

Speech training aids for hearing-impaired individuals: I. Overview and aims

LYNNE E. BERNSTEIN, PhD, MOISE H. GOLDSTEIN, Jr., ScD, and JAMES J. MAHSHIE, PhD
*Speech Processing Laboratory, Johns Hopkins University, Baltimore, MD 21218 and Department of
Audiology, Gallaudet University, Washington, DC 20002*

Abstract—Prelingual profound deafness typically results in aberrant or unintelligible speech production. For approximately 70 years, researchers and engineers have attempted, with little success, to provide electronic aids for speech training. Recent computer and signal processing technology has provided the impetus for several groups to implement new speech training aids. Following a review of deaf speech characteristics, several current computer-based aids are described. Included among those reviewed are two interrelated speech training aids which resulted from collaboration among the authors.

Key words: *acoustic/physiological signals, computer-based speech training aids, electroglottograph, fundamental frequency, hearing impaired, microphone, pneumotachograph, Speech Practice/Training Stations.*

INTRODUCTION

In the United States each year, 1,000 to 2,000 children are born with profound deafness or experience profound hearing impairment before they begin to learn speech and language (41). These prelingually, profoundly deaf children benefit only minimally from hearing aids (2). As young children (i.e., ages 18 to 48 months), they typically have aberrant vocalizations and virtually no productive

vocabulary. As older children and adults, the quality of their speech is often not adequate to permit fluent interaction with hearing people (30). The use of electronic speech training aids to improve their speech has long been a goal.

This paper discusses speech characteristics of the deaf; reviews computer-based speech training aids; and describes the development of 2 interrelated computer-based aids developed at The Johns Hopkins University. A description of the hardware implementation of those training aids is given in a companion paper by Ferguson, Bernstein and Goldstein (this volume [11]). A second companion paper by Mahshie, Alquist-Vari, Waddy-Smith, and Bernstein describes software, and clinical experience with the Johns Hopkins aids (also this volume [32]).

BASIC ISSUES

Speech intelligibility of deaf speakers

Several investigations have attempted to determine the characteristics and intelligibility of speech by the deaf. Osberger and McGarr (43) suggest that while differences in the frequency of occurrence of various speech segments (i.e., consonants and vowels) are reported across studies, overall consistency in the quality of segmental productions has been observed. Deaf speakers of English typically do not make use of the full inventory of vowels (of which there are approximately 15). Among the most commonly used vowels are the midvowels / Λ , α /

Address all correspondence and requests for reprints to Lynne E. Bernstein, PhD, Center for Auditory and Speech Sciences, Mary Thornberry Bldg., Gallaudet University, 800 Florida Avenue NE, Washington, DC 20002.

and the low front vowels /æ, ε/ (25,47,50,52). This pattern of production is the result of a tendency to substitute more central vowels (i.e., those produced with the tongue in a more neutral position), for those requiring more extreme articulatory gestures (1,36). An exception to this generalization was reported by Carr (8), whose 5-year-old hearing-impaired subjects produced a wider range of vowels than those observed by other researchers.

Analyses of various speech inventories reveal a consistent pattern of consonant production. Generally, hearing-impaired children tend to produce front consonants /b, p, m, w/ more often than back consonants (e.g., /g,k,h/), probably because the front consonants are more visible for lipreading. The more commonly observed speech errors include: confusion of the voiced-voiceless distinction (e.g., /b/ confused with /p/); substitutions of one consonant for another; inappropriate nasality; difficulty in producing consonant clusters (e.g., "spr" in "spring"); and, omission of word-initial and word-final consonants (6,19,21,27,28).

In addition to articulatory patterns that differ from normal, deaf children also tend to exhibit atypical voice fundamental frequency, duration (or rhythm), intensity, and voice quality patterns. Prosodic characteristics such as stress and intonation are normally conveyed through control of voice fundamental frequency, duration, and amplitude. Among the most common temporal aberrancies are: slower than normal speaking rate (20); prolonged speech sound segments (7); and, more pauses of greater duration than produced by normal-hearing speakers (4,23). In addition, deaf speakers typically fail to produce temporal distinctions that are commonly used by normal-hearing individuals to mark consonant voicing (e.g., the distinction between /b/ and /p/), and lexical stress (e.g., the distinction between "the produce" and "to produce").

Among the most common disruptions affecting the deaf speaker's fundamental frequency are: use of average fundamental frequency that is higher than normal (18); use of a restricted fundamental frequency range leading to a monotonous quality (20); and production of occasional pitch breaks (35). In addition, deaf speakers often exhibit difficulty in producing fundamental frequency patterns that signal lexical stress as well as in grammatical distinctions (e.g., question versus statement) (44).

Perhaps one of the most noticeable aspects of the speech of deaf speakers is its characteristic vocal quality. While the specific factors contributing to such atypical quality are unclear, common voice quality descriptors used to characterize deaf speech are "breathy voice," "tense voice," and "nasal quality."

In addition to recording the various speech errors made by the deaf, investigators have realized that it is important to determine the effects of the various speech errors on intelligibility. In general, negative correlations have been reported between segmental errors and intelligibility: as the number and types of errors increase, intelligibility decreases (21,46).

However, attempts at determining the effect of suprasegmental errors on speech intelligibility have led to somewhat equivocal findings. For example, Hudgins and Numbers (21) report a correlation of 0.73 between speech rhythm and speech intelligibility, which is similar in magnitude to the correlation they report between total consonant errors and intelligibility, and higher than the correlation reported between vowel errors and intelligibility. Others have reported lower correlations between speech timing errors and speech intelligibility (29).

A clear picture has yet to emerge concerning the role of aberrant phonatory characteristics in speech intelligibility. Some research has suggested that inadequate phonatory control, such as intermittent phonation, pitch breaks, loudness breaks, and excessive changes in fundamental frequency, are strongly correlated with speech intelligibility (46). Ling (29) and others (40,48) argue that control of phonation and respiration, and the basic speech postures that underlie suprasegmental speech characteristics, is fundamental to correct production of both segmental and suprasegmental speech characteristics.

In summary, it appears that both segmental and suprasegmental speech errors are common among deaf speakers, and that these errors are related to reduced speech intelligibility.

Auditory feedback in speech production development

The normal-hearing child, or the aided child with hearing loss in the mild-to-severe range, receives auditory feedback for his/her speech production. This feedback is in the service of both speech and language acquisition. The extreme difficulty with

which the profoundly deaf child (with no additional handicapping conditions) achieves intelligible speech and language can be attributed to a lack of adequate auditory feedback. This premise is supported by the fact that some prelingually, profoundly deaf individuals do achieve intelligible speech, but that achievement is typically the result of prolonged individual speech training with a skilled therapist (40). Included among successful speakers are several blind and deaf individuals who communicate via the Tadoma methods (9). Tadoma users, who receive speech information by placing their hand on the face and neck of the talker, can receive and process speech at low to normal rates. These individuals, who achieve speech communication via the somesthetic channel, have achieved intelligible speech despite profound deafness and blindness and have served as "existence proofs" in justifying the implementation of various technologies for speech training. However, it is worth noting that there are only about 10 to 20 Tadoma users and that extensive one-on-one training is required.

Computer-based speech training aids

A number of comprehensive reviews of speech training aids for the deaf have appeared over the past several years (5,26). The overview here is concerned with computer-based aids, but a statement by Braeges and Houde (5) helps put the development and use of all electronic speech training aids for the deaf into perspective:

A speech display which would be useful in teaching speech to the hearing-impaired has been the goal of applied speech science for the past five decades, and the number and variety of aids that have resulted from these efforts are overwhelming. Since the beginning of the modern electronic era (1920), there have been more than 100 different speech training aids developed. Almost all of these have been considered, by their developers, to be significant contributions in the area of speech training. However, few of them have been formally evaluated. Very few have had a significant impact on teaching speech to the deaf, and none have come into widespread use (p. 222).

Braeges and Houde (5) outline some reasons for this outcome, including "erroneously high expectations of both teachers and engineers," and the absence of "clinically developed and tested procedures for using speech aids" (p. 222-223).

Other sources of failure can be suggested. For example, although developmental considerations indicate that the most effective use of training aids must be made during the years of childhood, and especially during the preschool period of language acquisition, most aids appear to have been designed without regard for the specific linguistic, cognitive, and attentional attributes of young children. In early devices, speech information usually was displayed in a manner similar to that used by engineers and speech scientists, such as, time by frequency plots for voice pitch (e.g., the Visipitch from Kay Elemetrics), or spectrograms (47).

Another problem has been the restricted accessibility to speech training aids outside of therapy. Although a student might progress during a therapy session, carryover is typically minimal. In order to effect carryover, extensive practice is required. Osberger, Moeller, and Kroese (42) point out that, "Often, a child is seen for individual speech therapy only once or twice a week for a brief session, or the child receives instruction with a large group of other children" (p. 146). Thus, even if therapy involves a potentially effective speech training aid, its benefits are likely to be limited if that aid is available only when the therapist works with the child.

Problems encountered in using speech training aids may also be the result of placing laboratory equipment in the hands of individuals who do not have specific technical expertise. In their 1973 discussion of speech training aids, Nickerson and Stevens (40) note that, "Some of the devices that have been developed have been rather difficult to use because they require careful and frequent adjustment" (p. 448). Until relatively recently, the use of computers has also involved the necessity for technical expertise. However, the current widespread use of personal computers provides a far different context for developing speech training aids from any that existed until now. The personal computer is a machine that has been engineered for use by individuals without specific technical expertise and has become highly familiar to both children and adults.

Evaluation of speech training aids

In general, speech training aids have undergone only limited clinical evaluations. Some commercially-marketed aids appear not to have been evaluated

at all. Evaluation is needed in at least 4 areas to determine: 1) whether and how speech is improved as a result of work with the aid; 2) whether and how the therapist benefits from use of the aid; 3) whether or not design of the aid takes into account perceptual, cognitive, and attentional characteristics of the user for whom it is intended (for example, the ease with which the aid is used at various age or developmental levels must be determined); and, 4) whether the signal processing capabilities of the aid help the user develop the desired speech characteristics (5).

Providing a review of computer-based speech training aids involves the use of product descriptions and other materials not ordinarily cited in publications. The aids described below were selected for discussion on the basis of the availability of published or presented information that address the points of evaluation mentioned above. Only in the case of the Bolt, Beranek and Newman experimental system (40) have published reports dealt, to some extent, with all the main areas of evaluation listed above.

The Bolt, Beranek and Newman System. The first computer-based speech training aid was developed around a Digital Equipment Corporation PDP-8E minicomputer (40). This was an experimental system, and no commercial system resulted directly from its development. The system consisted of 3 sensors (voice-microphone, accelerometer on the throat, and accelerometer on the nose); a preprocessor, the computer, and various output displays. The preprocessor included a pitch extractor, a spectrum analyzer, and a nasal detector.

Several visual displays were designed to appeal to school age children: 1) a "ball game"; 2) a "vertical spectrum"; and, 3) a cartoon face (39). The ball game software was written so that a "ball" could be made to expand or contract as a function of loudness. The same ball could be driven through a hoop by control of voice pitch. The vertical spectrum appeared as a changing 2-dimensional shape, in which frequency was displayed symmetrically along the y-axis, and amplitude along the x-axis. A cartoon face was used to display voicing, fundamental frequency, loudness, and "s"- or "z"-detection by varying individual attributes of the cartoon as a function of the various speech features. A time-by-speech-attribute display was also implemented, providing amplitude, fundamental frequency, voicing,

nasalization, and second formant frequency as a function of time, presented in a manner that would be familiar to engineers. Nickerson, Kalikow, and Stevens (39) state: "A disadvantage in the use of time functions is the fact that many features cannot be represented simultaneously in an integrated fashion. Showing several time functions in parallel on the same display is a possibility; it is not clear, however, that the viewer can make effective use of such a display" (p. 127). An additional limitation of such displays is that they make no provision for the cognitive/attentional characteristics of children, conforming rather to formats used in the laboratory.

Data collected on use of the system showed that improvements were made to varying degrees along all the dimensions for which visual information was provided (3). Improvement was greater for specifically trained utterances than for spontaneous or elicited untrained speech. The system was used for providing diagnostic information, and students used the system extensively both with and without supervision. It was concluded that a computer-based system could be a valuable tool if used in an "effective speech program" with "adequate teacher preparation."

The IBM-France Speech Training Project. A speech training aid has been under development at IBM-France. In 1979, it was placed in the National Institute for Deaf Children in France. The stated goal for the aid was "to visualize the child's voice," and several of the graphic displays involved presentation of acoustic parameters on plots with time versus a second dimension, such as intensity or pitch (22). Some software was written for speech training by means of playing games in which the child must exercise voice pitch control to move an object around the computer monitor. The developers suggested that the best software designers might well be the children's teachers, and so developed only a limited variety of display software.

Several technical displays are available as part of the system; for example, displays of linear predictive coding and autocorrelation coefficients. This software is considered "too complex for a deaf child, but might prove useful for the training of speech teachers, or for teaching basic acoustics concepts" (22). To our knowledge, this system has not become a commercial product, nor are clinical evaluations available, although the designers have presented several technical reports at conferences (10).

Matsushita Electric Industries and Rion Companies Aid, Japan. The explicit goal of the Japanese system is to provide a training aid that provides information about "all" acoustic-phonetic speech characteristics (38). Like the Johns Hopkins aid, whose goals were developed independently, the Japanese aid is comprised of 2 interrelated systems, one for the clinic and the other for the home. The aid extracts: speech intensity, pitch, spectrum, the voiced versus voiceless distinction, neck vibration, noise vibration, nasality, tongue position, expiratory airflow in front of the mouth, plosiveness, and fricativeness. Sensors are attached by adhesive to the nose and throat of the user. A microphone is mounted in front of the mouth on a headset. An artificial palate is used to detect tongue position and movement. A hand-held sensor detects airflow. The graphics for this system provide simultaneous displays of one or more of the input signals. For example, tongue position is shown by a 2-dimensional graph of the palate and voice pitch is displayed as a time-by-frequency plot.

Training is based on stored models that conform to the desired articulatory or phonatory targets. The trainee can observe the similarity of his/her productions to those of the stored models in terms of the display parameters. Also, a small flower opens or shuts as a graded indication of a "goodness" metric calculation.

The systems were used at the National Rehabilitation Center for the Disabled in Japan. Training involved subjects between the ages of 18 and 20 years. The investigators (38) report improvements in intelligibility for subjects who worked with the system and a therapist; for those who worked with the system alone, however, no adequate description of evaluation methods is given.

The Indiana Speech Training Aid (ISTRA). Currently, a project is underway at Indiana University to develop an aid based on speaker-dependent speech recognition (24). The ISTRA project builds on earlier work (42) at Boys Town Institute for Communication Disorders in Children (50). The Indiana effort is based in part on principles of behavioral psychology that suggest that intelligible speech can be learned through reinforcement of successive syllable, whole-word, or phrase approximations to the desired speech behavior. The model for this approach is the speech therapist providing the child with information about the "goodness" of each utterance.

The aid uses an Interstate Voice Products Vocalink SRB speech recognition board. Use of the board for speech training involves the therapist who works to elicit the child's best speech tokens. Tokens are stored as templates, and subsequent drills use a goodness-of-fit score between the stored target and each utterance produced during the drill. The child is given visual feedback in the form of, for example, a bar graph in which the length of the bar corresponds to the overall goodness of the utterance.

The ISTRA project has involved evaluation of the goodness-of-fit metric, as well as clinical work with a small group of children (24). The goodness-of-fit metric was compared with judgments by a panel of listeners. The average inter-judge correlation was 0.77. The correlation between human judgments and goodness-of-fit scores was 0.78. These results are interpreted as evidence that the speech recognition board is an adequate substitute for a therapist in making determinations of the overall goodness of utterances during speech training drills.

Testing of ISTRA has involved hearing-impaired children and normal-hearing, functional-misarticulating children. Results suggest speech improvements for both trained and untrained words for both the hearing-impaired and normal-hearing children.

The Orometer at the University of Alabama in Birmingham. Fletcher and his co-workers at the University of Alabama in Birmingham have developed a system called the "orometer" based on the notion that "appropriate measures and visual displays of articulatory actions can serve as an alternative speech-learning modality, to parallel auditory-vocal learning of speech by hearing persons" (p. 526) (13).

The orometer uses computer processing of input signals to generate visual displays. Signals available to the orometer are derived as follows: 1) Position, configuration, and movement of the tongue are obtained from optical sensors placed along the midline of a pseudopalate. The pseudopalate is a thin acrylic plate shaped to the individual wearer's palate. The optical sensors (as many as 8), are pairs of miniature narrow-beam light-emitting diodes and phototransistors. 2) The pattern of tongue contact against the roof of the mouth and the teeth is obtained by using an array of as many as 96 metal electrodes in a grid pattern on the pseudopalate. A 10 kHz common signal is applied to the speaker's

wrist and the tongue contact completes the signal path. 3) The positions and movements of the lips and lower jaws are monitored with the aid of a video camera and special signal processors. The camera detects light reflectors or light-emitting diodes on the speaker's lips and face, and on cantilevers attached to the upper and lower teeth. Position and movement of up to 16 light sources can be determined. 4) Acoustical signals from 2 microphones relay data directly to an audio tape recorder, and to a 32-channel filter bank for computer storage and subsequent production of digital spectrograms (12).

Two reports provide information about the clinical use of the system to modify the speech of deaf individuals. Both are case studies. In the first study (16), an 18-year-old profoundly deaf male was fitted with a pseudopalate. Following pre-training tests, the subject was given about 20 hours of direct articulatory instruction and practice. Post-testing was done immediately following training, and again 10 months later. The conclusion was that feedback from the linguapalatal contact patterns can lead to a more accurate production of place and manner cues; however, intelligibility was still very low after training. The 10-month post-test indicated that gains were maintained over the intervening period.

The second case study involved a 3-and-a-half-year-old deaf child (15). A video display was used to present images of tongue position resulting from the child's vocalizations and those of a clinician. Training consisted of nine 20-minute sessions. The findings were generally positive, indicating that the training provides transfer of information from visual perception to motor performance. Furthermore, the subject was able to use an adult model to learn timing control and articulatory gestures.

A study of 10 children with severe-to-profound hearing impairment and 10 control children with normal hearing investigated interplay between lip position and lip-positioning skill (14). Given visually displayed lip position, both groups achieved lip-position targets with high accuracy, suggesting that the feedback was adequate for the hearing-impaired group to perform about as well as the hearing children, despite great differences in speech motor practice.

The Gallaudet University Speech Training System. A project was initiated at Gallaudet University to explore the use of a computer-based system for

assessing, monitoring, and modifying phonatory behavior associated with speech production by the deaf (31). Feedback for non-visible phonatory behavior was considered clinically important because those behaviors involving adjustment of non-visible structures (such as the larynx) are more likely to be inaccurately produced than those involving more visible structures (such as the lips).

The system incorporates 3 components: 1) transducers for detecting the extent of vocal fold contact (the electroglottograph), airflow rate (the pneumotachograph), and the acoustic waveform (a microphone); 2) a specially-designed parameter extraction device for extracting fundamental frequency, laryngeal articulation and vocal quality parameters from the transducer signals; and 3) a PDP 11/34a computer system. The system provides deaf speakers with feedback concerning their phonatory behaviors, rather than feedback based on the acoustic consequence of such behavior. The rationale for this strategy is that providing feedback for speech behaviors themselves is more direct, and therefore likely to be more effective for eliciting new speech patterns.

The system also provides detailed assessment information. For example, it was possible to obtain a detailed analysis of fundamental frequency (F_0) characteristics of a speaker during connected speech. The system permitted characterization of F_0 mean and variance, as well as description of distribution skewness, and extent of F_0 modulation. Such measures were considered important indicators for both diagnostic purposes and as a means of assessing change associated with intervention.

Training results obtained with college-age deaf students showed that deaf speakers are able to use feedback to modify fundamental frequency characteristics of their speech and to learn about the appropriate physiological adjustments of the larynx associated with production of a voiced-versus-voiceless segmental distinction (e.g., /b/ versus /p/) (33,34). While the results were encouraging, and suggested that such an approach to correcting aberrant phonatory aspects of speech production are feasible, the system was not designed for use with young children; nor was it practical to replicate for use outside the laboratory.

Systems for Vowel Training. We know of 2 current efforts to develop computer-based aids for vowel production. One, referred to as the "Vowel

Corrector,” is at the University of Nijmegen in The Netherlands (45). This aid provides feedback in terms of position of vowels on a 2-dimensional plot. Position of speech tokens on the plot is obtained using a statistical technique known as discriminant analysis, which is applied to the output of 15 one-third-octave filters between 250 and 6,400 Hz. Report of an informal test of the Vowel Corrector indicates that children, 8 to 14 years of age, were able to interpret the display and enjoyed working with it.

A second aid is being developed at Old Dominion University (52). This aid makes use of an analog speech parameter extractor to obtain pitch, loudness, and the output from a 16-channel bandpass filter bank. Vowel spectra are converted, via a spectral principal-components analysis (53), to a display based on a correspondence between principal-components and levels of red, green, and blue. As speech is analyzed in real time, each 20-msec-sample is displayed as a color bar whose height corresponds with the amplitude in that sample. The color bars flow across the screen from left to right as new ones are added. Testing with normal-hearing adults suggests that vowels can be identified in terms of their color patterns, and talkers can produce vowels that match target color patterns.

The Johns Hopkins Speech Training Aid

During the past several years, an effort has been under way in the Speech Processing Laboratory of The Johns Hopkins University to develop 2 interrelated computer-based speech training aids for profoundly deaf children. One of the aids, the Speech Training Station (STS) was designed for use in a school or clinic. The other aid, the Speech Practice Station (SPS), was designed to be used by the deaf child at home. The rationale for design of a home system was that deaf children need much greater opportunity to receive guided practice and feedback than they can possibly receive in typical school or clinic therapy.

The STS was designed with the intention of providing therapists and children with information from acoustic (i.e., microphone) and physiologic measures derived from instruments such as an electroglottograph (EGG) and pneumotachograph (PTG). It was posited that the normal-hearing child depends greatly on audition in developing control of phonatory and articulatory activity. A role for

speech training is, therefore, to provide information to substitute for the information normally available through the auditory channel.

The STS and SPS were intended to be used by children as young as 2 to 3 years of age. As suggested above, these young children typically have little or no vocabulary and frequently no control over phonation. A method in wide use for speech training with these children was devised by Ling (29), who outlines a series of stages through which the child is guided, with each stage composed of the training of “subskills.” The first 2 stages are concerned with suprasegmental characteristics of speech: spontaneous vocalization and vocalization on demand, control over voice duration, production of repeated syllables on a single breath, control of vocal intensity, and control of vocal pitch. Later stages build on earlier ones and focus on segmental speech characteristics. Games for the STS and SPS were devised to complement therapy following Ling’s approach.

Physiologic and acoustic measures can be used to achieve estimates of some activities underlying suprasegmental characteristics, such as rate and quality of phonation (17), and control over the breathstream. Training of suprasegmental speech characteristics with the STS and SPS is through use of simple voice-driven computer games. The goal of the software design was to present the child with familiar visual images that are voice-controlled—for example, a balloon rising and falling in relation to loudness—rather than technical displays, such as a time-by-intensity plot. However, some software does incorporate time by voice pitch displays. Several games provide direct feedback in real time for characteristics such as voice pitch, intensity, duration, and rhythm.

Other games were designed to give the child feedback *after* each vocalization. These games were intended to promote automaticity, since the desired goal is for the child to be able to control articulation without the direct visual feedback. The software for training suprasegmental speech characteristics is described in detail in Mahshie *et al.* in this volume (32).

Although most work on the Johns Hopkins aids has, to date, involved development of software and hardware for training suprasegmental speech characteristics, the overall design of the aids takes into account the processing requirements for complex

tasks such as speech recognition. A detailed account of the system is given in Ferguson *et al.* in this volume (11).

Software has been written to facilitate setting up home practice sessions. After use of the STS in the school, parameters are set for use of the SPS in the child's home. The child is given a floppy disk containing those parameters. Records of home practice are stored on the same disk and returned to the school by the child.

Development of the STS and SPS at The Johns Hopkins University. A basic tenet for development of the STS and SPS is that a successful speech training aid can be achieved only through the participation of a group of individuals with expertise in clinical practice, speech science, and engineering, and with the participation of deaf children. A team of such individuals is responsible for the STS and SPS.

Engineering for the STS and SPS took place in the Speech Processing Laboratory at Johns Hopkins University with one full-time engineer and several graduate and undergraduate electrical engineering students. Children began to participate as soon as the first versions of training games were written in software. A STS was placed in the Kendall Demonstration School at Gallaudet University where it was used on a regular basis (32). A second STS remained in the Speech Processing Laboratory and was used for engineering development and therapy sessions involving children from the Baltimore, Maryland area (32). By bringing children into the laboratory on a regular basis to work with a therapist, the engineers and investigators were able to constantly evaluate the evolving system. Problems with using software or hardware were immediately apparent and the nature of speech training and deaf speech was demonstrated to nonclinical personnel.

During the various design phases of the STS and SPS, all of the team members were brought into decision processes. Thus, clinicians, investigators, and engineers met to discuss design goals and possible implementations on a regular basis.

Evaluation. Mahshie *et al.* (this volume [32]) report in detail on the clinical experience obtained with the STS in the laboratory at Johns Hopkins University and at the Kendall School. Included in the report is an evaluation of SPS home trials. Two conclusions in that report are: 1) that the children benefitted from use of the speech aids; and, 2) that

speech practice at home is feasible, and might result in speech activity unlikely to occur otherwise.

DISCUSSION

Attention in this paper has been focused on the capabilities of various computer-based speech training aids. However, engineering efforts alone are unlikely to result in intelligible speech by the deaf. It is anticipated that if these aids are to be effective, they must be used with a therapist working within a curriculum. Development of curricula requires carefully planned and executed clinical investigations. Thus, the currently favorable technological climate must be regarded as only the necessary, but not sufficient, context for development of speech-training aids for the deaf.

Introduction of training aids that use sophisticated signal analyses based on knowledge of acoustic phonetics and/or speech physiology, implies the need for therapists with adequate understanding of acoustic phonetics and speech physiology of deaf speech. Introduction of such aids also implies the need for education of those who are in a position to purchase training aids for clinics and school systems. We believe that these needs cannot be addressed adequately in the laboratory alone but must be addressed by the larger professional community.

ACKNOWLEDGMENTS

The work reported here was supported by NIH/NINCDS (NS-4-2372).

REFERENCES

1. **Angelocci A, Kopp G, Holbrook A:** Some observations on the speech of the deaf. *Volta Rev* 29:156-170, 1964.
2. **Boothroyd A:** *Sensory Capabilities of Hearing-impaired Children*, R.E. Stark (Ed.), Baltimore, MD: University Park Press, 1974.
3. **Boothroyd A, Archambault P, Adams RE, Storm RD:** Use of a computer-based system of speech training aids for deaf persons. *Volta Rev* 77:178-193, 1975.
4. **Boothroyd A, Nickerson R, Stevens K:** *Temporal Patterns in Speech of the Deaf— A Study in Remedial Training*. North Hampton, MA: Clarke School for the Deaf, 1974.

5. **Braeges JL, Houde RA:** Use of speech training aids. In *Deafness and Communication: Assessment and Training*, D.G. Sims, G.G. Walter and R.L. Whitehead (Eds.), Baltimore, MD: Williams & Wilkins, 1982.
6. **Brannon JB Jr:** The speech production and spoken language of the deaf. *Language and Speech* 9:127-135, 1966.
7. **Calvert DR:** Some acoustic characteristics of the speech of profoundly deaf. Unpublished doctoral dissertation, Stanford University, 1961.
8. **Carr J:** An investigation of the spontaneous speech sounds of five-year-old deaf born children. *J Speech Hear Disord* 19:22-29, 1953.
9. **Chomsky C:** Analytic study of the Tadoma Method: Language abilities of three deaf-blind subjects. *J Speech Hear Res* 29(3):332-347, 1986.
10. **De Benedetto MD, Destombes F, Merialdo B, Tuback JP:** Phonetic recognition to assist lipreading for deaf children. In *Proceedings of the International Congress on Acoustics, Speech and Signal Processing*, New York: IEEE Press, 1982.
11. **Ferguson JB, Bernstein LE, Goldstein MH:** Speech training aids for hearing-impaired individuals: II. Configuration of the Johns Hopkins Aid. *J Rehabil Res Dev* 25(4), 1988 (this issue).
12. **Fletcher SG:** Seeing speech in real time. *IEEE Spectrum* 19:42-45, 1982.
13. **Fletcher SG: Dynamic orometrics:** A computer-based means of learning about and developing speech by deaf children. *Am Ann Deaf* 128:525-534, 1983.
14. **Fletcher SG:** Visual feedback and lip-positioning skills of children with and without impaired hearing. *J Speech Hear Res* 29:231-239, 1986.
15. **Fletcher SG, Hasegawa A:** Speech modification by a deaf child through dynamic orometric modeling and feedback. *J Speech Hear Disord* 48:178-185, 1983.
16. **Fletcher SG, Hasegawa A, McCutcheon MJ, Gilliom JD:** Use of linguapalatal contact patterns to modify articulation in a deaf adult. In *Advances in Prosthetic Devices for the Deaf: A Technical Workshop*, D.L. McPherson and M. Schwab (Eds.), Rochester, NY: NTID Press, 1979.
17. **Fourcin AJ:** Laryngographic assessment of phonatory function. *The American Speech-Language-Hearing Association, ASHA Reports* 11:116-127, 1981.
18. **Gilbert H, Campbell M:** Speaking fundamental frequency in three groups of hearing-impaired individuals. *J Commun Dis* 13:195-206, 1980.
19. **Gold T:** Speech production in hearing-impaired children. *J Commun Disord* 13:397-418, 1980.
20. **Hood RB:** Some physical concomitants of the perception of speech rhythm of the deaf. Unpublished doctoral dissertation, Stanford University, 1966.
21. **Hudgins AWF, Numbers FC:** An investigation of the intelligibility of the speech of the deaf. *Genetic Psychology* 44:1-48, 1942.
22. **IBM France:** IBM France Scientific Centre, Paris, France, 1984.
23. **John JE, Howarth JN:** The effect of time distortions on the intelligibility of deaf children's speech. *Language and Speech* 8:127-134, 1965.
24. **Kewley-Port D, Watson CS, Cromer PA:** The Indiana Speech Training Aid (ISTRA): A microcomputer-based aid using speaker-dependent speech recognition. Presented at *American Speech-Hearing-Language Foundation Computer Conference*, Houston, TX, 1987.
25. **Lach R, Ling D, Ling A, Ship N:** Early speech development in deaf infants. *Am Ann Deaf* 115:522-526, 1970.
26. **Levitt H, Pickett JM, Houde RA (Eds.):** *Sensory Aids for the Hearing Impaired*. New York: Institute of Electrical and Electronics Engineers, IEEE Press, 1980.
27. **Levitt H, Smith CR, Stromberg H:** Acoustical, articulatory and perceptual characteristics of the speech of deaf children. In *Proceedings of the Speech Communication Seminar*, G. Fant (Ed.), New York: Wiley, 1976.
28. **Levitt H, Stromberg H:** Segmental characteristics of the speech of hearing-impaired children: Factors affecting intelligibility. In *Speech of the Hearing Impaired: Research, Training, and Personnel Preparation*, I. Hochberg, H. Levitt and M.J. Osberger (Eds.), Baltimore, MD: University Park Press, 1983.
29. **Ling D:** *Speech and the Hearing-Impaired Child: Theory and Practice*. Washington, DC: The Alexander Graham Bell Association for the Deaf, Inc., 1976.
30. **Maasen B:** Artificial corrections to deaf speech: Studies in intelligibility. Doctoral dissertation. Katholieke Universiteit te Nijmegen, Netherlands, 1985.
31. **Mahshie JJ:** A computerized approach to assessing and modifying the voice of the deaf. In *Proceedings of the 1985 International Congress on Education of the Deaf* (in press).
32. **Mahshie JJ, Alquist-Vari D, Waddy-Smith B, Bernstein LE:** Speech training aids for hearing-impaired individuals: III. Preliminary observations in the clinic and home. *J Rehabil Res Dev* 25(4), 1988, this issue.
33. **Mahshie JJ, Hasegawa A, Mars M, Herbert E:** Fundamental frequency characteristics of deaf speakers. *J Acoust Soc Am* 75:S58, 1984.
34. **Mahshie JJ, Herbert E, Hasegawa A:** Use of airflow feedback to modify deaf speakers' consonant voicing errors. *American Speech-Language-Hearing Association* 26:10, 1984.
35. **Martony J:** On the correction of the voice pitch level for severely hard of hearing subjects. *Am Ann Deaf* 113:195-202, 1968.
36. **Monsen RB:** Normal and reduced phonological space: The production of English vowels by deaf adolescents. *J Phonetics* 4:29-42, 1976.
37. **Monsen RB, Leiter E:** Comparison of intelligibility with duration and pitch control in the speech of deaf children. *J Acoust Soc Am* 57:S69, 1975.
38. **Murata N, Yamada Y, Sugimoto T, Hirosawa K, Shibata S, Yamashita S:** Speech training aid for people with impaired speaking ability. In *Proceedings of the International Congress on Acoustics, Speech, and Signal Processing*, New York: IEEE Press, 1986.
39. **Nickerson R, Kalikow DN, Stevens KN:** Computer-aided speech training for the deaf. *J Speech Hear Disord* 41:12-132, 1976.

40. **Nickerson R, Stevens KN:** Teaching speech to the deaf: Can a computer help? *Trans of Audio and Electro-acoustics* AU21(5):445-455, 1973.
41. **NINCDS (National Institute of Neurological and Communicative Disorders and Stroke).** *Human Communication and Its Disorders: An Overview.* Bethesda, MD: National Institutes of Health, Public Health Service, 1969.
42. **Osberger MJ, Moeller MP, Kroese JM:** Computer-assisted speech training for the hearing impaired. *Acad Rehabil Audiol* 14:145-158, 1981.
43. **Osberger MJ, McGarr N:** Speech production characteristics of the hearing-impaired. In *Speech and Language: Advances in Basic Research and Practice* (Vol. 8), N. Lass (Ed.), New York: Academic Press, 1982.
44. **Phillips N, Remillard W, Bass S, Pronovost, W:** Teaching of intonation to the deaf by visual pattern matching. *Am Ann Deaf* 113:239-246, 1968.
45. **Povel D-J, Wansik M:** A computer-controlled vowel corrector for the hearing impaired. *J Speech Hear Res* 29:99-105, 1986.
46. **Smith CR:** Residual hearing and speech production in deaf children. *J Speech Hear Res* 18:795-811, 1975.
47. **Stark RE:** Phonatory development in young normally-hearing and hearing-impaired children. In *Speech of the Hearing-Impaired: Research, Training, and Personnel Preparation*, I. Hochberg, H. Levitt, and M.J. Osberger (Eds.), Baltimore, Maryland: University Park Press, 1982.
48. **Stevens K, Nickerson R, Rollins A:** Suprasegmental and postural aspects of speech production and their effect on articulatory skills and intelligibility. In *Speech of the Hearing-Impaired: Research, Training, and Personnel Preparation*, I. Hochberg, H. Levitt and M.J. Osberger (Eds.), Baltimore, MD: University Park Press, 1983.
49. **Sykes J:** A study of the spontaneous vocalizations of young deaf children. *Psychological Monographs*, 52:104-123, 1940.
50. **Watson CS:** Personal communication, 1987.
51. **West J, Weber J:** A phonological analysis of the spontaneous language of a four-year-old hard of hearing child. *J Speech Hear Disord* 38:25-35, 1973.
52. **Zahorian SA, Jargharghi AJ:** Color display of vowel spectra as a speech training aid for the deaf. *J Acoust Soc Am* 77:S97, 1985.
53. **Zahorian SA, Rothenberg M:** Principal-components analysis for low-redundancy encoding of speech spectra. *J Acoust Soc Am* 69:832-845, 1981.