TOUCHING VOICES WITH THE TACTILATOR*

Gunilla Öhngren

Abstract
The purpose of this study is to investigate whether the Tactilator can give vibrational support to speechreading and - if so - what cognitive abilities are needed to take advantage of this additional information. The Tactilator is a single-channel tactile aid. The talker holds a contact-microphone against the throat, and the deafened adult has a bone-conductor in his hand. The vibrations from the talker's larynx are transmitted wirelessly to the bone-conductor. Sixteen deafened adults were tested with speech tracking live, two sentence-based video-tests, and a word decoding video-test. The subjects were tested with or without the Tactilator in all conditions. The Tactilator-supported speechreading proved to be generally superior to speechreading only. Cognitive testing revealed that using the Tactilator was best predicted by a text-based phonological recoding test, a lipped mono-syllabic word decoding test, and a complex working memory test (R=.81; Adj.R=.57). The implications for solving the signal-to-noise ratio problem, by the use of a contact-microphone both for tactile-aid and hearing-aid users, is discussed.

INTRODUCTION
The deafened adult has lost the ability to perceive speech in the "normal" way, that is through the auditory channel. The deafened adult, therefore, has to rely mostly on information given by the visual channel, which means that s/he has to perceive speech by visual speechreading. Visual speechreading is a very difficult task as only 10-25% of speech sounds are visible on the lips (Woodward & Barber, 1960; Dodd, 1977; Jeffers & Barley, 1977 for a review; Amcoff, 1970 for Swedish stimuli). A study by Risberg & Lubker (1978) also indicated that many prosodic elements such as syllabic stress, juncture etc. are difficult to speechread.

Research has shown that three basic cognitive components characterise a skilled speechreader: visual word decoding, information-processing speed, and verbal inference-making (De Filippo, 1982; Gailey, 1987; Lyxell, 1989; Rönnberg, 1990). Unless the visual speechreader is extremely skilled in these abilities, speechreading for the deafened adult must be supplemented with additional information: This can be through the tactile sense by vibrators, or electrotactile stimulation; by electrical stimulation of the cochlea via a cochlear implant, or by additional visual information such as the Mouth-Hand-System /MHS/ (Borrild, 1972), Cued speech (Cornett, 1972) or by means of "Tecken Som Stöd" /TSS/ (Sign as support to speechreading), where signs support speechreading (Wolf, 1991).

Early attempts were made to help the deaf by substituting auditory information with tactual information. Gault (1924) originally built a speaking tube which transmitted the voice of the talker as air column vibrations and the deaf subject held the palm of his/her hand against the tube. The deaf subjects in the study made sufficient progress to encourage Gault to develop an electronic system called the Teletactor. The Teletactor was initially a single-channel, vibrotactile aid which was later

---

*This article is a part of my dissertation Touching Voices: Components of Direct Tactually Supported Speechreading; to be presented at the University of Uppsala, Nov. 6, 1992.
changed to a multichannel vibrotactile array. Although the results of this work were encouraging Gault’s program of research was discontinued, but the Teletactor was used extensively in US schools for deaf children in the 1930’s (Goodfellow, 1934). Work on multichannel tactile aids was continued by Wiener, & al. (1949), by Pickett (1963) and Pickett & Pickett (1963).

Over the last ten years cochlear implants have gradually emerged as a viable technical aid for deafened adults. The results obtained with cochlear implants with many deafened adults are very promising (Christman & Albert, in press). However, large individual variations in the obtained support is acquired regardless of the type of implant used. Some deafened will benefit very well, others rather moderately and still others will have no advantage at all of the implant. At present there is not possible to predict the outcome of the operation with any certainty from any preoperative testing. There are also some deafened adults who can not benefit from an implant due to medical contraindications or for personal reasons such as a refusal to have surgical intervention.

Interest in research into tactile aids has therefore been revived over the past decade. There are results which indicate that speech perception can be improved if speechreading is combined with various tactile devices (Boothroyd & Hnath-Chisholm, 1988; Plant, 1988; Weisenberger & Russel, 1989; Weisenberger, Craig, & Abbott, 1991). However, it appears that beneficial effects are only obtained after relatively long training per tools. Positive results have also been reported from three long-term users of tactile aids (Cholewiak & Sherrick, 1986; Plant, & al., 1984; Plant & Spens, 1986).

In a comparison of five commercially available tactile aids, Plant (1989) found that single-channel devices transmitted suprasegmental contrasts best, whereas a multichannel electro-tactile display performed best for segmental contrasts. Sherrick (in press) has compared performance with tactile aids under realistic conditions. He found that, at present, it appears that the most effective systems are multichannel vibrators laying out the speech signal in spatial patterns, representing frequency or voice formant relations. Lyxell, Rönberg, Andersson, & Linderoth (in press) examined the initial effect of using a single-channel vibro-tactile aid, – the Minivib (Spens & Plant, 1983). The results obtained, indicated no direct improvement in speechreading ability, when the aid supplemented the visual signal across a variety of speech perceptions tests. Thus, there are at present no results indicating that there is a tactile aid available, which – when combined with speechreading, is able to directly; that is without practice, replace auditory information when perceiving speech. An additional problem of currently available tactile aids, in common with hearing aids, is that they are all sensitive to background noise.

Some deaf-blind people use a special technique called the Tadoma method (Reed, et al., 1985). By this method the deaf-blind person perceives speech by placing his/her hand on the talker’s face, and monitors the facial activities associated with speech production. Typically, the Tadoma user places his/her thumb lightly on the lips of the talker, while the fingers fan out over the face and neck. Performance of Tadoma experts is roughly equivalent to that of normal hearing people listening in signal-to-noise ratios of 0-6 dB. For sighted subjects with impaired hearing, Tadoma presents a problem. Having the fingers spread out over the talker’s mouth will obviously interfere with the ability to speechread.
Viewing the problem of visual-tactile speech perception from the perspective of skill acquisition, there are examples in the literature of deaf people with a remarkable ability to understand speech via tactual information. Katz (1925/1989; p. 189) tells of Willetta Huggins, who could understand speech, when she placed her fingertips on the larynx, the thorax, or the head of the speaker. She was even able to understand speech when the vibrations were transmitted via a billiard cue-stick or a sheet of paper.

Plant & Spens (1986) describe a Swedish man, who after becoming profoundly deaf at the age of eight years learned to use the vibrations from the speaker’s larynx, as a supplement to speechreading. This method which involves placing his hand on the speaker’s throat is known as Tactiling (Öhngren, Rönberg, & Lyxell, 1992). The method gave the subject enough information enabling him to continue his education among hearing classmates to tertiary level (Rönberg, in press). This communication method, using a tactile supplement to speechreading makes it possible for him to communicate at a nearly normal speaking rate (Plant & Spens, 1986; Öhngren, Kassling, Risberg, & Söderlund, 1991). In a generalization experiment it was shown (Öhngren, Rönberg, & Lyxell, 1992) that other deafened adults could also make use of tactile information by placing their hands on a speaker’s throat. However, the possibility of social restraints when communicating with unfamiliar persons makes the method somewhat limited.

A technical device has been built to replace the deafened adult’s hand on the speaker’s throat. It consists of a contact-microphone which the speaker holds against the throat, an amplifier, and a bone-conductor, which the deafened adult holds in his/her hand. This device, the Tactilator, was tested on the subject described by Plant & Spens (1986). His results when using the Tactilator was equal to those obtained when his hand is placed directly on the speaker’s throat (Öhngren, Kassling, Risberg, & Söderlund, 1990). One aim of the present study was to examine whether other deafened adults could directly, i.e., without training, take advantage of the information provided by the Tactilator.

The second aim was to find out the cognitive architecture underlying a successful combination of visual speechreading with the use of the Tactilator, and to compare these cognitive abilities with what is required for visual speechreading only. This will also allow an investigation of whether the same cognitive factors that predict visual speechreading ability can also predict skilled users of the Tactilator, or if there are other predictors involved. Lyxell & Rönberg (1987a; 1987b; 1989) have shown that short-term memory indirectly supports speechreading and, verbal inference-making is a direct predictor of speechreading skill. The ability to decode words from lip movements is also a direct predictor of speechreading skill, as well as information-processing speed (Rönberg, 1990). However, when using the Tactilator, speechreading is combined with a sense that is not the “natural” sense for speech perception, and as a result other cognitive abilities may be activated.

METHOD

Subjects

Sixteen deafened adult, eight men and eight women, with a mean age of 52.9 years, (sd=13.2 years), took part in the present study. The subjects were paid for their participation in the study. They had been deafened for a mean of 18.1 years, (sd=14.8 years).
Apparatus

The block diagram of the Tactilator is shown in Fig. 1.

![Block Diagram of Tactilator]  

Fig. 1. A block diagram of the Tactilator.

The contact-microphone was connected to a Sennheiser wireless transmitter SK 1011 and, the bone-conductor was connected to a Sennheiser wireless receiver EK 1011. The experimenter held the contact-microphone against her throat on the side of the larynx and the subject held the bone-conductor in his/her hand, between the thumb and the index-finger. The information provided by the bone-conductor were fundamental frequency (F0) and amplitude variations. The experimenter in this study was a female native speaker of Swedish with an F0 range of 170–250 Hz. The combined F0 and amplitude variations give the subject a strong impression of the rhythm of the speech. It is, however, difficult to say if the subjects could use linguistic information in the F0-signal due to the poor frequency resolution of the tactile sense. The main information in the signal might therefore be time-intensity variations. Experiments by Breeuwer & Plomp (1986), Grant (1986), Risberg & Lubker (1978) have shown that this signal gives efficient support during speechreading. Most of the studies, however, have been made with auditory presentations using normally hearing subjects. The results might therefore not be applicable to deaf subjects and the tactile sense.

Materials

Speechreading tests

Speechreading performance was measured with and without the Tactilator. A live speech tracking test (De Filippo & Scott, 1978) and three video-recorded tests; one sentence-based test (Lyxell & Rönberg, 1989), one Swedish translation of the CID-sentence test (Owens, Kessler, Telleen, & Schubert, 1981), and one Swedish monosyllabic word decoding test (Liden & Fant, 1954), were administered in both conditions. The video-recorded test-material was recorded with a Philips VKR-9500 camera and a JVC Super VHS video-recorder. The face of the speaker on the TV-monitor was approximately life-size. During the recordings the signal from the contact-microphone was recorded into one channel of the video-tape. The signal from the air microphone was recorded on the other channel.

Speech tracking test. Speech tracking is a test procedure described by De Filippo & Scott (1978). It is widely used to measure functional communication performance. The experimenter reads from a book, sentence by sentence, and the subject has to repeat verbatim what is read. Typically the procedure lasts for five or ten minutes. The words correctly perceived in the test session are counted and divided by the time elapsed to give a words per minute (wpm) rate. There are uncontrolled variables in the test such as the experimenter’s speaking rate, the difficulty of the text, and the
way of solving a blockage, that is, when the subject can not understand what the speaker has said (Tye-Murray & Tyler, 1988). To minimize variability in procedure between different subjects the experimenter (a) first read a whole sentence to give the subject a chance to grasp the overall meaning of the sentence; (b) then repeated the first part of the sentence; (c) thereafter repeated the difficult word. If the subject was still unable to correctly repeat the word the experimenter wrote the word in block letters on a piece of paper as quickly as possible. The subjects were encouraged to guess and tell the experimenter the word as soon as s/he found a suitable word. The goal was to minimize the time spent on blockages (Spens, Grossspeltus, Öhngren, Plant, & Risberg, 1992). Each test session lasted 10 minutes and the book used was a simplified, easy reading version of Moa Martinson’s: Mor gifter sig (Mother is going to get married).

The speech tracking procedure has to be administered live voiced due to the need of interaction between the talker and the receiver. Hence, although the speech tracking procedure is of ecological relevance for evaluating speechreading ability when it is used as a test procedure, it can be argued that there are some shortcomings. The common procedure to control the presentation has been to video-record the speaker. Subjects participating in video-recorded tests, however, often complain that the test-situation is worse than in real life (Rönnberg, Öhngren, & Nilsson, 1983). Therefore, to obtain a controlled comparison with the speech tracking procedure with the Tactilitior all other speechreading tests were video-recorded with the same speaker as for speech tracking.

Sentence-based speechreading test. The test was developed by Lyxell and Rönnberg (1989). The subject’s task was to speechread 16 sentences, eight with and eight without tactile support. The sentences are subdivided into two topics: one based on a train trip and the other on a visit to a restaurant. Prior to the presentation of the stimulus sentences the subject was asked to read a frame-story, which instructed him/her to take a particular role in the conversation dyad, – the conductor on the train for the "train topic", and the waiter/waitress in the restaurant, and to speechread the counterpart. The sentences within each of the two blocks are unrelated to each other. For half of the sentences, the subject was also given an extra written contextual cue. This cue informed the subject more specifically about the possible content of the statement/question. The sentences in the test were simple, declarative Swedish sentences. A video-recorder was used to present the sentences. First, the subject speechread one sentence, then the experimenter stopped the tape and the subject reported what s/he had perceived. The answer was tape-recorded and later transcribed. The subject was encouraged to guess what s/he had not perceived. Then, after eight sentences, s/he had to speechread what was said in the second scenario, now using the bone-conductor and thus having information from the speaker’s larynx. Sentences from both topics were presented in each test session, one topic with the Tactilitior and one with speechreading only. The Tactilitior condition was counterbalanced across subjects. On the second test session, which was always the following day, a parallel version of the tests were administered. The topic which was tested with speechreading only the session before, was now tested with the Tactilitior and speechreading and vice versa. The dependent measure of speechreading ability was the proportion of words correctly perceived, averaged over the eight sentences x two sessions in each test condition, speechreading only and speechreading with the Tactilitior.
The Swedish version of CID-sentence test. A Swedish translation of two of the lists – Mac XIV A and Mac XIV B – of the CID (CHABA) Everyday Sentences was used. The lists form part of The Minimal Auditory Capabilities Battery (Owens, Kessler, Telleen, & Schubert, 1981). The subjects were encouraged to guess, and the proportion of correctly perceived words, averaged over the 20 sentences, constituted the dependent measure of speechreading ability.

Word decoding test. Two phonetically balanced word lists were used (Liden & Fant, 1954). Each list consists of 50 common monosyllabic words. Each word was preceded by the number of the word and a cue phrase: "Nu säger jag" (Now I am saying). The subject repeated orally which word s/he had perceived and the experimenter wrote down the answer. The subjects were again encouraged to guess. The proportion of correctly identified words, from the total number of 50 words, represented the measure of speechreading ability.

Cognitive tests
The Cognitive tests battery were administered by a computer program called TIPS; Text-Information-Processing-System (Ausmeel, 1988). The program permits textual items to be presented letter-, word- or line-wise in a fixed, moving, or growing text-window. Presentation rate; that is time per item and/or inter-item interval, can be manipulated using the program's parameter menu. The same subroutine is used in all kinds of access tests and no source of variation in precision emanates from the software per se. The order of test presentation is automatically re-randomized by TIPS, and is unique for each subject.

The description of the cognitive tests used is relatively brief. For a more detailed description the reader should consult Rönberg, Arlinger, Lyxell, & Kinnefors (1989).

Long-Term Memory Access Tests
Physical matching. Thirty-two pairs of letters were presented and the subject's task was to decide whether the two letters were identical or not. The letter-pair had to have the same name, with either the same upper case or lower case format, to be classified as correct. Half of the pairs were identical. Different letter contrasts were used in four trials using ab, nh, de, and rt. The letter-pair was presented on the computer screen (Apple Lisa). A fixed text-window was employed within which the letter-pair was presented for 2 s. The reaction-time was measured from the onset of the 2-s interval, and the 2-s interval also served as the maximum response time. After the response (i.e., striking a predefined key) another 2-s inter-item interval commenced before presentation of the next word.

Name matching. The task was to match the name of the letter. Sixty-four responses were collected in four trials. Every trial consisted of 16 responses, half of which the letter-pairs were the same with respect to the name. The letter contrasts used were the same with respect to the name as in the physical matching test, and the presentation of the items was the same as well.

Lexical access. The task was to decide whether a string of letters constituted a real word or not. Fifty true words and 50 lures were used. All the true words were familiar, Swedish three-letters words according to Allén (1970). The presentation characteristics were the same as for the other long-term memory access tests.
Semantic access. The task for the subject was to decide whether a word belonged to a certain predefined semantic category or not. There were four categories: "colours", "occupations", "diseases", and "parts-of-the-body", and within each category there were 24 items. Twelve items belonged to the category and 12 items were lures. The presentation of the items were the same as for the other long-term memory access tests.

The latency data were based on the average of each subject's time for yes/no responses in each test of long-term memory access. The accuracy data were based on number of correct yes/no responses in each test.

Short-Term Memory Tests

Digit span. A series of digits was presented in a fixed text-window at a rate of one digit per 0.8 s with an inter-item interval of 0.075 s. The subject's task was to recall the digit series orally in the correct serial order. The experimenter registered the response, pressed a button and the next sequence of digits were presented. After a practice session the first span size consisting of three digits in three different sequences was employed. The next span size was four digits, eventually, ending with a span of eight digits. The response interval was maximised to 2 minutes, but for most subjects around 10 s was sufficient.

Word span. The same procedure as with digits were used with words. Entire words, semantically unrelated and phonologically dissimilar, were presented.

Reading span. The first task was to read a sentence, which was presented word-by-word at a rate of one item per 0.8 s with an inter-item interval of 0.075 s. The sentence varied from three to six words, and the subject was asked to say out loudly "yes" if the sentence was meaningful and "no" if it was absurd. "The girl sang a boat" is an example of an absurd sentence. 1.75 s were allowed for the yes/no response before the next sentence appeared on the monitor. The second task was to recall the final word of each sentence in the correct, sentence-wise, serial order after each sequence of sentences. Three sequences of sentences per span size were used, from span size three to span size six. As in the other span tests, the experimenter registered the answers and pressed a button to start the next sequence of sentences as soon as the subject had responded. The response interval was limited to 2 minutes.

For all span tests, the total number of items recalled correctly, divided by the maximum number, i.e., 99 digits, 99 words, 55 final words, across all span sizes, constituted the scoring criterion.

Recency. Eight 12-item lists of unrelated words were presented at a rate of one item per second in a fixed text-window. During a 2-minutes response interval the subject wrote down as many words as he/she was able to recall. Recency was computed as the proportion of words that were recalled from the final four words in each of the eight lists. The average recency of the eight lists represented the dependent measure.

Sentence completion test. Twenty-four sentences were subdivided into three blocks of eight sentences. Each block had a specific topic – a restaurant, a travelling on a train scenario and a clothing shop scenario – which was given in print. The subject was informed about the global topic within which a statement/question could have been made as well as which role, waiter/waitress; conductor; and shop-hand, s/he was to take. For half of the sentences within each block a cue was given. The cue informed the subject more specifically about the possible content of the statement or
the question. All sentences were simple, declarative Swedish sentences. The subject first read the contextual frame after which the experimenter exposed the incomplete sentence for 7 s. The subject thus had to orally complete the sentence based on one or two sources of information: the topic and a cue. The response interval was 25 s per sentence. The answers were tape-recorded and later transcribed by the experimenter. The number of the correct words divided by the number of maximum words in each sentence constituted the basis for the proportion used. The average proportion based on the 24 sentences was used as the actual data for each subject.

The rhyme tests. Four lists of word-pairs were presented at a rate of one pair per second in a fixed text-window. The first list consisted of 50 pairs of monosyllabic and bisyllabic common Swedish words. In the second list there were 50 pairs of bisyllabic words only and in each pair of words there was one "real" word and one nonword. The third list contained 30 monosyllabic pairs of nonwords and the fourth 30 bisyllabic pairs of nonwords. The subject's task was to judge whether or not the pair of words presented on the screen rhymed. The response time was 5 s. The latency data were based on the average of each subject's yes/no response in each list, and the accuracy data were based on the number of correct/incorrect judgements concerning if the pair of words rhymed or not.

General Procedure

When the subject arrived for the first test session, the experimenter introduced the Tactilator. None of the subjects had any experience with the Tactilator prior to the present study, although five of the subjects had previous experience with other tactile aids. For approximately five minutes, the experimenter talked with the subject to provide her/him with some experience with the Tactilator's vibratory patterns.

The eight women and the eight men were randomly assigned to four different test orders with two men and two women assigned to each order. The orders were as follows: First, there was always a speech tracking session. If the subject started with speechreading with the Tactilator, then the next day, the session began with speechreading only and vice versa. After the speech tracking procedure there was a short break and then the video-recorded testing began. The video-recorded tests were administered so that they appeared in a different order, with or without the Tactilator. The first two test sessions were always on two consecutive days. After each session the subject was asked for his/her subjective impressions of the Tactilator.

After about a week the subject participated in the third test session, in which the subject was introduced to the computerised cognitive test battery TIPS (Text-Information-Processing-System) (Ausmeel, 1988). After a break for coffee, the guessing/inference making tests were administered, and finally a vocabulary test was administered (The F-test, Psykologiförlaget, Stockholm). This was done to control the possibility that the subjects had a less than normal understanding of written Swedish words.

Design

The overall design was a 2x2 within-subjects experimental design. The first factor refers to the communication factor; speechreading only and speechreading with the Tactilator. The second factor is the test modality; live (speech tracking) vs. video-recorded tests (sentence-based and a word-decoding tests).
Fig. 2. A comparison between speed-reading results using speed-reading only and speed-reading with the Tachlloon.
RESULTS

A separate one-way ANOVA, based on words per minute in the live speech tracking test, revealed a clear main effect, $F(1/15)=107.87, p<.0001, MSe=3.13$, using the Tactilator as an aid to speechreading. The main effect is also present in the separate ANOVA, based on the proportions correctly perceived items in each of the three video-recorded tests, $F(2/30)=23.42, p<.000, MSe=.02$. No other effect reached significance. All subjects gained from the Tactilator in the speech tracking condition (See Fig. 2).

The second purpose was to examine whether there are individual cognitive differences which may offer a plausible explanation for the robust effect of Tactilator-supported speechreading, and if the results deviate from earlier results concerning cognitive skills in visual speechreading (Lyxell, 1989; Lyxell & Rönberg; 1989; Rönberg, 1990). As the speech tracking test is considered to be the most ecological measure of a deafened adult's communication ability, and the video-recorded tests in this study served as a control for the experimenter's live presentation, it was deemed appropriate to correlate the speech tracking results with the cognitive results.

From the correlational pattern possible cognitive predictors were chosen, not only for their significant correlations with Tactilator supported speechreading, but also for being the best predictors belonging to different theoretical groupings given by earlier research (Lyxell, & Rönberg, 1989; Rönberg, 1990). Also, input data to the correlational analysis were constituted by values averaged across conditions within each particular test. As a result of applying the above mentioned criteria the predictors chosen were the best recoding, visual decoding, information-processing speed, inference-making and memory tests (see Table I).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Recording</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Decoding</td>
<td>-0.043</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Speed</td>
<td>0.359</td>
<td>-0.352</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Inference making</td>
<td>0.106</td>
<td>-0.383*</td>
<td>-0.301</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Tactilator tracking</td>
<td>-0.53*</td>
<td>0.512*</td>
<td>-0.504*</td>
<td>-0.342</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6. Reading span</td>
<td>0.032</td>
<td>-0.091</td>
<td>-0.346</td>
<td>0.536*</td>
<td>0.302</td>
<td>1</td>
</tr>
</tbody>
</table>

Table I. Correlation matrix for predictors of using the Tactilator in speech tracking (n=16).

The first step in the multi-regression analyses was to choose predictors from the theoretical groupings which significantly correlated with the speech tracking with the Tactilator criterion viz.; a decoding test (the word test), a speed test (semantic access), and a recoding test (the rhyme test). The multiple-regression analysis showed that the best predictors ($R=.71$, adj.$R^2=.48$) were the recoding test (beta=$-.44$, $p<.06$) and the visual decoding test (beta=$.43$, $p<.06$). This three-variable multiple-regression analysis did not include the information-processing speed test (i.e., semantic access) as a significant contributor to the solution. Next step, then, was to attempt to accomplish a three-variable solution with three variables as significant contributors. We therefore had to relax the criterion regarding significance for simple correlations
with the criterion (see Table 1). Thus, two additional regression analyses were carried out with either the inference-making test or the complex memory test included as the third variable. Of these two analyses, the best multi-regression analysis showed (R=.81, Adj.R2=.57) that the cognitive predictors of Tactilator-supported speech tracking was (a) the speechread mono-syllabic word decoding test (beta=.52, p.<.01), (b) the phonological test measuring accuracy in judging if two bisyllabic non words rhymed (beta=-.52, p.<.01), and (c) the complex working memory test, i.e., the reading span test (beta=.37, p<.05).

The information-processing speed test which was significantly correlated with the criterion, did not come out as a predictor in the first regression analysis. Part of the reason as to why this test did not reach significance was that its intercorrelation with decoding was relatively high, although not significant. This may implies that decoding is indirectly supported by information-processing speed (cf. Rönnberg, 1990).

Finally, as a control, a similar multi-regression procedure with speechreading only was carried out (R=.81, Adj.R2=.56). The best cognitive predictors were (a) the word decoding test (beta=.62, p<.00), b) the phonological test (beta=-.42, p<.03), and (c) the complex memory test, (beta=.38, p<.07).

In sum: The results clearly suggest that deafened adults improve their speechreading with the Tactilator without training. The most relevant cognitive predictors for making use of the information given by the Tactilator together with speechreading are (a) a phonologically based recoding test, (b) a visual decoding test, and, (c) a working memory test. The same predictors hold true for speechreading only, although the order strength of the predictors are different. By and large, the present data replicates the work done by Lyxell & Rönnberg (1987; 1990).

DISCUSSION

The results will be discussed in two separate sections. Firstly, the outcomes of the Tactilator support in the different speechreading tests will be evaluated and discussed. Secondly, the results from the cognitive tests will be presented, and a simple structure of cognitive predictors of speechreading with the Tactilator will be discussed.

Tactilator supported speechreading

Most of the earlier work concerning tactile aids has been carried out using artificially deafened normally hearing persons using headphones with masking noise. The present study used 16 deafened adults from a variety of vocational backgrounds with an age range from 26 to 71 years. Further, the study was conducted on persons who had been without auditory information for an extended time-period. As a result of these factors, and given that the effects of the Tactilator were systematic and strong, it should be possible to generalise the results to other deafened adults.

The direct (i.e., without training) effect of the Tactilator for all the subjects in the speech tracking condition has rarely been obtained in earlier studies with other tactile aids (Axelsson, Berenstaaf, Hansson, & Spens, 1987; Lyxell, Rönnberg, Andersson, & Linderoth, 1991), and might depend on how the speech signal is processed. Earlier research has shown that multichannel tactile aids, which lay out the speech signal as a spatial pattern representing frequency or voice formant relations, are the most effective systems, but there is no evidence of direct effects. In single-channel
aids the speech signal is often processed with filtering, amplitude compression etc. The frequency of the signal is fixed and bears no relation to speech frequencies.

When Tactiling, vibrations from the speaker’s larynx are transmitted to the receiver’s hand, which is held on the speaker’s throat (Öhngren, Rönnberg, & Lyxell, 1992). There is no signal processing. The only limitations are those imposed by the tactile sense which has its maximum sensitivity around 250 Hz. Variations in the speaker’s F0, as well as variations in amplitude will be the main components in the percept. Similarly, when using the Tactilator no processing of the information from the larynx takes place. The larynx microphone merely registers the vibrations from the larynx.

Another important factor concerning speech processing is that both commercially available and experimental tactile aids are based on the same premise as conventional hearing aids. That is, the speech signal has to pass the speaker’s vocal tract and travel through the air into a microphone, where different processing mechanisms take place. In the Tactilator, the speech signal is taken directly from the speaker’s vocal tract, and that signal is transmitted via radio frequency directly to the vibrator. No changes in the sound waves, on their way to the microphone, or any disturbing noise, interfere with the signal. Thus, the Tactilator concept represents a totally different way of solving the technical transmission of speech from talker to receiver.

One can argue that using a larynx microphone would be disturbing for the talker in the communicative situation. When talking to a deafened adult, however, both the normal hearing speaker and the deafened adult often experience problems. As a talker, it is frustrating to notice that you can not be understood when talking with a deafened adult. If the deafened adult provided the speaker with a microphone, and asked her/him to hold it against her/his throat and speak normally, and if it was easier to be understood, then the inconvenience of holding a larynx microphone against the throat would probably be less, than the frustration of not being understood.

One way of viewing the results from a theoretical standpoint is based on the so called ecological realism approach (Shaw, Turvey, & Mace, 1982). Here, it is claimed that the availability of information in normal patterns of light and sound is not meagre and ambiguous, but instead meaningful and rich without limit. One can then expect that each individual extends and improves his/her perceptual repertoire depending on individual needs and opportunities. In Gibson’s (1979) work it is clearly featured that the “achievements of a perceptual system are susceptible to maturation and learning. The information that is picked up becomes more and more subtle, elaborate, and precise with practice. One can go on learning to perceive as long as life goes on” (Gibson, 1979, p. 245).

The inspiration behind this study was the man who uses the hand to support speechreading, and who perceives speech at a nearly normal conversation rate. He has become an expert during 45 years usage of visual information together with tactile information. He has discovered and uses information that is of more direct relevance for speech understanding.

In this study the subjects were able to make use of the available information directly. Runeson (1988) has called for a skill-oriented perspective in perception research. The present research was done from a skilled model with naive subjects. All of them improved their speech understanding using the combined visual and tactual
information in the same way as the skilled model – but not of course to his level of excellence. This may give ground for an interpretation based on direct "smart" perceptual mechanisms. Runeson (1977) argues that natural systems are sensitive to some invariant combination of "elementary" properties. The invariants in speech perception, when information from the visible speech movements and the vibrational pattern given by the Tactilator are combined, are evidently easier to perceive compared to when other tactile aids with pre-processed speech are combined with visual speechreading. The direct robust effect on naive users and the excellence of the skilled user, suggests a useful invariant embedded in the information available. One necessary prerequisite for a Gibsonian type of interpretation is thus the direct effect of the Tactilator (i.e., without practise), but it is by no means a sufficient condition for this type of interpretation. One complicating aspect is that individual differences in using the Tactilator also are related to cognitive abilities.

Cognitive predictors of Tactilator supported speechreading

One of the two strongest cognitive predictors, of using the Tactilator information, was a phonologically based test, in which the subjects had to determine whether bisyllabic non-words rhymed or not. To be able to do that, one has to have access to an articulatory control process, based on inner speech (Baddeley, 1990). This process is also capable of processing written material, converting it into a phonological code and registering it in the phonological store. This phonological store is capable of holding speech-based information for about two seconds. By reading off the memory trace from the phonological store into the articulatory control process the memory trace can be refreshed. This process is also underlying subvocal rehearsal. Baddeley (1990) argues that the phonological loop must have been developed in the process of the evolution of speech production and comprehension.

Given that we have built-in in the perceptual system a phonological base for perceiving speech (Conrad, 1979; Rönberg, 1987), it might be easier for the system itself to do the selection of the useful information for speech perception through the visual and tactile sense. Interestingly, the putative access to the phonological loop via visual-tactile speechreading seems also to be validated by the subjective reports of auditory experience. The subjects were, in the post-experimental interview, asked to report how they felt the vibrations from the Tactilator. It was obvious that the tactile impression was easy to compare with auditory information. Impressions such as "I saw the Tactilator as a hearing aid I wanted more sound in;" "The Tactilator gives a form of auditory sensation which makes the speaker more concrete. It is no longer a 'dumb-movie', suggest such a perceptual-cognitive communality."

Concerning the cognitive predictors at large we replicated the critical predictors obtained in previous studies (Lyxell, 1989; Rönberg, 1990). One theoretical implication of this result is broadly that skill in speechreading and speechreading with the Tactilator rests on the same cognitive factors. The results also suggest that information might be processed in the same way when tactile information supplements speechreading as when there is only visual speechreading. One key to this may be the fact that relatively smooth phonological processing (i.e. inner speech) is important both to visual – tactile and visual speechreading. We further suggest that effective lexical access, and hence decoding of speech gesture seems to be mediated by a phonological route (Rönberg, 1987; Rönberg, 1990).
Lyxell (1989) found that visual word decoding, information-processing speed, and verbal inference-making characterized the skilled speechreader. In this study, the first predictor for being a good speechreader was similarly a word decoding test. In the Lyxell study, however, the reading span test constituted an indirect predictor, and the verbal inference making test was the direct predictor. In this study the reading span test was a direct predictor of both being a skilled speechreader and for being able make use of the vibrations from the Tactilator together with speechreading. One possible explanation for the disagreement might be the different kind of test used. In the Lyxell study there was a sentence-based test (Lyxell, & Rönning, 1987) administered, but in this study the main criterion was speech tracking. Compared to sentence-based speechreading, speech tracking demands a capacious working-memory. In the sentence-based test, the speechreader knows the context, but one particular sentence does not necessarily relate to the next. In speech tracking the story is successively built up, sentence by sentence, and it is easier for the speechreader to understand if s/he remembers the previous parts of the story, that is, s/he has to store and process information in working memory.

Finally, Rönning (1990) argues that there is a basic cognitive speed function in speechreading. This study only support this in an indirect way, as decoding was a main predictor of being able to use Tactilator supported speech tracking.

**Clinical Applications**

1. When transmitting vibrational information from the speaker’s larynx to the receiver’s hand, the perceived variations in fundamental frequency perceived are enough to give a direct, that is without training, support to speechreading to all subjects in this study. This suggest that, when transducing auditory information into tactual information, the speech processing in the aid should be as simple as possible.

2. The phonologically based, recoding test, the decoding test, and the complex working-memory test, constituted the best predictors of being able to use the vibrations from the Tactilator. Crucial for the deafened adults is thus to have access to the phonological loop (i.e., inner speech), and therefore, in spite of the deafness, s/he has to continue using his/her own speech.

3. The deafened adults represents a very small group in our society. When it comes to hearing impairment, however, nearly 10% of the population are hearing impaired and around 200 000 of them need to use hearing aids. For most hearing aid users the hearing aid will give enough information when used in silent surroundings. A problem for all hearing aid users is, however, how to separate the speech sounds from noise when being in a noisy environment. The Tactilator concept applied to hearing aids would improve the signal-to-noise ratio, and thus, improve speech understanding for the hard of hearing.

**ACKNOWLEDGEMENTS**

The author would like to thank Arne Risberg, Jerker Rönning and Geoff Plant for their comments on this paper, by which it was greatly improved. This research was supported by a grant from the Swedish National Board for Technical Development.

**REFERENCES**


