COMPARISON OF THE SPEAK (SPECTRAL MAXIMA) AND MULTIPEAK SPEECH PROCESSING STRATEGIES AND IMPROVED SPEECH PERCEPTION IN BACKGROUND NOISE


Melbourne
Victoria

As more is known about speech processing for Cochlear Implant patients, results should continue to improve. It now appears possible that Cochlear Implant patients may, in some instances, reach performance levels that are better than those obtained by most severely deaf people who use hearing aids.

During the last 15 years, our research has resulted in improvements in speech perception for multi-channel cochlear implant patients. This is shown in Figure 1, where it can be seen that the open set sentence scores for electrical stimulation alone were 18% for the inaugural FOF2 processor, 34% for the FO/F1/F2 or WSPIII processor, and 60% for the present commercially available device, the Multipeak speech processor. Figure 2 summarizes those speech processing strategies. The FOF2 speech processor codes acoustic amplitude as electrical charge and codes the fundamental or voicing frequency as pulse rate. The pulses are presented non-simultaneously to different electrodes which code mid-frequency spectral peaks, or second formants, as place of stimulation. The improved FO/F1/F2, or WSPIII, processor also codes the low frequency spectral peaks, or first formants, as place of stimulation. A further improvement occurred when we also presented the outputs from three high frequency filters to three fixed basal electrodes, and this is the Multipeak speech processor.

In order to improve speech processing for cochlear implants further, we have developed a very fast and flexible speech processor which allows us to test a range of alternative strategies. This speech processor, developed in the Department of Otolaryngology at the University of Melbourne, uses advances in digital signal processing technology. It can carry out 40 million operations per second and can be programmed with any type of speech processing strategy, which is shown in Figure 3. We have
FlC. 3. The Research Speech processor which uses a digital signal processing chip which was developed by the Human Communication Research Centre, Department of Otolaryngology, University of Melbourne.

THE SPECTRAL MAXIMA SPEECH PROCESSING STRATEGY

Filterbank Analyser

Apical Electrode Array

Basal

FIG. 4. A diagram of the spectral maxima speech processing strategy.

used this speech processor to evaluate the alternative strategy shown in Figure 4. With this strategy we select the six maximal outputs from a bank of sixteen band pass filters, and code these frequencies as place of stimulation. However, instead of presenting the fundamental or voicing frequency as rate of stimulation, the electrodes are stimulated at a constant rate and the variations in speech amplitude convey the fundamental or voicing frequency. This strategy has been called the Spectral Maxima Speech Processor or SMSP.

We compared SMSP and the Multipeak strategies on a small group of our research patients and the results for three patients are shown in Figure 5. These initial results showed that for electrical stimulation alone, speech perception was better for the SMSP device. For the sentence tests, the SMSP results were close to 100% correct. The two speech processors were also compared to see how well they performed in the presence of background noise. The results are shown in Figure 6 for four patients at signal-to-noise ratios varying from 20 down to 5dB. The noise was Multipeak babble, and average sound levels were taken for the speech and noise signals. The SMSP performed better than the Multipeak MSP processor in the presence of background noise.

In view of the encouraging results with the SMSP speech processor, Cochlear Pty Limited have developed SMSP as a commercial device called SPEAK. SPEAK is basically the same as SMSP. The main difference is that SPEAK extracts six spectral maxima from the outputs of 20, rather than of 16, band pass filters. The SPEAK strategy has been compared with the present Multipeak strategy in a clinical trial for the FDA in the US. This has been undertaken on patients in Melbourne, Sydney and Centres in the US. The results for Melbourne patients are similar to those obtained overall, and are shown in Figures 7 and 8. The results are

SPEAK Field Trial (Melbourne)

(n = 12)

% correct

S.I.T. sentences in quiet

FIG. 6. Open-set sentence scores for electrical stimulation alone for different signal-to-noise ratios in four patients using the Multipeak and Spectral Maxima speech processors.

FIG. 5. Closed-set consonant and vowel and open-set sentence scores for electrical stimulation alone using the multipeak and Spectral Maxima speech processor.

FIG. 7. The open-set sentence scores for electrical stimulation alone in 12 Melbourne patients using the MSP (Multipeak) and SPEAK speech processors in quiet.
The patterns of electrode stimulation for speech sounds for both the Multipeak and SMSP Processors are shown in Figures 9 and 10. The pattern is more similar to speech sound for the SPEAK strategy. The improvements in speech perception seen with the SPEAK strategy may therefore be due to better representation of the speech signal.

The improvement in Cochlear Implant Speech Processing with the SPEAK strategy means that we can now consider patients for implantation who have some residual hearing and use a hearing aid. This is possible as implant results are now, on average, better for implantation than for a hearing aid in people with a severe to profound hearing loss.

To illustrate this, we would like to present the case of a patient, Mrs CP. This woman used a hearing aid in her better ear, and obtained a 60% aided sentence score in this ear. She had trouble communicating in a noisy environment and was limited in the use of the telephone at work. As a result of our recent advances in cochlear implant speech processing, we can now say that, on average, she would do better with a cochlear implant and the SPEAK strategy. However, a decision to operate was made difficult by the fact she had advanced cochlear otosclerosis. We needed to know how this condition would affect her chances of success. To help make the decision we reviewed the results for our otosclerotic patients and found them to be, if anything, better than the average, unless the disease limited us in the number of electrodes we could use. This could occur from bone growth in the basal turn reducing the number of electrodes inserted, or it could be from bone absorption in the region of the horizontal position of the facial nerve, requiring electrodes to be switched off to prevent facial nerve stimulation.

In the case of Mrs CP, standard CT scans were difficult to interpret in order to see whether bone growth or bone absorption would affect the outcome. For that reason we made a 3D reconstruction of the cochlear CT scans, and found this to be very helpful in providing a clear picture of cochlear pathology. As a result of our preoperative evaluations we were able to advise the patient that there would be a reasonable chance of success if she had a cochlear implant. The patient had a cochlear implant on 24 August, 1993. Post-operatively, she obtained a 100% open-set sentence score for electrical stimulation alone and now communicates effectively at work. She is better able to use the telephone, and uses the implant in preference to her hearing aid. We are also undertaking studies on her to see if we can combine auditory information from the hearing ear, and information from the cochlear implant in the other ear, in order to maximise her speech perception.

As the ability to hear with a hearing aid or a cochlear implant in the presence of background noise is very important, we have been undertaking research to see how to design microphones in order to obtain a clearer speech signal. This method is called adaptive beam forming and enables a patient to hear a clearer speech signal if one is facing the speaker and noise is coming from one side. In this case the signal in the right ear will be $S+N_1$ and the left ear $S+N_2$. The signal goes into a summer which has an output $2S+N_2+N_1$ and a subtractor which has an output $N_2-N_1$. The next stage (FIR) adjusts the $N_2-N_1$ to approximate $N_2+N_1$ so that one should finish up with the signal $2S$ or $S$.

Results using the automatic beam forming system for
cochlear implant patients has been compared with simply summing the inputs from two microphones. In quiet, sentence results were the same, but at a 0dB signal-to-noise ratio the adaptive beam forming results were much better. A 0dB signal to noise ratio is hard to hear even for normal hearing listeners, as it means that the background speech noise is at the same intensity as the speech being listened to.
Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Clark, Graeme M.; Whitford, L.; Van Hoesel, R.; McKay, C. M.; McDermott, H. D.; Seligman, P.; Vandali, A.; Pyman, B. C.; Cowan, R. C.

Title:
Comparison of the SPEAK (Spectral Maxima) and multi-peak speech processing strategies and improved speech perception in background noise

Date:
1995

Citation:

Persistent Link:
http://hdl.handle.net/11343/27447