Veterans Administration

Journal of Rehabilitation Research and Development, Vol. 23 No. 1 1986
pages 95-99

Speech Communication for the Deaf: Visual, Tactile, and Cochlear-Implant[a]

J. M. PICKETT

Rehabilitation Engineering Center for the Deaf and Hearing Impaired
Gallaudet College
Washington, D.C. 20002

Abstract — A review is given of current research and development on electronic devices to aid speech communication for the deaf. Visual and tactile displays are compared with stimulation of hearing via electrodes implanted in the cochlea. Specific comparative performance data are given for cochlear electrical implants versus tactile aids.

INTRODUCTION

This paper reviews recent results in research on new aids for the profoundly deaf. Work in this field concentrates on speech communication, the most crucial problem affecting the education, vocations, and acculturation of the deaf. Persons deafened very early in life, before learning to speak and understand, fail to develop easy speech communication. Lipreading is a very unreliable way to perceive speech messages; the speech information that is visible is estimated to be only about one-third of the necessary information according to Hutton, 1969 (1). The speech of most early-deafened deaf persons is not effectively intelligible.

Aids using visual presentation of speech information have been developed, for speech feedback in voice training as well as for aiding speech reception. Tactile aid development has focused more on reception. Auditory implanted electrodes provide a very rudimentary form of hearing via surgical approaches to the inner ear (cochlea). A detailed history and review of sensory aid developments to 1980 will be found in Levitt et al. (2); a compendium of recent research in implants was published in 1983 by Parkins and Anderson (3). Current work is summarized here and comparisons are drawn between tactile and electroauditory aids.

VISUAL AIDS

The most advanced visual speech aid, the Autocuer, has been developed and pretested with trained deaf viewers at Gallaudet College; see Cornett et al., 1977 (4). The Autocuer uses a 16-channel spectrum analyzer on a microchip to analyze and classify incoming speech sounds. It flashes a symbol representing each syllable of received speech via eyeglasses worn by the deaf lipreader. The system picks out the vowels and consonants in the speech stream and pairs each consonant with the vowel that follows it. This procedure defines consonant-vowel syllables as the recognition units. The system does not attempt to identify the consonant-vowel syllables precisely but it classifies each as one of a category of nine syllables. The viewer can usually determine which syllable of the

[a]This study was supported by the National Institute of Handicapped Research and the Gallaudet Research Institute.
The paper was presented at the Second International Conference on Rehabilitation Engineering, Ottawa, June 1984.

nine was actually spoken from watching the speaker's lip and tongue movements, which are viewed concurrently through the eyeglasses display together with the category-cue symbols. Subjects' performance of 80 percent to 95 percent correct words received under simulated use of the system warranted a program of field tests in 1983.

Electronic displays of speech acoustic patterns have been adapted to teaching the deaf to produce more-intelligible speech. In one aid, the time-varying spectrum of speech is displayed as frequency tracks of the spectral peaks on a video monitor (the Speech Spectrographic Display by Kay Elemetrics of Pinebrook, New Jersey). These instant spectrograms of the pupil's speech seen by the trainee on the monitor are used to train deaf students to speak more intelligibly. This process is being tested at the Rochester School for the Deaf and the National Technical Institute for the Deaf, a division of the Rochester Institute of Technology, Rochester, New York.*

Instead of displaying all of the speech spectrum, workers at other research centers have built indicators of particular acoustic features such as the pitch of the voice (fundamental frequency), the voice intensity envelope, or the general shape of the sound spectrum. (Examples are the indicators available from Speciall Instruments and the Kay Visipitch.)

Indicators of the positions and movements of the tongue and larynx are also under study as feedback devices for speech training. An indicator may operate from the acoustic signal or directly from the speakers' articulators themselves. An example of a direct indicator is the artificial palate developed at the University of Alabama at Birmingham; the palate holds an array of electric contacts in the mouth, and those touched by the tongue during speaking are indicated on a display. Thus a trainee can see how the tongue articulates. This aids in finding the correct tongue placement for producing such consonants as t, d, s, and z, or such vowels as ee and ay, as described by Fletcher in 1982 (5).

Another articulation sensor is the Laryngograph, which monitors the impedance across the larynx between two plates held against the throat. The Laryngograph was developed and is being tested by Prof. A. Fourcin at the Phonetics Department of University College, London, and in our Center at Washington. The impedance varies in waveform with the degree of contact between the vocal folds, or vocal cords. The rapidly changing configurations of the contact between the folds are correlated with voice characteristics. The Laryngograph wave may be processed to provide information for display of several aspects of voice production. Voice characteristics that may be displayed include voice pitch, the coordination of voice, and voice qualities such as clarity and carrying power.

Some speech-training research uses a computer system of speech recognition. A deaf person may occasionally pronounce a word or a phrase very well, but most of the time may not: the problem is the deaf trainee's inability to compare, directly and repeatedly, his or her own production with the model. The recognition system supplies a visible feedback of that comparison. The system can store suitable speech patterns and, for a training session, can be programmed with the pattern of a word the trainee finds difficult. Then the trainee can practice by pronouncing that word into the system, which compares each successive attempt with the stored model and presents a graphic feedback of the degree of match. This concept is being tested by Dr. Osberger at the Boys' Town National Institute for Communication Disorders in Children, Omaha, Nebraska. Scott Instruments manufactures a speech-training system for correcting word pronunciation.*

Tactile Aids to Speech Reception

Tactile displays have been favored over visual as aids to speech reception, because the eyes are often occupied with other tasks while one is receiving speech messages. However, the tactile sense modality is not as well understood as vision and may have limitations that dictate concurrent lip-reading to obtain a usable level of communication with a small, wearable, tactile aid: see Loomis, 1981 (6).

Tactile speech systems have frequently employed the "vocoder" approach. The "channel" vocoder derives a stream of spectral data using a bank of filters to analyze the input speech. The tactile display consists of a row of stimulators, one for each filter band, arrayed along the skin surface. The stimulation intensity of each stimulator is controlled by the sound energy detected in its filter channel. Both vibratory and electrical stimulators have been used in different studies.

Engelmann and Rosov in 1975 (7) demonstrated the performance of a 23-channel vibrotactile vocoder in various word- and sentence-learning tasks with young adult normals and deaf children. The speaker ("teacher") sat beside the tactile-receiving child,

*Personal communication, 1983: R. Whitehead, National Technical Institute for the Deaf, One Lomb Memorial Drive, Rochester, New York 14623.

*Personal communication, 1982: Scott, B; Scott Instruments, Denton, Texas

with no lipreading. First only a small set of five words was used, spoken in isolation, until the subject could identify the words correctly, at first or after prompting, at a criterion level of success. New words were then introduced for further training until criterion was again achieved.

The tactile reception learning of four deaf boys was studied. The best subject (8 years old) achieved 80 percent correct (60 of a pool of 75 words) in a test after 34 weeks of training; he then accelerated his progress to achieve 90 percent correct (122 of a pool of 135 words) in 13 more weeks of training.

One deaf boy was trained wearing his hearing aid in addition to receiving the speech via the tactile vocoder; he learned 35 words to 94 percent correct (tactile plus hearing) in 8 weeks. When tested with hearing only, he scored 65 percent correct, and 40 percent correct with tactile only, indicating that he had been able to integrate tactile with auditory information to identify among his first 35 words. At 21 weeks he achieved 90 percent of 60 words and at the end of 26 weeks he scored 78 percent of 100 words using his hearing aid together with the tactile aid.

Engelmann and Rosov (7) concluded that "hundreds of corrected repetitions are required...to learn simple tactile discriminations" but that, given good learning conditions, the rate of learning will accelerate "once an initial set of 30 to 40 words has been mastered."

Brooks and Frost in 1983, reported a study of subjects who were learning to identify single common words presented on a 16-channel tactile vocoder (8). The subjects had normal hearing. Procedures were similar to those of Engelmann and Rosov (7). Daily sessions of about half an hour were held 5 days a week, continuing to add new 5-word sets to the training/testing pool. The subject with the most-extended training reached criterion on 150 words in 55 hours of training. With further training, 250 words were identified at 80 percent correct and the words were also found to be identifiable in sentences, as reported in Frost et al. in 1983 (9). In a test of tracking the meaning of a connected speech discourse (see test explanation below), the subject achieved a rate of 51 correct words per minute when lipreading with the aid of the tactile vocoder.

Thus it appears that tactile vocoders provide a limited but usable amount of speech information and that training of well-motivated subjects with systematic feedback procedures will enable tactile identification among a large set of words. Wearable tactile vocoders are not yet available, but development is under way to achieve wearability.

Two studies have recently examined the language behavior of very young deaf children with wearable tactile aids. Friel-Patti and Roeser, publishing in 1983 (10) carried out a study of the effects on communication by deaf children of wearing a tactile aid during class and therapy sessions. Four profoundly deaf children each wore a three-channel tactile belt for 10 to 11 hours per week for 16 weeks of their Fall preschool semester. All children also wore hearing aids at all times. Every third week each child participated in an individual half-hour communication-therapy session during which a 10-minute video-recorded sample was made of the child's production of signed and spoken communication while wearing the aids. These productions were analyzed as to duration, type, and content of communication. In the Spring semester, a no-tactile-aid condition was instituted for the same subjects. The same procedures were carried out except that the tactile aids had been removed and the children reverted to the use of hearing aids alone.

Over the Fall "tactile" semester, the duration of communicative productions from the children increased from a mean of about one-and-one-half minutes per 10-minute sample to about 4 minutes. During the following Spring semester, with the hearing aids only, the duration of communication decreased over the semester from a mean of about 3 minutes to a mean of 2 minutes per 10-minute sample. Informal comments also indicated that the tactile aid had had a beneficial effect on the amount of communicative expression by the children. Teachers and parents spontaneously reported that it had been "much easier" to elicit vocalization from the children when they were wearing the tactile aid.

Goldstein et al. reported in 1983 that a 3-year-old deaf child, who used a hearing aid but had only a five-word vocabulary of signed spoken words, greatly increased her communicative attention when she was fitted with a wearable single-channel tactile aid. She then gained a 400-word vocabulary over the following 7 months. This aid simply derived the rhythms of sound received and presented them on a single vibrator worn on the chest (11).

These two studies indicate that relatively simple tactile aids may provide an important increase in communication at a crucial period in the deaf child's education.

Electroauditory Implants and Tactile Aids.

Auditory implants are electrode systems fixed surgically in the cochlea so as to stimulate fibers of the auditory nerve. These fibers are found to be

stimulable in a very large majority of profoundly or totally deaf persons. Because such systems are not able to simulate the normal sound-to-neural conversion process, the sound perceived via implants is highly abnormal. However, it appears that basic auditory rhythms, and some low-frequency pitch information, are perceivable by many of the implanted patients, even with only a single electrode. Under research are multiple-electrode systems that result in extended ranges of usable frequencies. At present the very modest amount of speech information provided by implants seems roughly comparable to the information available through tactile aids. Let us briefly review some of these results here—a complete review will be found in (12).

A multiple-channel tactile system has been laboratory-tested using speech test materials similar to some of those used to test an Australian multichannel implant system: see Clark et al., 1983 (13). The two sets of data offer the opportunity to compare advanced implant and tactile systems. The tactile aid was the Multipoint Electro-Tactile Speech Aid (MESA) reported on by Sparks et al. in 1979 (14). MESA provided 36 channels of speech-spectral information to 36 tactile-stimulating locations arrayed horizontally along the abdomen. The Australian implant system employed 22 electrodes arrayed along a range of mid- to high-frequency fibers ending in the cochlea.

The speech tests were of two types: (i) consonant identification, and (ii) tracking connected discourse as described by DeFilippo and Scott in 1978 (15). Consonant identification was tested with closed sets of spoken syllables (the sets of choices were listed for the subject). In the tracking test the speaker reads a simple story, phrase by phrase, to the receiver, who sits opposite lipreading the speaker and repeating each word or phrase just spoken; if the receiver is wrong on a word, the speaker repeats his utterance until the receiver gets it all correct. The measure of tracking success is the rate of advancement through the story, in words-per-minute. (For a reference rate: if the receiver were using normal hearing, the rate would be simply half the normal oral reading rate, which is about 220/2 or 110 words per minute: see Scott et al., 1977 (16).

The consonant reception of the implant patients, through the implant alone, averaged 41 percent correct; this was boosted to 69 percent correct with added lipreading. The MESA tactile subjects obtained 40 percent correct through the aid alone and 75 percent correct with added lipreading. Thus it appears that the basic consonant information made available is similar in amount whether one uses tactile or implant aids as they exist today.

In the tracking task, the implantees averaged 36 words per minute while the tactile subjects averaged 52 words per minute.

Detailed examination of the individual scores indicated that, unaided, the tactile subjects were much better lipreaders of consonants than were the implantees. Possibly this is because the tactile subjects were normal hearing and were professional collaborators in the research—and thus might be expected to be more language-test-sophisticated than typical implant patients. It should also be noted that, although the tactile subjects received consonants through the aid alone as well as did the implantees, two of the subjects, who had only 15 hours of tactile experience, did not significantly integrate feeling with lipreading. The other three subjects had more previous experience with tactile reception and made complementary use of both tactile and lipreading information.

There are some further results using word- and sentence-tests with a simpler tactile aid developed at the Central Institute for the Deaf, reported by Scott et al. in 1977 (16) and DeFilippo in 1984 (17). These tactile results may be compared with test results reported by Dowell et al. in 1985 (18) for Australian implantees. The tactile aid produced both electrical and vibratory skin sensations. In its two-channel mode, only the overall speech vibration intensity and the occurrences of high frequency noise-like speech sounds (electrical) were provided. In a three-channel mode, information about the spectral spread over the high frequencies was presented via spread-versus-compact electrotactile sensations.

Considering first the implantees' performance in receiving monosyllabic words, only 6 percent were heard correctly through the implant alone but the implant information gave a large boost to success in lipreading the words, to 52 percent correct. The implant-aided lipreading of everyday sentences was highly successful (82 percent to 100 percent correct) except for one implantee at 36 percent. One implantee reported that he could use the telephone: he obtained 21 percent of sentences correct when tested over a phone circuit without lipreading (19). Recently the external processor of the Australian implant system was redesigned with the result that seven implantees who were retested showed a 30 percent improvement; two of them received the sentences 90 percent and 100 percent correct without lipreading; the other five averaged 50 percent correct.

The tactile subjects were not tested at all with

the aid alone, but their tactile-aided lipreading from a large set of monosyllabic words was 66 percent correct, somewhat better than the implant subjects. One of the tactile subjects was also similar to the implantees in aided lipreading of sentences—91 percent correct. These data suggest that a relatively small amount of coded tactile information may have as much aid-potential in lipreading as a relatively complex implant system.

From the point of view of the client, implant aids immediately provide, at the least, rough sensations of sound from the environment; however, implants entail major surgery and high associated costs. Tactile aids entail a learning period but are low-cost, easily replaceable aids. Both types of aid typically provide only a modicum of information about speech. However, the information received through these two aid modalities is partially non-redundant. This suggests that further gains might be obtained with a hybrid system combining a hearing implant and a tactile aid ■

REFERENCES

1. Hutton CL: Aural Rehabilitation. **In** Griffith, J (ed): Persons with Hearing Loss. Springfield, Illinois: C. C. Thomas, 1969.
2. Levitt H, Pickett JM, Houde RA (eds): Sensory Aids for the Hearing Impaired. New York: IEEE Press, 1980.
3. Parkins CW, Anderson SW (Eds): Cochlear Prostheses: An International Symposium. New York: Ann NY Acad Sci (vol 405) 1983.
4. Cornett R, Beadles R, Wilson B: Automatic Cued Speech. **In** Pickett JM (ed), Papers from the Research Conference on Speech-Processing Aids for the Deaf. pp 224-239. Washington, D.C.: Research Institute, Gallaudet College, May 1977.
5. Fletcher SG: Seeing speech in real time. IEEE Spectrum: 41-44, April 1982.
6. Loomis, JM: Tactile pattern perception. Perception 10:5-27, 1981.
7. Engelmann S, Rosov R: Tactile hearing experiment with deaf and hearing subjects. Exceptional Children 41: 243-253, 1975.
8. Brooks PL, Frost BJ: Evaluation of a tactile vocoder for word recognition. J Acoust Soc Am 74:34-39, 1983.
9. Frost BJ, Brooks PL, Gibson DM, Mason JL: Identification of novel words and sentences using a tactile vocoder (abstract). J Acoust Soc Am, Suppl 1, 74:S104-105, 1983.
10. Friel-Patti S, Roeser R: Evaluating changes in the communication skills of deaf children using vibrotactile stimulation. Ear and Hearing 4:31-40. 1983.
11. Goldstein M, Proctor A, Bulle L, Shimizu H: Tactile Stimulation in Speech Reception: Experience with a Non-auditory Child. **In** Hochberg I, Levitt H, Osberger M (eds): Speech of the Hearing Impaired, pp 147-166. Baltimore: University Park Press, 1983.
12. Pickett JM, McFarland WF: Auditory implants and tactile aids for the profoundly deaf. J Speech Hear Res 28:134-150, 1985.
13. Clark GM, Tong YC, Dowell RC: Clinical Results with a Multi-Channel Pseudobipolar System. **In** Parkins CW, Anderson SW (eds), Cochlear Prostheses: An International Symposium. New York: Ann NY Acad Sci (vol 405) pp 370-376, 1983.
14. Sparks DW, Ardell LA, Bourgeois M, Weidmer B, Kuhl P: Investigating the MESA (multipoint electrotactile speech aid): The transmission of connected discourse. J Acoust Soc Am 65:810-815, 1979.
15. DeFilippo CL, Scott BL: A method for training and evaluating the reception of ongoing speech. J Acoust Soc Am 63:1186-1192, 1978.
16. Scott B, DeFilippo C, Sachs R, Miller J: Evaluating with spoken text a hybrid vibrotactile-electrotactile aid to lipreading. **In** Pickett JM (ed), Papers from the Research Conference on Speech-Processing Aids for the Deaf, pp 106-118. Washington, D.C.: Research Institute, Gallaudet College, May 1977.
17. DeFilippo CL: Laboratory projects in tactile aids to lipreading. Ear and Hearing 5:211-227, 1984.
18. Dowell RC, Martin LFA, Clark GM, Brown AM: Results of a preliminary clinical trial on a multiple-channel cochlear prosthesis. Ann Otol Rhinol Laryngol 94:244-250, 1985.
19. Brown AM, Clark GM, Dowell RC, Martin LFA, Seligman PM: Telephone Use by a Multichannel Cochlear Implant Patient: An Evaluation Using Open-Set C.I.D. Sentences. Department of Otolaryngology, University of Melbourne, Australia (unpublished paper).