III. SPEECH AND HEARING DEFECTS AND AIDS

A. A SINGLE-TRANSDUCER VIBROTACTILE AID TO LIPREADING

Geoff Plant*

Abstract

Four deaf subjects were tested using a vibrotactile aid to lip-reading presenting voicing information and a cue to signal the presence of high frequency consonants. Testing at the level of consonant perception presented lipreading alone, and lipreading supplemented by the aid showed improvements in the perception of consonant voicing and manner of articulation in the aided condition. Testing at the word and sentence level showed differing results for the subjects completing the tasks. A congenitally deaf subject with a history of non-hearing-aid use showed no improvements in the aided conditions whereas another subject with a history of very successful hearing aid use evidenced improvements in the aided condition for both sets of materials. Testing at the level of connected discourse revealed improvements in the aided condition for two subjects but equivalent scores aided/unaided for the subject with limited hearing aid experience. Testing in the tactile-alone condition showed that the subjects were able to perceive the presence/absence of /s/ and word syllable number and type with a high degree of proficiency. The significance of the results and future research directions are discussed.

Introduction

A number of studies (Risberg, 1974; Risberg & Agelfors, 1978; Risberg & Lubker, 1978; Rosen, Fourcin, & Moore, 1981; Grant, Ardell, Kuhl, & Sparks 1985) have demonstrated that lipreading performance by both hearing and hearing-impaired subjects is greatly enhanced when the visual signal is supplemented by auditorily presented fundamental frequency (F0) information. The success of these studies has encouraged a number of research efforts aimed at investigating the presentation of F0 information via the tactile sense as an aid for profoundly deaf persons. Approaches adopted include the use of either an electrotactile (Grant, 1980) or a vibrotactile (Boothroyd, 1983) array presenting F0 changes as spatial changes, a single vibrator with the F0 range transposed into the frequency region where the skin is maximally sensitive to frequency change (Rothenberg & Moltor, 1979; Plant & Risberg, 1983), and a single vibrator presenting F0 as a fixed frequency (Plant, Macrae, Dillon, & Pentecost, 1984a). The studies by Grant (1980) and Plant & al. (1984a) both showed significant improvements in performance when the tactual signal supplemented lipreading. Plant & al. (1984a) reported significant improvements in the lipreading performance of an experienced vibrotactile aid user at the phoneme, word and sentence levels when F0 information was presented tactualy. They also found a significant improvement in aided connected discourse tracking scores. These results

* National Acoustic Laboratories, Sydney, Australia; guest researcher at KTH from Nov. 1985 - March 1986.
would appear to support the choice of $F_0$ as the speech parameter to be displayed via a tactual aid. This is not surprising as such an approach potentially offers much useful information to assist both speech perception and production by the profoundly deaf. Studies of visual consonant perception have highlighted the paucity of speech information available via lipreading alone. For example, although consonant place of articulation can be identified with a relatively high degree of proficiency, voicing distinctions cannot be detected (Woodward & Barber, 1960; Erber, 1974; Plant & Macrae, 1977). A tactual aid providing a reliable cue to consonant voicing should greatly increase the amount of segmental information available to the lipreader. Similarly, fundamental frequency change also serves to signal prosodic features such as emphatic stress. Again, this cannot be reliably detected via lipreading alone (Risberg & Lubker, 1978). Studies by Rothenberg & Molotor (1979) and Plant & Risberg (1983) have shown that, with training, emphatic stress in sentences can be detected at a relatively high performance level via the tactual sense. At the speech production level studies of both adventitiously and congenitally deaf speakers have revealed deficiencies in $F_0$ control. Plant & Hammarberg (1983) in a study of the speech of three deafened Swedish subjects reported a restricted $F_0$ range and an inability to use $F_0$ change to signal emphatic stress in sentences. Studies of the speech of the congenitally deaf (see Osberger & Mc Garr, 1982) have shown that the average range of pitch variations is "more reduced than those of speakers with normal hearing" (Osberger & Mc Garr, 1982, p. 249). Osberger & Mc Garr also report that some acoustic studies of the speech of the hearing impaired "provide evidence that a lack of voice-onset-time contributes to the perception of (a) voiced-voiceless confusion" (Osberger & Mc Garr, 1982, p. 236). They cite a study by Monsen (1976) which found that hearing impaired speakers who were unable to produce the voiced-voiceless contrast "were considerably less intelligible than those who produced the voicing distinction" (Osberger & Mc Garr, 1982, p. 237).

This paper describes results obtained by four deaf subjects using a single-transducer vibrotactile aid presenting $F_0$ cues and information as to the presence/absence of high frequency speech energy. Results obtained with an earlier version of this aid adjusted to provide voicing duration and intensity cues presented either tactually to a deafened subject (Plant, Macrae, Dillon, & Pentecost, 1984a) or auditorily to normal hearing subjects (Plant, Macrae, Dillon, & Pentecost, 1984b) showed the effectiveness of the approach. Lipreading scores of normal hearing subjects rose by 11 percentage points when the auditorily presented information supplemented lipreading. This is statistically significant and represents more than a doubling of the scores obtained via lipreading alone. The results obtained by the deafened subject are summarized earlier in this introduction.
Fig. 1. Block diagram of the experimental aid used in this study.
Method

A block diagram of the experimental equipment is presented in Fig. 1. The aid consists of a pitch extractor which either drives a 115 Hz oscillator when voicing is present in the speech signal (Frequency Constant) or the extracted F₀ contour is transposed to the frequency range 40-140 Hz (Frequency Variable). At this stage of the aid's development the latter setting can only be used successfully with male speakers. The intensity of the presented signal can either be constant (Intensity Constant) or determined by the intensity of the speech in the frequency range 500-1000 Hz (Intensity Variable). A high-pass filter set at 5 kHz triggers a 300 Hz oscillator if significant energy is present in the speech signal. The intensity of this signal is determined by the energy in the frequency band. In practice the oscillator is triggered by /s, z, /, ʒ, tʃ, dʒ/ and to a lesser extent by /t/. The output signal is fed to a power amplifier which enables the subject to vary the intensity of the vibrator. The output transducer used is that developed for the Sentiphone (Traunmüller, 1980) and is designed to be hand held.

Subjects

Four deaf subjects have been trained and tested using the vibrotactile aid. Subject 1 was a 22 year-old female totally deafened by meningitis two months before the commencement of the training program. During the training sessions the experimental aid was used at all times. For daily use the subject was fitted with a Minivib (Spens & Plant, 1983) which she wore at all times. The testing and training procedures were discontinued following implantations of a multi-electrode cochlear implant (Mecklenberg & Brimacombe, 1985). This cochlear implant fitting has been extremely successful with the subject displaying good open-set speech recognition via audition alone. Subject 2 was a 36-year-old female with a congenital hearing impairment. The onset of total deafness for this subject occurred at age 24. This subject has had extensive experience with vibrotactile stimulations and has served as a subject in two previous studies of vibrotactile aids (Plant, 1982; Plant & al., 1984a). A Minivib was fitted to this subject approximately 12 months prior to the study described in this paper. This aid is worn consistently on a daily basis and the subject has expressed great satisfaction with the fitting. Subject 3 was a congenitally profoundly deaf male with excellent speech and language skills. This subject is an extremely successful hearing aid user but expressed a willingness to take part in the present study. The subject's experience with tactile stimulations was limited to the testing and training sessions. All testing and training with this subject was conducted with his hearing aid removed. In this condition the subject is unable to perceive speech via audition. When asked to compare the signals produced by his hear-
ing aid and that of the vibrotactile aid the subject reported that they were essentially the same but the vibrotactile signal was "lower and more muffled". Subject 4 was a 21 year old congenitally deaf male with a history of non-hearing aid use. He was educated at an oral school for deaf children during primary school (5-11 years) and then fully integrated at the high school level. This subject has excellent written language but unintelligible speech. He was referred for fitting of a vibrotactile aid in mid-1984. A Minivib was subsequently fitted, and a training program using the experimental aid was initiated on a weekly basis. The emphasis in the training program was directed towards using the experimental aid to improve the subject's speech production skills.

The decision to include this subject in the present study was made when it became obvious that he was able to successfully face the information provided via lipreading and the vibrotactile aid. Training in the aided perception of 20 consonants in an /aCa/ frame was used as part of the training program initiated with this subject. Analysis of the subject's scores after eight training sessions revealed a mean score of 61.6% with the subject's mean score for Sessions 1-4 being 53.1% and that for Sessions 5-8 being 70%. This latter score is considerably higher than that which could be reasonably attributed to lipreading alone. An analysis of the subject's feature perception for Sessions 1-4 and 5-8 showed improved perception of consonant voicing and the stop/continuant, nasal/oral, and sibilant/non-sibilant contrasts in the latter four sessions. This was a most unexpected result as the subject reported he had never derived any benefit from hearing aids, and it was felt that at the age of 20 it was unrealistic to expect him to develop bimodal perceptual skill. These results, however, indicated that the subject did receive assistance in lipreading from the vibrotactile aid at least at the level of consonant discrimination and it was decided to investigate whether this improvement could be generalized to higher level materials.

**Presentation**

All of the testing materials were presented live-voice by the experimenter, a 36 year-old male native speaker of Australian English, seated approximately 1 metre from the subject. The testing was conducted on a one-to-one basis in a sound treated test room. Both the experimenter and the subject wore head mounted microphones to ensure a constant lips-to-microphone distance and an optimal signal-to-noise ratio. The aid setting used by Subjects 1 and 2 was Frequency Constant/Intensity Variable. This setting was chosen as both the subjects were female and the aid would not have delivered a satisfactory signal for their own speech. This was not a problem with Subjects 3 and 4 and as a result the setting Frequency Variable/Intensity Constant was selected. Testing was usually conducted in sessions of approximately one hour. Subject 2, however, was tested in two hour sessions as she lived a considerable distance from the facility where the testing was conducted.
All of the subjects participated in the study on a voluntary basis although Subject 2 was paid a fee to offset her travel costs.

**Material**

The aim of the test program was to evaluate the vibrotactile aid's effectiveness as a lipreading supplement and to investigate the aid's ability to signal selected suprasegmental and segmental features of speech.

**Lipreading**

A number of lipreading tests were administered presenting materials ranging from nonsense syllables to connected discourse. The tests used were:

**Consonant identification**

A. Twenty consonants /p,b,m,t,d,n,k,g,f,v,s,z,h,d,z,t,s,w,r,l,j/ were presented in an [aCa] frame. Lists were prepared consisting of 40 test items presenting each of the consonants twice in a random order. Eight testing sessions were conducted using these materials. In each session one list was presented lipreading alone and one list lipreading plus the tactile aid. The materials were presented live-voice with the order of presentation and the test lists varied from session-to-session in an attempt to minimize the possibility of practice and order effects. The subjects were asked to write down each consonant as it was presented.

B. The 12 consonants /p,b,t,d,k,g,m,n,s,z,f,v/ were presented in an [aCa] frame. Lists consisting of 60 test items presenting each of the consonants five times in a random order were prepared prior to testing. Four testing sessions were conducted with one list presented lipreading alone and one list lipreading plus the tactile supplement in each session. The materials were again presented live-voice with presentation order and the lists used varied from session to session. A response sheet setting out the 12 alternatives was prepared and the subjects were instructed to both say and point to the consonant they thought had been presented. The subject's responses were recorded by the experimenter. An additional 4 lists were presented to subject 3 via the tactile aid alone. In this case the experimenter's face was hidden by a piece of card and the subject had to rely solely on the information provided by the tactile aid.

**Word identification**

The materials presented at this level were the Macquarie PE word lists (Clark, 1981) 1-4. The lists were presented live-voice over two testing sessions. Two lists were presented lipreading alone and two lists lipreading plus the tactile aid in Session 1. In Session 2 the lists were re-presented in the opposite sensory condition. The subjects were instructed to write down each word as it was presented.
Sentence identification

Lists 2:1, 1, and 2:3 of the SPIN Test (Kallikow, Stevens, & Elliott, 1977) were used to evaluate the subjects' performance at this level. Testing was conducted over two sessions with the materials presented live-voice. In Session 1 one list was presented lipreading alone and the other list was presented via lipreading plus the tactile aid. In Session 2 the same lists were presented but in the opposite sensory condition to that used in Session 1. The subjects' task was to write down the final word of each sentence as it was presented.

Connected discourse

Six testing sessions were used to evaluate the subjects' ability to perceive connected discourse materials. Subjects 1 and 2 were tested using De Filippo & Scott's (1978) tracking procedure. Two sessions of connected discourse tracking each of 10 minutes duration were presented at each testing session. The order of presentation for the two sensory conditions, lipreading alone and lipreading plus the tactile supplement, was altered from session to session in an attempt to minimize any order effects. At the completion of each testing session the number of words completed was calculated and divided by the time elapsed providing "an index of rate, or efficiency of communication" (De Filippo & Scott, 1978, p. 1186) expressed in words-per-minute (wpm). A modification of the tracking procedure was used with Subject 4. This was necessary as this subject's speech and manual communication skills were not sufficient to allow the use of conventional connected discourse tracking (CDT). A short story was rewritten and divided into 12 parts, each consisting of a total of 200 words. Each part was then divided into a number of short segments (around 6-12 words) and was presented to the subject segment-by-segment for identification. Each segment was presented as many times as the subject required. The entire segment was presented each time and no clue words or rewordings were used. The subject was asked to write down as much of the segment as possible. When the subject was satisfied with his written response, it was then corrected for the number of correct words and the subject was shown the written version of the correct form. At the completion of each part the number of words correctly identified was calculated and divided by two to give a percentage correct score. The time which elapsed for each part was also recorded. The number of words correctly identified was divided by the time elapsed to give a correct-words-per-minute (cwpm) score. The procedure was used over six testing sessions. In each session one part was presented lipreading alone and one part lipreading plus the tactile aid. The order of presentation was varied from session to session.
Tactual perception

Testing was also carried out to investigate the subjects' abilities to perceive selected speech features via the tactile aid alone. The tests examined the subjects' abilities to detect the presence or absence of /s/ in meaningful words and their abilities to detect word stress patterns.

Presence/absence of /s/

The following three tests were all presented to the subjects via the tactile aid alone without the benefit of lipreading cues.

Presence/absence of initial /s/

Forty word pairs differing only in the presence or absence of initial /s/ (for example "ski/key", "scare/care") were presented using an ABX procedure. The two alternatives were presented to the subject in a written form on a text sheet and each word was then said. One of the two words was then presented as the stimulus and the subject was asked to circle the word which s/he thought had been repeated.

Presence/absence of final /s/

Forty word pairs contrasting words differing only in the presence or absence of final /s/ (for example "bat/bats", "tap/taps") were also presented using the procedure outlined above.

Presence/absence of initial and/or final /s/

The final task required the subject to choose the stimulus word from a list of four contrasts. The words differed in that one did not contain /s/, one had an initial /s/, one a final /s/, and one had both an initial and a final /s/ (for example "care/cares/scare/scares"). The four contrasts were introduced in turn and one was then presented as the stimulus. The subject was expected to circle the word which s/he thought had been presented.

Word syllable number and type

The test used to evaluate the subjects' ability to identify syllable number and type was the Syllables in Words Test (Plant, 1984). The test consists of 40 items and contrasts monosyllabic trochaic, spondaic, and trisyllabic words. Twenty of the best items have been selected to provide relatively easy syllabic contrasts. In these cases the syllables are joined by either stop consonants or voiceless fricatives (for example "we/blackboard/bottle/addition"). The remaining 20 test items present more difficult contrasts in which the syllables are joined by voiced, continuant consonants (for example "top/yo-yo/winner/television"). The subject was given a text sheet setting out each of the contrasts and was then asked to circle which of the four words s/he thought was presented as the stimulus. This test was also presented without lipreading cues being available to the subject.
Results

The subjects' results for the task involving perception of the 20 consonants are presented in Fig. 2. It can be seen that in every test session aided performance was better than that obtained when the same materials were presented via lipreading alone. The subjects' mean scores over the eight testing sessions for the two sensory conditions lipreading alone and aided lipreading respectively were: Subject 1 - 47.2% and 60.9%, Subject 2 - 49.1% and 60.7%, Subject 3 - 48.75% and 74.7%, and Subject 4 - 43.75% and 69.7%. Overall this represents an improvement in the aided condition of 20.8 percentage points. Confusion matrices were drawn up from the subjects' responses and these were used to gain a detailed description of the subjects' abilities to perceive consonantal features. The results of this analysis are presented in Table I. It can be seen that the vibrotactile aid provides information assisting in the recognition of consonant voicing and consonant manner, especially in the identification of stops, sibilants, and nasals. For three of the four subjects there were improvements in the aided condition in the perception of the affricates and semi-vowels. Not surprisingly there was little difference in the subjects' performances for the two sensory conditions at the level of consonant place of articulation. All four subjects are able to identify consonant place with a high degree of accuracy via lipreading alone.

The results obtained by Subjects 2, 3, and 4 for the four testing sessions using the 12 consonant set are presented in Fig. 3. Unfortunately Subject 1 had left the testing program by the time these materials were introduced. Again, in all cases aided performance is well above that obtained in the lipreading alone condition. The subjects' mean scores across the four sessions for lipreading alone and aided lipreading respectively were: Subject 2 - 51.25% and 82.5%, Subject 3 - 52.1% and 86.2%, and Subject 4 - 55.4% and 80.4%. Overall this represents an improvement in the aided condition of 30 percentage points. Confusion matrices were again used to obtain a detailed analysis of the subjects' performance. The results of this analysis are presented in Table II and show large improvements in the perception of voicing and manner of articulation when the tactile aid supplemented lipreading. Analysis within the manner of articulation category revealed improved perception of stops, sibilants, and nasals in the aided condition.

Of special interest are the scores obtained by Subject 3 in the tactile alone condition. The subject's mean score for the four sessions was 35.8% indicating a level of performance well above that which could be reasonably attributed to chance (8.3%). A detailed analysis of the subject's performance in the tactile alone condition revealed that he was able to correctly identify voicing on 84.6% of presentations. The subject's overall score for manner of articulation (70%) also shows the effectiveness of the aid in presenting consonantal contrasts. The subject was able to correctly identify stops on 83.5% of presentations,
nasals on 70%, and sibilants on 80%. The subject's score for the non-
sibilant fricatives in contrast was only 20% which represents chance
level performance. Examination of Table II, however, reveals that /f/
and /v/ were always correctly identified by lipreading alone. The tac-
tile aid under investigation is designed to be a supplement to lip-
reading, providing information which cannot be perceived accurately
visually. Consequently, the low score obtained for fricatives identified
should not be seen as a deficiency in the tactile aid's performance.
The subject's score for place of articulation (41.7%) is mainly attri-
butable to the high score obtained for the alveolars (69.1%). The aid's
300 Hz signal is triggered by /s/, /z/, and /t/ and this enables percep-
tion of the alveolars of a level well above chance.

Fig. 2. Subject's performance for the visual alone (0)
and visual-tactual (x) perception of 20 conso-
nants in an /aCa/ frame across 8 testing sessions.

Fig. 3. Subject's performance for the visual alone
(0) and visual-tactual perception of 12 conso-
nants in an /aCa/ frame across 4 testing
sessions. Testing was also conducted with
Subject 3 in the tactual alone (☐) condition.
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L%</td>
<td>UV</td>
<td>L%</td>
<td>UV</td>
<td>L%</td>
</tr>
<tr>
<td>Overall</td>
<td>47.2</td>
<td>66.9</td>
<td>69.1</td>
<td>66.7</td>
<td>48.7</td>
</tr>
<tr>
<td>Voicing</td>
<td>67.5</td>
<td>80.4</td>
<td>71.9</td>
<td>82.8</td>
<td>72.2</td>
</tr>
<tr>
<td>Manner (overall)</td>
<td>65</td>
<td>76.2</td>
<td>68.7</td>
<td>85.2</td>
<td>71.6</td>
</tr>
<tr>
<td>Stone</td>
<td>63.5</td>
<td>77.1</td>
<td>56.3</td>
<td>92.7</td>
<td>80.3</td>
</tr>
<tr>
<td>Affricates</td>
<td>50</td>
<td>53.1</td>
<td>28.1</td>
<td>25</td>
<td>40.6</td>
</tr>
<tr>
<td>Fricatives</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Sibilants</td>
<td>29.2</td>
<td>31</td>
<td>85.4</td>
<td>85.4</td>
<td>32.1</td>
</tr>
<tr>
<td>Nasals</td>
<td>46.9</td>
<td>53.1</td>
<td>63.7</td>
<td>84.4</td>
<td>46.9</td>
</tr>
<tr>
<td>Semi-vowels</td>
<td>84.4</td>
<td>78.1</td>
<td>65.9</td>
<td>98.4</td>
<td>88.5</td>
</tr>
<tr>
<td>Place (overall)</td>
<td>90.6</td>
<td>91.9</td>
<td>50</td>
<td>89.4</td>
<td>95.3</td>
</tr>
<tr>
<td>Bilabials</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Labio-dentals</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Rounded labials</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Alveolars</td>
<td>86.5</td>
<td>92.7</td>
<td>88.5</td>
<td>92.7</td>
<td>90.6</td>
</tr>
<tr>
<td>Post-alveolars</td>
<td>65.7</td>
<td>62.5</td>
<td>56.2</td>
<td>43.7</td>
<td>89.6</td>
</tr>
<tr>
<td>Palatal velars</td>
<td>97.9</td>
<td>97.9</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Glottal</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

**TABLE 1.** The results obtained by the individual subjects and the group means for the task involving perception of 20 consonants in an /ae/ frame via lipreading alone and aided lipreading across eight testing sessions.
Word identification

The four lists of Clark's (1981) Macquarie PB-words were prescribed to Subjects 3 and 4. The scores obtained by the subjects for the individual lists in the two sensory conditions are presented in Fig. 4. Subject 3's mean performance for the aided condition (51.5%) is markedly better than that obtained when the lists are presented via lipreading alone (32%). Subject 3, however, had the same mean score (17%) for both the unaided and aided conditions. This would indicate that although Subject 4's performance for the consonant recognition tasks indicated he was able to fuse inputs from the two modalities, this ability has not, as yet, been generalized to the word level.

![Fig. 4. Performance by Subjects 3 and 4 for visual alone (O) and visual-tactual perception of the 4 lists of the Macquarie PB Word Lists (Clark, 1981).](image)

Sentence identification

Lists 2:1 and 2:3 of the SPIN Test were presented to Subjects 2, 3, and 4 and the results are given in Fig. 5. The mean scores for the subjects in the unaided and aided conditions respectively were: Subject 2 - 54% and 67%, Subject 3 - 45% and 63%, and Subject 4 - 19% and 16%. Again, Subject 4's performance indicates that his ability at the syllable level has not as yet been generalized to higher level tasks.

Connected discourse tracking (CDT)

The CDT scores obtained by Subjects 1 and 2 are presented in Fig. 6. These show a clear superiority for the aided condition with only one instance where the scores obtained were equivalent (Session 3 for Subject 1). In all other 11 sessions the aided score exceeds that obtained unaided. The mean tracking rates for the two subjects across the six sessions unaided and aided respectively were: Subject 1 - 26.25 wpm and 35.25 wpm and Subject 2 - 40.3 wpm and 56.2 wpm.

Tactual perception

Presence/absence of /s/

The subjects' scores for the three subtests involving /s/ detection are presented in Fig. 8. Subjects 2, 3, and 4 completed all subtests but Subject 1 did not attempt Subtest 3. These results provide convincing proof that not only does the aid's processing scheme enable
<table>
<thead>
<tr>
<th></th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>Combined</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>51.25%</td>
<td>82.5%</td>
<td>52.1%</td>
<td>86.25%</td>
</tr>
<tr>
<td>Voicing</td>
<td>65.4%</td>
<td>63.75%</td>
<td>90.8%</td>
<td>68.3%</td>
</tr>
<tr>
<td>Manner (overall)</td>
<td>82.5%</td>
<td>96.7%</td>
<td>85%</td>
<td>95%</td>
</tr>
<tr>
<td>Stops</td>
<td>80%</td>
<td>95%</td>
<td>90%</td>
<td>95.8%</td>
</tr>
<tr>
<td>Fricatives</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Sibilants</td>
<td>82.5%</td>
<td>100%</td>
<td>85%</td>
<td>97.5%</td>
</tr>
<tr>
<td>Nasals</td>
<td>72.5%</td>
<td>95%</td>
<td>55%</td>
<td>85%</td>
</tr>
<tr>
<td>Place (overall)</td>
<td>99.6%</td>
<td>99.6%</td>
<td>99.6%</td>
<td>100%</td>
</tr>
<tr>
<td>Bilabials</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Labio-dentals</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Alveolars</td>
<td>99%</td>
<td>99%</td>
<td>100%</td>
<td>100%</td>
</tr>
<tr>
<td>Velars</td>
<td>100%</td>
<td>100%</td>
<td>97.5%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table II. The subjects' scores for the task involving perception of 12 consonants in an [aCa] frame via lipreading alone, lipreading plus the tacyile aid, the tactile aid alone (subject 2 only).
Fig. 5. Performance of Subjects 2, 3, and 4 for the visual alone (0) and visual-tactual perception (X) of 3 lists of the SPIN Test (Kallikow, Stevens, & Elliott, 1977).

Fig. 6. Tracking rates in words-per-minute obtained by Subjects 1 and 2 using De Filippo & Scott's (1978) Connected Discourse Tracking procedure in visual alone (0) and visual-tactual condition (X).
Fig. 7. Tracking rates (correct words-per-minute) and percentage correct scores for Subject 4 obtained visual alone (O) and visual-tactual (X) using the modified tracking procedure.

Fig. 8. The subjects' scores for the 3 subtests investigating perception of /s/. The numbers refer to the individual subjects.
detection of /s/ initially and finally but the results of Subtest 3 indicate that the subjects are also able to locate the temporal position of /s/ within the word. This would appear to be a much more complex task than that presented in Subtests 1 and 2 which merely required the subject to specify whether /s/ was present or absent. The mean scores obtained for the three subtests were 88.7%, 57.5%, and 83.3% respectively.

Word syllable number and type

The scores obtained by Subjects 1, 3, and 4 for the Syllables-in-Words test were 77.5%, 67.5%, and 62% respectively (Mean = 69.2%).

Discussion

The results obtained in this study indicate that the experimental aid offers much information which can serve as a very useful supplement to lipreading. The subjects' performances on the two consonant perception tasks showed that many of the confusions confronting the lipreader can be overcome by an aid presenting F₀ information as either a fixed or variable frequency. Features which cannot be perceived visually, for example voicing and nasality, can be reliably detected with the tactile aid. Analysis of the aided performance of all four subjects confirms this as do the results obtained by Subject 3 for the consonant perception task presented tactually. Reliable information as to consonant voicing and nasality would also seem to offer much useful information for the development of speech production skills. The subjects also evidenced much improved recognition of the sibilant consonants in the aided condition. The use of a frequency change to signal the presence of the sibilant consonants appears to be an effective method; a view supported by the excellent results obtained by the subjects for the /s/ detection tasks. The great importance of /s/ in signalling grammatical functions in English has been outlined by Rudmin (1983). A reliable means of sibilant detection such as is provided by the present aid would thus appear to be a most useful feature for both speech perception and production training.

The scores obtained for the tasks involving word and sentence perception also show the potential value of the aid. Subject 3's score for both word and sentence recognition rose by almost 20 percentage points in the aided condition. This subject appears to be able to use the information provided by the aid with a high degree of proficiency. It may be that his extensive experience in processing limited auditory information through his hearing aid may account for this ability. Subject 2 was also a successful hearing aid user prior to the onset of deafness and this experience may help explain her ability to use the vibrotactile aid. Conversely, Subject 4 appears unable at this stage in his training to use the tactual information as a supplement to lipreading of higher level materials. This subject's exposure to bisensory
inputs, however, is very limited and it may be that with additional training and experience this subject will be able to better use the tactual information as a supplement to lipreading.

The CDT scores obtained by Subjects 1 and 2 represent considerable improvements in performance in the aided condition. One problem with CDT, however, is that the scores obtained with a familiar speaker may not reflect a generalized skill. In an attempt to determine whether the improvements reported here were attributable to speaker familiarity, six sessions of CDT were conducted with Subject 2 using a female speaker who had not previously worked with the subject. The subject's mean unaided and aided scores for these six testing sessions were 46 wpm and 58.7 respectively. Thus there would appear to be few differences in performance with a familiar or an unfamiliar speaker. Subject 4's performance for the modified tracking procedure again showed no difference between aided and unaided performance but this may be overcome with more training.

One aspect of this testing program which must be taken into account is the limited exposure the subjects have had with the experimental aid in its present form. The most experienced subject (S2) has had only approximately 50 hours experience spread over a 12 month period. Despite this extremely short exposure period all the subjects evidenced skill in bimodal perception at the level of consonant perception and all but Subject 4 showed improvements in performance with higher level materials. Reed, Rabinowitz, Durlach, Braida, Conway-Fithian, & Schultz (1985) reporting on the speech perception abilities of a number of deaf-blind users of Tadoma, a tactile method of speech communication, hypothesized that the extensive training, the training the subjects had received, was a major contributing fact to their exceptional performance. Reed & al. (1985) cite the example of one of their subjects who was trained using the Tadoma method "for both receptive and expressive language one-on-one for several hours per day from age 5 to 21" (Reed & al., 1985; p. 254). They further note that "when attempts are made to compare Tadoma and artificial displays on normal subjects with comparable amounts of training, the superiority of Tadoma is greatly diminished" (Reed & al., 1985; p. 254). In order to adequately evaluate the aid described in this paper a wearable version needs to be developed and fitted to a number of subjects for daily use. Extensive training on a one-to-one basis will also be needed to ensure that the subjects receive optimal benefit from the aid. The excellent results reported for Tadoma users (Reed, Rubin, Braida, & Durlach, 1978; Reed, Doherty, Braida, & Durlach, 1982; Reed, & al., 1985; Norton, Schultz, Reed, Braida, Durlach, Rabinowitz, & Chomsky, 1977) provide an "existence proof for communication of speech through the tactile sense" (Reed & al. 1985; p. 247) and should encourage the development of tactile aids for both sighted and blind deaf persons. There would appear to be strong similarities between the sensory information available to the deaf-blind user of Tadoma and the sighted user of the aid discussed in this paper.
Reed & al. (1985) believe that the physical characteristics sensed by the Tadoma user are (i) laryngeal vibration, (ii) air flow, (iii) "in-out and up-down movements of the lips" (Reed & al., 1985; p. 251), and (iv) jaw movements. Sighted subjects should be able to detect (iii) and (iv) via lipreading alone while (i) and to a lesser extent (ii) may be produced by the tactile aid. If a wearable form of the aid can provide reliable cues to voicing and manner of articulation, it should prove to be a very useful aid to both receptive and expressive communication. One problem which will need to be confronted in the development of a wearable aid is the site of stimulation. In the studies conducted with this aid the output transducer has always been hand held. While this is acceptable in an experimental aid, alternative stimulation sites such as the wrist or sternum would appear to be better suited for every day use. Sherrick (1985) has noted that "we still face the seemingly trivial problem of whether the hands are uniquely suited to processing tactile signals, or if other sites do as well when properly addressed" (Sherrick, 1985; p. 219). It may be that the most acceptable solution to this problem, at least in the short term, would be to mount the output transducer on either the wrist or sternum and encourage the user to place his/her hand on it in communication situations. In this way the possibility of transfer to the alternative stimulation point may be assisted while ensuring that there is no information loss in the short term. In order to ensure presentation of an optimal signal even in noisy situations the use of a radio frequency input to the wearable aid should be considered. This would appear to be necessary if a reliable F0 signal is to be presented. The use of radio frequency input would also ensure that the presence of sibilant consonants is accurately signalled even at a distance.

The study has not answered the problem of whether F0 should be presented as a fixed or variable frequency. In part this is due to the varying amounts of the testing program completed by the individual subjects. One component of the test program completed by all subjects was the 20 consonant recognition task. At this level there appear to be few, if any, consistent differences between Subjects 1 and 2 who used the fixed frequency and Subject 3 and 4 with the variable signal. Where comparisons can be made with higher level materials, there again appear to be no differences in performance resulting from the two processing strategies. The problem of limited exposure to the alternate displays, however, should be taken into consideration. It may be that the only reliable way to determine whether one processing scheme is superior would be to provide training using both approaches with individual subjects and measure performance over time. Plant & Risberg (1983) in a study using an earlier version of the present aid found that an extended training period was required before their subject was able to reliably detect F0 contour changes in simple sentences. Obviously the development of a wearable device providing both processing strategies would greatly assist in determining whether one approach is to be preferred.
Fig. 9. Subject 4's production of the test sentence "I can see a cut" and "I can see a cart" recorded prior to the start of training (A) and 3 months (C) after commencement of training.
Subjects could be encouraged to experiment with both strategies in specific circumstances. The subjects' subjective reports coupled with objective testing should provide the information necessary to determine whether one approach is superior.

A final area which needs to be considered is the use of the aid for speech production training. As previously noted the emphasis in the training program provided to Subject 4 was towards using the experimental aid to improve his speech production skills. In an attempt to determine whether any improvement resulted from training spectrograms were made, pre-training of the subject saying each of the words "sheep, ship, cart, cut, dog, door, and show" in the carrier phrase "I can see a ???". Additional recordings were made three months and nine months after the commencement of training. Figs. 9 and 10 present examples of the subject's speech at these three times. The areas which appear to be most improved are the long/short vowel distinction and enhanced sibilant production. An example of Fig. 9, for example, shows very little difference in duration for [a] and [æ] pre-training. In both recordings made after the commencement of training, however, there is a long/short vowel rate of approximately 2:1. Similarly in the pre-training recording the sibilant [s] in "see" does not appear, but does, albeit weakly, in the later productions. In Fig. 10 it can be seen that the subject produced the vowels [i] and [I] with relatively normal length differences pre-training. The most obvious errors in the subject's speech in this case were the absence of [s] in "see" and [ʃ] in "ship" and "sheep". In the two latter recordings, however, both [s] and [ʃ] are present although again rather weakly. It would appear that the aid has assisted in the development of more normal speech patterns although it should be noted that the subject's speech is still relatively unintelligible. It is hoped to continue training with this subject on a long-term basis, and measurements will be made to determine the efficiency of the approach adopted.

The results reported in this paper indicate that the experimental aid offers useful information for enhanced speech communication. The production of a wearable version of the aid should provide many useful insights into its full potential and it is hoped that this will be the next step in its development.
I can see a ship I can see a sheep

Fig. 10. Subject 4's production of the test sentences "I can see a ship" and "I can see a sheep" recorded prior to the start of training (A) and 3 months (B) and 9 months (C) after commencement of training.
References


