INTRODUCTION

This paper briefly describes the history of speech processing developments leading to the presently available Speak processing strategy. The similarities and differences of the Speak and Continuous Interleaved Sampling (CIS) strategies are then discussed and some recent key experimental observations are examined as a guide to potential future coding strategies. Key issues for future coding strategies and implant designs are the number of electrodes and stimulation rates in use. Consideration of these issues has led to development of a prototype implant to be used for advanced speech-processing research.

HISTORY

At the University of Melbourne, the first speech processing experiments were carried out in 1979. The speech processor used a set of 10 band-pass filters with each filter independently controlling a sequence of biphasic current pulses delivered to a corresponding electrode, with the intention of modeling the neural activity of a normal cochlea. However, the interactions between electrodes led to large and unpredictable fluctuations in loudness, and this concept was quickly abandoned.

It was then decided that only 1 electrode should be stimulated at a time, as doing this would make it impossible for the current pulse delivered by 1 electrode to interfere directly with the pulse delivered by any other electrode. This form of stimulation was originally described as "sequential stimulation" and has more recently also become known as "interleaved stimulation."

In 1979, it was anticipated that the cochlear implant would be an assistance to lipreading, so the first simple strategy (F0F2) was used to deliver information that cannot be seen on the lips. A single electrode was selected according to the mid-spectral peak of the second formant (F2), which is formed by structures of the vocal tract largely hidden from view. This selection was based on the theory of tonotopic organization of the cochlea, in which low tones are perceived in more apical positions and higher-frequency tones in more basal positions. The pulse rate was set to the fundamental frequency (F0) of voiced speech or at a randomly generated rate in the absence of voiced speech.

With increasing experience and development of speech processor technology the trend was to deliver further spectral information on the available array of electrodes. In 1985, the F0F1F2 strategy added the selection of a second electrode to represent the energy of the first formant of speech (F1). In 1989, the Multipeak (Mpeak) strategy added more high-frequency information from three fixed filters, allowing a selection of up to 4 electrodes to be stimulated sequentially to represent the sound spectrum.

The Speak coding strategy was introduced in 1994 and delivers the greatest amount of spectral information currently available with a cochlear implant system. With Speak, groups of electrodes are selected for sequential stimulation. These groups represent the most energetic frequency peaks in the sound spectrum.

When considering the electrode selection made with the Speak coding strategy for a hypothetical vowel, the pulse rate to the selected electrodes depends on several factors, but ranges from 250 to 300 pulses per second (pps). This is in the same range as the rate used to represent unvoiced sound in the earlier strategies. Current spread from adjacent closely spaced electrodes will tend to stimulate many nerves in common, with the result that the stimulation rate of an individual nerve is the sum of the rates of all the electrodes to which the nerve is responding. Given that any spectral peak is represented by 2 to 4 electrodes, the effective pulse rates will range from 500 pps to approximately 1,200 pps.

The left panel of the Figure shows the 20-channel Speak electrodogram for the word "choice." The vertical axes show the electrode numbers and the corresponding frequency scale. Each vertical bar indicates the stimulation of an electrode. Speak provides a more comprehensive and complete model of the spectrum than did the earlier coding strategies. When considering the temporal capability of the Speak strategy it is important to recognize that although the rates available on individual electrodes are lower than with the CIS strategy, the density of the electrodes spaced at 0.75 mm is providing a combined neural stimulation rate dependent on the sound energy available within a particular spectral peak. For example, the vowel transition on the Speak electrodogram is represented on 3 or 4 adjacent electrodes providing an effective neural stimulation rate of 750 pps to 1,200 pps to nerves at the center of this transient electrode cluster.

The right panel of the Figure shows the electrodogram for a 6-electrode CIS strategy, again for the same word, "choice." The CIS method of speech processing was developed for patients who use only 6 intracochlear electrodes. The 6 electrodes were spread over the same distance as the 22 electrodes of the Nucleus implant. Consequently, these 6 widely spaced electrodes have less potential to represent fine-resolution spectral information than closer-spaced electrodes. In optimizing the 6-electrode CIS strategy for adults, Wilson et al. used pulse rates in the range 800 pps to 2,500 pps. They found that using a fixed stimulus rate and a strictly sequential electrode sequence, subjects would report an annoying buzz if the stimulation rate was below approximately 800 pps, and that in
some cases performance was improved if the rate was increased to levels at which the buzz was no longer perceived.

With Speak, the effective neural stimulation rates are within the optimum range found for CIS. However, with Speak it remains to explore the effective neural stimulation rates at the highest rates of the CIS strategy.

FUTURE DIRECTIONS

There is other evidence that perhaps a higher number of electrodes may contribute to improved performance and that future directions of investigation should include more sophisticated electrode selection at higher rates of effective neural stimulation.

McDermott and McKay investigated stimulating adjacent electrode pairs and found that not only could the percepts produced by many adjacent electrode pairs be reliably distinguished, but also that stimulating 2 adjacent electrodes produced an intermediate place-pitch perception. This result offers some explanation for the clinical benefit observed with Speak in comparison to Mpeak. It also supports the implant design trend toward more closely spaced electrodes as opposed to fewer, widely spaced electrodes.

Von Wallenberg et al (this suppl, section 16) undertook studies at Medizinische Hochschule Hannover to investigate patients' abilities to discriminate electrode percepts. A modified Nucleus implant that had 20 intracochlear electrodes and 2 extracochlear electrodes was used. A comparison was made between electrode discrimination ability using bipolar and monopolar stimulation pulses, and it was found that discrimination ability was similar for both modes.

As monopolar stimulation has a much greater current spread than bipolar stimulation, this result of equal discrimination ability suggests that the place-pitch percept may depend not only on the tonotopic location of the nerves in the cochlea, but also on the relative firing times of the nerves in response to each stimulation as well. This suggests that complex stimuli sequences might be used with arrays composed of even more closely spaced electrodes.

Pyman and Clark (this suppl, section 18) have described the mechanical and surgical considerations behind a stimulator that has been developed to prototype stage. This stimulator is connected to the Nucleus 22-electrode array and has in addition 2 extracochlear electrodes. It is capable of combined stimulus rates exceeding 15,000 pps. The combination of high rate and large number of closely spaced electrodes allows investigation of the effects on performance of complex stimulus sequences designed to take advantage of intermediate place-pitch percepts and of the effect on perception of relative nerve firing times.

In addition, the prototype implant includes neural response telemetry circuits, described by Carter et al (this suppl, section 2) that provide a valuable tool for investigating nerve firing patterns and may provide information to allow objective adjustment of the system.

CONCLUSION

A history and analysis of strategies has been presented and some key experimental results have been examined that together open new avenues for investigation: closely spaced electrodes in combination with higher rates and complex stimulus sequences to take advantage of intermediate place-pitch perception, and the effect of relative neural firing time. These capabilities are now combined with the ability to investigate nerve firing patterns with the addition of neural response telemetry. These additions to the implant design have the dual objectives of improving speech coding strategy performance and providing objective information on neural activity to assist optimum device fitting.

REFERENCES

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Patrick, J. F.; Evans, A. R.

Title:
Implant designs for future coding strategies

Date:
1995

Citation:

Persistent Link:
http://hdl.handle.net/11343/27467

File Description:
Implant designs for future coding strategies

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.