J Am Acad Audiol 7: 396–405 (1996)

Initial Evaluation of an Interactive Test of Sentence Gist Recognition

Nancy Tye-Murray*†
Shelley Witt†
John Castelloe†

Abstract

The laser videodisc-based Sentence Gist Recognition (SGR) test consists of sets of topically related sentences that are cued by short film clips. Clients respond to test items by selecting picture illustrations and may interact with the talker by using repair strategies when they do not recognize a test item. The two experiments, involving 40 and 35 adult subjects, respectively, indicated that the SGR may better predict subjective measures of speechreading and listening performance than more traditional audiologic sentence and nonsense syllable tests. Data from cochlear implant users indicated that the SGR accounted for a greater percentage of the variance for selected items of the Communication Profile for the Hearing-Impaired and the Speechreading Questionnaire for Cochlear-Implant Users than two other audiologic tests. As in previous work, subjects were most apt to ask the talker to repeat an utterance that they did not recognize than to ask the talker to restructure it. It is suggested that the SGR may reflect the interactive nature of conversation and provide a simulated real-world listening and/or speechreading task. The principles underlaying this test are consistent with the development of other computer technologies and concepts, such as compact disc-interactive and virtual reality.

Key Words: Audiologic tests, cochlear implant, interactive testing, laser videodisc, repair strategies, sentence gist recognition (SGR)

Abbreviations: CD = compact disc, CPHI = Communication Profile for the Hearing-Impaired, CST = Connected Speech Test, SGR = Sentence Gist Recognition, SHHH = Self-Help for Hard of Hearing Persons

When testing speech recognition skills, audiologists often use tests that present nonsense syllables, words, or sentences. The client's task is to repeat verbatim the test stimuli. Some investigators suggest that traditional audiologic tests may not always index how well clients perform in conversational interactions, for at least three reasons (e.g., Cox et al, 1987; Tye-Murray and Tyler, 1988; Binnie, 1991; Erber, 1992). First, most audiologic test lists present unrelated speech stimuli. In typical conversations, utterances are related by linguistic and often situational context. Second, clients usually must repeat what they hear verbatim. In everyday conversation, they more often listen for the gist of the message and not for the purpose of repeating every word (Tye-Murray et al, 1992; see also Bransford and Frank, 1971, and Sachs, 1967, for related research). Finally, clients usually interact with their conversational partners during the give and take of an ongoing conversation. If a client does not recognize an utterance, the individual can ask the partner to repeat or modify the message. Clients who repair communication breakdowns effectively might experience fewer difficulties in conversations than standard audiologic test results would predict; clients who do not repair breakdowns very well might experience more difficulties.

Nonsense syllable tests present additional problems to those listed above. They have low face validity because clients are not required to recognize syllables such as /apa/ and /ada/ in everyday conversations. Nonsense syllable tests also do not require clients to organize streams of information into words or larger linguistic

*Central Institute for the Deaf, St. Louis, Missouri; †Department of Otolaryngology—Head and Neck Surgery, University of Iowa Hospitals and Clinics, Iowa City, Iowa

Reprint requests: Nancy Tye-Murray, Central Institute for the Deaf, 909 South Taylor St., St. Louis, MO 63110

units or process information with the same sequential rapidity that is necessary when attending to ongoing speech. Clients cannot use suprasegmental information for recognizing the test items either.

Some investigators have created new test procedures to circumvent some of the problems associated with traditional audiologic test procedures (Cox et al, 1987, 1988; Montgomery, 1991; Erber, 1992). Cox and her colleagues developed the Connected Speech Test (CST). The CST presents sets of related sentences spoken conversationally by one talker. The client learns beforehand the topic of each set of sentences and then attempts to repeat verbatim each sentence. Performance is scored by the percentage of key words correct. Erber (1992) designed an adaptive hierarchical procedure. Subjects hear a sentence repeated up to five times in order to determine the minimal perceptual conditions that must exist in order for a subject to repeat the sentence verbatim, and also to index how many times a sentence must be repeated in order for it to be correctly recognized. The adaptive hierarchy progresses as follows: the sentence is presented in an audition-only condition, the sentence is repeated in an audition-only condition, the sentence is repeated with clear speech (e.g., slow speaking rate, inserted pauses), the sentence is repeated with clear speech, with one important word presented in an audition-plus-vision condition, and, finally, if the subject is still unable to repeat the sentence verbatim, it is presented again with clear speech, with all words presented in an audition-plus-vision condition. Erber's procedure is interactive because the subject's responses influence what happens next in the test procedure. However, the subject cannot choose which repair strategy to use, and repair strategies are limited to repetition and clear speech. Although these newer test paradigms bypass some of the shortcomings of more traditional audiologic tests, they still present speech recognition tasks that differ from real-world conversational tasks.

In this investigation, we evaluated a computerized interactive procedure, the Sentence Gist Recognition (SGR) test, which may mimic a real-world speechreading task more closely than many previous test measures. The SGR consists of sets of related sentences that are supported by topical and environmental contextual cues. Clients are not asked to repeat sentences verbatim but, rather, must recognize a picture illustrating a test item. When clients cannot recognize a message, they may interact with the talker via a computer interface and ask for additional information until recognition is achieved.

The two experiments in this investigation had three purposes. The first purpose was to describe the performance of a group of adult multichannel cochlear implant users on a range of audiovisual speech recognition tasks and self-assessment items and to relate perceived handicap to audiometric performance. Previous research suggests that most cochlear implant users can recognize some speech in an audiovisual condition (e.g., Tyler et al, 1988b; Dorman, 1993), and ability to recognize nonsense syllables correlates with an ability to recognize words in sentence contexts (Tyler et al, 1988a). Research has also shown that traditional word recognition tests may not fully index the amount of listening handicap experienced by hearing-aid users in everyday conversations (Rowland et al, 1985; Matthews et al, 1990; Pedersen and Rosenhall, 1991). In this investigation, we determined whether sentence recognition tests reflect the handicap experienced by cochlear implant users when they attempt to speechread, and whether SGR scores might correspond better with subjective measures than scores from traditional audiologic tests.

The second purpose was to determine whether the SGR indexes similar speech recognition abilities as more traditional audiologic measures, despite the differences in test tasks. Verbatim repetition, especially in the case of a nonsense syllable test, depends more upon analytic speech recognition skills, whereas recognition of the gist of a sentence (as with the SGR) may depend more on synthetic skills. Conceivably, some clients may be more analytic perceivers, whereas other clients may be more synthetic perceivers.[1] If the SGR indexes different skills, then it may be worthwhile to include a measure like this in an audiologic test battery, especially if audiometric test results are intended to guide development of an aural rehabilitation plan for a particular client.

The final purpose was to examine use of repair strategies in a structured task, and to explore how usage relates to successful repair of communication breakdown. Previous research

[1]Analytic skills relate to clients' abilities to recognize individual sounds in a message. Synthetic skills relate to their abilities to recognize the gist of a message, even though they might not identify every sound or word (Jeffers and Barley, 1971).

suggests that clients usually ask the talker to repeat a message following communication breakdown (Tye-Murray, 1992; Tye-Murray et al, 1990, 1993). They are unlikely to ask the talker to alter the structure of the message. However, some evidence suggests that asking a talker to repeat a message is less effective than asking the talker to say something different. For instance, sometimes clients are more likely to recognize the message if the talker rephrases it or elaborates than if the talker repeats it (Gagné et al, 1991; Tye-Murray et al, 1994). In this investigation, we examined these issues more fully.

In Experiment I, the SGR was administered in an audition-plus-vision condition. In Experiment II, it was administered in an audition-only condition and, thus, posed a more challenging speech recognition task.

EXPERIMENT I

Subjects

Forty multichannel cochlear implant users (22 female) served as subjects. Participation in this study was a requirement of the Iowa Cochlear Implant Project. Subjects ranged in age from 26 to 75 years, with a mean of 55.1 years (SD = 15.1). Duration of deafness ranged from 2 to 45 years, with a mean of 13.8 years (SD = 12.4). Fifteen subjects used the Cochlear Corporation Nucleus device, sixteen used the Richards Ineraid device, and nine subjects used the Clarion Mini-Med device. The duration of device usage ranged from 3 to 88 months, with a mean of 42.7 months (SD = 26.7).

Materials

The SGR. The SGR consists of 50 sentences stored on a laser videodisc. The speech materials were filmed by a professional film crew. The sentences relate to one of six topics: an evening party, a shoe store, a restaurant, an office lounge, a movie theater, and a home living room. The sentences range in length between five and eight words. They contain commonplace words.

The sentences are spoken by a total of 13 talkers (8 female). With some exceptions, the talkers were selected because they were available to participate in the rather time-consuming filming procedures. Five talkers were selected because they actually are the people they portray in the topic sentence sets: a shoe salesman, a concession stand worker, a 7-year-old girl, an office receptionist, and an office supervisor. The talkers span ages ranging from childhood to middle age. The talkers were not selected because *a priori* they were identified as enunciating exceptionally well. This selection of talkers was designed to increase the face validity of the SGR. During the course of a normal day, most clients communicate with both males and females who vary in age and also vary in how easy they are to speechread.

SGR testing for this experiment proceeded as follows. A subject sat before a high-resolution computer touchscreen 14" monitor. The topic of the sentence set was written across the monitor for a brief period of time. Afterwards, a 10-sec real-time motion film clip was presented. This clip established a context for the subsequent set of sentences. For instance, one clip showed a group of people sitting in an office lounge. A woman walked toward the group and faced one of the persons. This woman then appeared on the screen and spoke a sentence as if she were talking to someone in the group. One sentence was "Please lend me a quarter for pop." Only the woman's head and shoulders appeared as she spoke. Immediately afterwards, six pictures appeared on the screen, one of which corresponded to the sentence. The pictures had some items in common, so that the subject typically must have recognized more than one word to respond correctly.

The subject's task was to touch the appropriate picture in the response screen. If correct, the next sentence item occurred. If incorrect, five repair strategies appeared on the touchscreen that conveyed the following commands: "Say it again," "Speak a key word," "Elaborate the message," "Simplify the message," and "Rephrase the message." The repair strategies were presented both orthographically and graphically (see Tye-Murray et al, 1988). Whichever strategy the subject touched happened right away. For instance, the talker reappeared on the monitor and rephrased the original sentence. The six pictures then reappeared. The client's first response was covered by an "X" and was not available as an option. The subject then had to choose another picture. This sequence continued until the subject responded correctly. Chance performance, the probability of correctly responding on the first trial, was 17 percent.

After the client viewed the 12 sentences related to an office lounge setting, another film clip occurred to establish a context for the next set of sentences. The sequence of events described above repeated until the test concluded. The results were stored on floppy disc and could be printed at the conclusion of the test. The results included the following information: the number of correct responses selected after the first presentation, the second presentation, the third presentation, and so forth; a listing of repair strategies selected after each incorrect response; and whether, after having been presented with a particular repair strategy on a particular presentation, the subject responded correctly. Audiologic performance was determined by computing the average number of presentations needed to respond correctly. Thus, if a subject typically had to select two repair strategies following the initial presentation of an item, that subject would score 3.0. With this scoring procedure, how well subjects selected appropriate repair strategies influenced the results. For instance, a subject who selected a repair strategy that was optimally helpful (e.g., perhaps selecting the "rephrase" repair strategy after not understanding a single word of the initial presentation) likely scored better than a subject who selected repair strategies poorly (e.g., perhaps selecting the "repeat" repair strategy after not understanding a single word).

Before testing began, the subject viewed two practice sets of sentences, along with two film clips that established a context for each set. The example test sentences, the repaired versions of the sentences, and example descriptions of the film clips appear in Appendix 1. The hardware used to present the test is described in Appendix 2.

During test development, subjects with normal hearing completed a preliminary version of the SGR test to ensure that the pictures adequately represented the corresponding test items. The percentage of pictures correctly identified after the first presentation ranged from 94 to 100 percent, with an average of 97 percent (SD = 2.6). When subjects missed an item, they tended to be the same items across subjects. Thus, two changes were made to the preliminary version of the SGR test. First, for two test items, the six-split picture response frames were replaced with new frames that contained pictures that were easier to distinguish. Second, text was superimposed under specific pictures in two different six-split picture frames. The word *menu* was written under a picture that showed a restaurant menu, and the word *cream* was written under a picture of a woman passing a container filled with cream.

Speech Recognition Test Materials. Subjects completed the SGR and two additional, more traditional, audiologic tests: a sentence test and a nonsense syllable test. The sentence test was comprised of the 50 primary sentences presented in Tye-Murray et al (1990). In this test, six different talkers speak a total of 50 sentences. Only the heads and shoulders of the talkers appear on the monitor. The materials are stored on laser videodisc and presented with the same hardware as the SGR (see Appendix 2). The test is scored by the percentage of words repeated verbatim.

The nonsense syllable test was the Iowa Laser Videodisc Consonant Confusion Test (Tyler et al, 1986). This test contains 13 consonants presented in an [aCa] context, spoken by a male talker. Only the talker's head and shoulders appear on the monitor. The materials are stored on laser videodisc and presented on a 14" computer monitor. The test is scored by the percentage of consonants correct.

Self-Assessment Measures. Subjects completed two self-assessment measures, the Communication Profile of the Hearing-Impaired (CPHI; Demorest and Erdman, 1986) and the Speechreading Questionnaire for Cochlear-Implant Users (Tye-Murray and Purdy, 1992). Only the four items relating to speechreading success in a noisy environment were selected from the CPHI for analysis (presented in Appendix 3), since the other items do not relate directly to subjective impressions of speechreading success in the presence of background noise.

The Speechreading Questionnaire for Cochlear-Implant Users is similar to the questionnaire used to assess members of Self-Help for Hard of Hearing Persons (SHHH) in Tye-Murray et al (1992), but has been modified to assess cochlear-implant users rather than hearing aid users. The questionnaire includes items specifically designed to assess speechreading performance with four different types of talkers (familiar and unfamiliar, female and male), as well as other items that assess other domains such as vision-only speech recognition. Two items (presented in Appendix 4) assess performance in a noisy environment and were selected for analysis. Only the responses for unfamiliar talkers were analyzed, since the audiologic tests contained utterances from talkers who were unfamiliar to the subjects. The following definition of speechreading appears on the top of every page of the Speechreading Questionnaire: "the ability to read someone's lips while listening at the same time."

Procedures

Subjects completed the CPHI and the speechreading questionnaire in the clinic either the evening before or the evening after testing. One subject completed both questionnaires in his own home two nights before his test session.

The test session consisted of the SGR and the two audiometric tests (a sentence test and a nonsense syllable test). The three speech recognition tests were presented in an audition-plus-vision condition with six-talker babble. The speech stimuli were calibrated at 65 dB SPL with a +8 dB signal-to-noise ratio.

Each subject completed all of the test materials in a random order on the same day.

Five subjects did not complete the entire test battery and were therefore excluded from any data analysis involving the particular tests that were not completed.

Results

The means and standard deviations for each audiologic test appear in Table 1 (N = 35), along with the minimum and maximum scores achieved. Scores for the SGR represent the average number of presentations necessary for a correct response. The lower the number, the fewer presentations were needed. Subjects demonstrated a wide range of performance, especially for the nonsense syllable and sentence tests.

Pearson correlations revealed that performance on the SGR was correlated with both performance on the nonsense syllable test ($r = -.743$, $p < .0001$) and the sentence test ($r = -.784$, $p < .0001$). Performance on the nonsense syllable test and the sentence test was also correlated ($r = .650$, $p < .0001$). These results suggest that subjects who perform well on one audiologic measure are likely to perform well on another measure, regardless of response task.

In order to relate performance on the audiologic tests to subjective impressions of speechreading performance, a composite score for each subject for the CPHI items was computed by averaging together the responses to the four select items. Alpha measures the correlation between each item and the remaining items in the composite group. Alpha was computed for the four CPHI items and found to be .87, indicating a high level of internal consistency among the item responses (Carmines and Zeller, 1979). A composite score for the speechreading questionnaire was also computed by averaging the responses to the two select items. Alpha for this set of responses was even higher, .95.

To assess the predictive abilities of the audiologic test measures for the CPHI and speechreading questionnaire composite scores, one- and two-variable regression models were fit to the data to determine the R-squares. The R-square values are estimates of the proportion of variance in the response variable accounted for by a predictor variable or set of predictor variables. In this set of analyses, the audiologic test measures were the predictor variables and the CPHI and speechreading questionnaire composite scores were the response variables. The analysis revealed that the SGR accounted for 42 percent of the variance in the CPHI composite scores, whereas the sentence test accounted for 37 percent of the variance and the nonsense syllable test accounted for 33 percent. The SGR accounted for 23 percent of the variance in the speechreading questionnaire responses, while the sentence test accounted for 11 percent of the variance; the nonsense syllable test also accounted for 11 percent. The sentence test and the nonsense syllable test combined accounted for 43 percent of the variance in the CPHI composite score and 14 percent of the variance in the speechreading questionnaire composite score. In short, the SGR was somewhat more predictive for both composite scores (i.e., CPHI and speechreading questionnaire) than either the sentence test or the nonsense syllable test.

To investigate whether the audiologic measures provided redundant information about subjective impressions of speech reading abilities, p values of partial regression coefficients in several multiple regression models were computed. In each model, the partial regression coefficient for each predictor variable, controlling the other predictor variables, was computed (when a partial regression coefficient

Table 1 Scores for the Audiologic Tests (N = 35)

Test	*Mean*	*SD*	*Minimum Score*	*Maximum Score*
Nonsense Syllable Test	66	20	19	96
Sentence Test	62	22	20	93
SGR	1.3	(.32)	2.1	1.0

A score of 1.0 on the SGR represents no errors, scores for the nonsense syllable test indicate percentage of phonemes correct, and scores for the sentence test indicate percentage of words correct.

is significant [$p < .05$], the predictor variable has significant explanatory power not provided by the other predictor variables). The analysis revealed that the SGR provided significant information about the CPHI that was not provided by the sentence test ($p = .0259$) or the nonsense syllable test ($p = .0126$). The SGR also provided significant information about the speechreading questionnaire that was not provided by the sentence test ($p = .0291$) or the nonsense syllable test ($p = .0349$).

Selection of repair strategies and the effects of selection on performance were examined using all 40 subjects. The repair strategies that occurred after a first incorrect response were examined in order to determine which repair strategies were most popular and which repair strategies were most effective. Table 2 indicates that the repeat repair strategy was selected most frequently, followed by the rephrase repair strategy. Subjects rarely selected the key word strategy.

We determined whether subjects were more likely to select the correct response when five of the six picture choices remained after receiving a sentence repetition, or were more likely to respond correctly after receiving a message rephrasing or elaboration. The remaining repair strategies were not considered because they were so rarely selected, and therefore may be unrepresentative of effectiveness. Since subjects had widely varying numbers of repair strategies, a weighting scheme was used in a paired-comparison t-test of pairs of repair strategies, where 1 = the first repair strategy, 2 = the second repair strategy, $n1i$ = number of times the first strategy was used by a subject on a given item presentation, and $n2i$ = number of times the second repair strategy was used by a subject on a given item presentation.

Each subject was weighted according to the following formula:

$$\text{weight } i = \frac{n1i \, (\times) \, n2i}{n1i + n2i}, \text{ if } n1i + n2i > 0$$

The results of the weighted t-test indicated that subjects were not significantly more likely to respond correctly after a sentence rephrasing ($t = 1.41$, $p > .05$) or a sentence elaboration ($t = -2.194$, $p > .05$) than following a sentence repetition.

Table 2 Number of Times Repair Strategy Selected after First Incorrect Response

Repair Strategy	*Times Selected (%)*
Repeat	171 (41)
Simplify	54 (13)
Rephrase	103 (25)
Elaborate	75 (18)
Key Word	15 (4)

EXPERIMENT II

Purpose of Study

The purpose of Experiment II was to evaluate the SGR in an audition-only condition so that the test would be more difficult, and subjects would be required to use more repair strategies.

Subjects

Thirty-five multichannel cochlear implant users, 18 males and 17 females, served as subjects. Twenty-two subjects from Experiment I served as subjects in Experiment II. A minimum of 3 months elapsed between participation in the two experiments. Subjects ranged in age from 28 to 76 years, with a mean of 51.6 years (SD = 15.7). Duration of deafness ranged from 2 months to 67 years, with a mean of 13.8 years (SD = 14.3). Nine subjects used the Cochlear Corporation Nucleus device, 12 subjects used the Richards Ineraid device, and 14 subjects used the Clarion Mini-Med device. Duration of device usage ranged from 6 to 100 months, with a mean of 45.4 (SD = 32.9).

Materials

Test materials consisted of the Iowa Laser Videodisc Consonant Test (nonsense syllable) and the CPHI.

Procedures

Subjects completed the CPHI, the nonsense syllable test, and the SGR during the same test session. The nonsense syllable test and the SGR were presented in an audition-only condition with six-talker babble. The speech stimuli were calibrated at 70 dB SPL with a +8 dB signal-to-noise ratio. Six items from the CPHI were selected for data analysis (see Appendix 3). These items index speech recognition in an audition-only condition in the presence of background noise.

Nine subjects did not complete the entire test battery and were therefore excluded from any data analysis involving the particular test not completed.

Results

On average, subjects needed 2.0 presentations on the SGR (SD = .60) to correctly recognize an item and scored 40.5 percent on the nonsense syllable test (SD = 22.1). Scores on the SGR ranged from 1.24 presentations to 3.32 presentations. Scores on the nonsense syllable test ranged from 7.0 percent to 85.0 percent. The same set of analyses performed in Experiment I were performed in Experiment II. Pearson correlations suggest that performance on the SGR was correlated with performance on the nonsense syllable test ($r = -.716$, $p < .0001$).

A composite score, computed by averaging together the responses to the six selected items from the CPHI, was used to relate performance on the SGR and the nonsense syllable test with subjective impressions of speech recognition. Alpha was found to be .72, indicating a moderately high level of internal consistency among items. A regression model was fitted to the data to assess the predictive power of the SGR and the nonsense syllable test. The analysis revealed that the SGR accounted for 57 percent of the variance in the CPHI composite score while the nonsense syllable test accounted for only 20 percent.

P values of partial regression coefficients were computed to determine whether the SGR and the nonsense syllable test provide redundant information about the subjective measures. The SGR was shown to provide significant information about the CPHI composite score that was not provided by the nonsense syllable test ($p < .0001$).

Selection of repair strategies was examined using all 35 subjects. The repeat repair strategy was selected most frequently (39% of the time), followed by the rephrase strategy (31% of the time). The key word repair strategy was the least popular (6% of the time).

DISCUSSION

The fact that SGR is most predictive of subjective measures and provides significant information beyond what the sentence and nonsense syllable tests provide indicates that this procedure is promising and warrants further study. Certainly, it can be argued that the SGR has greater face validity than more traditional audiologic tests because it requires gist recognition, provides situational context, and allows the subject to interact with the talker.

The selection of repair strategies during the SGR corresponds with previous research showing that persons with hearing impairments typically use the repeat repair strategy rather than alternative strategies following a communication breakdown (Tye-Murray, 1992; Tye-Murray et al, 1992, 1993). The repeat, rephrase, and elaboration repair strategies were found to be equally effective. This finding agrees with Tye-Murray et al (1990) but does not agree with Gagné and Wyllie (1989) and Kelsay et al (reviewed in Tye-Murray, 1994). These latter investigators showed the repeat repair strategy to be less effective. One interpretation of this discrepancy is that relative repair strategy effectiveness varies with the test task or test materials. Gagné and Wyllie (1989) used word and phrases rather than recorded sentences, as in this investigation. Kelsay et al examined repair strategies used during an interactive task where parents instructed their hearing-impaired children to perform activities with objects. Another possibility for the discrepancy might be that effectiveness varies depending on how much of the original utterance was understood. For instance, if none of the message was recognized, rephrasing might be the most effective repair strategy. If all but one or two words were recognized, the repeat repair strategy might be the most effective strategy. When repair strategy selections are averaged over test items, as in the present investigation, this variation may result in all strategies appearing equally effective. On the other hand, investigations that employ stimuli with one or few words, or investigations that permit interaction with a live talker, may yield different results since the range of how much of the original utterance can be recognized is much shorter due to fewer number of words, or because the live talker receives feedback from the subject concerning how much of the original utterance was recognized, and selects repair strategies accordingly.

The present attempt to incorporate interactive technology into audiologic testing is not unique. For instance, Dempsey et al (1992) described an automated version of the continuous discourse tracking task (DeFilipo and Scott, 1978), where text can be repeated if a subject does not provide a verbatim repetition of a text segment. Computer technologies and concepts, such as Compact Disc-Interactive and virtual reality, continue to evolve. As such, interactive procedures will likely become more commonplace in audiologic settings in the near future. New test procedures may increasingly reflect the interactive nature of conversation and present simulations of real-world environmental listening and speechreading conditions.

Additional work is needed with the SGR. It is important to standardize the SGR on a larger group of subjects and perhaps develop equivalent test lists. Adapting the procedure so that it can be used with compact disc (CD) technology may also be worthwhile. Many audiologic clinics employ CD players for routine speech detection and speech discrimination testing. If the SGR were adapted for a CD player, clinicians would not be required to purchase costly laser videodisc equipment. A primary obstacle in adapting the procedure for a CD player is to identify adequate visual substitutes for the film clips that establish situational contexts. Perhaps slides displayed on the test booth wall, or pictures from a picture test manual, would suffice. Finally, performance on the test was good. Perhaps more difficult test items, especially for occasions when testing occurs in an audition-plus-vision condition, might yield scores that fall more in the mid-range of performance.

Acknowledgment. This work was supported (in part) by a research grant awarded to the Department of Otolaryngology—Head and Neck Surgery, University of Iowa, (number 2 P50 DC 00242) from the National Institute of Deafness and Other Communication Disorders, National Institutes of Health; grant RR00059 from the General Clinical Research Centers Program, Division of Research Resources, NIH; the Lions Clubs International Foundation; and the Iowa Lions Foundation.

REFERENCES

Binnie CA. (1991). New perspectives in audiological rehabilitation. In: Studebaker GA, Bess FH, Beck LB, eds. *The Vanderbilt Hearing-aid Report II*. Parkton, MD: York Press, 233–243.

Bransford JD, Franks JJ. (1971). The abstraction of linguistic ideas. *Cog Psychol* 2:331–350.

Carmines EG, Zeller RA. (1979). *Reliability and Validity Assessment*. Beverly Hills, CA: Sage.

Cox RM, Alexander GC, Gilmore C. (1987). Development of the connected speech test (CST). *Ear Hear* 8:119S–126S.

Cox RM, Alexander GC, Gilmore C, Pusakulich KM. (1988). Use of the connected speech test (CST) with hearing-impaired listeners. *Ear Hear* 9:198–207.

De Filippo CL, Scott BL. (1978). A method for training and evaluating the reception of ongoing speech. *J Acoust Soc Am* 63:1186–1192.

Demorest ME, Erdman SA. (1986). Scale composition and item analysis of the Communication Profile for the Hearing Impaired. *J Speech Hear Disord* 29:515–535.

Dempsey JJ, Levitt H, Josephson J, Porrazzo J. (1992). Computer-assisted tracking simulation (CATS). *J Acoust Soc Am* 92:701–710.

Dorman M. (1993). Speech perception by adults. In: Tyler R, ed. *Cochlear Implants: Audiological Foundations*. San Diego: Singular, 145–190.

Erber NP. (1992). Adaptive screening of sentence perception in older adults. *Ear Hear* 13:58–60.

Gagné J-P, Stelmacovich P, Yovetich W. (1991). Reactions to requests for clarification used by hearing-impaired individuals. *Volta Rev* 93:129–143.

Gagné J-P, Wyllie KM. (1989). Relative effectiveness of three repair strategies on the visual-identification of misperceived words. *Ear Hear* 10:368–374.

Jeffers J, Barley M. (1971). *Speechreading*. Springfield, IL: Charles C. Thomas.

Matthews LJ, Lee F, Mills JH, Schum DJ. (1990). Audiometric and subjective assessment of hearing handicap. *Arch Otolaryngol—Head Neck Surg* 116:1325–1330.

Montgomery AA. (1991). Aural rehabilitation: review and preview. In: Studebaker GA, Bess FH, Beck LB, eds. *The Vanderbilt Hearing Aid Report II*. New York: York Press, 223–231.

Pedersen K, Rosenhall U. (1991). Correlations between self-assessed hearing handicap and standard audiometric tests in elderly persons. *Scand Audiol* 20:109–116.

Rowland JP, Dirks DD, Dubno JR, Bell TS. (1985). Comparison of speech recognition-in-noise and subjective communication assessment. *Ear Hear* 6:291–296.

Sachs JS. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Percept Psychophys* 437–442.

Tye-Murray N. (1992). Preparing for communication interactions: The value of anticipatory strategies for hearing-impaired listeners. *J Speech Hear Res* 35:430–435.

Tye-Murray N. (1994). Communication strategies training. In: Gagné J-P, Tye-Murray N, eds. *Research in Audiological Rehabilitation: Current Trends and Future Directions. J Acad Rehabil Audiol Monogr Suppl* 27:193–208.

Tye-Murray N, Knutson JF, Lemke J. (1993). Assessment of communication strategies use: questionnaires and daily diaries. In: Newman CW, Jacobson GP, eds. *Seminars in Hearing: Self-assessment Scales in Audiology*. New York: Thieme, 338–353.

Tye-Murray N, Purdy SC. (1992). *The Speechreading Questionnaire for Cochlear Implant Users*. Iowa City, IA: Department of Otolaryngology, University of Iowa Hospitals.

Tye-Murray N, Purdy SC, Woodworth G. (1992). The reported use of communication strategies by members of SHHH and its relationship to client, talker, and situational variables. *J Speech Hear Res* 35:708–717.

Tye-Murray N, Purdy SC, Woodworth G, Tyler RS. (1990). Effects of repair strategies on visual identification of sentences. *J Speech Hear Disord* 55:621–627.

Tye-Murray N, Tyler RS. (1988). A critique of continuous discourse tracking as a test procedure. *J Speech Hear Disord* 53:226–231.

Tye-Murray N, Tyler RS, Bong B, Nares T. (1988). Using laser videodisc technology to train speechreading and assertive listening skills. *J Acad Rehabil Audiol* 21:143–152.

Tye-Murray N, Witt S, Schum L, Sobaski C. (1994). Communication breakdowns: Partner contingencies and partner reactions. *J Acad Rehabil Audiol* 27:107–133.

Tyler RS, Preece JP, Tye-Murray N. (1986). *The Iowa Phoneme and Sentence Tests* [laser videodisc]. Iowa City, IA: Department of Otolaryngology—Head and Neck Surgery, University of Iowa.

Tyler RS, Tye-Murray N, Gantz BJ. (1988a). The relationship between vowel, consonant and word perception in cochlear-implant patients. In: Banfai P, ed. *Proceedings of the International Cochlear Implant Symposium, Duren, West Germany, 1987*. West Germany: Fachbuchhandlung, 627–632.

Tyler RS, Tye-Murray N, Lansing C. (1988b). Electrical stimulation as an aid to speechreading. *Volta Rev Monogr Speechreading* 90:119–148.

APPENDIX 1

Example Test Items, Repaired Versions of Each Sentence, and Descriptions of Film Clips Presented in the SGR

The repaired versions of each sentence are in the following order: "Say it again," "Simplify the message," "Rephrase the message," "Elaborate the message," and "Speak a key word."

Example Test Sentences

Film clip = evening party

A person stands in the entrance of a door wearing a dress and a jacket. A hostess approaches.

1. Let me take your jacket.
   — Let me take your jacket.
   — Give me your jacket.
   — I will take your jacket.
   — Give me your jacket. I will hang up your jacket.
   — jacket
2 Your dress is very pretty.
   — Your dress is very pretty.
   — Your dress is pretty.
   — You are wearing a pretty dress.
   — You look nice. Your dress is pretty.
   — dress
3. Do you want a glass of wine?
   — Do you want a glass of wine?
   — Do you want some wine?
   — Have some wine.
   — Would you like some wine? Here is a glass.
   — wine

Film clip = living room

A woman walks into a living room, sits down on a couch and takes off her shoes. A man walks into the living room with a cup of coffee and also sits down.

1. I went shopping for groceries.
   — I went shopping for groceries.
   — I bought groceries.
   — I went to the supermarket.
   — I went to the supermarket. I did the shopping.
   — shopping
2. I bought some chicken for dinner.
   — I bought some chicken for dinner.
   — I bought chicken.
   — We will have chicken for dinner.
   — I bought some food for tonight. We will have chicken.
   — chicken
3. What are we having for supper?
   — What are we having for supper?
   — What is for supper?
   — What are we eating tonight?
   — When are we eating? What are we having for supper?
   — supper

APPENDIX 2

Hardware Used to Present the Test Battery

1. IBM PS/2 8516 touch display model 001
2. Pioneer LD-V8000 laser disc player
3. Crown D-75 Dual Channel Power Amplifier
4. One BOSE Music monitor speaker

APPENDIX 3

The following shows test items selected from the CPHI (Demorest and Erdman, 1986) and their corresponding numeric code in the original questionnaire. Subjects assigned one of the following responses to indicate how effectively they communicate in each situation described in the items: rarely, sometimes, half the time, often, and almost always.

Experiment I

Item 2. You are at a social gathering with music or other noise in the background.
Item 8. Someone's talking to you while you are watching TV or listening to the stereo.
Item 12. You are at a dinner party for several other people.
Item 16. You are hearing a conversation at a social gathering while others are talking nearby.

Experiment II

Item 1. Someone in your family is talking to you while you are driving or riding in a car.
Item 4. You are at work and someone is talking to you from another room.
Item 6. You are talking on the telephone when you are at work.
Item 8. Someone is talking to you while you are watching TV or listening to the stereo.
Item 11. You are at home talking on the telephone.
Item 15. You are at home and someone is talking to you from another room.

APPENDIX 4

The following are test items selected from the Speechreading Questionnaire for Cochlear Implant Users (Tye-Murray and Purdy, 1992). Subjects were assigned a score number from 1 (completely disagree) to 100 (completely agree) beside each conversant listed beneath each item to indicate how strongly they agreed with the statement.

1. I am at a meeting. Someone across the table says something to me. I am wearing my cochlear implant. I can speechread the person if the person is:
   1–100____ an unfamiliar male
   1–100____ an unfamiliar female
2. I am talking to several people. I am wearing my cochlear implant. I can speechread if the speaker is:
   1–100____ an unfamiliar male
   1–100____ an unfamiliar female