Teaching speech skills to deaf children by computer-based speech training

Öster, A-M.

journal: STL-QPSR
volume: 36
number: 4
year: 1995
pages: 067-075

http://www.speech.kth.se/qpsr
Teaching speech skills to deaf children by computer-based speech training

Anne-Marie Oster

Abstract
Prelingually deaf children very rarely develop speech spontaneously. Speech is almost always developed through training, using the visibility of phonetic features, reading, tactile sensation and, if possible, residual hearing. These inputs severely limit speech perception potential and cause deaf children to make unavoidable articulatory deviations that affect their ability to signal meaning differences in spoken language. Through visual feedback, it is possible for a deaf child to establish an orosensory motor control of his speech. Computer-aided speech training presents meaningful visual feedback of distinctive contrasts that are not visible via speech reading and consequently difficult to produce correctly. IBM's "SpeechViewer II" has been used in the schools for deaf children as well as in the pre-school training of severely hearing-impaired children in Sweden. When it is used as a supplement to traditional speech training methods offered by the therapist, it has been shown to be an objective and effective tool in helping to increase the intelligibility of the speech of deaf children.

Introduction
The hearing child acquires speech and language spontaneously through a combination of hearing and speech reading. Auditory and visual input helps him to imitate, control and compare the vocal output of other speakers with his own production. Children who are born with a severe auditory deficit have no acoustic speech target to imitate and to compare their own production with. Therefore, they develop no spontaneous speech but a built-up speech through vision, tactile sensation and, if possible, residual hearing. They have to rely on the limited visibility of phonetic features in learning oral speech and on orosensory motor control in maintaining speech movements. An efficient speech training with deaf children should offer the children possibilities to perceive invisible speech parameters as well as to imitate and compare their vocal output with that of the teacher. Many training aids have been developed, where a speech element was shown as a visual signal and correct behaviour was indicated by a special light on a lamp or by the deflection of a meter needle. However, this type of technical aids has never been used frequently in the speech clinic. The reasons for the limited use may be that the visual feedback was difficult to understand, not natural, delayed, unattractive and did not motivate the children. Furthermore, due to technical problems, lack of pedagogical manuals and evaluations, the teachers were not motivated to use them as standard procedures.

However, computer-based speech training has capabilities of offering the children immediate and meaningful visual feedback and might make it easier for the teacher to instruct and explain what is wrong and what is correct and through motivational and frequent training establish an intelligible production. In most schools for deaf children in Sweden, computer-based visual speech training has become a standard and valuable complement to the regular speech-training activities. The system used is the IBM's
“SpeechViewer II”, which consists of 15 interactive programs aimed at assisting the child in achieving awareness and control over various speech attributes such as voicing, timing, pitch and loudness, as well as over refined articulation and prosody. The child inputs speech through a microphone and feedback is given immediately by a variety of graphical designs and game-like strategies synchronised with optional auditory playback. The programs are grouped into sections for three areas of use: displays for basic awareness of selected dimensions of speech, such as pitch, loudness, timing, and the presence or absence of voicing; skill-building voice-controlled games and exercises for pitch, timing and vowel production; and speech patterning with graphical presentation of the speech signal for analysis and training of refined articulation and prosody.

Results of computer-based speech training with severely hard-of-hearing children

The work on SpeechViewer II with deaf children up to 16 years of age has proved to be successful, worthwhile and very efficient, especially in the instruction phase of training. The children have appreciated it and have derived benefit from the use of the system both on the segmental and suprasegmental level. The positive effect of a systematic computer-based training of prosodic contrasts was shown in a study by Öster (1989). Two prelingually deaf children with difficulties in producing certain phonological contrasts in Swedish were trained systematically with the system during eight weeks in order to evaluate its efficiency.

One of the children (child I) was a 15-year-old deaf boy, with some residual hearing in the low frequencies. He had difficulties in producing quantity differences between phonologically long and short vowels. This is an important contrast in Swedish which is realised as a difference in duration and, for some pairs, also with a difference in vowel quality. Furthermore, the following consonant has opposite quantity, i.e., long vowels are followed by short consonants and vice versa. The child did only control the phonological contrast between long and short /o/ before training. His realisations of short /æ/ and /u/ were always produced long. In some cases he did not control the vowel quality. It was not the intention to train vowel quality in this study but, in some cases, the pronunciation improved or became more stable after the duration training. The other child (child II) was a 13-year-old girl, who had difficulties in producing distinctions between voiced and voiceless velar stops. Her hearing threshold was within the region of vibration.

The children were video recorded before and after training when they read minimal word pairs containing the phonological contrasts of vowel-length and voicing. Short sentences in which the target words were included were also recorded to study the pronunciation of the words in isolation compared to the pronunciation of the words in running speech. Untrained word pairs which contained the same contrasts were also recorded to study the generalisation effects. Narrow phonetic transcriptions were made of the video recordings using the symbols of the International Phonetic Alphabet and some diacritical marks which have been developed for the transcription of babbling and phonetic development in early infancy.
The children were trained with the speech patterning program “Pitch and Loudness” for about ten minutes twice a week during eight weeks. This is a graphical presentation of the speech signal where the voiced/voiceless contrast is clearly indicated by different colours. Voiced sounds are red and voiceless sounds are green. Discrimination between long and short vowels is visible through the differences in duration of the red colour that indicates voiced vowels (Figure 1).

Figure 1. Display of the speech patterning program “Pitch and Loudness”. In the upper portion, the teacher’s pattern of the word pairs /haka/ (chm) and /hacka/ (chop) is shown and in the lower portion the child’s production is shown. It can be seen that the child only mastered the long quantity of the vowel /a/.

It was easy to explain to the children what was deviant in their production by comparing their speech with the speech of the therapist on a split screen display. Figure 2 shows the subjective assessment of correctly produced quantity of long and short vowels by child I before and after training. The result shows, in spite of the limited amount of data, that the child learned to produce the short versions correctly after training. Short vowels improved more than long vowels. An improvement in the production of durational contrast between vowels after systematic training was found in all vowels except for /y/. An improvement was also found in running speech and in untrained words similar to that found in trained words.

Child II also improved in producing voicing contrasts with the help of the computer training. Before training, /g/ was pronounced without voicing in medial position. In final position, /g/ was omitted. However, after training, /g/ was pronounced as a voiced velar stop in both medial and final positions. This child also produced the differences in voicing between velar stops when she read the words in isolation as well as when she read the words in running speech. As for child I the improvement in trained words was also found in untrained words.
Correctly perceived quantity of long Swedish vowels

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Before Training</th>
<th>After Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Å</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>A</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>I</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>E</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Y</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Ö</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>U</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Mean</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

n = 2 5 8 8 4 2 4 2

Correctly perceived quantity of Swedish short vowels

<table>
<thead>
<tr>
<th>Vowel</th>
<th>Before Training</th>
<th>After Training</th>
</tr>
</thead>
<tbody>
<tr>
<td>O</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Å</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>A</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>I</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>E</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Y</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Ö</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>U</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Mean</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

n = 2 5 8 8 4 2 4 2

Figure 2. Subjective evaluation of correct duration of Swedish long and short vowels before and after training. The vowels are represented by orthographic symbols.

A structural speech training program to be used with the children who have some verbal skills is in progress. Training is done on six levels: from isolated vowel (V1), single syllable on the basis of visemes (CV1), repeated syllables (CV1CV1) with different stress patterns, alternated syllables (CV1CV2) for automatisation of prosody,
words (CVCV) for a meaningful training and, finally, short phrases made up of the CVCV-words. The program is designed as a protocol where the child's mastery of breath control, intensity, pitch, duration and voice quality on all levels is listed by the teacher according to the data statistics provided by the system, for example, duration of voiced/unvoiced portions, mean, range and standard deviation of pitch and intensity. It is also the aim of this work to try to follow and assess the phonological development of deaf children who are relying on visual information.

**Results of computer-based voice training of a severely hard-of-hearing pre-school child**

Two years ago the system was introduced in the pre-school training of the Danderyd Hospital in Stockholm. The ambition was to take care of early skills during the sensitive period for learning and individually train such abilities as respiration, loudness, pitch and voice quality before the children join the speech clinic of a school for deaf children. Below, experiences and results of computer-based voice training will be reported of a five-year-old, prelingually deaf boy (D), who uses sign language for communication. His pure-tone averages were 78 dB in the right ear and 102 dB in the left ear. His phonation was too high and monotonous around 700 Hz, which he could not perceive or control by himself. This was very disturbing for those closest to him and he was constantly reduced to silence and faced with irritation. His teacher wanted him to learn and establish a natural pitch and a voice to be proud of. During training, D wore his hearing-aids. By way of introduction, it was important to get D to understand that what happened on the screen was dependent on his phonation and pitch.

We started the training with one of the simple pitch-awareness modules to increase awareness of vocal pitch and quantify his pitch range. A helicopter changes position vertically as pitch variations occur and two white marks indicate minimum and maximum pitch attained (Figure 3). D's pitch was high but his range was low: 688-756 Hz.

![Figure 3. Display of the module "Pitch". The helicopter shows D's pitch and range.](image)

---

**Figure 3. Display of the module "Pitch". The helicopter shows D's pitch and range.**
D observed the helicopter on the highest floor but he was unable to lower it. The feedback made him aware of his vocal pitch but it did not instruct him how to lower it. Instead, we used "the Pitch and Loudness Patterning", where time is represented on the horizontal axis and pitch along a vertical frequency scale in Hz. The split screen in Figure 4 displays the teacher's input of a sustained /a:/ with a natural pitch in the upper portion and D's phonation during 3.8 sec with a mean of 604 Hz in the lower portion of the screen.

![Figure 4. Display of the speech patterning program "Pitch and Loudness". In the upper portion, the teacher’s phonation during 4.4 seconds is shown. In the lower portion, D’s sustained spontaneous phonation during 3.8 seconds is shown.](image)

The next step was to get D to vary his pitch. He touched the teacher’s larynx and observed her pitch variation in the upper part of the screen many times over. In the lower part of the screen in Figure 5, the positive result of D is shown during 5 sec.

![Figure 5. Display of the pitch variation training with the speech patterning program "Pitch and Loudness".](image)
The third step was to get D to lower his pitch. Figure 6 shows, in the upper part of the screen, the pattern of the teacher when she repeatedly lowered her pitch at the same time as D touched her larynx. The lower part shows that during the first 3 seconds D's pitch was very high but suddenly it dropped to 327 Hz. He was stunned and went back to the high and varied pitch pattern to get control over his voice. From now on D varied between high and low pitch and at a given sign by the teacher he immediately lowered his pitch.

Figure 6. Display of the training to get D to lower his pitch with the speech patterning program "Pitch and Loudness".

Figure 7 shows a typical training session of repeated phonation with a natural pitch. To vary the training and strengthen his control of pitch D also tried skill-building...

Figure 7. Display of a typical training session of sustained and repeated phonation with a natural pitch.
programs like the one in Figure 8, where he had to produce correct pitch variations to control the vertical movements and sustained voicing to control the horizontal movement of an object toward targets arranged in a curve.

Figure 8. Display of the skill-building program “Pitch”. D monitored the figure with his voice towards the gold pieces by varying his pitch between 275 and 325 Hz.

Figure 9 shows that D learned and established a lower pitch, more natural for his age, and that he got awareness and control over his voice. The upper portion of the screen shows the spontaneous phonation two weeks after D finished training. The average pitch was 266 Hz. The lower portion of the screen shows his spontaneous phonation eight weeks after training. The average pitch was 263 Hz and his voice was a soft, nice voice of a five-year-old little boy.

Figure 9. The upper portion shows D’s spontaneous phonation two weeks after finishing his training and the lower portion shows his spontaneous phonation eight weeks after finishing training.
Conclusion

The work with the SpeechViewer shows that computer-aided speech training may well be used as a valuable expansion of traditional speech training of severely hard-of-hearing children. However, it must be pointed out that even the best computer program could never replace a therapist, only assist in his/her work. Computer-based speech training should be a complement to traditional methods and is a powerful tool for those therapists who in addition to mastering the technique are competent in articulatory and acoustic phonetics.

One explanation of the positive training result is the very instructive and pedagogical information that this program provides in the form of a direct, meaningful, motivational, natural and objective visible feedback. It gets the children to understand what is wrong and what is correct in their production. By means of the speech patterning module, it is easy to show and instruct the child about the different distinctive features that underlie phonological contrasts in language. Training with minimal word pairs, which only differ with respect to one opposition, has proved to be an appropriate and efficient training strategy.

Another very important advantage with computer-based speech training is the objective evaluation of the child's speech. The speech therapist has the difficult role of encouraging and motivating the child at the same time as she/he must criticize and evaluate the child's attempts. The “SpeechViewer” and other visual speech training aids give the therapist and the child better possibilities to cooperate.

The system also gives the speech therapist increased flexibility in training since the task levels and performance criteria offer many choices. Much of traditional training is combined in one piece of equipment, built to display only one or a few speech parameters or features, which makes the selection of the most suitable training easy. Furthermore, efforts and improvements in training can easily be registered and documented since the information on the screen can be printed out or saved in files.

There are some possible drawbacks with computer-based speech training that are worth consideration. The cost of the system is high, and to some extent new training techniques need to be developed. There is also always a risk that the training will be adjusted to the technical aid and not to the child's needs.

References