Comparison of Two Computerized Speech Training Systems: *SpeechViewer* and *ISTRA*

John Ryalls
Ecole d'orthophonie et d'audiologie
Université de Montréal

Résumé:

Deux systèmes d'intervention informatisée (Aide pour l'entraînement de la parole ISTRA de l'Université Indiana et SpeechViewer d'IBM) ont été comparés. On a vu que chacun de ces systèmes est approprié pour une population spécifique: plus jeune ou plus atteinte dans le cas de SpeechViewer et capable de lire de courts mots dans le cas d'ISTRA. Pour cette raison, ces deux systèmes sont plus complementaires que compétitifs.

Since the 1950's there have been a series of attempts to use machines to provide a visual substitute for the auditory feedback that is deficient or lacking in the hearing-impaired speaker. Overviews of such speech training systems can be found in Pickett (1972), Levitt (1972), Strong (1975), Braeges and Houde (1983), and Bernstein, Goldstein, and Mahshie (1988). Speech training systems developed before the late 1970's tended to use spectrographic-type displays; systems since that time tend to make increasing use of computer-graphic displays that are more inviting to children but not directly related to speech characteristics. Decreasing costs of micro-computers coupled with the development of software appropriate for speech training promise an increasing computerization of the speech-language pathology clinic throughout the 1990's.

Although the speech-language pathologist can never be replaced by the computer, there are a number of ways in which computer-based speech training may complement human intervention, ways that are advantageous to both client and clinician (see, for example, Mahshie, Vari-Alquist, Waddy-Smith, & Bernstein, 1988; Watson, Reed, & Kewley-Port, 1989). One problem that presently exists with computer-assistance in the speech-language pathology clinic, however, is the dearth of information available on the clinical effectiveness of such computer-based speech training systems. While we have not yet been able to conduct such tests of clinical effectiveness, we have been able to evaluate two systems at the Université de Montreal in the laboratory setting. In the interest of increasing information in this area, we will present results of our evaluation, although the comments that follow are essentially laboratory statements and may not apply directly to clinical situations.

The two computerized speech training systems that will be compared here are: (1) IBM's SpeechViewer and (2) Indiana University's ISTRA (Indiana Speech Training Aid). Although there are certainly other interesting speech training systems, notably one developed jointly at Johns Hopkins University and Gallaudet University (Mahshie et al., 1988), these two were selected for review here for their performance, cost effectiveness, and availability. SpeechViewer was made available through IBM of Canada in Montreal, and the author was able to discuss SpeechViewer's clinical application for hearing-impaired children with Muriel Mischook of McGill University's Project for Hearing-Impaired Children. We were able to obtain ISTRA software by virtue of being a Beta-test training site, and purchased the necessary speech recognition board with our own funds. The author also had been able to observe the clinical use of ISTRA previously in the Speech and Hearing Clinic at Indiana University. Both of these computer-based speech training installations offer systematic clinically-based speech training through game-like drill programs. First, a short description of each will be presented, followed by a comparison of the two.

SpeechViewer

IBM launched SpeechViewer as a commercially available product in their "Independence Series" in North America in the fall of 1988. (A French-language version also will be available in Québec at the time of this article's publication.) SpeechViewer uses its own special IBM A/D (analog/digital) card and runs on the PS/2 (Personal System 2). The SpeechViewer brochure states that:

> This clinical tool is based on ten years of research at the IBM France Scientific Center and the help of clinicians who have used prototypes in 29 countries on 6 continents. A group of 12 clinical modules is designed to complement your most successful therapy methods. The diversity of modules makes SpeechViewer appropriate for all age ranges and for many speech disor-

ders. It will help you present well-defined speech stimuli, help your clients make progress with each response, and provide your clients with motivational feedback that will hold their interest.

Children receive colorful and captivating visual feedback on such important speech production characteristics as: (1) voiced/voiceless contrasts; (2) fundamental frequency; (3) vocal attack; (4) vocal intensity; and (5) vowel quality. In one display, for example, if the child's vowel production is good, an animated monkey climbs up the coconut tree and pushes a coconut down, which then falls with an appropriate whistling sound. In another game, the child guides a marker through a maze by means of sustained production of a particular vowel, for each of the four possible directions.

Clinicians can present visual models for their clients to imitate and some basic speech measures, such as fundamental frequency and intensity, can be effected in a straight forward and user friendly manner.

McGill University's Project for Hearing-Impaired Children has found a positive response for SpeechViewer from the young hearing-impaired children that they see in clinic (Muriel Mischook, personal communication). A review of Speech-Viewer can be found in a previous issue of this journal (Thomas-Stonell, 1989). Some additional information on clinical applications of SpeechViewer can be found in Cauvin, Matteodo-Peyracchia, and Maulet (1988) and Destombes (1987). IBM (1989) has published an internal report of the results of four test sites' use of *Vocalization*—SpeechViewer's experimental prototype. Technical information on Speech-Viewer prototype hardware and software is presented in Crepy, Denoix, Destombes, Rouique, and Tubach (1983).

ISTRA

Indiana University's ISTRA system builds on earlier work conducted at Boys Town Institute for Communication Disorders in Children (Osberger, Lippman, Moeller, & Kroese, 1981; Kewley-Port & Watson, 1987). The system runs on a standard IBM PC/XT/AT computer using a commercial voice recognition board built by International Voice Products of Tustin, California. A hard disk, although not necessary, is a very useful accessory. ISTRA is described by its developers in the following manner:

> Ongoing clinical trials of the ISTRA system have demonstrated effective improvement in speech production. The theoretical approach is first to form templates from a child's current best productions of a word and then to use the score generated by matching new utterances to these templates as feedback to indicate the goodness of articulation (Kewley-Port, Watson, Maki, & Reed, 1987; p. 372).

ISTRA provides a game interface to computerized speech quality judgments. Such a system is ideally suited for repeated drill on particular words and phonemes. The speaker produces utterances into a microphone which are then compared to stored acoustic templates. Feedback to the speaker regarding success in replicating good speech quality characteristics is provided in terms of progress with a given computer game. The child interacts with the machine through a series of drills in which the feedback indicates the quality of each prompted word that is spoken in game formats, such as Target, Baseball, and Moonride. For example, in Baseball if the child's production matches very well with his own stored best productions, the child sees a player run completely around the baseball diamond, the word "Homerun" appears on the screen, and his score is increased. If the production is poor, the child receives a strike. The criterion for success in these games is automatically adjusted by the program according to the speaker's continued success at matching speech targets. At the Indiana University clinic where ISTRA is presently undergoing clinical evaluation, hearing-impaired children of at least nine years of age can typically undertake drill work by themselves once training words have been selected and templates formed with the speech-language pathologist.

ISTRA has undergone clinical testing in single-subject design experiments at Indiana University, and preliminary results support the success of ISTRA in improving speech production (Kewley-Port et al., 1987; Watson, Reed, Kewley-Port, & Maki, 1989; Kewley-Port, Watson, Elbert, Maki, & Reed, 1989).

Comparison

While many of the principles of the two systems are basically the same, such as the provision of training on speaker-dependent templates, the systems are still rather different. In fact, they are probably best conceived of as systems that aim at somewhat different clinical populations. SpeechViewer may be more appropriate for the younger and/or more severely impaired speaker. It is intended to create an interest in the sound environment, to train voice control, and improve vowel quality. Training for consonant phonemes is limited to the voicing distinction, as is training of whole word utterances. It is at the level of whole word utterances that ISTRA training really begins, although it also can be used to train isolated vowel utterances. Since some current speech training methods for the hearing-impaired (Ling, 1976) train voice and vowel sounds earlier than consonants, syllables, and whole words, SpeechViewer and ISTRA together can be viewed as providing a complementary continuum of training all the way from voice training to perfecting the pronunciation of whole words and rudimentary sentences.

ISTRA games are somewhat more sophisticated than those for SpeechViewer in that difficulty levels are automatically adjusted. ISTRA drill games thus are probably more appropriate for the older child. Since ISTRA employs written prompts, there is a presupposition that the child can at least read simple words, although it is perfectly conceivable that the clinician offer the child spoken models.

SpeechViewer's games do not automatically adjust according to children's performance, but the difficulty criterion can be manually adjusted by the clinician. This renders SpeechViewer a little less interactive with the speaker. It also should be noted that ISTRA's speech recognition card appears to offer tighter acoustic criteria than the card used by IBM. For example, SpeechViewer seems to have some trouble with the /i, y/ vowel distinction in French at least with the multi-speaker example vowel templates provided by IBM.

SpeechViewer also provides some basic acoustic speech measures such as fundamental frequency, duration, and amplitude that are not obtainable with ISTRA; however, the spectral measures provided by SpeechViewer are not sufficient for most research purposes. The acoustic measures accessible with ISTRA appear in a speech coded form (dubbed an "Istragram" by the Indiana team), and since they do not directly reflect standard speech measures, they are not useful to most speech-language pathologists.

Both of these systems, although obviously developed with the hearing-impaired child in mind, could be applied to other clinical populations, such as the adult neurologically-impaired speaker. It should be pointed out, however, that some of SpeechViewer's graphics (such as the clown face that demonstrates amplitude and voicing) may appear a little juvenile to the adult. ISTRA's games also may require increased complexity to sufficiently challenge the mildly impaired adult speaker. Both systems also might be adapted for pronunciation drills with the second language learner.

SpeechViewer employs a handheld microphone that is conveniently passed between clinician and patient; however, it may be difficult to keep the mouth-to-microphone distance constant in the case of young children. ISTRA uses a headset that keeps the mouth-to-microphone distance standard and is more comfortable than holding a microphone for any length of time. The microphone is a little less intrusive in this manner.

One advantage that ISTRA has over SpeechViewer is that it automatically keeps clients' records and progress information as part of the manner in which the clinician interacts with the software. This aspect of clinical practice is ignored by the present version of SpeechViewer. But SpeechViewer has the advantage of offering basic speech measures, such as fundamental frequency and amplitude; in this respect SpeechViewer functions very much like a Visi-pitch, but has the advantage that measures are obtainable in a much more user-friendly environment.

At the present time ISTRA has already undergone some controlled clinical evaluations, the result of which are published in Watson, et al., (1989) and Kewley-Port et al., (1987). Clinical evaluation of SpeechViewer, by comparison, remains descriptive.

Although there is no substitute for actual clinical use of these computer systems, some of the basic characteristics of each of these systems have been compared in this article. Controlled clinical tests designed to demonstrate the effectiveness of each of these products are necessary. It is inconvenient that, for the time being at least, each of these systems uses a different graphics system that prevents implementation of both on the same AT machine, which many clinicians already possess. This is lamentable since these two systems are actually more complementary than competitive. While each system can still be improved in some obvious ways, both still receive a very high laboratory rating for their likely clinical utility. Computers hold a promising future for speech-language therapy and may represent one means for meeting the demand for increasing speech-language services.

Address all correspondence to:
John Ryalls
Ecole d'othophonie et d'audiologie
Faculté de médecine
Université de Montréal
C.P. 6128, Succ. A
Montréal, Québec H3C 3J7

References:

Braeges, J., & Houde, R. (1982). Use of speech training aids. In G. Walter & R. Whitehead (Eds.), *Deafness and Communication: Assessment and Training*. Baltimore: Williams & Wilkins.

Bernstein, L., Goldstein, M., & Mahshie, J. (1988). Speech training aids for hearing-impaired individuals: I. Overview and aims. *Journal of Rehabilitation Research and Development, 25* (*4*), 53-62.

Cauvin, M.-C., Matteodo-Pey racchia, L., & Maulet, M. (1988). Voix et parole, quelle aide apporte l'informatique? *Audiologia Italiana, 5(3)*, 200- 208.

Crepy, H., Denoix, B., Destombes, F., Rouique, G., & Tubach, J.-P. (1983). Speech processing on a personal computer to help deaf children. In R.A. Mason (Ed.), *Information Processing 83: Proceedings of the IFIP 9th World Computer Congress*, Paris, France.

Destombes, F. (Ed.) (1987). Réunion des utilisateurs des programmes IBM d'aide ! l'education de la parole. Compagnie IBM France, Etude F.120.

IBM Personal System/2 SpeechViewer (1989). Four case studies. IBM Corporation internal report.

Kewley-Port, D., Watson, C., Maki, D., & Reed, D. (1987). Speaker-dependent speech recognition as the basis for a speech training aid. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing* (372-375). Dallas, Texas.

Kewley-Port, D., Watson, C., Elbert, M., Maki, D., & Reed, D. (1989). The Indiana Speech Training Aid (ISTRA) II: Training curriculum and selected case studies. Unpublished manuscript.

Ling, D. (1976). *Speech and the hearing-impaired child: Theory and practice.* Washington D.C.: A. G. Bell Association for the Deaf.

Mahshie, J., Vari-Alquist, D., Waddy-Smith, B., & Bernstein, L. (1988). Speech Training aids for hearing-impaired individuals: III. Preliminary observations in the clinic and children's homes. *Journal of Rehabilitation Research and Development, 25 (4)*, 69- 82.

Osberger, M.J., Lippman, R., Moeller, M., & Kroese, J. (1981). Computer-assisted speech training for the hearing impaired. *Journal of Academic and Rehabilitative Audiology, 14*, 145- 158.

Thomas-Stonell, N. (1989). SpeechViewer review. *Journal of Speech-Language Pathology and Audiology, 13*, 59-60.

Watson, C., Reed, D., Kewley-Port, D. & Maki, D. (In press). The Indiana Speech Training Aid (ISTRA) I: Comparisons between human and computer-based evaluation of speech quality. *Journal of Speech and Hearing Research.*

Watson, C., & Kewley-Port, D. (1989). Computer-based Speech Training (CBST): Current status and prospects for the future. Unpublished manuscript.

Erratum

In the June issue of the *Journal of Speech-Language Pathology and Audiology* (Volume *13*, Number 2) an error was made in the article entitled, "Current Issues in Probe Tube Measurements," by Marshall Chasin. Figure 1 and 2 were inadvertently switched. The legend for Figure 1 applies to Figure 2, and likewise, the legend for Figure 2 applies to Figure 1. The Editor regrets the confusion this error may have caused to readers.