Computer-aided testing and training of the hearing impaired

Mizuno, C. and Risberg, A.

journal: STL-QPSR
volume: 25
number: 2-3
year: 1984
pages: 109-118

http://www.speech.kth.se/qpsr
II. SPEECH AND HEARING DEFECTS AND AIDS

A. COMPUTER-AIDED TESTING AND TRAINING OF THE HEARING IMPAIRED
Chiharu Mizuno* and Arne Risberg

Abstract
In the evaluation of different strategies for recoding of speech in aids for the deaf it is necessary to compare different solutions in well controlled training experiments. In these situations it is often advantageous to use self-instructed training. A system has been developed where a microcomputer controls a random-access cassette tape recorder. A set of different training and testing programs have been developed. Results are given from some preliminary tests with the equipment. The experience from these experiments are positive. The subjects can select the training material and the type of training they would like to work with. They can work in their own pace and get an immediate feedback. The system will be used for testing and training in our work on tactile recoding of speech and in work with cochlear implants.

Introduction
In developing new technical aids for the deaf, the speech signal is often recoded to a new acoustical signal, to an electrical signal that stimulates the auditory nerve (cochlear implant) or to a signal that is received by the sense of touch (Risberg, 1982). To evaluate the selected recoding strategy, it is necessary to train subjects to perceive speech through the system. This training must be made in a very controlled way to facilitate comparisons between different coding strategies. One attractive possibility in this situation is to use programmed training based on some type of machines. In the earlier work performed at the Dept. of Speech Communication and Music Acoustics on speech recoding aids, two machines of this type were developed (Spens and Risberg, 1968). These machines were based on a standard twin track tape recorder. On one track the stimuli were recorded and on the other a code signal that identified the stimulus. The subject answered on a set of pushbuttons. They could repeat the stimulus if they wanted and they got an immediate feedback of the correct answer.

The experiences from these devices were positive. However, only very simple training programs could be used and the result could only be obtained as the number of correct responses. A much more flexible solution is to use a microcomputer-controlled system (Yoshimoto, 1982; Mizuno, et al., 1981). A system of this type intended for speechreading training has been described by Cronin (1979). The system is based on the use of a slightly modified video tape recorder. On the tape is recorded a time-code that makes it possible to find and play any part of the tape. Audiovisual training is often wanted as most speech recoding aids for the deaf are primarily speechreading aids. In some situations, however, auditory stimulation is sufficient and as a random-access cassette tape recorder recently was put on the market by Tandberg, it was

decided to start developing programs for computerized testing and training. Some preliminary results from this work are presented here.

**Technical system**

Fig. 1 shows a block diagram of the equipment and the experiment situation. The microcomputer used is a conventional 8 bit microcomputer, Luxor ASC800. The tape recorder is a Tandberg TCC 530 (Miles and Moulder, 1984). This tape recorder can be controlled by ASCII-character commands through a series port (RS232C). All the normal functions of the recorder can be controlled by the computer. A command can also be given to play any given part of the tape between two counter settings. The counter has a "basic accuracy of +/-0.5 sec referred to playback speed". The counter is automatically set to zero when the tape is rewound.

Programs that are written in Basic send the start and end addresses of any desired stimulus through a series line. Different training programs have been tried, and the subject chooses the stimulus material and the type of training or testing program he would like to work with through a menu. The signal from the tape recorder is either presented through headphones or to the technical system that is tested. The subject gets the stimuli from the different transducers and gives his/her answers on the keyboard, according to the instructions displayed on the screen. The result is presented as number of repetitions made of a stimulus, percent correct, as a confusion matrix, etc. At present, only a small number of different training and testing programs have been tried.

**Example of programs**

Fig. 2 shows a flow chart of the training and testing programs. Based on this basic flow chart, the following different programs have been developed.

**Program for making training programs**

The stimuli are recorded on a cassette tape. The tape is placed in the TCC 530 recorder and rewound. The counter is then automatically set to zero. Counter readings for the beginning and end of the different stimuli are noted and a list is prepared. To compensate for the limited accuracy of the counter, a margin of about 2 sec is added to these values. When the program is run, it asks for the name of the program, the number of stimuli, the names of the stimuli, and the counter values for the beginning and the end of the different stimuli. These data are read into the program.

**Stimulus selection program**

The maximum number of stimuli in a program is at present 18 but this number can, of course, be increased. It most situations it is, however, not possible to work with so many stimuli, at least not in the beginning of the training. A program has therefore been made which
Fig. 1. Block diagram of the system and the experiments.
Fig 2. Flow chart of training and testing programs.
allows the subject to choose a limited number of stimuli from the given set. He can, based on a previous result, select the stimuli that were most often confused.

**Listening program**

The subject can choose the stimulus he would like to listen to from a list of the stimuli selected by the program above. The subject decides when to stop training. The result is given as the number of times that the different stimuli have been chosen.

**Test program with feedback**

The selected stimuli are presented randomly a selected number of times. The subject can repeat the stimuli before answering and when he has given his answer, he is informed whether he was right by the comment "You are right!". If he is wrong, he is informed about the correct answer and the stimulus is repeated. The result is given as a confusion matrix and the number of repetitions made for each stimulus.

**Test program**

The selected stimuli are presented randomly a selected number of times. No feedback is given. The test result is presented as a confusion matrix and percent correct.

**ABX-program**

The stimuli A and B are presented after each other and then either A or B is randomly presented. The task of the subject is to identify the stimulus as A or B. The test result is given as a list of correct stimuli, the given answer, and percent correct. If the test is repeated, the correctly perceived stimulus pairs are automatically removed.

**AB-program**

Either stimulus A or B is presented. The task of the subject is to identify the stimulus as A or B. The result is given as a list of correct stimuli, a list of the given answers, and percent correct.

**Examples of results**

In the following some examples are presented from results that have been obtained in the preliminary tests.

1. **Confusion of Swedish vowels**

   The equipment was used to demonstrate the effect of a hearing impairment on the speech perception ability at an Open House demonstration at the Dept. of Speech Communication and Music Acoustics. The nine Swedish vowels, lowpass filtered at 1 kHz, were presented to 37 normal hearing Swedes. In this test, each vowel was presented once in a random order. The number of correct answers ranged between two and eight. The
vowel /i/ was often chosen instead of /y/. The highest scores were usually obtained by young people about 20 years old, and also by a trained phonetician. The direct presentation of the confusion matrix after the test run was very illustrative.

2. Training in identifying Swedish vowels and consonants

The nine Swedish vowels and the 18 Swedish consonants were used in a training program for two English-speaking girls, 10 years (subject 1) and 11 years (subject 2) old. Table I shows the results for the nine vowels. In the consonant text, subject 2 got 94% correct answers for the second test session, while subject 1 got 83% for the third test session. As the Swedish vowel system differs quite a lot from the English one, while the two consonant systems are very similar, these results are understandable.

3. Identifying Swedish numerals through a tactile vocoder

A tactile vocoder which has been developed at Hokkaido University (Ifukube et al., 1978; Ifukube, 1982). In this device, the information is transmitted by a matrix of 16 rows of three vibrating needles on a fingertip. The speech signal is analyzed in a 16-channels filter bank (range from 200 Hz to 4 kHz). Each row of needles is excited at a fixed frequency of 200 Hz with an amplitude proportional to the energy extracted from the filters.

The Swedish numerals from zero to nine were selected to measure the ability of three normal hearing people to learn to identify the tactile stimuli. The stimuli where spoken in the same manner and by the same speaker as in the comparison of different tactile systems made by Spens (1980). The subjects selected the type of training sequence they wanted to work with. Usually they started with the listening program and then continued with the test program with feedback. The session was then finished with the test program without feedback. In this program fifty stimuli were randomly presented.

Fig. 3 shows the results in percent correct for the three subjects. The confusion matrix in Fig. 4 shows the result obtained by subject 3 in session 4. This subject has previous experience from this type of tactile display. The numerals /1/ and /7/ were pronounced as monosyllables and the other as bisyllables. The numerals /6/ and /8/ have a very similar acoustic structure with a distinct two-syllabic pattern. The numerals /0,2,3,4,5,9/ are all two-syllabic but the syllabic structure is less marked than for the numerals /6/ and /8/. Theoretically, it ought to be possible to obtain spectral information from the tactile matrix that enables a good distinction between the stimuli. The results do, however, not show this. This is probably due to the very limited training time. One problem is the nature of the tactile display. The finger must always be placed in the same way and with the same pressure and this is difficult to learn. Masking factors might also be an explanation of the poor result.
Table I. The learning process of nine Swedish vowels by two English speaking children.

<table>
<thead>
<tr>
<th>Session nr</th>
<th>Percent correct</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Subject 1</td>
</tr>
<tr>
<td>1</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>44</td>
</tr>
<tr>
<td>3</td>
<td>33</td>
</tr>
<tr>
<td>4</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>44</td>
</tr>
<tr>
<td>6</td>
<td>56</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
</tr>
</tbody>
</table>

Fig. 3. Learning curves for three subjects in an experiment where they tried to learn to identify ten Swedish numerals presented from a tactile vocoder.

Fig. 4. Result for subject S1 from training to identify the Swedish numerals 0-9 presented by a tactile vocoder.
4. Identifying the Swedish numerals by a cochlear implant subject

The same test material as used in the evaluation of the tactual vocoder was used by a patient that recently had obtained a cochlear implant with an extracochlear single channel electrode (Hochmair-Desoyer et al., 1981). The result shown in Fig. 5 was obtained after a very short training. It is apparent that even in this case the gross time pattern is easily identified but the spectral differences between the numerals /0,2,3,4,5,9/ are more difficult to learn. This test was, however, made about a month after the implantation.

5. Identification of the phonemes /s/, /t/, and /n/ by a cochlear implant-subject

In the previous experiment it was noted that the subject had some difficulty in distinguishing between the phonemes /s/, /t/, and /n/. A training tape was therefore prepared that contained the Swedish words /lo:sa/, /lo:ta/, and /lo:na/. Fig. 6 shows the result of the training with this material. It apparent that the subject has no difficulties in identifying /n/ but that the differences between /s/ and /t/ cannot be perceived from the beginning. In session 3 she randomly chooses between the two stimuli and the number of repetitions before given an answer is high. In session 4 she has apparently to some extent learnt the distinction and in the last two sessions she never repeats a stimulus and gets almost all answers correct. At the bottom of the figure the percent correct obtained at each session is shown. In a following test program the test material was increased to nine words with the same basic difference. The test result showed that there was a carry over from the training with the three words.

Discussion

In comparison with the classical method where a therapist is involved in presenting the stimuli and checking the answers, the described system gives the advantage of being automatic and of allowing the subject to work in his own pace and to listen to the stimulus as many times as he likes. Moreover, he gets an immediate feedback for each stimulus and at the end of the test session he gets his total score. Thus, it is possible to apply this system in the field of rehabilitation and education as well as in research. Young people do not seem to have anything against computers and they are interested in this type of training. On the other hand, the use of computer-assisted training in hospitals, with older subjects, will probably create some problems. Since teachers and therapists as a rule do not have time to learn to program a computer, it is necessary to develop self-instructing programs, that make it possible for them to program the computer with the training material they would like to use. This problem should be solved by a close contact between programmers and users of the programs. A program of this type has been written.
Fig. 5. Identification of the Swedish numerals 0-9 by a subject who used a single channel cochlear implant.

Fig. 6. Confusion matrices and learning curve for the training of the distinction between the phonemes /s/, /t/, and /n/ by a cochlear implant subject.
In the present equipment the type of programs that can be developed are somewhat limited by the access time of the tape recorder. This problem can be solved by the use of an audio-disk recorder. Work is in progress to use a video recorder in the same system. This will allow us to develop training and testing programs that include lipreading. Even here a video-disk system will be the final solution.

Acknowledgments

This work has been made possible by a scholarship given to Chiharu Mizuno by the Japanese Ministry of Education, Science and Culture, who is very grateful for the priceless opportunity to work during ten months in Sweden. Thanks are also due to everybody at the Dept. of Speech Communication & Music Acoustics for their heartily friendship and sincere cooperation.

References


