

# **Development of a Speechreading Test on the Utilization of Contextual Cues (STUCC): Preliminary Findings with Normal-Hearing Subjects**

Jean-Pierre Gagné

*Hearing Health Care Research Unit, Elborn College  
The University of Western Ontario, London, Ontario, Canada*

Kenneth G. Tugby

*Audiology Department  
Northwestern Health Unit, Fort Frances, Ontario*

Jocelyne Michaud

*Cedar Audiology, Sudbury, Ontario*

A preliminary investigation leading to the development of a Speechreading Test on the Utilization of Contextual Cues (STUCC) is described. A total of 396 prospective test items were administered to a group of 30 normal-hearing subjects. Test items consisted of two sentences: an introductory and a test sentence. Each test sentence was paired with an unrelated and a related introductory sentence. Difference scores (for cognate pairs of related and unrelated test sentences) were computed and the internal consistency of the STUCC was examined. The 198 pairs of test sentences yielded a Coefficient Alpha of .75. The elimination of 94 pairs of test sentences increased the Coefficient Alpha to .94.

Speechreading refers to "the process of perceiving spoken language using vision as the sole source of sensory evidence" (Boothroyd, 1988, p. 77). One factor known to enhance speechreading performance is the provision of contextual cues (Hanin, 1988; Sanders, 1982; Summerfield, 1983). Several studies have demonstrated that situational cues (i.e., relevant information obtained from the environment) serve to improve speechreading performance (Garstecki, 1976, 1979; Garstecki & O'Neill, 1980; Pelson & Prather, 1974). For example, Garstecki (1976) compared speechreading performance under three different experimental

---

Address correspondence to: Jean-Pierre Gagné, PhD, Hearing Health Care Research Unit, Elborn College, The University of Western Ontario, London, Ontario, Canada, N6G 1H1.

conditions: (a) without any situational cues, (b) with visual and non-verbal auditory cues unrelated to the main idea of the test sentence, and (c) with visual and auditory cues related to the test sentence. Significantly higher scores were obtained when the test sentences were accompanied by related visual and auditory cues than for the other two experimental conditions. Garstecki (1976) concluded that the provision of relevant situational cues improved speechreading performance.

Linguistic contextual cues in the form of phonological, lexical, syntactic, semantic, pragmatic, and topical constraints have been shown to enhance speechreading performance (Boothroyd, 1988; Summerfield, 1983). For example, visual-phoneme recognition scores obtained with test stimuli that consisted of English words (i.e., with lexical constraints) were found to be significantly higher than scores obtained with nonsense syllables or pseudowords (e.g., Haas, 1982; Hanin, 1988). Also, Hanin (1988) investigated the effects of syntactic, semantic, and topical constraints on speechreading. Visual-word recognition scores obtained with stimuli presented in a carrier phrase were compared to those obtained with stimuli presented in sentences of a known topic which provided additional linguistic cues. Initially, the 24 subjects were divided into two groups based on their speechreading experience. Half of the subjects were inexperienced (normal-hearing) speechreaders whereas the other subjects were cochlear implant recipients (with an acquired hearing loss) who had completed a 10-week aural rehabilitation program. The results of the investigation revealed that both groups of subjects obtained significantly higher scores for the sentences of a known topic. Moreover, the experienced subjects benefited more from the provision of the additional linguistic cues than the inexperienced subjects (Hanin, 1988, p. 58). Later, the same subjects were re-divided into two groups according to their speechreading competency (based on keyword recognition scores for sentence-length stimuli). When analyzed in this fashion, the results revealed that the competent speechreaders made better use of contextual cues than the poor speechreaders (Hanin, 1988, p. 83). The author concluded that linguistic contextual cues in the form of syntactic, semantic, and topical constraints are important determinants to speechreading performance.

Previous investigations have shown that speechreading performance was enhanced by the provision of semantic cues (Albright, Hipskind, & Schnuckers, 1973; Hanin, Kishon-Rabin, & Plant, 1987). Similarly, several investigators have reported that speechreading performance improved significantly when the subjects were informed of the topic of the test sentences before viewing the test stimuli (Hanin, Hnath, & Boothroyd, 1985; Smith & Kitchen, 1972). In summary, it is generally accepted that relevant situational and linguistic cues serve to enhance speechreading performance among adults with normal language competencies (French-St. George & Stoker, 1988; Stoker & French-St. George, 1984).

### **Inter-subject Variability**

The ability to extract and use situational and linguistic contextual cues varies

considerably. Two investigations serve to illustrate this point. First, in Garstecki's (1976) study, the range of scores obtained from individual subjects was greater when relevant situational cues were provided. For the young normal-hearing subjects presented with non-relevant situational cues, performance ranged from 15 to 35% correct, and for the sentences presented with relevant situational cues performance ranged from 45 to 75% correct. For the elderly hearing-impaired subjects presented with non-relevant situational cues, performance ranged from 5 to 25% correct, and for the sentences presented with relevant situational cues performance ranged from 15 to 55% correct. Later, Garstecki (1979) concluded that not all subjects were equally predisposed to benefit from situational cues.

As mentioned above, Hanin (1988) obtained visual-word recognition scores for test stimuli presented either in a carrier phrase or in sentences of a known topic. To measure the effects of syntactic, semantic, and topical constraints on speechreading, a difference score (DS) was computed for each subject by subtracting the score obtained for words presented in a carrier phrase from the scores obtained for the words presented in sentences. The DSs displayed by the group of 24 subjects ranged from  $-2.8$  to  $52.8$  (Hanin, 1988, p. 58). Moreover, the results revealed that speechreading experience or competence alone could not be used to predict a subject's proficiency to extract linguistic contextual cues from sentence-length stimuli. In addition, the results suggest that a subject's visual-recognition score for words presented in isolation cannot be used to predict accurately how much that individual will benefit from contextual cues available in sentence-length stimuli (Hanin, 1988, p. 114).

### **Rationale for the Present Investigation**

Presently there are no standardized procedures available to assess a speech-reader's ability to extract linguistic cues from a spoken message. Yet, it may be of interest to determine whether an individual aural rehabilitation client could benefit from a visual-speech perception training program (or activities) designed to optimize the use of linguistic cues that may be available in a spoken message (Garstecki, 1979). Several visual-speech perception training programs have incorporated activities designed to improve an individual's ability to extract and make use of situational and linguistic cues (e.g., Cherry & Rubinstein, 1988; De Filippo & Scott, 1978; Erber, 1988; Jeffers & Barley, 1971; Kaplan, Bally, & Garretson, 1987; Palmer, 1988; Perry, 1977). However, it is difficult to evaluate the effects of such training activities on speechreading performance. A standardized test designed to assess the use of linguistic cues could be used as one possible outcome measure to evaluate the relative benefit of those training programs.

### **Scope of Present Investigation**

A long-term research program was initiated to develop a speechreading test to assess the use of contextual cues. The test is named the Speechreading Test

on the Utilization of Contextual Cues (STUCC). The present report summarizes the results of an investigation that was completed prior to the formal development and standardization of the STUCC.

Several steps are involved in the process of test development and standardization. The process may include: (a) a definition of the objective(s) of the test and the establishment of its content validity, (b) the development of test items, (c) the selection of the test administration and scoring procedures, (d) an examination of the psychometric properties of the test (including the internal consistency and inter-form reliability), and (e) the establishment of the concurrent and/or construct validity of the test (Anastasi, 1982; Bilger, 1984; Cronbach, 1970; Nunnally, 1978; Sims, 1982).

The initial stages of test development usually require the administration of a large pool of prospective test items to a sample of subjects who are representative of the population for whom the test is intended (Bilger, 1984; Nunnally, 1978). The target population for the STUCC consists of adults with an acquired hearing loss. A total of 396 test sentences were developed as potential test items for the STUCC. A pilot study revealed that subjects required more than 3 hours to complete the test protocol required to develop the STUCC. The recruitment of a sufficient number of hearing-impaired adults (i.e., more than 100) who would volunteer for such an investigation was considered impractical. Moreover, based on previous speechreading experiments, it was surmised that a considerable number of the prospective test items developed for the STUCC would be too difficult to be included in the final test form. In order to minimize the time commitment required from a large number of hearing-impaired subjects, a decision was made to conduct a preliminary investigation with a group of normal-hearing subjects. The two major objectives of the investigation were: (a) to determine whether the test format as well as the test administration and scoring procedures designed for the STUCC could be implemented successfully in a speechreading test and (b) to reduce the pool of prospective test items that were initially constructed for the STUCC. The results of the investigation would be used to determine if the development and standardization of the STUCC should be pursued. It is recognized that any future investigations on the development and standardization of the STUCC will require the participation of hearing-impaired subjects.

## METHOD

### Subjects

Thirty adult females ranging from 18 to 30 years of age participated in the investigation. Most of the subjects were undergraduate students in communicative disorders. Their participation in this study was volunteered without remuneration. Hearing screening tests confirmed that all the subjects had normal hearing sensitivity (i.e., detection thresholds no greater than 25 dB HL [re: ANSI-1969] at .5, 1, 2, and 4 KHz). Also, a vision screening test confirmed

that all the subjects had normal (or corrected normal) visual acuity (i.e., 20/20 binocular vision as measured with a Snellen chart). None of the subjects had any formal training or experience with speechreading prior to their participation in the investigation.

### **Development of the STUCC**

*Background information.* The Speech Perception in Noise test (SPIN; Bilger, Nuetzel, Rabinowitz, & Rzezchowski, 1984; Kalikow, Stevens, & Elliott, 1977) influenced the development of the STUCC. Kalikow et al. (1977) showed that *auditory* keyword recognition scores were significantly higher when the test stimuli were embedded in sentences that provided some related linguistic cues than when the keywords were presented in neutral sentences (i.e., sentences that did not provide any additional linguistic cues). Bilger (1984) administered all 10 forms of the original SPIN test to a group of 128 adults with a sensorineural hearing loss. Performance was scored separately for the 250 low-predictability test items and the 250 high-predictability test items. In order to normalize the data, the low- and the high-predictability score obtained for each subject were transformed into Z-scores. The data indicated that there was a linear relationship between the low- and the high-predictability test items (the slope of the y-intercept was .998). Moreover, the results revealed that when the SPIN test was administered in an auditory-only sensory modality there were no individual differences in the ability to use contextual cues.

Gagné, Seewald, and Stouffer (1987) used a videotape recording of some SPIN sentences to investigate the effects of linguistic cues in a visual speech perception task. Visual-keyword recognition scores were obtained from 48 normal-hearing adults who were shown 200 test items. Half of the test items consisted of high-predictability sentences. In the remaining test items, the same keywords were presented in low-predictability sentences. The mean keyword recognition score obtained for this group of normal-hearing subjects was 6% ( $SD = 12\%$ ). The results failed to show a significant difference in visual-keyword recognition scores between the two types of sentences. These results were attributed (at least in part) to a floor effect that was due to the difficulty of the test items. The authors concluded that the videotape recordings of the SPIN sentences that were used for the investigation were not appropriate to measure the use of linguistic cues in speechreading. Further scrutiny of the data revealed that only 39 keywords were identified correctly by at least 25% of the subjects. Of those 39 test items, there was a higher proportion of keywords presented in a high-predictability context than low-predictability context. This finding suggested that with some modifications it might be possible to incorporate some SPIN sentences into a speechreading test.

The test paradigm used in the STUCC was somewhat similar to the Related Sentences Test which was developed to measure visual memory (De Filippo, 1982). Each item of the Related Sentences Test consisted of two sentences. The first sentence (presented without sound) was the test sentence. The second

sentence, which was related to the test sentence, was spoken (without sound) and simultaneously captioned at the bottom of the television monitor. De Filippo (1982) observed that the test scores were better when the test sentences were paired with a second (related) sentence than when they were presented in isolation. De Filippo's use of captions ensured that the linguistic cues provided by the second sentence would be available to all subjects regardless of their speech-reading proficiency. A similar strategy was incorporated into the STUCC.

*Test stimuli.* Each test item of the STUCC consisted of two sentences: an introductory sentence followed by a test sentence. The introductory sentences were generated by the investigators. They consisted of either declarative or imperative sentences that were between 6 and 9 syllables in length. The test sentences consisted of some modified high-predictability sentences taken from the eight forms of the SPIN test (Bilger et al., 1984). All the proper names contained in the SPIN sentences were changed to personal pronouns. Also, for consistency, the two high-predictability items from the SPIN test that consisted of interrogative sentences were excluded from the STUCC. Thus, 198 modified SPIN sentences were included in the initial version of the STUCC.

Each test sentence was presented once in a related context and once in an unrelated context (see Table 1). In the related context, the introductory sentences were designed to provide linguistic contextual cues that would facilitate the visual recognition of the keywords (used for scoring) that appeared in the accompanying (following) test sentence. The same 198 introductory sentences were used to generate the unrelated test items. Each introductory sentence was randomly paired with a test sentence. The pairs of sentences were examined to ensure

**Table 1**

Example of Related and Unrelated Test Items Included in the STUCC.  
Keywords are Underlined.

| UNRELATED SENTENCES   | RELATED SENTENCES   |
|---|---|
| <i>You need special watering tools. *</i><br>I <u>cut</u> my <u>finger</u> with a <u>knife</u> . **           | <i>I was careless with the sharp blade. *</i><br>I <u>cut</u> my <u>finger</u> with a <u>knife</u> . **         |
| <i>We had a clear view of the stage. *</i><br>The <u>baby</u> <u>slept</u> in his <u>crib</u> . **            | <i>The mother sang a lullaby. *</i><br>The <u>baby</u> <u>slept</u> in his <u>crib</u> . **                     |
| <i>He was unable to pay his bills. *</i><br>He took a <u>bath</u> in the <u>tub</u> . **                      | <i>The shower curtain was torn. *</i><br>He took a <u>bath</u> in the <u>tub</u> . **                           |
| <i>The lineman didn't throw the switch. *</i><br>The <u>bride</u> <u>wore</u> a <u>white</u> <u>gown</u> . ** | <i>The groom was dressed in a black suit. *</i><br>The <u>bride</u> <u>wore</u> a <u>white</u> <u>gown</u> . ** |
| <i>Please send this letter by plane. *</i><br>This <u>camera</u> is out of <u>film</u> . **                   | <i>I can't take your photograph. *</i><br>This <u>camera</u> is out of <u>film</u> . **                         |

\*Introductory Sentences: captioned and spoken without sound.

\*\*Test Sentences: spoken without sound.

that the introductory sentence did not provide any significant linguistic contextual cues related to the accompanying test sentence. The initial version of the STUCC was comprised of 396 prospective test items (i.e., 198 test sentences presented once in an unrelated context and once in a related context).

**Scoring.** Performance on the STUCC is determined by computing the correct visual-word recognition score for a number of pre-selected keywords that appear in the test sentences. The following criteria were used to select the keywords: (a) only one- or two-syllable words and (b) only common nouns, adjectives, adverbs, or verbs (except the verbs: *to be*, *have*, *give*, and *take*). In addition, all the words that met the above criteria were eliminated as keywords if they appeared in any of the introductory sentences used in the STUCC. The initial version of the STUCC was comprised of 1242 keywords (i.e., 621 keywords from related test items and 621 keywords from the unrelated test items). Three scores may be computed from the STUCC: (a) a visual-keyword recognition score for unrelated test items, (b) a visual-keyword recognition score for related test items, and (c) a difference score. The difference score is computed by subtracting performance for the unrelated test items from the performance for the related test items. The difference score is considered to be a measure of the influence of the additional linguistic contextual cues provided by the related introductory sentences.

**Test Recording.** A videotape camera (NEC, model V30) and a professional videocassette recorder (JVC, model BR-64000) were used to record the test items developed for the STUCC. The 396 test items were recorded in a random order (based on the principle of sampling without replacement). Subsequently, a professional editing recorder (JVC, model BR-86000) and a microcomputer (Sony, model SMC-70G) were used to caption the introductory sentences and to format each test item as follows: (a) a test item identification number appeared on the television monitor for three seconds, (b) a written caption of the introductory sentence was displayed (in white) at the bottom of a black television screen for three seconds, (c) the talker appeared on the screen and spoke the introductory sentence while the written caption continued to be displayed on the monitor, (d) the captioned sentence was removed and the talker spoke the test sentence (without any caption), and (e) the talker disappeared from the screen and the monitor was left blank for approximately 10 s (i.e., the response-interval). The edited version of the STUCC was copied onto a Beta-1 videocassette recorder/player (Sony SLO-325) for test administration.

During the recording sessions the talker (a female adult) wore a black turtle-neck sweater and sat in front of a plain grey backdrop. The camera was positioned approximately 1.2 m directly in front of the talker. The videotape image included the talker's head and shoulders. Floodlights were used to remove head shadows and to optimize the visibility of the talker's face. The talker was instructed to look directly into the camera and to articulate each sentence slowly and clearly without any exaggeration of mouth movements.

### Procedure

The test items were divided into four forms. Three test forms consisted of 100 items each, whereas the fourth form consisted of 96 test items. The subjects were tested in groups of five. Each group completed four test sessions over a period of 4 weeks. Only one test form was administered during a test session. A Latin-square based design for incomplete counterbalancing was used to control for learning and order effects (Doehring, 1988). Specifically, counterbalancing for order of presentation of the test forms was achieved with four groups of subjects (i.e., each test form was presented in each test position). Two other unique presentation orders were selected for the remaining two groups of subjects.

The STUCC was played on a 19-inch color television monitor (Sony trinitron, model MPV-1910). The test was presented without sound. The subjects were seated in a semi-circle (at a 0-45° viewing angle) approximately 1.6 m from the television monitor. They were given a response-form and told that each test item consisted of two sentences. They were informed that the first sentence would be captioned as well as spoken (without sound) and that the second sentence would be spoken only (without caption or sound). Also, they were informed that the second sentence might or might not be related to the first sentence. The subjects were instructed to write-down as much of the second (test) sentence as they could. Finally, the subjects were encouraged to guess.

Each keyword was scored separately. To receive credit for a keyword, the subject's response had to be identical to the orthographic representation of the keyword as it appeared on the scoring template that was prepared from the recorded test stimuli. No credit was given if the response differed from the scoring template. Specifically, the morphological markings of the subject's response had to match those of the keywords. For example, no credit was given if the subject responded "spray" for the keyword "sprays." Finally, for test sentences where the responses included more than one keyword, the keywords provided by the subjects had to appear in the appropriate order to be scored correctly.

## RESULTS AND DISCUSSION

### Comparison of Unrelated and Related Test Items

The mean visual keyword recognition score for the unrelated test items was 15.21% correct ( $SD = 14.33$ ) and the mean visual recognition score for the related test items was 33.48% correct ( $SD = 13.65$ ). A one-tailed  $t$ -test for matched samples revealed that the visual keyword recognition scores were significantly higher for the related test items than for the unrelated test items [ $t(29) = 108.35$ ;  $p < .01$ ]. The results indicated that the subjects benefited from the additional contextual cues available in the introductory sentences that were related to the test sentences. It was concluded that the test format incorporated into the STUCC could be used to measure the use of contextual cues in speech-reading.



### Item Analysis

A difference score (DS) was computed for each test sentence. The DS was obtained by subtracting the mean percent correct keyword recognition score for a test sentence presented in an unrelated context from the mean percent correct keyword recognition score for the same test sentence presented in a related context. The DSs for the cognate pairs of sentences ranged from -17 to 83. The mean DS was 39.14% (median = 17%,  $SD = 9.65$ ). A mean DS of zero or a negative difference score was obtained for 17 of the 198 pairs of test sentences.

A reliability program (SPSS/PC+, 1988) was used to examine the internal consistency of the test items. Specifically, item-total correlations were computed between the DSs of cognate pairs of test sentences and the total DS for all 198 cognate pairs of test sentences. When the 198 pairs of test sentences were considered (i.e., all 396 test items), the resulting Coefficient Alpha ( $\alpha$ ) was .75. In order to improve the internal consistency of the STUCC the test items that yielded negative item-total correlation coefficients or coefficients of less than .1 were eliminated as potential test items. This procedure resulted in the elimination of 94 cognate pairs of test sentences (i.e., 188 test items). The remaining 104 cognate pairs of test sentences resulted in a Coefficient  $\alpha = .94$ . Descriptive statistical analyses based on those 208 test items revealed that the mean performance for the unrelated test items was 10.16% (median = 5.53;  $SD = 7.70$ ; range = 58). The mean performance for the related test items was 31.5% (median = 27.75;  $SD = 15.54$ ; range = 79). The mean DS was 21.33% (median = 20.00;  $SD = 13.44$ ; range = 65.54). The psychometric properties of those 104 cognate pairs of test sentences (i.e., 208 test items) are consistent with the test characteristics sought in the development of the STUCC. Hence, those test items were retained as potential test items for the STUCC and will be used to finalize the development and standardization of the STUCC (a copy of the test items may be obtained from the first author).

Several factors were considered in the elimination of test items from the STUCC. First, a major objective of this preliminary investigation was to reduce the number of test items for the STUCC. A smaller pool of test items will reduce test administration time. This should facilitate the recruitment of hearing-impaired subjects required for the next phase of test development. However, it was important to retain a sufficient number of test items as it is likely that additional items will be discarded during the next phase of test development. Also, a sufficient number of test items should be retained because it would be advantageous to develop multiple equivalent test forms of the STUCC.

Second, test items that yielded very low item-total correlation coefficients were discarded because they did not contribute substantially to the internal reliability of the STUCC. A priori, there were no reasons to believe that hearing-impaired individuals would perform better than normal-hearing (college-age) subjects on those test items. Moreover, the Coefficient  $\alpha$  that resulted from the elimination of those test items (i.e., .94) indicates that the internal consistency

of the revised STUCC was high. Those findings support further development of the STUCC. It should be noted that the internal consistency of the STUCC could have been improved further by eliminating additional pairs of test sentences from the pool of remaining test items. In fact, a systematic removal of test items, based strictly on item-total correlation coefficients, revealed that the elimination of an additional 53 cognate pairs of test sentences would yield a Coefficient  $\alpha = .955$ . However, optimizing Coefficient  $\alpha$  was not considered the main purpose of the present investigation.

### CONCLUSIONS

The results of the present investigation should be interpreted with caution. First, only 30 subjects participated in the experiment. Moreover, all the subjects were normal-hearing and most were university students. Hence, the present findings are not necessarily applicable to the population for whom the test is intended (i.e., hearing-impaired adults with an acquired hearing loss). Nevertheless, these preliminary findings suggest that further development of the STUCC is warranted.

The results revealed that the visual-keyword recognition scores for the related test items were significantly higher than those obtained for the unrelated test items. These findings suggest that the introductory sentences that accompanied the test sentences in the related sentence pairs provided contextual cues that served to enhance speechreading performance. The present findings are consistent with those of previous studies that investigated the effects of contextual cues on speechreading performance (Albright et al., 1973; Garstecki, 1976, 1979; Garstecki & O'Neill, 1980; Hanin, 1988; Hanin et al., 1985; Hanin et al., 1987; Pelson & Prather, 1974; Smith & Kitchen, 1972).

The mean performance obtained for the unrelated test items was lower (i.e., 15%) than keyword recognition scores (for sentences) reported in other speechreading investigations (e.g., Bannister & Britten, 1982; Spitzer, Leder, Milner, Flevaris-Phillips, & Giolas, 1987; Wilson, Dancer, & Stamper, 1982). Several factors could account for this finding. First, there are some indications that the test materials used in the present investigation were difficult to speechread. Gagné et al. (1987) reported that the mean visual-keyword recognition scores for high-predictability sentences of the SPIN test was approximately 8% for a group of 48 normal-hearing subjects. Second, the test paradigm of the STUCC may contribute to the low performance observed for the unrelated test items. In the present investigation, the subjects were aware that the test items consisted of either related or unrelated pairs of sentences. However, the natural tendency may have been to "expect" the introductory sentences to *always* provide contextual cues. This perceptual bias may have resulted in lower speechreading performance for the unrelated pairs of sentences. The results of the present investigation do not make it possible to evaluate the effect of the introductory sentences for unrelated test items. Given the relatively low level of performance observed

from a group of normal-hearing adults, most of whom were university students, it is likely that the clinical application of the STUCC may be restricted to adults with an acquired hearing loss.

It is well known that speechreading performance is influenced by the intelligibility of the talker (Kricos & Lesner, 1982; Lesner & Kricos, 1981). In the present investigation, the talker (an audiologist) was aware of the characteristics of intelligible visible-speech. She articulated each sentence clearly without any exaggeration. However, it is not possible to assess the effects of the talker's intelligibility because all the test items were presented by the same person. It is important to note that the present results apply only to the specific videotape recordings used in the present investigation.

The test format incorporated into the STUCC appears to be suitable for clinical application. None of the subjects expressed any difficulty with the test instructions or the test procedure. However, some issues related to the test format require further investigation. First, as noted above, the influence of the introductory sentences on speechreading performance for the unrelated test items remains to be determined. At the present time there is no evidence that performance for the unrelated sentences would be comparable to performance for test sentences presented without an introductory sentence. Second, the test format used in the STUCC does not make it possible to isolate the effects of each type of linguistic cues (i.e., syntactic, semantic, pragmatic, and topical constraints) available in the related test items. Hence, the STUCC should be considered a screening test of the effects of contextual cues in speechreading. Further investigations are required to ascertain the contribution of specific types of linguistic contextual cues provided by the related introductory sentences. Third, the STUCC relies on Difference Scores to quantify the effects of contextual cues in speechreading. The test format makes it possible to observe directly the effects of the introductory sentences on speechreading performance. However, there have been reports that the reliability of test scores is reduced when Difference Scores are used to measure performance (Hummel-Rossi & Weinberg, 1975, cited in Montgomery & Demorest, 1988). The reliability of the STUCC (i.e., test-retest and/or equivalent form reliability) remains to be established.

As a result of the present investigation 94 of the original 198 pairs of test sentences were eliminated from the STUCC. The elimination of those test items resulted in a Coefficient  $\alpha$  that would be acceptable for a clinical test. Moreover, the amount of time required to complete the STUCC was reduced by almost half. It is expected that the shorter test will make it easier (and more efficient) to recruit the subjects required for the next phase of test development and standardization. The remaining items of the STUCC should be administered to a large group of hearing-impaired adults with an acquired hearing loss. The results obtained from a sub-sample of the subjects (approximately two-thirds of the total number of subjects) should be used to perform an item analysis. The test items retained for the STUCC should be used to determine the psychometric properties of the STUCC. If a sufficient number of items are retained (and if the internal

consistency of the test is sufficiently high) it may be possible to develop two equivalent test forms. The final version of the STUCC will be cross-validated by comparing the results obtained from the sub-sample of subjects to the data obtained from the remaining third of the subjects (Bilger, 1984).

Longer-term objectives of the present research program include an investigation of the construct and predictive validity of the STUCC. For example, the sensitivity of the test to accurately measure an individual's use of contextual cues during everyday conversations could be investigated. Also, the use of the STUCC to identify hearing-impaired individuals who could benefit from intervention programs designed to improve their use of contextual cues could be evaluated.

### ACKNOWLEDGEMENTS

This work was supported by a Health System-Linked Research Unit Grant from the Ministry of Health (Ontario) and a research grant from Unitron Limited. We wish to acknowledge M.E. Jennings for her contribution to the development of the STUCC and D. Ling, R.C. Seewald, and D.E. Greenaway for their comments on earlier drafts of the manuscript. Also, we would like to thank M.E. Demorest and A.A. Montgomery for their comments and suggestions. Finally, we are grateful to S. Noh and especially V. Masterson for their help with the data analysis. A version of this report was presented at the annual conference of the American Speech-Language-Hearing Association held in Seattle, Washington, November, 1990.

### REFERENCES

- ANSI. (1969). ANSI S3.6-1969, "Specifications for audiometers" (American National Standards Institute, New York).
- Albright, P., Hipskind, N.M., & Schnuckers, G.H. (1973). A comparison of visibility and speechreading performance on English and Slurvian. *Journal of Communication Disorders*, 6, 45-52.
- Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Macmillan Publishing Co., Inc.
- Bannister, M.L., & Britten, C.F. (1982). Linguistically based speechreading assessment. *Journal of Communication Disorders*, 15, 475-479.
- Bilger, R.C. (1984). Speech recognition test development. In E. Elkins (Ed.), *Speech recognition by the hearing-impaired* (pp. 2-15). Rockville, MD: American Speech-Language-Hearing Association.
- Bilger, R.C., Nuetzel, J.M., Rabinowitz, W.M., & Rzezchowski, C. (1984). Standardization of a test of speech perception in noise. *Journal of Speech and Hearing Research*, 27, 32-48.
- Boothroyd, A. (1988). Linguistic factors in speechreading. In C.L. De Filippo & D.G. Sims (Eds.), *New reflections on speechreading* [Monograph]. *The Volta Review*, 90(5), 77-87.
- Cherry, R., & Rubinstein, A. (1988). Speechreading instruction for adults: Issues and approaches. In C.L. De Filippo & D.G. Sims (Eds.), *New reflections on speechreading* [Monograph]. *The Volta Review*, 90(5), 289-301.
- Cronbach, L.J. (1970). *Essentials of psychological testing* (3rd ed.). New York: Harper and Row.
- De Filippo, C.L. (1982). Memory for articulated sequences and lipreading performance of hearing-impaired observers. *The Volta Review*, 84(3), 134-146.
- De Filippo, C.L., & Scott, B.L. (1978). A method for training and evaluating the reception of ongoing speech. *Journal of the Acoustical Society of America*, 63, 1186-1191.
- Doehring, D.G. (1988). *Research strategies in human communication disorders*. Boston, MA: Little, Brown and Company.
- Erber, N.P. (1988). *Communication therapy for hearing-impaired adults*. Abbotsford, Victoria,

- Australia: Clavis Publishing.
- French-St. George, M., & Stoker, R.G. (1988). Speechreading: An historical perspective. In C.L. De Filippo & D.G. Sims (Eds.), *New reflections on speechreading* [Monograph]. *The Volta Review*, 90(5), 17-31.
- Gagné, J.-P., Seewald, R.C., & Stouffer, J.L. (1987, November). *List equivalency of SPIN forms in assessing speechreading abilities*. Paper presented at the annual convention of the American Speech-Language-Hearing Association, New Orleans, LA.
- Garstecki, D.C. (1976). Situational cues in visual speech perception by geriatric subjects. *Journal of the American Audiology Society*, 2(3), 99-106.
- Garstecki, D.C. (1979). The use of situational cues in visual communication. In M.A. Henoch (Ed.), *Aural rehabilitation for the elderly* (pp. 109-121). New York: Grune and Stratton.
- Garstecki, D.C., & O'Neill, J.J. (1980). Situational cue and strategy influence on speechreading. *Scandinavian Audiology*, 9, 147-151.
- Haas, W.H. (1982). Stimulus predictability and speechreading performance. *The Volta Review*, 90, 156-162.
- Hanin, L. (1988). *The effects of experience and linguistic context on speechreading*. Unpublished doctoral dissertation, The City University Graduate School, New York.
- Hanin, L., Hnath, T., & Boothroyd, A. (1985). *Lipreading with and without knowledge of sentence topic* (Unpublished Report No. RCI13). New York: City University Graduate School.
- Hanin, L., Kishon-Rabin, L., & Plant, G. (1987). *Perception of real and nonsense sentences: Comparison of speechreading and auditory perception of low-pass filtered speech* (Unpublished Report No. RCI23). New York: City University Graduate School.
- Hummel-Rossi, B., & Weinberg, S.L. (1975). Practical guidelines in applying current theories to the measurement of change, Part I: Problems in measuring change and recommended procedures. Ms No. 916, *Journal Supplement Abstract Service*, 5.
- Jeffers, J., & Barley, M. (1971). *Speechreading (lipreading)*. Springfield, IL: Charles C. Thomas Publisher.
- Kalikow, D.N., Stevens, K.N., & Elliott, L.L. (1977). Development of a test for speech intelligibility in noise using sentence materials with controlled word predictability. *Journal of the Acoustical Society of America*, 61(5), 1337-1351.
- Kaplan, H., Bally, S.J., & Garretson, C. (1987). *Speechreading: A way to improve understanding*. Washington, DC: Gallaudet University Press.
- Kricos, P.B., & Lesner, S.A. (1982). Difference in visual intelligibility across talkers. *The Volta Review*, 84, 219-225.
- Lesner, S.A., & Kricos, P.B. (1981). Visual vowel and diphthong perception across speakers. *Journal of the Academy of Rehabilitative Audiology*, 14, 252-258.
- Montgomery, A.A., & Demorest, M.E. (1988). Issues and developments in the evaluation of speechreading. In C.L. De Filippo & D.G. Sims (Eds.), *New reflections on speechreading* [Monograph]. *The Volta Review*, 90(5), 193-214.
- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Palmer, L. (1988). Speechreading as communication. In C.L. De Filippo & D.G. Sims (Eds.), *New reflections on speechreading* [Monograph]. *The Volta Review*, 90(5), 33-42.
- Pelson, R.O., & Prather, W.F. (1974). Effects of visual message-related cues, age, and hearing impairment on speechreading performance. *Journal of Speech and Hearing Research*, 17, 518-525.
- Perry, A.L. (1977). A lipreading curriculum for adults. *The Volta Review*, 79, 86-92.
- Sanders, D.A. (1982). *Aural rehabilitation* (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.
- Sims, D.G. (1982). Hearing and speechreading evaluation for the deaf adult. In D.G. Sims, G.G. Walter, & R.L. Whitehead (Eds.), *Deafness and communication: Assessment and training* (pp. 141-154). Baltimore: Williams and Wilkins.
- Smith, R.C., & Kitchen, D.W. (1972). Lipreading performance and contextual cues. *Journal of Communication Disorders*, 5, 86-90.

- Spitzer, J.B., Leder, S.B., Milner, P., Flevaris-Phillips, C., & Giolas, T.G. (1987). Standardization of four videotaped tests of speechreading ranging in task difficulty. *Ear and Hearing*, 8, 227-231.
- SPSS Inc., M.J. Norusis (1988). *SPSS/PC+ advanced statistics*. V2.0. Chicago, IL: SPSS Inc., 1988.
- Stoker, R.G., & French-St. George, M. (1984). Factors influencing speechreading performance: Research findings. In W.H. Northcott (Ed.), *Oral interpreting: Principles and practices* (pp. 95-122). Baltimore: University Park Press.
- Summerfield, Q.S. (1983). Probabilities and limitations of lipreading. In M.E. Lutman & M.P. Haggard (Eds.), *Hearing science and hearing disorders* (pp. 132-179). New York: Academic Press.
- Wilson, S., Dancer, J., & Stamper, J. (1982). Visual equivalency of Harris' revised CID everyday sentence lists. *The Volta Review*, 82, 267-273.