. Are there differences in listener comprehension scores between related and unrelated sentence stimuli for each cue condition?
3. Are there meaningful relationships between intelligibility and comprehension scores within each of the cue conditions?

Method

The present study was part of a larger project (Hustad, 1999) and is the second paper in a two-part series. Data presented in the first paper were based on intelligibility measures in which listeners transcribed what they thought speakers said in a word-for-word fashion under each of four cue conditions (no cues, topic cues, alphabet cues, combined cues) and each of two stimulus cohesion conditions (unrelated sentences and related sentences forming a cohesive narrative) for a total of eight experimental tasks. In the present study, data were obtained from eight additional experimental tasks in which listeners answered factual comprehension questions about individual stimulus sentences to which they had not been exposed previously for each of the four cue conditions and each of the two stimulus cohesion conditions. Methods followed procedures published in Hustad and Beukelman (2001).

Speakers With Dysarthria

Four female speakers with severe dysarthria secondary to cerebral palsy (CP) contributed speech samples for this study. Two had mixed athetoid spastic quadriplegia and associated mixed spastic-hyperkinetic dysarthria. Two had spastic CP, one with diplegia and one with quadriplegia, and associated spastic dysarthria. Speakers were required to meet the following criteria: (a) speech intelligibility between 15% and 25% as measured by the Sentence Intelligibility Test (Yorkston et al., 1996), (b) native speakers of American English, (c) age between 19 and 46 years, (d) able to produce connected speech consisting of at least eight consecutive words, and (e) able to repeat sentences of up to eight words following a verbal model. Specific details regarding perceptual characteristics for each speaker, as determined by a certified speech-language pathologist, are provided in Table 1.

Table 1. Perceptual characteristics of speakers with dysarthria.

	Speaker			
Characteriatic	1	2	3	4
Age	19 years	24 years	46 years	42 years
Speech Diagnosis	mixed spastic-hyperkinetic dysarthria	mixed spastic-hyperkinetic dysarthria	spastic dysarthria	spastic dysarthria
Articulatory Characteristics	slow, labored imprecise increased syllable and word duration prolonged sounds and intervals irregular articulatory breakdown	slow, labored, imprecise increased syllable and word duration prolonged sounds and intervals irregular articulatory breakdown	slow, labored, imprecise	slow, labored,, imprecise poor word segmentation irregular syllable segmentation
Respiratory Characteristics	audible inhalation reduced inhalatory and exhalatory control	shallow inhalation	reduced inhalatory and exhalatory control	reduced inhalatory and exhalatory control
Resonatory Characteristics	mild hypernasality	mild hypernasality	mild nasal air emission	moderate hypernasality moderate nasal air emission
Prosodic Characteristics	short phrases inappropriate pauses	short phrases inappropriate pauses	short phrases	short phrases equal syllable stress
Rate of Speech	24 words per minute	23 words per minute	38 words per minute	35 words per minute
Intelligibility on SIT	24%	16%	17%	15%

Development of Speech Stimuli

Speakers with dysarthria produced 16 narrative passages, each of which consisted of 10 sentences. Passages were designed to represent contexts common to adult speakers of American English (i.e., vacation, wedding, sporting event, buying a car). Individual sentences within each narrative were developed to be meaningful in isolation and to contribute information to the narrative that was independent of all other sentences.

All sentences were declarative in nature, with most noun and verb phrases following a subject-verb or subject-verb-object syntactic structure. The length and content of stimulus material were equated on several different linguistic parameters, including number of words per sentence, number of words per narrative, number of syllables per sentence, type-token ratio of each passage, and reading level of each narrative. See Table 2 for specific details. Although semantic and syntactic predictability of stimulus material were not formally controlled, narratives and their constituent sentences were developed to be predictable in nature, following standard American English conventions.

Development of Topic Cues

Topic cues were developed for each narrative and constituent sentence to reflect the main idea of the passage. One topic cue was provided for each narrative, for a total of 16 different topic cues. Topic cues were designed to reflect the entire narrative as well as each individual sentence within the narrative. To determine the appropriateness of sentence-topic pairs, 10 independent judges rated each of the 160 pairs in the corpus. Sentences and topics underwent modifications until 90% of judges (9 of 10) independently rated each sentence-topic pair as acceptable. See the Appendix for sample sentences and associated topic cues.

Table 2. Linguistic characteristics of speech stimuli.

Characteristic	Number
Narratives	16
Sentences per narrative	10
Topics per narrative	1
Words per narrative	65
5-word sentences per narrative	2
6-word sentences per narrative	3
7-word sentences per narrative	3
8-word sentences per narrative	2
Different words per narrative	49
Type token ratio for each narrative	.75
Average number of syllables per word for each narrative	1.4
Total number of syllables per sentence for each narrative	9.0–9.1
Average number of words per sentence for each narrative	6.5
Number of 1-syllable words per narrative	45
Number of 2-syllable words per narrative	15
Number of 3-syllables words per narrative	4–5
Number of 4-syllable words per narrative	0–1
Reading level for each narrative	5.7

Development of Comprehension Questions

Comprehension of individual sentences was assessed through responses to short-answer wh-type questions. For each sentence within each narrative, an open-ended question that was specific and detail oriented with respect to factual information presented in the target sentence was developed. Questions were explicitly phrased so as not to provide any leading information regarding target responses for any sentence. To assure that correct answers to comprehension questions could not be guessed without knowledge of the referent sentence, each question was pilot tested. This was accomplished by independently presenting all questions and the topic cue associated with the referent sentence, but not the referent sentence itself, to 10 independent judges who were asked to write down the best or most likely answer to each question. Questions were presented in sequential order within each narrative to optimize guessability under the assumption that correct responses to questions would be more easily guessed in narrative context when listeners were presented with associated topic cues than in random sentence context. Any question for which the answer was guessed correctly by one or more judges was discarded and a different question was developed to replace it. A series of 10 different judges responded to any new questions for which answers were previously guessed correctly. This procedure continued until each question was verified as "unguessable," as indicated by 100% of judges (10 of 10) failing to guess the target response correctly. Examples of comprehension questions and their associated referent sentences are provided in the Appendix.

Recording Speech Samples

Speakers were recorded in an acoustically treated environment using digital audio tape (DAT) (44.1 kHz sampling rate; 16-bit quantization) and a head-mounted microphone positioned 5 cm from the speaker's mouth. Speakers produced each stimulus sentence following the experimenter's model. Orthographic representations of stimulus sentences were also provided on a computer screen placed in front of the speaker. Speakers were instructed to speak naturally, so that productions did not sound as though they were being read. On several occasions, individual speakers were asked to repeat target sentences to obtain natural-sounding productions that characterized their habitual speech more accurately.

Judgments of naturalness were made subjectively by the experimenter and were based on previous experience interacting with each speaker.

Stimulus Tape Preparation

Digital recordings were transferred to a personal computer, edited to remove extraneous comments, and normalized to a peak amplitude of 69 dB. All 160 digital audio files for each speaker were then duplicated four times (once for each cue condition) and matched with video images associated with each cue condition (no cues, topic cues, alphabet cues, and combined cues).

Cue Conditions

For the topic-cues condition, a video image showing the orthographic representation of the topic was associated with the spoken production of each sentence. Each topic was shown for the duration of the spoken production as indicated by individual waveforms.

For the alphabet-cues condition, a video image of an alphabet board arranged in ABC fashion was presented. During the spoken production of each sentence, individual graphemes representing the first letter of each word in the sentence were highlighted on the alphabet board for the full duration of each word. Over the course of each sentence, listeners saw a series of highlighted initial graphemes presented in real time and corresponding to the spoken production of each word. Word boundaries were identified through visual and auditory inspection of the acoustic waveform for each sentence. The presentation duration for each word-initial grapheme varied among speakers and among words because of individual speaker differences.

For the combined-cues condition, listeners were presented with visual images of both alphabet and topic cues as described for each condition above. For speech stimuli in the no-cues condition, listeners were presented with a plain blue background shown in conjunction with the habitual speech signal and no other visual information.

Randomization and Counterbalancing

To prevent learning effects and order effects associated with cue conditions and speakers, a Latin Square counterbalancing scheme (Campbell & Stanley, 1963; Cook & Campbell, 1979; Kirk, 1995) for permutations of cue-condition presentation order was employed. Because there were 24 permutations for the four cue conditions, 24 different stimulus tapes were made—each reflecting a unique cue-condition order and speaker-presentation order. On each stimulus tape, individual speakers appeared once in a given cue condition so that groups of listeners heard four different speakers, each associated with a different cue condition.

On each of the 24 stimulus tapes, individual sentences occurred in only one condition and were repeated twice in that condition. During the first presentation listeners heard all 10 sentences in succession and were instructed to listen only. During the second presentation of the same sentences, listeners were instructed to answer comprehension questions following each sentence. Across intelligibility and comprehension experiments, listeners were presented with a total of 160 different sentences (40 sentences produced by each speaker). Half of these sentences were presented as cohesive narratives, and half were presented as unrelated sentences across intelligibility and comprehension tasks and their associated cue conditions. For each tape, the following procedures were followed: (a) Eight narratives were designated for the related sentences tasks—4 narratives were used for intelligibility tasks, 4 narratives were used for comprehension tasks. (b) The remaining 80 sentences were quasi-randomly assigned into 8 lists of 10 sentences so that no more than 2 sentences from the same narrative were assigned to each list—4 lists were used for intelligibility tasks; 4 lists were used for comprehension tasks. (c) Narratives and random sentence lists were assigned to comprehension or intelligibility tasks. (d) Two lists of random sentences (one for comprehension tasks and one for intelligibility tasks) and two cohesive narratives were assigned to each cue condition. (e) Different speakers were assigned to each cue condition. These procedures were replicated for each of the 24 stimulus tapes. In this way, length, complexity, and predictability of stimulus material were equated across cue conditions, stimulus cohesion conditions, and speakers so that each cell of the research design reflected data obtained from a quasi-random sampling of different sentences or narratives.

Experimental Sequence

On each of the 24 videotapes, the following sequence of events occurred: (a) orthographic instructions for the forthcoming task, directing listeners to watch and listen only; (b) orthographic number for the first sentence; (c) audio target sentence with associated cues provided orthographically; (d) orthographic number for the second sentences; (e) audio target sentence with associated cues provided orthographically. Items a–e were repeated until the initial preview of each of the 10 sentences was completed. Following this first preview of all 10 stimulus sentences, the following items were presented: (a) orthographic instructions directing listeners to answer the comprehension question presented during the interval between each sentence, (b) orthographic number for the first sentence, (c) audio target sentence with associated cues provided orthographically, (d) orthographic comprehension question for the first sentence, (e) orthographic number for the second sentence, (f) audio target sentence with associated cues presented orthographically, (g) orthographic comprehension question for the second sentence. Again, items b–g were repeated until all 10 sentences and their associated comprehension questions had been presented. The final experimental sequence for each tape was exported from computer to digital videotape, maintaining first-generation-quality audio and video signals following NTSC broadcast-quality standards.

Nondisabled Listeners

Seventy-two nondisabled individuals served as listeners. 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, or cognitive disabilities per self-report. All listeners were either currently attending college or graduate school or had completed college or graduate school. As such, college-level literacy skills were assumed. Listeners had a mean age of 21 years (SD = 2.464). Gender composition was 8 males and 64 females. Gender was not a variable of interest; therefore no effort was made to balance the number of male and female listeners.

Experimental Task

Presentation of Stimuli

Small groups of different listeners were randomly assigned to each of the 24 stimulus tapes. Digital audio-video tapes were presented in a quiet listening environment. During the experiment, listeners were seated at desks that were approximately 4–6 feet away from a 25-inch television monitor positioned at eye level (approximately 3.5 feet from the ground). One external speaker, positioned at listeners' chest level (approximately 2.5 feet from the ground) directly below the television, and a digital video cassette player were both attached to the television monitor. The average signal-to-noise ratio, measured from where listeners were seated, was approximately 30 dB SPL. (Average signal level was approximately 65 dB SPL; peak ambient noise level was approximately 35 dB SPL.)

Administration Instructions

Listeners were informed that during this experiment, they would hear four women with cerebral palsy who were producing meaningful and grammatically correct sentences. Some of the sentences would form a short story, and some of the sentences would be completely unrelated. In addition, listeners were told that they would see four different types of cues (no cues, topic cues, alphabet cues, and combined cues—not in this order), one associated with each speaker. Specific information regarding each cue condition was provided before presentation of that condition. Listeners were told that there would be two presentations of each set of stimulus sentences. During the first presentation of 10 sentences, they were told to watch and listen without writing anything on their answer sheet. During the second presentation of the same 10 sentences, listeners were told to answer the question presented on the video monitor, taking their best guess if they were unsure of the correct answer. Listeners were informed that the purpose of the study was to determine whether particular kinds of information helped people, like themselves, understand these speakers better. Finally, the experimenter explained that listeners could take as much time as necessary to answer each question and that breaks would be available between experimental tasks, if desired.

Scoring and Reliability

Responses to each comprehension question were scored in a binomial fashion by the first author as either correct or incorrect. Responses were counted as correct if they accurately provided the specific information requested from the referent sentence. Responses were counted as incorrect if the information provided by the listener was not accurate and specific with respect to the question and its referent sentence. For statistical analysis, data were pooled across speakers within each cue condition and experimental task. This was necessary because there were unequal numbers of observations among cue conditions for individual speakers, which would bias any statistical comparisons examining individual results.

Because scoring responses to comprehension questions was somewhat subjective, intra- and interjudge agreement was very important. To assess intrajudge agreement, half of the sample (data from 36 listeners) was re-scored by the first author 12 months after the initial scoring was completed. In addition, interjudge agreement was assessed by having a different judge score one quarter of the sample (data from 18 listeners). Using the formula percent agreement = (agreements / agreements + disagreements) $\times$ 100, reliability was 95% and 90% respectively. These values indicate strong intra- and interjudge agreement for scoring of the dependent measure.

Experimental Design

A 2 $\times$ 4 within-subjects repeated measures design (Kirk, 1995) was employed for this study. Accordingly, data from each of the 72 listeners made up each cell of the design. One of the within-subjects repeated measures was stimulus cohesion, with its two categories of related sentences and unrelated sentences. The other within-subjects repeated measure was cue condition, with its four categories: no cues (NC), topic cues (TC), alphabet cues (AC), and combined cues (CC).

Results

A planned contrast approach to ANOVA was employed in which only the contrasts of interest were subjected to statistical analysis because research questions of interest were specified a priori (Hertzog & Rovine, 1985; Kirk, 1995; Marascuilo & Levin, 1983; Marascuilo & Serlin, 1988; Seaman, Levin, & Serlin, 1991). This conservative approach to ANOVA employs fewer statistical tests and thus has a reduced probability of Type I error relative to traditional ANOVA models. The experiment-wise alpha level for the present study was set at .05 and was partitioned using the Holm Sequentially Rejective Bonferroni Test (Holm, 1979; Kirk, 1995; Seaman et al., 1991). Statistical results for each planned contrast are shown in Table 3 and will be presented according to the four groups of questions addressed in this study.

Table 3. Planned contrasts for comprehension (CCRS = combined cues related sentences, ACRS = alphabet cues related sentences, TCRS = topic cues related sentences, NCRS = no cues related sentences, CCUS = combined cues unrelated sentences, ACUS = alphabet cues unrelated sentences, TCUS = topic cues unrelated sentences, NCUS = no cues unrelated sentences).

Contrast	Mean difference	*df*	Standard Error for contrast	*t*
CCUS–ACUS	.2111	71	.0262	8.055*
CCUS–TCUS	.2111	71	.0283	7.463*
CCUS–NCUS	.3013	71	.0233	12.946*
ACUS–NCUS	.0901	71	.0195	4.621*
ACUS–TCUS	.0000	71	.0228	0.000
TCUS–NCUS	.0914	71	.0189	4.752*
CCRS–ACRS	.1542	71	.0356	4.331*
CCRS–TCRS	.2764	71	.0302	9.145*
CCRS–NCRS	.3819	71	.0334	11.420*
ACRS–NCRS	.2278	71	.0350	6.507*
ACRS–TCRS	.1222	71	.0331	3.685*
TCRS–NCRS	.1056	71	.0229	4.595*
NCRS–NCUS	.0429	71	.0214	2.001***
TCRS–TCUS	.0583	71	.0221	2.640**
ACRS–ACUS	.1806	71	.0302	5.972*
CCRS–CCUS	.1236	71	.0330	3.738*

* statistical significance at $p < .001$
** statistical significance at $p < .01$
*** statistical significance at $p < .05$

Cue Conditions and Comprehension of Unrelated Sentence Stimuli

Results of statistical analyses follow. Combined cues resulted in higher comprehension scores than alphabet cues (t = 8.05, $p < .001$), topic cues (t = 7.46, $p < .001$), and no cues (t = 12.95, $p < .001$). No cues resulted in lower comprehension scores than both alphabet cues (t = 4.621, $p < .001$) and topic cues (t = 4.752, $p < .001$). Alphabet and topic cues did not differ significantly. Descriptive statistics are provided in Table 4, and pooled speaker data are displayed graphically in Figure 1.

Cue Conditions and Comprehension of Related Sentence Stimuli

Statistical results indicated that combined cues resulted in higher comprehension scores than alphabet cues (t = 4.33, $p < .001$), topic cues (t = 9.14, $p < .001$), and no cues (t = 11.41, $p < .001$). No cues resulted in lower comprehension scores than alphabet cues (t = 6.51, $p < .001$) and topic cues (t = 4.59, $p < .001$). Finally, alphabet cues yielded higher comprehension scores than topic cues (t = 3.685, $p < .001$).

Stimulus Cohesion

Statistics showed that cohesive narratives resulted in higher comprehension scores than unrelated sentences for no cues (t = 2.01, $p < .05$), topic cues (t = 2.64, $p < .01$), alphabet cues (t = 5.97, $p < .001$), and combined cues (t = 3.74, $p < .001$). This difference was not significant for topic cues.

Figure 1. Mean percent comprehension (±*SD*) by cue condition for related and unrelated sentences.

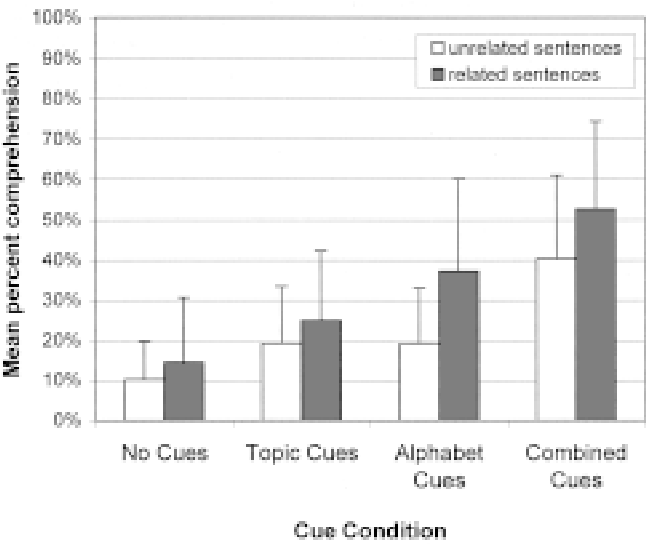


Table 4. Comprehension data for individual speakers by cue condition (NC = No Cues, TC = Topic Cues, AC = Alphabet Cues, CC = Combined Cues) and sentence cohesion (unrelated and related sentences).

		Speaker 1		Speaker 2		Speaker 3		Speaker 4		All speakers (weighted)	
Cue Condition		unrelated sentences	related sentences	unrelated sentences	related sentences	unrelated sentences	related sentences	unrelated sentences	related sentences	unrelated sentences	related sentences
NC	*M*	9.23	10.00	8.00	5.00	6.88	16.88	14.83	23.48	10.15	14.44
	SD	7.60	10.80	6.96	7.61	7.93	13.02	11.96	20.80	9.55	16.18
	Range	0–20	0–40	0–20	0–20	0–20	0–50	0–40	0–70	0–40	0–70
	N	13	13	20	20	16	16	23	23	72	72
TC	*M*	15.79	25.79	13.68	18.42	25.26	33.68	22.67	21.33	19.17	25.00
	SD	11.21	19.24	14.22	15.00	17.44	13.42	11.63	17.67	14.51	17.12
	Range	0–40	0–80	0–50	0–60	0–70	10–60	10–40	0–50	0–70	0–80
	N	19	19	19	19	19	19	15	15	72	72
AC	*M*	14.00	33.00	10.56	30.00	28.82	46.47	24.71	40.59	19.17	37.22
	SD	10.95	20.55	9.98	21.69	13.64	27.60	12.81	19.52	13.81	22.90
	Range	0–40	0–80	0–30	0–70	0–50	0–100	10–50	10–80	0–50	0–100
	N	20	20	18	18	17	17	17	17	72	72
CC	*M*	34.50	44.50	38.67	45.33	48.00	59.00	39.61	61.18	40.28	52.64
	SD	19.05	17.61	23.56	17.67	20.42	23.37	19.83	23.42	20.83	21.75
	Range	0–60	20–80	10–80	20–80	10–80	10–90	0–70	20–90	0–80	10–90
	N	20	20	15	15	20	20	17	17	72	72

Descriptive Differences Among Cue Conditions for Individual Speakers

Descriptive data for individual speakers suggests similar patterns of results among speakers. Similar to intelligibility data reported by Hustad and Beukelman (2001), there were two consistent observations among individual speakers and cohesion conditions: (a) the combined-cues condition seemed to result in the highest comprehension scores; and (b) the no-cues condition seemed to result in the lowest comprehension scores for listeners. Individual speaker data are displayed in tabular form for related and unrelated sentences in Table 4 and in graphic form in Figure 2.

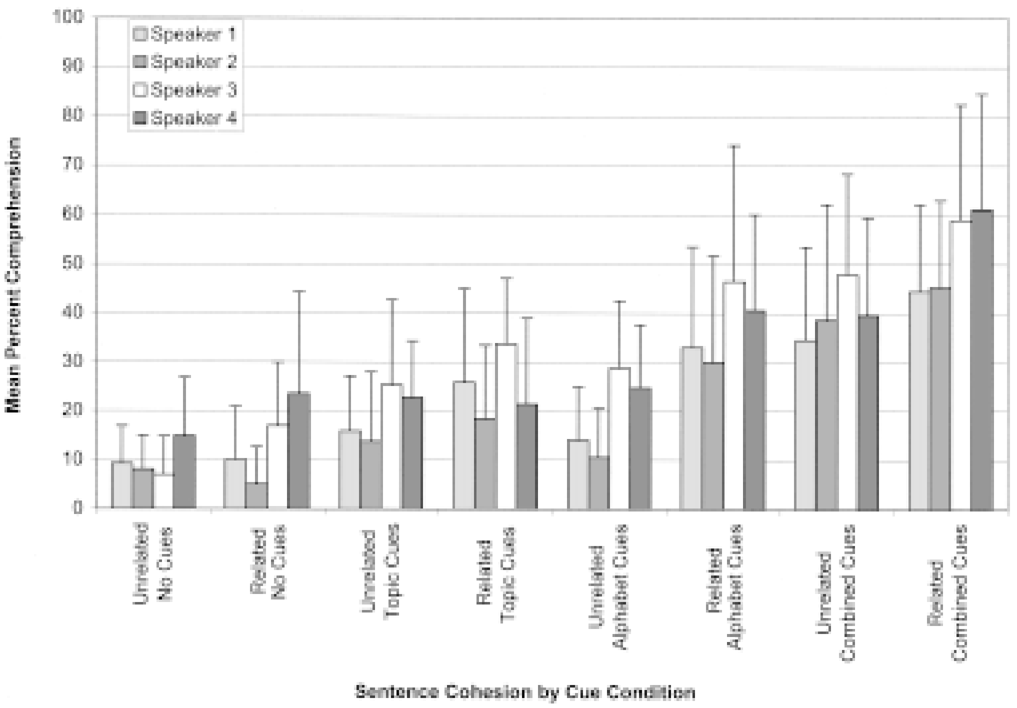
Relationship Between Intelligibility and Comprehension Measures

A series of Pearson product moment correlation coefficients were computed for intelligibility (percent of words identified correctly) and comprehension (percent of comprehension questions answered correctly) data, which were obtained from the same speakers and listeners. The relationships between intelligibility and comprehension of unrelated sentences by cue condition were as follows: no cues (r = –.152, p = .203), topic cues (r = .432, p < .001), alphabet cues (r = .129, p = .280), and combined cues (r = .157, p = .188). The relationships between intelligibility and comprehension of cohesive sentences by cue condition were as follows: no cues (r = .113, p = .343), topic cues (r = .304, p < .01), alphabet cues (r = .548, p < .001), and combined cues (r = .249, p < .05).

Discussion

Speech processing is said to involve at least two components: decoding or surface-level mapping of the acoustic signal onto phonemes and words and higher level integration of words onto meaning, ultimately leading to language comprehension (Duffy & Pisoni, 1992; Marslen-Wilson, 1989). The present study examined listener comprehension of dysarthric speech under four conditions involving different, explicitly provided linguistic-contextual cues and two stimulus cohesion conditions. Comprehension was measured by listeners' ability to answer postperceptual, open-ended, simple, fact-based questions for each sentence produced by speakers with severe dysarthria. It was assumed that this type of task would require higher level integration of information because listeners were required to respond to information presented rather than simply to transcribe what they heard. The present study is the second in a two-part series examining effects of linguistic

Figure 2. Mean percent comprehension (+*SD*) by individual speaker and cue condition for related and unrelated sentences.



cues in the form of experimentally imposed AAC strategies on listener processing of dysarthric speech. Findings from listener comprehension measures bear some similarity to findings from intelligibility measures reported by Hustad and Beukelman (2001), particularly regarding the effects of cue conditions on comprehension of related sentence stimuli. However, findings regarding the effects of cue conditions on comprehension of unrelated sentences and the effects of stimulus cohesion on comprehension are different from those observed for intelligibility data. In addition, the relationship between intelligibility and comprehension measures was examined in the present paper. Specific results are discussed below.

Effects of Cue Conditions on Comprehension

Combined Cues and No Cues

In general, for both related and unrelated sentence stimuli, results showed that the no-cues control condition resulted in lower comprehension scores than any other cue condition, and the combined-cues condition resulted in higher comprehension scores than any other cue condition. These findings are consistent with those for analogous intelligibility results reported by Hustad and Beukelman (2001) and suggest that when listeners are provided with explicit word-specific and explicit contextual cues, both intelligibility and comprehension of the message are optimized.

Alphabet and Topic Cues

Results regarding the effects of topic and alphabet cues differed for comprehension of related and unrelated sentences. When listeners were presented with related sentences, alphabet cues enhanced comprehension more than topic cues. Duffy and Pisoni (1992) suggest that the human language-comprehension system may use context to compensate for perceptual processing or decoding problems like those that may be experienced in the presence of a degraded speech signal. For related sentences forming a narrative, two types of context were available to listeners: implicit top-down context that builds from sentence to sentence within the narrative and explicit contextual cues that were superimposed on the dysarthric speech signal. When listeners were presented with related sentences forming a narrative, both types of context were readily available to facilitate comprehension. Alphabet cues may have had a more powerful effect than topic cues on comprehension of related sentences because alphabet cues provided listeners with new information that they were not able to infer through implicit top-down context. In the alphabet-cues condition, listeners were provided with explicit orthographic information regarding the first letter of each constituent word in the utterance, which seems likely to facilitate the decoding aspects of speech processing. Results examining analogous intelligibility data for related sentences confirm that alphabet cues consistently yielded higher intelligibility scores than did topic cues (Hustad & Beukelman, 2001). Ultimately, if decoding is enhanced, it would seem logical to expect comprehension to benefit as well.

For unrelated sentences, neither topic nor alphabet cues enhanced comprehension differentially. This may be because listeners did not have the benefit of implicit top-down knowledge to the same extent as they did in the related-sentences conditions, and as a result both cue conditions provided listeners with novel information that was equally useful in processing the speech signal. This finding is different from that revealed by analogous intelligibility data for unrelated sentences in which alphabet cues yielded higher scores than did topic cues. The discrepancy between comprehension and intelligibility findings may be due to the differing nature of the tasks employed in each of the two experiments and may reflect increased difficulty with higher level processing relative to lower level decoding when implicit contextual information (i.e., derived from cohesive narratives) is constrained.

Stimulus Cohesion

Results of the present study showed that when listeners were provided with related sentences forming a cohesive narrative, their comprehension scores were an average (across cue conditions) of 10% higher than when they were provided with unrelated sentences for each of the cue conditions. However, the magnitude of this difference varied between 4% (no cues) and 18% (alphabet cues). Again, this finding would seem to be attributable to the availability of top-down implicit knowledge regarding the context of the message that ultimately improves comprehension in the related-sentences condition. Provision of cues seemed to increase the gap between related and unrelated sentences, with descriptive data for each cue condition suggesting a greater difference than that observed in the no-cues condition. This observation suggests that contextual cues—whether implicit, as in related sentences, or explicit, as in information provided via AAC strategies—make an important difference in listener's ability to comprehend messages produced by speakers with severe dysarthria.

Stimulus cohesion results were very different for intelligibility data reported by Hustad and Beukelman (2001) in which the only difference between related and unrelated sentences occurred with alphabet cues. Again, this discrepancy may be related to the differing nature of the tasks facing the listener. Theoretically, in intelligibility tasks emphasizing decoding, listeners have not integrated constituent messages into higher level meaning and therefore may not yet have had the opportunity to benefit from the implicit context provided by related sentences. However, when listeners receive alphabet cues associated with intelligibility tasks, the decoding gains from these cues seem to allow listeners to benefit from top-down implicit information to a greater extent.

Relationship Between Intelligibility and Comprehension

Correlation coefficients examining the relationship between comprehension and intelligibility for the same speakers and listeners showed virtually no relationship in the no-cues condition for both related and unrelated sentences. These results are consistent with those found by Giolas and Epstein (1963). The finding that intelligibility and comprehension measures do not relate to one another suggests that intelligibility and comprehension measures employed in the present study tap different perceptual phenomena. Furthermore, this result indicates that complete and accurate decoding (as measured by intelligibility scores) is not requisite for comprehension.

The relationship between comprehension and intelligibility was somewhat altered when supplemental cues were available to listeners. With unrelated sentences, the relationship between comprehension and intelligibility was relatively strong and significant for topic cues, demonstrating that higher comprehension scores were linearly associated with higher intelligibility scores. This finding highlights the importance of explicit contextual knowledge in processing severely dysarthric speech, particularly in the absence of linguistic cohesion associated with narratives. Similarly, with related sentences, the relationship between comprehension and intelligibility was relatively strong and significant for alphabet cues, suggesting the importance of explicit cues that facilitate decoding when inferential knowledge regarding the topic is present.

Limitations and Future Directions

The present study represents one of the first attempts at measuring listener comprehension of dysarthric speech and contrasting comprehension findings with analogous intelligibility data. As such, many limitations exist. First, the authors attempted to separate comprehension and intelligibility on the basis of task demands placed on listeners. That results for intelligibility and comprehension data were different, particularly regarding the effects of stimulus cohesion, would seem to suggest different perceptual phenomena were captured by the two measures. There are many tools that could be used to measure comprehension, and the present study examined only responses to simple fact-based questions associated with individual referent sentences. In the future, research should examine listener comprehension of higher level content, such as inferential questions, main ideas, humor, and idioms, produced by speakers with dysarthria. In addition, different dependent measures, such as forced choice questions, story re-telling, and open-ended questions pertaining to the entire narrative rather than to individual sentences, should be explored. Finally, different scoring systems, such as qualitative rating scales and consensus ratings based on groups of judges, should be explored to increase the sensitivity of dependent measures employed.

Clinical Implications

The results of the present study reflect carefully controlled experimental findings and as such should be generalized to clinical situations with caution. Taken together with results from Hustad and Beukelman (2001), findings suggest that both intelligibility and comprehension of severely dysarthric speech are markedly improved when listeners are provided with explicit contextual cues that supplement speech. In particular, combined cues, involving topic and alphabet cues, increase both comprehension and intelligibility more than alphabet or topic cues alone. In addition, listener comprehension of dysarthric speech is optimized when cohesive narratives are presented, suggesting that speakers should refrain from making rapid topic switches whenever possible. Additional research examining the clinical implementation of these speech supplementation strategies is necessary to determine their impact on real communication between speakers with dysarthria and their listeners.

Acknowledgments

This research was supported, in part, by funds from the Barkley Memorial Trust. The authors wish to express special thanks to the speakers with dysarthria who participated in this study and to Cara Ullman for preparation of stimulus tapes.

References

Beliveau, C., Hodge, M., & Hagler, P. (1995). Effect of supplemental linguistic cues on the intelligibility of severely dysarthric speakers. *Augmentative and Alternative Communication, 11,* 176–186.

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

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

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Boston: Houghton Mifflin Co.

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

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation design and analysis issues for field settings*. Boston: Houghton Mifflin Co.

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

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

Drager, K. D., & Reichle, J. E. (2001). Effects of age and divided attention on listeners' comprehension of synthesized speech. *Augmentative and Alternative Communication, 17,* 109–119.

Duffy, S. A., & Pisoni, D. B. (1992). Comprehension of synthetic speech produced by rule: A review and theoretical interpretation. *Language and Speech, 35,* 351–389.

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.

Giolas, T., & Epstein, A. (1963). Comparative intelligibility of word lists and continuous discourse. *Journal of Speech and Hearing Research, 6,* 349–358.

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

Higginbotham, D. J., Drazek, A. L., Kowarsky, K., Scally, C., & Segal, E. (1994). Discourse comprehension of synthetic speech delivered at normal and slow presentation rates. *Augmentative and Alternative Communication, 10,* 191–202.

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics, 6,* 65–70.

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

Hustad, K. C. (1999). *Effects of context on decoding and comprehending dysarthric speech*. Unpublished doctoral dissertation, University of Nebraska–Lincoln.

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. (2000). Integrating AAC strategies with natural speech in adults with chronic speech intelligibility challenges. In D. Beukelman, K. Yorkston, & J. Reichle (Eds.), *Augmentative communication for adults with acquired neurologic disorders* (pp. 83–106). Baltimore: Paul H. Brookes.

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

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

Klatt, D. (1979). Speech perception: A model of acoustic-phonetic analysis and lexical access. *Journal of Phonetics, 7,* 279–312.

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

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

Marslen-Wilson, W. D. (1989). Access and integration: Projecting sound onto meaning. In W. D. Marslen-Wilson (Ed.), *Lexical representation and processing* (pp. 3–24). Cambridge, MA: MIT Press.

Paris, C., Gilson, R., Thomas, M., & Silver, N. C. (1995). Effect of synthetic voice intelligibility on speech comprehension. *Human Factors, 37,* 335–340.

Ralston, J., Pisoni, D., Lively, S., Greene, B., & Mullenix, J. (1991). Comprehension of synthetic speech produced by rule: Word monitoring and sentence-by-sentence listening times. *Human Factors, 33,* 471–491.

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

Weismer, G., & Martin, R. (1992). Acoustic and perceptual approaches to the study of intelligibility. In R. Kent (Ed.), *Intelligibility in speech disorders* (pp. 67–118). Philadelphia: John Benjamins Publishing Co.

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

Yorkston, K., & Beukelman, D. (1978). A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of Communication Disorders, 11,* 499–512.

Yorkston, K., Beukelman, D., & Tice, R. (1996). *Sentence intelligibility test for windows*. [Computer software]. Lincoln, NE: Tice Technology Services.

Received July 9, 2001

Accepted March 4, 2002

DOI: 10.1044/1092-4388(2002/043)

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

Appendix. Sample Lists of Related Sentences, Unrelated Sentences, Topic Cues, and Comprehension Questions.

Topic	Sentence	Comprehension Question
Purchasing a vehicle	Jason needed to buy a car.	What did he need?
Purchasing a vehicle	He wanted a new car.	What did he prefer?
Purchasing a vehicle	He considered two different models.	What did he do?
Purchasing a vehicle	Four wheel drive was a desired feature.	What feature did he want?
Purchasing a vehicle	Jason liked the large pickup trucks.	What kind did he like?
Purchasing a vehicle	Sport utility vehicles were his favorite.	What did he like best?
Purchasing a vehicle	He did not have much money to spend.	What was his problem?
Purchasing a vehicle	He bargained with a salesman for two hours.	What did he do?
Purchasing a vehicle	The final price was within his budget.	What happened?
Purchasing a vehicle	A used Jeep was what he purchased.	What did he get?
Ocean voyage	Their route was carefully planned.	What was done?
Acquiring a new home	Katherine and David wanted to buy a house	What did they want?
Vacation at the seashore	They have a cottage on the ocean.	What do they have?
Military life	Tom is in the Navy.	What is he in?
Beginning a new school year	School starts on August first in Maine	When does it begin?
Travel problems	Connecting flights are often hard to catch.	What is difficult?
Independence day	On July fourth, most people celebrate	What do people do?
Relocating to a new city	The Browns recently moved to Boston	What did they do?
Sports outing	Jeffery and Jacob are sports fanatics.	What do they like?
Wedding plans	Kim and Eric are getting married in May.	What are they doing?