EVALUATION OF A TWO-FORMANT SPEECH-PROCESSING STRATEGY FOR A MULTICHANNEL COCHLEAR PROSTHESIS

R. C. DOWELL, BSC, DIP AUD, P. M. SELIGMAN, BE, PHD; P. J. BLAMEY, PHD; G. M. CLARK, PHD, FRACS

Initial results with the two-formant speech-processing strategy (FOF1F2) confirm the advantage of a multichannel cochlear prosthesis capable of stimulating at different sites within the cochlea. The successful presentation of two spectral components by varying the place of stimulation leads to the possibility of presenting further spectral information in this manner. Because virtually all multichannel implant patients demonstrate good “place” (electrode site) discrimination, these more refined coding strategies should lead to benefits for the majority of implanters. Already, with the F0F1F2 strategy, we have a system that appears to provide some effective auditory-alone communication ability for the average patient.

Since clinical trials began with the Nucleus 22-channel cochlear prosthesis in 1982, the same speech-processing strategy has been used for all patients. This strategy codes the amplitude of speech signals onto the current level of stimulation, the fundamental frequency or voice pitch onto the rate of stimulation, and the second formant (or formant) onto the rate of stimulation. These acoustic features were chosen as conveying the most intelligibility, given that only three features are to be coded. The optimal method of coding these parameters was determined from intensive psychophysical research on a small number of multichannel implant patients. Subsequent clinical testing has indicated that all three of these acoustic parameters are conveyed effectively to patients and that the speech-processing strategy provides some understanding of unknown speech material without lipreading for almost all patients. Results to date show a mean score for open set standardized sentence material of 40% for patients with 12 months’ experience with the Nucleus prosthesis (R. C. Dowell, et al, unpublished observations).

However, it is known that full speech intelligibility is not possible with second formant (F2) information only, and that the first formant (F1) is also critical to speech understanding. Thus, if an effective method could be found to present F1 information without affecting the other acoustic parameters, the speech understanding of patients could be improved significantly. This paper presents some initial results with a new speech-processing strategy that provides F1 information in addition to the fundamental frequency (F0) and F2.

CODING STRATEGY

Psychoacoustic investigations have shown that electrical stimuli presented at two separate sites within the cochlea simultaneously can elicit two-component hearing sensations in deaf patients. This suggests that it may be possible to code two spectral components of running speech using such two-electrode stimuli. With the present system (F0F2), only one site within the cochlea is stimulated at any particular time. The place of stimulation, and hence the pitch of the elicited sound sensation, varies with the F2 of the signal. The new coding scheme (F0F1F2) uses an estimate of F1 to control the position of a second stimulation site within the cochlea. This strategy has been implemented by modifying a number of standard speech processors.

One of the problems of simultaneous electrical stimulation using more than one electrode pair is current interaction. The two electrical signals can combine causing large fluctuations in the perceived loudness depending on the relative positions and amplitude of the signals. In implementing the new coding strategy, this problem has been avoided by offsetting the two pulse trains by 800 μs. The Nucleus cochlear prosthesis typically uses pulse widths of 200 μs/phase, which means that the two currents can never interact because they do not overlap temporally. The perceived loudness of two-electrode stimuli can, in this way, be controlled easily and requires only a knowledge of each individual electrode’s loudness growth function (ie, the same information required for the standard F0F2 strategy). If current interaction was not avoided in this way, loudness information for each of the possible 200 combinations of electrodes would be needed. The large amount of psychoacoustic testing required would make such a system clinically impractical.

METHOD

To assess the effectiveness of the F0F1F2 strategy seven patients were fitted with new programmable speech processors modified for the presentation of F1 information. These patients had been using F0F2 strategies for periods ranging from 3 to 18 months. They were not selected in any way other than by their availability for testing and the ability of all seven patients to discriminate percepts elicited at different electrode sites within the cochlea (approximately 10% of patients are unable to discriminate electrodes because of early onset of deafness or degeneration of cochlear structures). Patients were assessed with the F0F2 strategy using the Central Institute for the Deaf (CID) open set everyday sentence test. This was presented live voice at average peak levels of 75 dBA. Patients had not been trained on this material and no contextual clues were given. Each patient then had 2 to 3 weeks of take-home experience with the F0F1F2 strategy and was restested with a different list of CID sentences. In addition, four of the patients were addressed on both occasions using continuous discourse tracking with auditory input alone. Tracking rates were assessed over a 10-minute period. This task required the patients to repeat verbatim an unknown text read out loud by a clinician. Responses were corrected using a hierarchy of strategies; however, no lipreading or visual cues were used at any time.

RESULTS

The CID sentence test scores for each patient are shown in Table 1. All seven patients scored substantially better on this test with the new F0F1F2 strategy. The mean improvement was highly significant for the group as a whole (mean difference = 32.4%, t = 5.79, df = 6, p < 0.01). The mean scores were 30.4% for F0F2 and 62.9% for F0F1F2.
TABLE 1. CENTRAL INSTITUTE FOR THE DEAF
EVERYDAY SENTENCE TEST

<table>
<thead>
<tr>
<th>Patient</th>
<th>FO/F2</th>
<th>F0/F1/F2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>54</td>
<td>90</td>
<td>+36</td>
</tr>
<tr>
<td>2</td>
<td>38</td>
<td>70</td>
<td>+32</td>
</tr>
<tr>
<td>3</td>
<td>31</td>
<td>54</td>
<td>+23</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>62</td>
<td>+46</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>26</td>
<td>+16</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>38</td>
<td>+18</td>
</tr>
<tr>
<td>7</td>
<td>44</td>
<td>100</td>
<td>+56</td>
</tr>
<tr>
<td>Mean</td>
<td>30.4</td>
<td>62.9</td>
<td>+32.4</td>
</tr>
</tbody>
</table>

$t = 5.79, df = 6, p < 0.01$.

Scores for the continuous discourse tracking test (Table 2) also showed large improvements for each patient with the new strategy. Mean tracking rates were 11.75 words per minute (wpm) for FoF2 and 30.5 wpm for FoF1F2, and this was highly significant (mean difference = 18.75 wpm, $t = 7.55, df = 3, p < 0.01$).

TABLE 2. CONTINUOUS DISCOURSE TRACKING RATES: AUDITORY INPUT ONLY

<table>
<thead>
<tr>
<th>Patient</th>
<th>FO/F2</th>
<th>F0/F1/F2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>24</td>
<td>42</td>
<td>+18</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>29</td>
<td>+16</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>20</td>
<td>+15</td>
</tr>
<tr>
<td>7</td>
<td>5</td>
<td>31</td>
<td>+26</td>
</tr>
<tr>
<td>Mean</td>
<td>11.75</td>
<td>30.5</td>
<td>+18.75</td>
</tr>
</tbody>
</table>

$t = 7.55, df = 3, p < 0.01$.

DISCUSSION

These results indicate that after a relatively small amount of experience with the FoF1F2 strategy, patients show substantial improvement in speech understanding without lipreading. The mean sentence score obtained with the new scheme of 62.9% is enough to allow effective interactive auditory communication for all of these patients. Not only the best patients improve (eg, patients 1 and 2), but patients performing only at average levels for the FoF2 strategy are improved to the extent that speech tracking at reasonable rates is possible without lipreading (eg, patients 4 and 7).

These results are consistent with those obtained with an acoustic model of the cochlear prosthesis. The acoustic model studies indicated that better speech recognition was obtained at the continuous discourse level because of the addition of F1 information and that this information improved scores for vowel recognition but had little effect on consonant recognition. Further investigations are now being undertaken with implant patients to evaluate these additional factors relating to the new strategy.

It may come as something of a surprise that the patients' subjective reactions to the new speech-processing strategy were at first generally negative. They immediately identified two components in speech input and in most cases interpreted the lower pitch component (F1 signal) as background noise interfering with the voice. It appears that after many months of listening experience with the FoF2 strategy, the patient's attention was directed to the F2 component of the signal to the exclusion of F1. However, after a short time, all patients adapted to the new scheme and reported improvement in the clarity and naturalness of speech, better recognition of environmental sounds, and improved communication ability, particularly without lipreading.

REFERENCES

Author/s:
Dowell, R. C.; Seligman, P. M.; Blamey, P. J.; Clark, Graeme M.

Title:
Evaluation of a two-formant speech-processing strategy for a multichannel cochlear prosthesis

Date:
1987

Citation:

Persistent Link:
http://hdl.handle.net/11343/27246