A study using nonsense syllables has shown that a multiple-channel cochlear implant with speech processor is effective in providing information about voicing and manner and to a lesser extent place distinctions. These distinctions supplement lipreading cues. Furthermore, the average percentage improvements in overall identification scores for multiple-channel electrical stimulation and lipreading compared to lipreading alone were 71% for a laboratory-based speech processor and 125% for a wearable unit.

INTRODUCTION

Patients with a profound or total hearing loss do not get satisfactory help with a hearing aid. Consequently, there has been interest in producing hearing artificially by electrical stimulation of residual auditory nerve fibers, and recent studies have shown patients benefit by single channel or multiple-channel stimulation. This has been demonstrated by a variety of tests which have included the Utley and Norton speechreading measure, the Craig lipreading inventory, closed-set phoneme lists, vowel identification, consonant identification, phonetically-balanced word lists, and everyday sentence lists.

In order to evaluate the multiple-channel cochlear implant and speech processor developed at the University of Melbourne, we have considered it desirable to use nonsense syllables with a vowel-consonant-vowel (VCV) structure as a test procedure. We have done this as these syllables have been used to evaluate and improve the speech processing strategy used with our device, and consonant confusion tests using these syllables have been used to assess communication skills in normal and hard of hearing subjects.

METHODS

Two patients with a total hearing loss were implanted with multiple-channel receiver-stimulator devices on August 1, 1978, and July 17, 1979, respectively. The first patient had a sudden hearing loss following a head injury 18 months prior to surgery. Pure-tone audiometry under headphones and free-field revealed no hearing at any frequency, and the patient consequently received no benefit from a hearing aid. Polytome X-rays of the temporal bone demonstrated a fracture involving the superior semicircular canal, and electrical stimulation of the promontory was positive. The second patient had a progressive hearing loss. Pure-tone audiometry under headphones showed no hearing in the left ear or operated ear, and in the right ear a threshold of 105 dB HTL at 0.5 kHz and 90 dB HTL at 0.25 kHz. Speech audiometry under headphones at intensities of 120 dB SPL gave 0% scores. This patient received no benefit from a hearing aid. Electrical stimulation of the promontory was also positive.

The receiver-stimulator units were implanted in a bed created in the mastoid bone. They provided ten channels of stimulation, with independent control of the pulse rate and current level on each channel. Electrical stimulation was produced by biphasic current pulses with the leading phase negative and each phase fixed at approximately 180 μs. The current levels were varied from 70 μA in fifteen 70 μA steps to approximately 1 mA. Power and data were transmitted through the skin by inductive coupling. The electrode arrays from the receiver-stimulator units were inserted into the scala tympani of the basal turn of the cochlea through the round window, and passed for a distance of 20-25 mm. The electrode arrays had 20 individual electrodes so that ten channels of stimulation could be obtained, using alternate electrodes as a common ground. The array was fabricated from Teflon-coated platinum wires with a bare diameter of 0.025 mm. The wires were enclosed in a Silastic tube (Dow Corning) with an outside diameter of 0.6 mm, and connected to 0.3 mm strips of platinum foil. These strips were then wrapped around the Silastic tube, and the free ends welded together.

Following the operation on the first patient, electrical stimulus thresholds for 1000 pulse stimuli were obtained and these varied from 67 μA to 402 μA for the different electrodes, and there was a dynamic range of approximately 450 μA. A series of psycho-physical tests was then carried out. These showed that he experienced a change in vowel color and sharpness with the position of the electrode, a variation in pitch with pulse rate, and an increase in loudness with current level. It therefore appeared that a possible speech processing strategy was to transform the spectral emphasis of a speech signal to electrode position, the fundamental frequency to pulse rate, and the speech signal energy to current level.

This strategy was implemented initially using a laboratory-based speech processor. In the speech processor the voicing frequency of speech (FO) and its energy (AO) were extracted using a 400 Hz low-pass filter, and an energy threshold detector of AO was used to determine whether voicing was present or not. In the presence of voicing the pulse rate on electrodes was made proportional to FO. If voicing was absent, a constant low pulse rate was used as it produced a sensation described as “rough” which was similar to a “noise” sensation previously experienced by the patients when they had hearing. In addition, the frequency of the dominant spectral peak in the midfrequency range, called the sec-
second formant frequency (F2) and the energy (A2) associated with this peak were estimated from the output of a 750-4000 Hz band-pass filter. For a given F2 estimate, an electrode was selected from a predetermined F2-to-electrode transformation map, which had been constructed by ranking the electrodes from dullest to sharpest, and assigning frequency subbands to these electrodes in an order from lowest to highest. Similarly, the current level was determined on the basis of an A2-to-current level map, obtained by plotting current identified as calculated, in addition, the proportion of voicing, manner and place distinctions correctly made was determined. It was considered important to assess the value of the implant in helping the patients to make voicing, manner and place distinctions, as this would be a measure of how effective the device is in supplementing lip reading cues. Lip reading cues are mainly based on place distinctions, for example, the bilabial explosives p and b have a different place of articulation from the lingua-alveolars t and d, and this can be readily seen on the lips. Therefore, if adequate voicing distinctions are provided by electrical stimulation, this will greatly assist in communication. Furthermore, manner distinctions will be of help if these are not readily visible on the lips, and could help distinguish the plosive b from the nasal m.

A wearable speech-processor (Fig. 1) was also constructed to provide similar processing strategies to the laboratory-based device. It was tested in the same way although the stimulus conditions were not varied and the order was EV, VO, EO. It was considered that this would 1) maximize the fatigue effect for EO, and 2) there would be a better practice effect for VO following EV. Presenting the conditions this way would therefore be more likely to give better results for VO rather than EO. Furthermore, due to staff changes another audiologist carried out the male voice tests. The results were also plotted as confusion matrices.

**RESULTS**

The results of the nonsense syllable tests for the laboratory-based speech processor are summarized in Table 1. The order of EO, VO, EV had no significant effect on the results, and test-retest reliability was good, especially for patient 1. Furthermore, it can be seen from Table 1 that when tested with female and male voices, the use of the multiple-channel cochlear implant led to a very significant improvement in both patients’ abilities to speech.

---

**TABLE 1. IDENTIFICATION OF CONSONANTS: LABORATORY-BASED SPEECH PROCESSOR**

<table>
<thead>
<tr>
<th>Proportion Correct</th>
<th>Improvement with ES plus Vision vs Vision Alone (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test</strong></td>
<td><strong>Voice</strong></td>
</tr>
<tr>
<td>Patient 1</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Patient 2</td>
<td>F</td>
</tr>
<tr>
<td></td>
<td>M</td>
</tr>
<tr>
<td>Average</td>
<td></td>
</tr>
</tbody>
</table>

* *b, m, p, d, n, t, g, k.*

Improvement is defined as the difference between the implant plus vision and vision alone scores divided by the vision alone scores and expressed as a percentage.

Proportion correct - Number of phonemes correctly identified divided by number of times each phoneme was presented (20 stimuli).

**ES** - Electrical stimulation.

---

**Fig. 1. Photograph of patient 1 using the wearable speech-processor.**

---

-- CLARK ET AL. --
The results of testing the wearable speech processor on patient 1 are summarized in Table 3. The test-retest reliability was good. Table 3 shows that when the wearable speech processor was used in conjunction with lipreading, the scores were significantly better than with lipreading alone. The improvements were 156% for the female speaker and 92% for the male. When the results were analyzed to determine the proportion of voicing, manner, and place distinction correctly identified it was found (Table 4) that voicing and manner distinctions were significantly better with multiple-channel electrical stimulation alone compared to lipreading alone, and the scores for electrical stimulation plus lipreading were either as good as electrical stimulation alone or better. Furthermore, although the place distinctions with electrical stimulation were not as good as those with lipreading, they were nevertheless significant, and in the combined electrical stimulation and lipreading situation they were better than with lipreading alone.

**DISCUSSION**

The findings in this study show that a multiple-channel cochlear implant and speech processor which extracts the voicing frequency and energy, and the frequency and energy of the dominant spectral peak in the midfrequency range are effective in providing help with voicing and manner distinctions, and to a lesser extent with place information. Help with voicing and manner distinctions should supplement lipreading cues in understanding running speech. Furthermore, it was also shown that the overall scores were significantly better when using multiple-channel electrical stimulation in conjunction with lipreading compared to lipreading alone, and the improvements varied from 44 to 92% for the laboratory-based speech processor, and from 92 to 156% for the wearable processor. The improved performance of the wearable processor may have been due to a learning effect, and the fact that a different male voice was used for the tests. The performance was nevertheless significant and indicates that the laboratory-based speech processor strategy had been adequately realized in the wearable unit. It is also of interest that the results for voicing, manner and place distinctions were slightly better in patient 1 for the wearable speech processor when compared with those for the laboratory-based speech processor.

**REFERENCES**


Author/s:
Clark, Graeme M.; Tong, Yit Chow; Martin, Lois F.; Busby, Peter A.; Dowell, Richard C.; Seligman, Peter M.; Patrick, James F.

Title:
A multiple-channel cochlear implant: an evaluation using nonsense syllables

Date:
1981

Citation:

Persistent Link:
http://hdl.handle.net/11343/28742

File Description:
A multiple-channel cochlear implant: an evaluation using nonsense syllables

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.