Cued Speech Journal VI 1998 p71-87
© 1998 National Cued Speech Association, Inc.

Tactual Cued Speech as a Supplement to Speechreading

Lorraine A. Delhorne, Joan M. Besing, Nathaniel I. Durlach, and Charlotte M. Reed

Introduction

The Cued Speech method devised by Cornett (1967) has proven to be a highly effective means of supplementing the information available through speechreading. For example, highly trained deaf receivers of Cued Speech are able to achieve nearly perfect reception of cued conversational sentences (e.g., Nicholls & Ling, 1982; Uchanski et al., 1992). The success of this method, combined with recent advances in speech recognition technology, motivates the development of an automatic cueing system.

The goal of an automatic cueing system is to allow persons who are deaf or hard of hearing to receive cues, equivalent to those used in manually Cued Speech, without the speaker needing to know the Cued Speech system. Such a device would employ signal processing (e.g., automatic speech recognition) to derive phonetic or feature information from the acoustic speech signal of a speaker, convert this information to a set of cues, and display these cues in real time to the speechreader.

In both manual Cued Speech and in previous work on the development of automatic cueing systems, the supplementary cues to speechreading have been presented through visual displays. In Cornett's Cued Speech system, the eight consonant categories are coded by handshape and the four vowel categories by hand location. These manual cues are received visually and integrated with cues from speechreading. In previous work on an automatic cueing system (Cornett, Beadles, & Wilson, 1977), cues derived from automatic speech recognition were displayed visually through an optical system. This system, which was worn like a pair of eyeglasses, was similar in design to the Upton Eyeglass Lipreading Aid (Upton, 1968). Through this display, cues were encoded through seven light emitting diodes. The projected image of this display was superimposed on the face, near the lips, when the speechreader moved his/her own head in the correct position.

In more recent research towards the development of an automatic Cued Speech system (Bratakos, 1995), synthetic cues derived from images of manual cue productions were superimposed on uncued recordings of sentences. The cues, which appeared as slightly smaller images of a hand, were presented discretely (i.e., without transitions) and at a slightly faster rate than manually Cued Speech. Experimental results obtained with experienced receivers of manual Cued Speech indicated that, when the synthetic cues were perfectly transcribed onto the video display, performance was equivalent to that achieved through manual Cued Speech.

Although both manual and synthetic displays of Cued Speech have typically relied on the visual presentation of cues, consideration may also be given to the use of the tactual sense for this purpose. Arguments may be made for potential advantages of tactual compared to visual displays of cues to supplement speechreading. Visual displays must be focused on the face of the talker and projected in the receiver's field of vision such that the cues can be integrated with those derived from speechreading. The presentation of the supplemental cues through a separate modality may allow for better performance on the individual components. Provided that good integration of the two types of information can be achieved, tactual displays of cues may lead to good performance and in some respects be more easily realized than visual displays.

Evidence of the successful integration of tactual cues with speechreading is provided by previous studies employing a variety of tactile displays derived from acoustic-based representations of the speech spectrum (e.g., see Weisenberger, 1992; Bernstein, 1992; Reed & Delhorne, 1995; Besing, Reed, & Durlach, 1995). In addition, some previous studies have been concerned with the tactual reception of cues employed in Cued Speech. Reed et al. (1992) studied the tactual reception of the manual cues associated with Cued Speech as a supplement to the Tadoma method of speechreading employed by some deaf-blind individuals. Using “Cued-Tadoma”, the consonant and vowel identification scores of normal subjects (who were deprived of visual and auditory cues) improved by roughly 25 percentage points over scores for Tadoma alone. The use of “Cued Tadoma” as a therapeutic tool with deaf-blind children has also been reported (e.g., Lasensky & Lane, 1984; Cornett 1985; Lasensky & Danielson, 1987). Finally, Cornett et al (1977) have reported briefly on a study in which cues were transmitted tactually using a keyboard. The output of the keyboard was a 300 Hz vibration delivered by piezoelectric crystals to the fingertips. The subject watched the speaker's face and simultaneously received the keyboard tactile stimulation. Near perfect performance was achieved for the reception of nonsense CV syllables.

The purpose of the current study was to continue to examine the feasibility of the tactual presentation of cues employed in manual Cued Speech as a supplement to speechreading. A pilot study (Experiment 1) was conducted to examine performance under conditions of speechreading alone and speechreading supplemented by tactual reception of the manual cues of Cued Speech. Based on the encouraging results obtained in the pilot study, follow-up studies were conducted to compare performance through visual and tactual reception of cues (Experiment 2A) and to investigate the effects of speaking rate (Experiment 2B).

I Experiment 1 - Pilot Study

Subjects

Six normal-hearing subjects, all of whom were teachers or transliterators of Cued Speech, participated in this study. The length of time that they had been using Cued Speech ranged from 2 years to 16 years (mean = 8.7 years). None of the subjects had any previous experience with the tactual reception of Cued Speech.

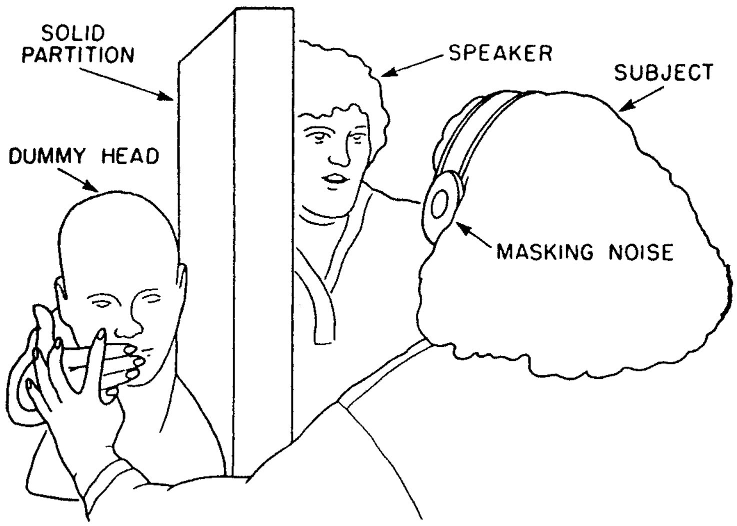

Figure 1 Illustration of the experimental setup.

Procedure

A one-hour session was conducted with each subject to evaluate the reception of nonsense syllables and sentences through Speechreading Alone (SA) and Speechreading plus Tactual Cues (S+TC). Two normal-hearing teachers of Cued Speech (T1 and T2) served as speakers. Each subject was assigned a speaker and remained with that speaker throughout the session. The experimental setup is shown in Figure 1. The speaker and the subject sat at a table facing one another. The subject wore earplugs and was presented with speech-shaped masking noise under headphones to eliminate any auditory cues. A dummy head was positioned to the right of the speaker (approximately an arm's length away) and was blocked from the subject's view by a solid partition. This head was employed in the S+TC conditions such that the speaker transmitted the manual cues using the dummy head as a spatial reference (rather than her own body). The subject received the cues by placing his/her hand loosely on the hand of the speaker and following the movements of the speaker's hand. Thus, the subject received the cues by feeling the sender's hand while the cues were being formed, and at the same time observing the sender's face for speechreading cues. For the SA conditions, cueing was not employed, and the subject simply observed the face of the talker for speechreading cues.

Test materials were placed on a typing stand in front of the speaker. The speaker read the test materials and the subject was instructed to repeat orally any part of the stimulus that they had received either through speechreading alone in the SA condition or through speechreading combined with tactual cues for the S+TC conditions. The subjects' responses were recorded by one of the experimenters. Before the actual testing began in the S+TC conditions, the speaker and the subject spent approximately 5 minutes finding the most comfortable positions for their hands and feeling the various handshapes and positions without using visual cues. All sessions were recorded on audiotape for later analysis.

Materials

Isolated syllables. Materials were CVC nonsense syllables. For each of the syllables (C1-V-C2), C1 was randomly selected from the full set of 24 English consonants used in the initial position; C2 was randomly selected from the set of 21 consonants that can appear in the final position; and V was randomly selected from the full set of 16 English vowels and diphthongs. Two 25-item lists of CVC syllables drawn from the above set were randomly generated. Subjects were tested with one list of 25 CVCs in the SA condition and the second list of 25 CVCs in the S+TC condition. All subjects received the same CVCs in the same conditions and in the same order of presentation. The subject was instructed to provide a CVC response following the presentation of each syllable. Correct-answer feedback was not provided.

Sentences. The sentence materials included lists from the CID Everyday sentences (Davis & Silverman, 1970), CUNY sentences (Boothroyd, Hanin, & Hnath, 1985) and Harvard sentences (IEEE, 1969). CID sentences are highly contextual and considered typical of everyday, conversational speech. Each CID list contains 10 sentences whose length ranges from 2 to 13 words per sentence. Each list contains 50 key words used in scoring. The CUNY sentences, also conversational, though slightly more difficult, contain 12 sentences per list with each list ranging from 3 to 14 words per sentence for a total of 102 words per list. All words in each list are scored for correctness. The Harvard sentences are considerably more difficult (i.e., less internally predictable) than the other sentence materials. Each sentence contains 5 key words and lists were constructed to be phonetically balanced. Scoring is based on key-word identification. All sentence testing was done in an open-set format (i.e., no topic was given prior to the presentation of a sentence). Examples of the CID, CUNY, and Harvard sentences are provided in the Appendix.

Each subject received the same list(s) of sentences for the same conditions and in the same order of presentation. For each of the two conditions (SA and S+TC), one list of CID sentences, two lists of CUNY sentences, and one list of Harvard sentences was presented. Following a sentence presentation, subjects were instructed to repeat orally any part of the sentence they had received. Correct-answer feedback was not provided.

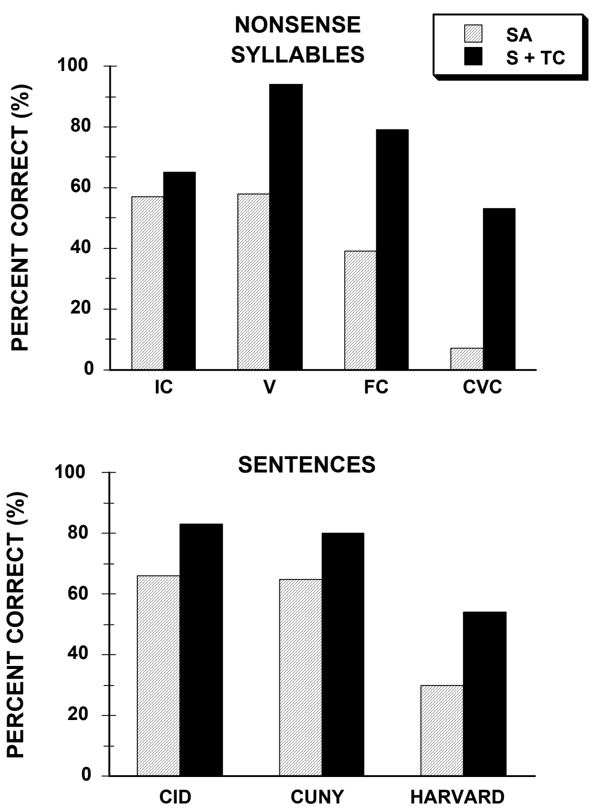

Figure 2 Percent -correct results of pilot study (Experiment 1) averaged across the six subjects. Hatched bars represent speechreading alone (SA) performance, and solid bars the tactual Cued Speech (S+TC) performance. Nonsense syllables (upper panel) were scored for percent correct of individual phoneme (IC=initial consonant, V=medial vowel, FC=final consonant) as well as whole syllable (CVC) identification. Sentences (lower panel) were scored according to key word (CID and Harvard) or total-word (CUNY) identification.

Results

Isolated Syllables. Results from the nonsense syllable testing averaged across the 6 subjects are shown in the top panel of Figure 2. The syllables were analyzed for percent correct of individual phonemes as well as for complete syllables. In all cases performance in the S+TC condition was better than in the SA condition. For complete syllables (CVC), the SA score was 7% correct compared to 53% for the S+TC. For individual phonemes, the presence of the tactual cues had a particularly large effect on the identification of the medial vowel (V) (58% for SA compared to 94% for S+TC) and the final consonant (FC) (39% vs. 79%). A smaller effect was seen with the initial consonant (IC) identification where there was an 8 percentage point improvement when cues were added to speechreading (57% correct for SA compared to 65% for S+TC).

Sentences. Performance on the reception of sentences is shown in the bottom panel of Figure 2. As with the nonsense syllable testing, subjects scored consistently higher when the cues were added to speechreading. Average scores for SA and S+TC were 66% vs. 83% for CID sentences, 65% vs. 80% for CUNY Sentences, and 30% vs. 54% for Harvard sentences. The presentation of tactual cues to supplement speechreading resulted in an average improvement over speechreading alone of roughly 16 percentage points for conversational sentences and 24 percentage points for the less contextual materials. The results of the pilot experiment demonstrated that with very little practice, subjects experienced in the use of visual Cued Speech were able to show substantial improvements to speechreading alone through the tactual reception of the manual cues associated with Cued Speech.

Several studies designed to follow up these preliminary results are described in Experiment 2 below.

II. Experiment 2 - Follow-up Studies

Follow-up studies were conducted with two of the six subjects who participated in the pilot study. Subjects S1 and S2 and both speakers (T1 and T2) traveled to our laboratory for 2 days of further testing. S1, a 30 year old male, learned to cue 9 years prior to our testing and worked as a Cued Speech transliterator for 8 of the 9 years. S2, a 25-year-old female, learned to cue approximately 2 years prior to testing and had been training and evaluating transliterators for 1.5 years.

A. Comparison of Visual and Tactual Reception of Cues

Due to time limitations in the pilot study, visual Cued Speech scores were not obtained for comparison with scores obtained using tactual reception of cues. In this follow-up experiment, two subjects were tested on the reception of nonsense syllables and Harvard sentences under three modes of presentation: SA, S+TC, and S+VC (speechreading combined with normal visual reception of cues). In addition, the number of trials obtained on each experimental condition was increased over that obtained in the pilot study to provide a more reliable estimate of performance.

Procedure

Tactual Session. For both the S+TC and SA conditions, the experimental setup was the same as that used in the pilot study (Figure 1) with the exception that the test materials were projected on a wall behind the subject rather than being displayed on a typing stand in front of the speaker. In this way the speaker could easily read the test materials without looking down. As in the pilot study, one speaker was assigned to each subject: T1 served as speaker for S1 and T2 served as speaker for S2.

Visual Session. Data for both S+VC and SA conditions were obtained using videotaped recordings of the speech material (with no audio signal). These materials were recorded in a professional studio by T1 and T2 on an earlier visit to our laboratory (see Bratakos, 1995); however, only T2's recordings were used for this test session. The videotaped recordings included both cued and non-cued materials. The subjects sat approximately 4.5 ft. from a 20-inch color monitor. After each syllable or sentence was shown, the subjects recorded their responses in writing.

Materials

Isolated Syllables. The CVC syllables were selected, generated, and randomized in the same manner as in the pilot study. A total of 1438 syllables (SA=719, S+TC=719) were presented to each subject in the tactual session, and a total of 300 syllables (SA=150, S+VC=150) in the visual session.

Sentences. Harvard sentence materials were used for these sessions and included a total of 8 lists (SA=4; S+TC=4) for the tactual session and a total of 12 lists (SA=6; S+VC=6) for the visual session.

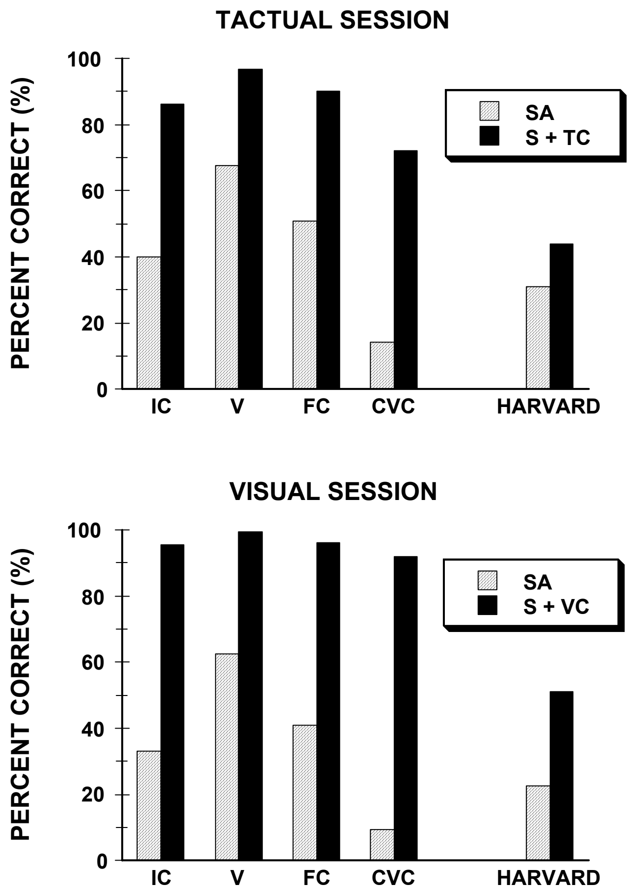

Figure 3 Results averaged across the two subjects of Experiment 2A are shown for tactual session (upper panel) and visual session (lower panel). In upper panel, hatched bars represent scores for speechreading alone (SA) and solid bars represent scores for tactual Cued Speech (S+TC). In lower panel, hatched bars represent SA scores and solid bars represent scores for visual Cued Speech (S+VC). Test materials and methods of scoring are as described for Figure 2.

Results

Results averaged across the two subjects are presented in Figure 3 for both the tactual (upper panel) and visual (lower panel) sessions. The scores obtained in the tactual session for nonsense syllables are generally higher than those obtained in the pilot study (compare with top panel of Figure 2). In particular, scores on the S+TC condition were higher for total CVC syllable recognition as well as for identification of each individual component. The increased number of trials, and increased exposure to the nonsense-syllable task, in the follow-up study appears to have led to greater proficiency in this task. Results on the Harvard sentences, however, did not indicate any improvements as a result of the additional exposure to this task in the follow-up study. In fact, the Harvard S+TC score for the two subjects tested in the follow-up study was roughly 5 percentage points lower than their scores in the pilot study. This result may have arisen from list-to-list variations in the Harvard sentences or from within-subject variability.

Results obtained for the visual reception of nonsense syllables through Cued Speech indicate nearly perfect performance on this task. Compared to the nonsense syllable scores on the S+TC condition, scores on the S+VC condition were slightly higher with the largest difference observed for total syllable recognition (S+VC=92% vs. S+TC=72% correct). For Harvard sentences, an improvement to speechreading of roughly 28 percentage points (SA=23%; S+VC=51%) was observed in the visual session compared to 13 percentage points (SA=31%; S+TC=44%) in the tactual session. In general, the SA scores obtained with live-voice presentation in the tactual session exceeded those obtained with videotaped materials in the visual sessions. Scores on the aided conditions, however, indicated slightly higher performance on S+VC (51%) compared to S+TC (44%) conditions. Nonetheless, with very little practice on the tactual reception of cues, subjects obtained benefits to speechreading that approached the benefits derived from visual reception of cues.

B. Effects of Speaking Rate.

A question that arose during the pilot study was whether the speakers spoke at slower rates (and perhaps more clearly) when presenting the cued material. A slower speaking rate in the S+TC condition may have made the speechreading task easier and thus could account for some of the improvement seen in the S+TC condition. Using the audiotaped recordings of each test session in the pilot study, measurements were made of the speakers' rate of speech while presenting the CUNY sentences in the SA and S+TC conditions. The average speaking rate for the non-cued material was roughly 4.5 syllables/sec (or 215 words/min) compared to an average rate of 3.4 syllables/sec (or 164 words/min) for cued materials. Thus, the production of cues appears to have resulted in a speaking rate that was roughly 1 syllable/sec (or 51 words/min) slower than that for uncued speech.

A special condition was designed to determine whether the slower speaking rate for the cued vs. uncued conditions in the pilot study had an effect on the results. In this condition (SA_{S+TC}), the speaker presented material as she did in the S+TC condition; however, the subjects were not given access (visually or tactually) to the cues and attended only to the speaker's face. (Thus, it was actually a speechreading alone condition). The speechreading scores obtained for this condition were compared to the speechreading scores obtained when materials were presented in the normal SA condition (where no cues were being produced).

Procedure

The physical setup for this experiment was basically the same as in Experiment 2A, except that the subjects were assigned different talkers. Therefore, T1 served as speaker for S2 and T2 served as speaker for S1.

Materials

Four lists of Harvard sentences were presented under each of two conditions: SA and SA_{S+TC}.

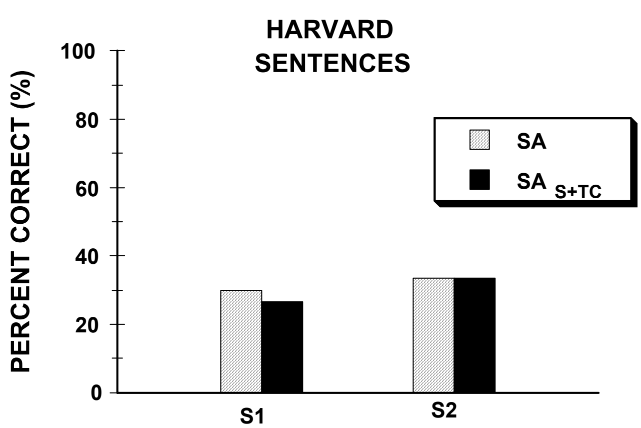

Figure 4 Performance of Individual subjects comparing the effects of speaking rate. The hatched bars represent scores for speechreading when cues are not presented (SA) and the solid bars represent scores for speechreading when the tactual cues are presented but the subjects do not have access to those cues (SA_{S+TC})

Results

Results of each subject's performance in the SA and SA_{S+TC} conditions averaged across the two speakers are shown in Figure 4. The results indicate that the speaking rate difference of 1 syllable/sec in the cued vs. uncued conditions did not affect speechreading scores. For S2, scores were identical for the two conditions, and for S1, the scores from the two conditions were within four percentage points of each other. These results are in agreement with other studies (e.g., Frisina & Bernero, 1958; Byers & Liberman, 1959; Black, O'Reilly, & Peck, 1963; Clarke & Ling, 1976) which demonstrated that rate of speaking does not affect speechreading ability when the rates do not deviate greatly from normal.

Summary and Conclusions

The results of the current study indicate that subjects who were familiar with the Cued Speech system were able to supplement their speechreading scores through the tactual reception of the manual cues associated with this method. That is, by placing a hand over the hand of a speaker who was producing cues, and at the same time attending to the speaker's face for speechreading cues, subjects showed substantial benefits to speechreading without any specific training in the tactual reception of the cues. Improvements to speechreading observed with the tactual reception of cues were only slightly inferior to those obtained by these same subjects for the standard visual reception of Cued Speech.

The visual Cued Speech scores reported in Figure 3 for Harvard sentences are somewhat lower than those reported by Uchanski et al. (1992) and Bratakos (1995) using the same set of videotaped cued materials. In these two previous studies, an average improvement of roughly 55 percentage points was observed for Cued Speech reception of Harvard sentences over speechreading alone, compared to roughly 28 percentage points in the current study. This difference in performance on difficult connected-speech materials most likely arises from different levels of experience in the reception of Cued Speech. The subjects in the Uchanski et al. and Bratakos studies were all deaf individuals who were highly experienced in the reception of Cued Speech. While the subjects in the current study were highly familiar with Cued Speech as teachers and transliterators, they were not as experienced in the reception of speech through this method. Future studies will include experienced deaf receivers of Cued Speech.

Future research will be directed towards the development of an artificial tactual display of cues for use in an automatic Cued Speech system. This research will draw upon current advances in the use of speech-recognition technology for automatic Cued Speech (Bratakos, 1995) and developments in devices for the stimulation of the kinesthetic and vibrotactile aspects of the tactual sensory system (Tan, 1996).

Appendix

Sentence reception test materials. Key words, as specified by the authors of these tests, are underlined.

CID Sentences (Davis & Silverman, 1970)

CID Sentence List A

1. Walking's my favorite exercise.
2. Here's a nice quiet place to rest.
3. Our janitor sweeps the floors every night.
4. It would be much easier if everyone would help.
5. Good morning.
6. Open your window before you go to bed.
7. Do you think that she should stay out so late?
8. How do you feel about changing the time when we begin work?
9. Here we go.
10. Move out of the way.

CUNY Sentences (Boothroyd et al., 1985)

CUNY Sentence List 1

1. Have you eaten yet?
2. How many of your brothers still live at home?
3. How many years of school did it take to become a nurse?
4. Where can I get my suit cleaned?
5. Cats are easy to take care of because you don't have to walk them.
6. She just moved into a three room apartment.
7. I like to play tennis.
8. We couldn't fly home yesterday because of the big snowstorm.
9. Remember to get plenty of rest and drink lots of fluids.
10. Carve the turkey.
11. Make sure you deposit that check.
12. Please make sure that you practice a lot before your next piano lesson.

Harvard Sentences (IEEE, 1969)

Harvard Sentence List 1

1. The birch canoe slid on the smooth planks.
2. Glue the sheet to the dark blue background.
3. It's easy to tell the depth of a well.
4. These days a chicken leg is a rare dish.
5. Rice is often served in round bowls.
6. The juice of lemons makes fine punch.
7. The box was thrown beside the parked truck.
8. The hogs were fed chopped corn and garbage.
9. Four hours of steady work faced us.
10. A large size in stockings is hard to sell.

References

Bernstein, L. E. (1992). The evaluation of tactile aids. In I. R. Summers (Ed.), *Tactile Aids for the Hearing Impaired* (pp. 167-186). London: Whurr Publishers.

Besing, J. M., Reed, C. M., & N.I. Durlach (1995). A comparison of auditory and tactual presentation of a single-band envelope cue as a supplement to speechreading. *Seminars in Hearing, 16*, 316-327.

Black, J. W., O'Reilly, P. P., & Peck, I. (1963). Self-administered training in lipreading. *Journal of Speech and Hearing Disorders, 28*, 183-186.

Boothroyd, A., Hanin, L., & Hnath, T. (1985). *A sentence test of speech perception: Reliability, set equivalence, and short term learning* (Report RC110). City University of New York (New York, NY).

Bratakos, M. S. (1995). *The effect of imperfect cues on the reception of cued speech*. Unpublished master's thesis, Massachusetts Institute of Technology.

Byers, V. W., & Liberman, L. (1959). Lipreading performance and the rate of the speaker. *Journal of Speech and Hearing Research, 2*, 271-276.

Clarke, B., & Ling, B. (1976). The effects of using Cued Speech: A follow-up study. *Volta Review, 78*, 23-34.

Cornett, R.O. (1967). Cued Speech. *American Annals of the Deaf, 112*, 3-13.

Cornett, R. O., Beadles, R., & Wilson, B. (1977). Automatic Cued Speech. In *Research Conference on Speech-Processing Aids for the Deaf,* (pp. 224-239). Washington, DC: Gallaudet College.

Cornett, R. O. (1985). Update on Cued Speech. *Cued Speech Annual, 1*, 3-8.

Davis, H., & Silverman, S. R. (1970). *Hearing and Deafness*. New York: Holt, Rinehart, & Winston.

Frisina, D.R., & Bernero, R.J. (1958). A profile of the hearing and speech of Gallaudet College students. *Volta Review, 60*, 316-321.

IEEE (1969). *IEEE recommended practice for speech quality measurements*. Institute of Electrical and Electronics Engineers, Inc.

Lasensky, J.A., & Lane, P.A. (1984). Case Study – Cued Speech with the Deaf-Blind. *ASHA, 26*, 91 (A).

Lasensky, J. A., & Danielson, P. M. (1987). Case Study: Cued Speech training for a deaf-blind ten-year-old. *Cued Speech Annual, 3*, 73-74.

Nicholls, G. H., & Ling, D. (1982). Cued Speech and the reception of spoken language. *Journal of Speech and Hearing Research, 25*, 262-269.

Reed, C. M., Rabinowitz, W. M., Durlach, N. I., Delhorne, L. A., Braida, L. D., Pemberton, J. C., Mulcahey, B. D., & Washington, D. L. (1992). Analytic study of the Tadoma method: Improving performance through the use supplementary tactual displays. *Journal of Speech and Hearing Research*, 35, 450-465.

Reed, C. M., & Delhorne, L. A. (1995). Current results of a field study of adult users of tactile aids. *Seminars in Hearing, 16*, 305-315.

Tan, H. Z. (1996). *Information transmission with a multi-finger tactual display*. Unpublished doctoral dissertation, Massachusetts Institute of Technology.

Uchanski, R. M., Delhorne, L. A., Dix, A. K., Braida, L. D., Reed, C. M., & Durlach, N. I. (1994). Automatic speech recognition to aid the hearing impaired: Prospects for the automatic generation of Cued Speech. *Journal of Rehabilitation Research and Development, 31 (1)*, 20-41.

Upton, H. W. (1968). Wearable eyeglass speechreading aid. *American Annals of the Deaf*, *113*, 222-229.

Weisenberger, J. M. (1992). Communication of the acoustic environment via tactile stimuli. In I. R. Summers (Ed.), *Tactile Aids for the Hearing Impaired* (pp. 83-109). London: Whurr Publishers.

Acknowledgements

This research was supported by research grant numbers 2-R01-DC-00126-16 and T32-DC-00010 from the National Institute on Deafness and Other Communicative Disorders, National Institutes of Health. The authors would like to thank Susan Russell, Robin Kittleman, Earl Fleetwood, Melanie Metzger, Betsy Kipila, Barbara Williams, Mary Ann Day, and Susan Price for their participation as subjects or talkers in the study. We would also like to thank Dr. R. Orin Cornett for his helpful comments on an earlier version of this paper. Finally, we are grateful to Robert Priest for the illustration used in Figure 1.

Permission to copy: The National Cued Speech Association, Inc, the copyright owner of this document allows single copies to be made for personal use, or multiple copies for classroom use. This consent does not extend to other kinds of copying including (but not limited to) copying for general distribution, for advertising or promotional purposes, for creating new collective works, or for resale.