

CLINICAL APPLICATIONS OF COMPUTER-BASED SPEECH TRAINING FOR CHILDREN WITH HEARING IMPAIRMENT

Anne-Marie Öster

Department of Speech, Music and Hearing,
KTH, Stockholm, Sweden
email: annemarie@speech.kth.se

ABSTRACT

Computer-based visual speech training has become widely used both within medical and pedagogical rehabilitation in Sweden. When learning speech motor ability clear instruction is of great importance in order for the learner to realise what is deviant and what is correct in his or her pattern of behaviour. The correct behaviour should then be established and automated through extensive training for it to be transferred to untrained situations. This paper discusses teaching strategies and reports positive training results obtained through the visually contrastive feedback that this modern technical speech training aid offers.

1. TECHNICAL AIDS FOR SPEECH TRAINING

Traditional speech training with profoundly hearing-impaired children is based on methods that help the children to learn speech by looking at the therapist's face and lips, through residual hearing or by feeling the therapist's face, throat and expiration air, etc. to establish an oro-sensory motor control of their speech movements. Efficient speech training should offer the children possibility to perceive invisible speech articulation as well as to imitate and compare their vocal output with that of the therapist. Many training aids have been developed where the information of speech elements is given through visual or tactile feedback to replace the auditory feedback used by hearing children during the development of the speech faculty. Speech was presented through vibrators at the fingertips and other parts of the body or as a visual signal to indicate various elements of speech. In early systems correct behaviour was indicated by a special lamp or by the deflection of a meter needle. Some systems used oscilloscopes to reveal more detailed information of the acoustic patterns of speech. Unfortunately, this type of technical aids was never evaluated within a pedagogical programme. Despite the fact that many of these aids have reported to improve the speech of some children, the use of them was limited. This was probably due to the fact that the visual feedback provided by this type of technical speech training aid was difficult to understand, unnatural, delayed, unattractive, and had no motivational impact on the children.

A significant contribution was made by Nickerson & Stevens (1976) when they developed the first computer-based speech

training system. Since then, the development of the micro-computer technique has offered more advanced computer-aided speech training programs that have enhanced the possibility for profoundly hearing-impaired children to develop intelligible speech (Youdelman & Levitt, 1991; Arends, 1993; Javkin, 1994). In several pre-schools and schools for deaf and hard-of-hearing children in Sweden, computer-based visual speech training has become a standard and valuable complement to the regular speech training activities. This is in part due to an ongoing collaboration between the therapists and the Hearing Technology Group at KTH. The system used is IBM's SpeechViewer II. The therapists report that the system is an objective and effective tool in assisting them in increasing the intelligibility of the children's speech as it gives them better possibilities to co-operate with the children. The immediate and relevant visual feedback makes it easier for the therapist to instruct the child and to explain what is wrong and what is correct in his or her production.

1.1. Important Demands on a Visual Speech Training Aid

Our results and experiences from computer-based visual speech training with profoundly hearing-impaired children of various ages have shown that in order to be efficient and enhance the possibility for a child to develop intelligible speech, a visual speech training aid has a number of important requirements.

- Clear instructions and pedagogical manuals must be created and made available for use with different groups of children.
- The visual feedback of the child's voice and articulation should be shown immediately and without delay.
- The aid must be acceptable to the therapist as well as to the child, which means that the aid must be attractive, interesting, easily comprehensible, easy to handle, and motivating.
- The visual pattern must be natural, logical and easily understandable. This means that training parameters as, e.g., pitch could be shown vertically as pitch variations occur; intensity through the size of an object that becomes larger as a sound becomes louder and smaller as a sound becomes softer; intonation and stress

through a continuous red curve; duration could be shown horizontally and voicing through a relationship between voicing and the change of a colour.

- The aid should provide a contrastive training, that is, the correct model of the therapist and the deviant production of the child are shown simultaneously and compared with each other.
- The aid should provide a flexible, individual, and structural speech and voice training and give an objective evaluation of the child's training results.

2. DESCRIPTION OF SPEECHVIEWER II

The system consists of a PC with a colour display. A microphone, an amplifier, and a speaker connected to a M-ACPA card allow the user to input, store and analyse speech and then display it and play it back. The software contains 15 interactive programs aimed at assisting a child in achieving awareness and control over various speech attributes such as voicing, timing, pitch, and loudness as well as refined articulation and prosody. Feedback is given immediately through a variety of graphical designs and game-like strategies synchronised with optional auditory playback. The programs are grouped into three sections for different areas of use: *awareness*, *skill building*, and *speech patterning*. The displays for basic awareness illustrate selected dimensions of speech such as pitch, loudness, timing, and the presence or absence of voicing. The skill building modules provide a game-like strategy to strengthen ability in pitch, timing, voicing, and speech sound production. Phonology exercises aim to develop the accurate production of one or more phonemes of a specific language. Whole words, short phrases, and single sounds produced by the child are matched against models by comparing spectra of the targets with the child's productions. The speech patterning programs display the speech signal as oscillograms, spectra and spectrograms for analysis and training of refined articulation and prosody. Throughout the system the *clinical management* allows the creation of sustained phoneme and speech segment models and client profiles as well as the reporting and management of client data.

With regard to inter-speaker variability problems, dialects, and deviant speech due to hearing disorders, a child's own "best production" should be stored as a target to establish a new and more correct behaviour through frequent training. The goal is accomplished when the child's best production becomes his or her most common production. This strategy is also used with the CISTA-Aid training (Youdelman, 1994).

3. IMPORTANT GUIDELINES FOR COMPUTER-BASED SPEECH TRAINING

An individual and efficient speech training must follow certain general guidelines, irrespective of age or the child's need for

speech training. Based on our work with SpeechViewer II we recommend the following training strategy.

3.1. Diagnosis of Individual Deviations as a Basis for Speech Training

The first stage involves an assessment of the deviations that should be corrected in order to increase the intelligibility of a child's speech. Many profoundly hearing-impaired children articulate several speech sounds correctly in isolation but have problems in producing them contrastively in various contexts. It is important to identify the unique characteristics of each child's phonological system through a phonological assessment. Valuable information will be provided about how successfully the child can produce speech contrasts in spoken language and the ways the phonetic elements used differ from those used in normal speech. Principles for an assessment of the phonological system of profoundly deaf children as a basis for speech training have been described by Öster (1995a).

Therapy and assessment are inseparable. All assessment is time-consuming and laborious. However, to prevent speech training from becoming a series of meaningless "articulatory gymnastic" sessions, the therapist must be prepared to spend time carrying out an appropriate diagnosis prior to therapy. Therapy and assessment are inseparable as assessment is required regularly during training. This approach is obvious in "*The Structural Speech Training Program*", which was developed within the project with hearing-impaired pre-school children. The program is designed as a protocol and is used with children who possess some verbal skills. Training is done on six levels starting with isolated speech sounds followed by repeated syllables where the consonant is chosen on the basis of visibility, repeated syllables with different stress patterns, alternated syllables for automatization of prosody, words containing the speech sounds of current interest for relevant training, and finally a short phrase containing the topical words. The child's mastery of breath control, intensity, pitch, duration, and voice quality on all levels is listed by the therapist according to the statistics provided by the system. The aim of this work includes following the phonological development in profoundly hearing-impaired children who rely on visual information to learn to speak.

3.2. Instructions That Aim at Awareness, Realisation, and Understanding

By means of the visual feedback given in the Speech Patterning Module "Pitch and Loudness", it is easy to make the child aware of deviant production and in what way it deviates and show him or her how to produce it correctly. The program displays pitch and loudness either individually or combined over time. A split screen offers an exact comparison of a child's production with a model of the therapist's correct production. With children who have some verbal skills the most effective speech material has shown to be minimal word pairs that contain that contrast which the child produces deviantly.

The voicing contrast between, e.g., /g/ and /k/ is clearly indicated with different colours. Voiced sounds are red and voiceless sounds are green. In the lower portion of Figure 1, the child's failure to produce a voicing contrast between final velar stops can be seen.

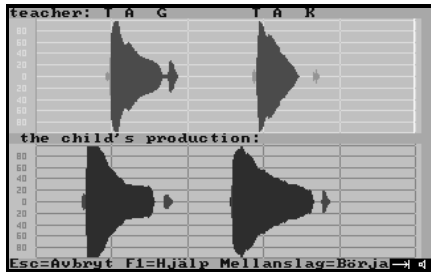


Figure 1: The upper portion shows the therapist's model of correctly produced voicing contrast. The lower portion shows an attempt by the child.

Discrimination between, e.g., long and short vowel quantity, which is an important phonological contrast in Swedish, is visible through a difference in duration of the red colour, signifying voiced vowels. Figure 2 shows the therapist's correct pattern of the minimal word pair /haka/ (chin) and /hacka/ (chop) in the upper portion, and a profoundly hearing-impaired child's production in the lower portion. The display shows that the child only mastered the long quantity of the vowel /a/.

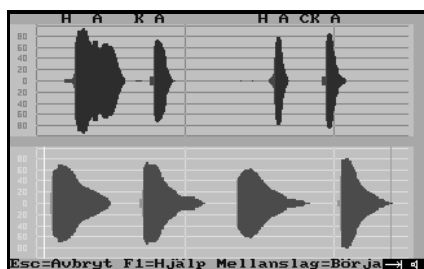


Figure 2: In the upper portion, the therapist's pattern of the word pairs /haka/ (chin) and /hacka/ (chop) is shown. In the lower portion, a prelingually hearing-impaired child's production is shown.

Figure 3 shows a training session of the stress pattern in the two homographs of the word /banan/ (track) and (banana), which are produced with different stress patterns. Pitch and loudness are combined over time in the program "Pitch and Loudness". The stressed syllable is visible in the therapist's model in the upper portion, through a rising pitch, longer duration, and higher intensity (indicated by thickness of the line). In the lower portion of the screen, the child's deviant stress pattern is shown.

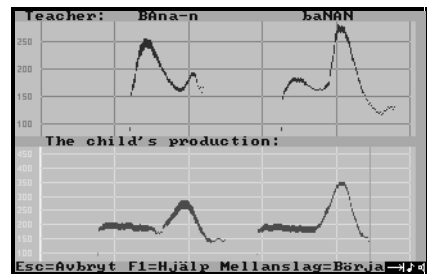


Figure 3: In the upper portion, the therapist's model of correct stress pattern is shown. The lower portion shows the child's production.

3.3. Contrastive Training to Establish Correct Production and to Obtain a New Pattern of Behaviour

When the child has been made aware of his or her deviant production and realises how it deviates, a continuous contrastive training should start, to obtain a new pattern of behaviour. Öster (1989) showed that the visual information helps the child to discover the correct place or manner of the speech sounds, thus allowing him or her to modify a deviant behaviour. In this study two 15-year-old profoundly hearing-impaired boys improved in producing quantity and voicing contrasts between velar stops after contrastive visual training with SpeechViewer I. The contrastive training consisted of minimal word pairs containing the phonological contrasts and short sentences in which the target words were included.

In Figure 4, a training session on the second level of the Structural Speech Training Program described above is shown. In the upper portion, the therapist's model of the repeated syllable /ma/ is recorded. In the lower portion, the imitated production of a profoundly hearing-impaired child is shown. It can be seen that the child produced the nasal /m/ deviantly.

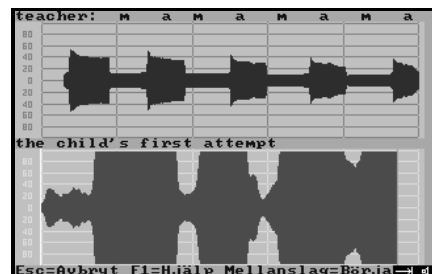


Figure 4: Display of loudness over time in "Pitch and Loudness". In the upper portion, the therapist's pattern of repeated production of the syllable /ma/ is shown. In the lower portion, a profoundly hearing-impaired child's deviantly produced nasal in the repeated production of the syllable is shown.

The instructive feedback provided through the contrastiveness of the visual patterns together with traditional assistance given by the therapist helped the child acquire correct production which is shown in Figure 5.

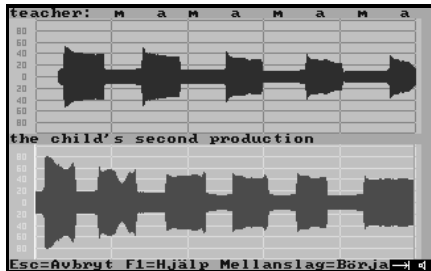


Figure 5: In the upper portion, the therapist's pattern of the repeated production of the syllable /ma/ is shown. In the lower portion, the child's correct production is shown.

Individual differences will determine the particular training material and order selected for training, since the training is based on existing skills. The treatment is influenced by the *Senory-Motoric Approach* devised by McDonald (1964). Bisyllabic drills are used, beginning with small changes of tongue movement between the production of the vowel and the consonant followed by progressively larger changes of movement.

The same positive effect was shown in a study by Öster (1995b). A 5-year-old prelingually deaf boy acquired a natural pitch after a contrastive visual voice training with the "Pitch and Loudness" program. Before training, his fundamental frequency was too high and monotonous, around 700 Hz, due to his lack of perception and control of his own voice. The child learned and established a lower, and for his age, natural pitch and gained awareness and control over his voice. In Figure 6, an example of the instructive feedback that aimed at lowering his fundamental frequency is shown. His spontaneous phonation eight weeks after training had an average pitch of 263 Hz.

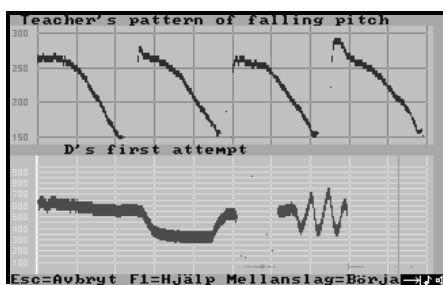


Figure 6: In the upper portion, the therapist's model of repeated lowered pitch is shown and in the lower portion, the child's positive training result is shown.

3.4. Further Training for Maintenance and Transfer to Untrained Material

When the child has acquired the target production, significant amounts of additional training must be invested to get him or her to maintain correct production and transfer skills to untrained situations. This is the most important element in the speech training program but the most difficult to carry out. The target production must be repeated and practised in a variety of contexts. The ultimate goal is a stabilisation of correct production and a system expansion.

4. DISCUSSION

It is important to point out that even the best computer program could never replace the therapist but only assist and facilitate his or her work. Computer-aided speech training is a complement to traditional methods and has a pedagogical value for the therapist who has a good knowledge of articulatory and acoustic phonetics as well as of the computer technique.

5. REFERENCES

1. Arends N. *The Visual Speech Apparatus*, Instituut voor Doven, Sint-Michielsgestel, The Netherlands, 1993.
2. Javkin, H. "A new speech training system: acoustic-articulatory data, video games, and synthesized model parameters", *Proc. of Second Intern. Symp. on Speech and Hearing Sciences, 81-92, Osaka, Japan, 1994*.
3. McDonald, E. *Articulation Testing and Treatment*, Pittsburgh, Stanwix House, 1964.
4. Nickerson, R., Stevens, K. "Teaching speech to the deaf. Can a computer help?", *IEEE Trans. Audio Electro-acoustics, AU-21, 445-455, 1973*.
5. Youdelman, K, Levitt, H. "Speech training of deaf students using a palatographic display", *Proc. of First Intern. Symp. on Speech and Hearing Sciences, 1-11, Osaka, Japan, 1991*.
6. Youdelman, K. "Computer applications in teaching speech to deaf children", *Proc. of Second Intern. Symp. on Speech and Hearing Sciences, 67-79, Japan, 1994*.
7. Öster, A-M. "Applications and experiences of computer-based speech training", *STL-QPSR 1, 59-62, 1989*.
8. Öster, A-M. "Principles for a complete description of the phonological system of deaf children as a basis for speech training", *Profound Deafness and Speech Comm., London: Whurr Publ. Ltd., 441-461, 1995a*.
9. Öster, A-M. "Teaching speech skills to deaf children by computer-based speech training", *Proceedings of 18th International Congress on Education of the Deaf, Tel-Aviv, Israel, 1995b*.