# The University of Melbourne/Nucleus Multiple-Channel Cochlear Implant

**Graeme M. Clark**
**Department of Otolaryngology,**
**The University of Melbourne, and**
**The Australian Bionic Ear & Hearing Research Institute,**
**East Melbourne**

*Abstract: The history of the development of the multiple-channel cochlear implant by a research team at the University of Melbourne, in collaboration with Nucleus Ltd, is reviewed and related to various strategies for speech processing and cochlear stimulation for the profoundly deaf. The results of clinical trials are summarised and evaluated.*

## REVIEW

In the 1960s the fundamental question was: could direct electrical stimulation of auditory nerve fibres help profoundly-totally deaf people understand connected or running speech? Profoundly-totally deaf people have very little residual hearing due to loss of the organ of Corti, and could therefore not be assisted by amplifying speech sounds with a powerful hearing aid.

As frequency information was known to be important in recognizing phonemes and words it was considered necessary to see how well electrical stimulation could reproduce the coding of frequency. Frequency was thought to be coded by spatial and/or temporal mechanisms. Spatially this was through the place or site of stimulation of the auditory centres. Temporally it was through the rate of nerve firing or more correctly the time intervals between nerve action potentials.

Studies specifically designed to investigate these questions on anaesthetized experimental animals showed that electrical stimulation could not reproduce the sustained firing seen with acoustic responses at rates above 200-500 pulses/s. The probability of the neurons firing at each pulse interval was greater for electrical stimulation than for sound (Clark, 1969). Furthermore, it was demonstrated that periodic electrical stimuli were encoded in the discharge of neurons up to a rate of 400-600 pulses/s (Merzenich et al., 1973). As the data from the above initial study (Clark, 1969) had been obtained from the auditory pathways in the anaesthetized animal, it could not be concluded the findings were necessarily applicable to the coding of frequency in the intact alert animal. To help resolve this matter a series of behavioural studies was undertaken on the experimental animal (Clark et al., 1972; Clark et al., 1973). These showed the animals had severe limitations discriminating different rates of stimulation above 200-600 pulses/s.

As our initial research on acute preparations and behavioural animals showed the limitations of using electrical stimulation to code frequencies above 500 pulses/s on a temporal or rate coding basis, it was considered that a single-channel implant would not be satisfactory for conveying important speech frequencies above 500Hz. It was necessary to convey frequencies up to approximately 4000Hz for adequate speech comprehension. For this reason, in 1971-72, we embarked on the development of a fully-implantable, multiple-channel (multiple-electrode) cochlear implant that would thus enable us to code frequencies above 500Hz (essential for speech intelligibility) on a place coding basis.

It was considered that the device needed to be fully-implantable, as a plug and socket could result in infection extending from the site where the socket emerged through the skin to the tissues within the body. This meant we had to design the electronic circuitry of the implantable receiver-stimulator unit so that a number of possible speech processing strategies could be evaluated on our patients. A fully-implantable system would not be as "transparent" as a plug and socket.

Having embarked on the development of a multiple-channel implant it was also necessary to study how best to site the electrodes within the cochlea and provide appropriate electrical stimuli, to both minimize trauma to the neurons we hoped to stimulate as well as localize the electrical current to discrete groups of nerve fibres for place coding. This involved computer modelling and animal experimental studies (Black and Clark, 1977; Black and Clark, 1978; Clark, 1973; Clark et al., 1975). The possibility of localizing the current to discrete groups of nerve fibres by using bipolar stimulation was demonstrated by Merzenich (1974).

Our first cochlear implant operation on a postlinguistically deaf adult with a profound-total hearing loss was carried out on 1 August 1978 (see Figure 1). Postoperatively we undertook psychophysical studies to explore his percepts for variations in place, rate and intensity of stimulation. They confirmed the limitations of discriminating rate of stimulation above 200-300 pulses/s seen previously with our studies on experimental animals. They showed that place pitch was different from rate pitch. It had timbre which varied from sharp to dull depending on whether a basal (high frequency) or apical (low frequency) electrode was stimulated respectively. The patient could also scale place pitch well for site of stimulation indicating that the current spread for common ground stimulation with our banded electrodes provided satisfactory current localization. Common ground stimulation, available with the prototype implant, resulted in current flowing from an active electrode to all the other

*Figure 1. A diagram of the Nucleus implantable receiver-stimulator and wearable speech processor unit for clinical trial by the FDA illustrating the components. Speech is picked up by the directional microphone and electrical signals sent to the speech processor for coding. The coded signals and power to operate the implant are transmitted through the intact skin by radio waves (inductive coupling). The implant decodes the signals and directs them to the electrodes in the array lying around the basal turn of the inner ear.*

electrodes on the array connected together electronically. Finally, studies on loudness showed this varied primarily with stimulus current level and the loudness growth due to increase in current was much steeper than for acoustical stimulation in normal hearing subjects.

Our first speech research involved presenting speech through a processor which modelled basilar membrane motion and the fine timing of auditory nerve firing. We did this in the belief that in spite of the limitations on the discrimination of rate of stimulation the closer we could get to simulating the spatio-temporal patterns of auditory nerve firing with speech sounds the more chance we would have of the patient understanding running speech. However, this strategy did not provide significant help, and this was thought to be due in part to unpredictable variations in loudness from summation of electrical fields with simultaneous stimulation on neighbouring electrodes.

The most important clue in developing a speech processing strategy to help our patient understand running speech came when he was asked to mimic the sounds he experienced when different single electrode sites were stimulated. He imitated the percepts as vowels and the vowel varied according to the place of stimulation. In explaining this relationship and in developing an appropriate speech coding strategy for multiple-electrode stimulation we considered it important to determine what equivalent acoustical signals could produce the spectral colours (/ɒ/, /ɔ/, /ɛ/, /ɪ/, /i/) or (/hot/, /cord/, /head/, /hit/, /seat/) perceived by the patient (Tong et al., 1979). We were helped in this regard by a study by Delattre et al., (1952) who had shown that single-formant synthetic vowels with formant frequencies at 720Hz, 2160Hz and 3000Hz were identified respectively as having vowel colours corresponding to /ɔ/, /ɛ/, /ɪ/. This suggested that the spectral colours produced by electrical stimulation on single electrodes could be approximately equivalent to those produced by acoustical signals with their spectral emphases centred at frequencies ranging from 720Hz to above 3000Hz. Research on our patient also showed that we could shift the vowel perceived from /ɔ/ to an /a/ or /cart/ by increasing the current level and we attributed this to a shift in the population of neurons stimulated and an averaging process referred to by Plomp (1976) for spectral components of acoustic stimuli.

We also carried out psychophysical studies using two electrodes stimulated simultaneously and at the same rate. The vowel perceived was different from that for each one stimulated separately. When we stimulated one electrode resulting in /ɒ/ and combined this with another resulting in /ɪ/ the percept was/ʌ/ or /hut . This suggested that an averaging process was taking place in the brain when the two populations of neurons were excited for spectral components in the two vowels. Another important finding was that when two nearby electrodes were stimulated simultaneously with rates which differed by approximately 5%-7% the patient perceived a consonant and following vowel. With the electrodes used this was /də/ or /dirt/. Data from the study suggested that beating was occurring, and that amplitude changes from the stimuli moving in and out of phase were responsible for the perception of the consonant in conjunction with the vowel. The stimuli were representing some of the acoustic cues for the coarticulation of consonants and vowels (Tong et al., 1979).

In deciding how to implement a speech processing strategy we considered we should extract one formant frequency initially, and code this as place of stimulation. We had seen that single electrode stimulation could produce vowel colours that were equivalent to those for the single formant representation of vowels. The relationship between a single formant representation of the first and second formants in vowels was, however, complex and not well understood. For that reason it was more appropriate to extract the second formant frequency as it had been shown from research by Potter et al., (1947), Cooper et al., (1952), Liberman et al., (1954), Fant (1956) and others to be the formant providing most speech information. It was coded as place of stimulation because it was in a high frequency range and our initial psychophysical studies on the patient had shown that high frequencies could be coded as place pitch.

The decision to extract the second formant frequency was a change in the direction of our research. It meant preprocessing of speech information, selecting important cues and determining how best to present the information to the nervous system. If electrical stimulation of the nervous system produced an information "bottleneck" as indicated by our animal experimental and human psychophysical data, it was important to be selective with the information presented through the "bottleneck". As a result of this decision we also extracted the fundamental or voicing frequency as it is not visible on the lips and its presentation would therefore be a great help as a lipreading aid. Voicing which is low in frequency was appropriately coded on a rate or temporal basis because our previous experimental animal and human psychophysical findings had shown that frequencies up to 300-500Hz could be coded this way.

In addition, it was considered important to present the coded fundamental frequency only to the individual electrodes presenting second formants on a place coding basis. This would ensure that the unpredictable variations in loudness that occurred with summation of electrical fields from simultaneous stimulation on two or more neighbouring electrodes with the physiologically-based processor would not occur. Finally, the intensity of the second formant was

coded as current level on each electrode used to code second formant frequencies.

This inaugural formant or cue extraction speech processing strategy was first implemented on a laboratory-based computer in 1978, and shown to provide the patient with approximately a 400% improvement in understanding running speech when using electrical stimulation combined with lipreading compared to lipreading alone. He was also able to understand some running speech when using electrical stimulation alone. A second postlinguistically deaf adult who had been profoundly deaf for 13 years rather than two years in the case of the first patient received a multiple-electrode cochlear implant on 13 July 1979. He obtained results that were similar to the first patient suggesting that the coding strategy might be generally applicable. It was also an important finding that even after 13 years of no auditory input he was still able to remember the sounds of speech and this showed his auditory pathways could process the coded speech information. In April 1980 the speech processing strategy was realized as a portable unit with dimensions 15x15x6.5cm and weight of 1.25Kg.

The laboratory-based speech processor consisted of three basic sections: speech signal parameter extraction; encoding of speech parameters to electrical stimulus parameters; and digital configuration of electrical stimulus data. Four speech signal parameters were estimated every 20 ms in the parameter extraction section. The fundamental (voicing) frequency (FO) and a low frequency energy measure (AO) were estimated by measuring the period and the average amplitude of the output waveform of a lowpass filter. The second formant frequency (F2) and its amplitude (A2) were estimated by measuring the zero crossings and average amplitude of the output waveform of a bandpass filter. The speech parameter estimates (FO, AO, F2, A2) were transformed to electrical stimulus parameters every 20 ms in the encoding section. Only one electrode was activated in any 20 ms time frame. For a given F2 estimate an electrode was selected according to a predetermined F2-to-electrode transformation map: the F2 frequency range was divided into nine subbands; each subband was assigned to a particular electrode. The subband of lowest frequency was assigned to the electrode with the dullest sensation, while the subband of highest frequency was assigned to the electrode with the sharpest sensation. The current level for the single-electrode pulse train was determined from A2. A 20 ms speech segment was classified as voiced if AO exceeded a pre-selected threshold, and unvoiced otherwise. For unvoiced speech segments, a constant low pulse rate for electrical stimulation was used. This low rate was used as it produced a sensation described as 'rough' which is the closest response to that of noise perceived by normal hearing subjects. For voiced speech segments, the pulse rate was proportional to FO, and was higher than the pulse rate used for the unvoiced segments. Given the electrode selection, current level and pulse rate for a 20 ms time frame, the digital configuration section of the speech processor formatted the electrical stimulus data, and transferred these data to the external transmitter unit (Tong et al., 1980a). The portable speech processor used the same overall coding strategies as the laboratory-based speech processor but analyzed speech every 10 ms rather than 20 ms.

Establishing the benefits of the inaugural FO/F2 speech processing strategy required setting up a series of speech perception tests in standardized conditions. This involved making sure that the test material was prerecorded, not previously presented, and administered in controlled conditions. To show the benefits of the implant in understanding running speech not only did we use the phonetically-balanced AB word test (Clark et al., 1981a) but also the CID everyday sentence test (Clark et al., 1981b) and then later the Tracking test (Martin et al., 1981). The testing was also carried out for electrical stimulation alone, lipreading alone and electrical stimulation combined with lipreading.

The results of the initial testing are reported in more detail elsewhere (Clark et al., 1981 a, b; Martin et al., 1981). In summary, however, with AB words the first patient obtained a score of 10% for words and 20% for phonemes using electrical stimulation alone, and an improvement for electrical stimulation combined with lipreading compared to lipreading alone of 300% (words) and 38% (phonemes). With CID sentences he obtained a score of 14% for electrical stimulation alone and a 386% improvement for electrical stimulation combined with lipreading versus lipreading alone. Similar results were obtained on the second patient.

Although the inaugural speech processor was very effective in providing help in understanding speech when used as a lipreading aid it was of more limited help when used alone without lipreading. Furthermore, its performance was significantly degraded when the signal-to-noise ratio was reduced. For these reasons we undertook further psychophysical research to help determine how well stimuli of relevance to speech comprehension were perceived. We also analyzed the speech perception results to determine which features were most effectively transmitted.

One important question was whether rate (temporal) or place coding was effective in conveying time varying frequency information which is a feature of speech signals. With speech there is a slowly varying fundamental or voicing frequency and consonants, in particular, have frequencies that change rapidly over a duration of about 20 ms (Clark et al., 1987).

Research to answer this question (Tong et al., 1982) showed that when the place of stimulation was varied across adjacent electrodes the two patients could discriminate the transition well for durations of 25, 50 and 100 ms. On the other hand, when varying pulse rates on a single electrode , there was a marked degradation in performance from 100 to 25 ms. This finding helped establish the validity of our speech processing strategy where variations in place of stimulation for F2 could be discriminated for the short durations of consonants. Nevertheless there was still a need that future research should aim at maximizing the amount of information transmitted on a place coding basis over short durations in time. Furthermore, an interesting sequel was that the research helped show that acoustic frequency discrimination probably takes place on a place coding basis, i.e. when we perceive a change in frequency this occurs through a shift in site of stimulation. With acoustic stimuli it is difficult to separate temporal from place coding, and the relative importance of each was not well established. With electrical stimulation place and temporal coding can be artificially separated, and it provided fundamental information on the coding mechanisms involved in frequency discrimination. Although the research on time varying rate of stimulation showed perceptual limitations for

durations of 25 and 50 ms, satisfactory discrimination at 100 ms was sufficient to code the slow variations of the fundamental frequency and confirmed the suitability of the FO/F2 strategy.

A further question of importance was the coding mechanisms for loudness, and in particular the extent to which loudness was a function of repetition rate. Basic physiological studies had shown that to acoustic stimuli the mean firing rate of units increase over a 20-50dB intensity range, and that the population of neurons stimulated also increases. The interrelation between the stimulus rate and excited population is not well understood. For our speech processing strategy it was necessary to know that if we coded the fundamental frequency as rate of stimulation to what extent would variations in rate produce changes in loudness, and could this be controlled if necessary to improve the speech processing strategy. The study was undertaken by comparing loudness changes for stimuli with single pulses per period (SPP), and stimuli with multiple pulses per period (MPP). With the MPP stimuli the mean pulse rate could be kept constant over time while the period of each group of pulses was shortened and the repetition rate increased. The study showed that the MPP stimulus produced approximately equal loudness with variations in repetition rate. The data from the study also indicated that loudness is a function of the physical variables, charge per pulse and overall pulse rate. Loudness did not necessarily increase with charge per unit time (Tong et al., 1983a).

Another question in establishing the effectiveness of our speech processing strategy was the extent to which interaction occurred between repetition rate and place of stimulation. This was relevant as the strategy involved presenting the fundamental frequency as rate and second formant as place of stimulation. Any significant interaction between repetition rate and place of stimulation would lessen the effectiveness of the strategy. The study was undertaken by asking the subject to categorize the stimulus as a "question" or "statement" on the basis of whether there was a rising or falling pitch produced by a slow change in rate of stimulation. The rate trajectories were varied on three separate electrodes from apical, middle and basal sites to see if there was an interaction between rate and place pitch in coding the fundamental frequency of speech. The SPP and MPP modes of stimulation were also compared to help determine whether variations in loudness or pitch changes alone were responsible for making "question" and "statement" judgements. The results firstly showed there was a significant interaction between rate and place pitch for the apical electrode but not so for the electrodes in the middle and basal regions. The interaction for the apical electrode probably occurred because a low stimulus rate occurred at an electrode site in a lower frequency region of the cochlea. The interaction was thought to be of secondary importance for speech perception in cochlear implant patients. Secondly, the results showed the lack of any significant difference between SPP and MPP modes of stimulation. This suggested that rate pitch alone and not variations in intensity were being used in coding changes in fundamental frequency.

As the FO/F2 speech processing strategy presented the fundamental frequency across electrodes when coding F2, it was important to not only study the interaction between rate and place pitch for individual electrodes, but for varying rate of stimulation across electrodes. The categorization performance for the presentation of the fundamental frequency across electrodes was compared with that on separate electrodes, and no significant difference seen. There was also no difference whether rate was varied across electrodes in an apical to basal or basal to apical direction. Furthermore, the categorization performance for electrical stimulation was similar to that reported by Fourcin et al., (1979) for acoustical stimulation with fundamental frequency trajectories superimposed on the syllables "oh". This correspondence between acoustical and electrical results suggested that the perception of pitch contours for electrical repetition rate in the presence of a variation in electrode position was similar to the perception of pitch contours for fundamental frequency superimposed on a variation in the spectral envelope of an acoustical signal.

Finally, a study was undertaken to confirm that the percepts for rate and place pitch were separate. The study showed that dissimilarites between the two pitch percepts could be best explained in two dimensions of perceptual space. This confirmed both our initial findings that rate and place pitch were different, and the lack of interaction between the fundamental frequency and stimulation on middle and basal electrodes. The finding also helped establish the validity of the FO/F2 speech processing strategy.

Understanding how the FO/F2 processor was effective in helping profoundly deaf people understand running speech and how to improve it, particularly for electrical stimulation alone, required not only the psychophysical studies referred to above but a detailed analysis of speech perception performance. An initial analysis of speech perception showed that vowel recognition for electrical stimulation alone (77% mean score) was better than for consonants (35% mean score) (Tong et al., 1980b). As consonants are more important than vowels for speech intelligibility, improvements in the perception of the former were obviously necessary. For this reason we analyzed the particular features of consonants perceived in a number of studies (Tong et al., 1980a; Clark et al., 1981c; Dowell et al., 1982). The results showed that for voicing, 30% information was transmitted with electrical stimulation alone, 0% with lipreading alone, and 47% with electrical stimulation combined with lipreading. These results confirmed that voicing, which is not visible on the lips, was being effectively transmitted by the FO/F2 speech processing strategy. A similar trend was seen for the nasals (/m/, /n/). On the other hand the reverse applied for affrication (/f/, /v/, /s/, /z/) and place (/b/, /d/, /g/, /p/, /t/, /k/). In the case of affrication the percentage information transmission was 9% with electrical stimulation alone, 58% with lipreading alone and 75% with combined electrical stimulation and lipreading. For place the information transmission was 14% with electrical stimulation alone, 70% with lipreading alone and 74% with electrical stimulation combined with lipreading. These results suggested there was a need for the transmission of more high frequency information for improving the perception of affrication and the transfer of more information over short durations for the phonemes conveying place information.

While the above studies were being undertaken to study how the FO/F2 strategy might be improved, work was carried out with Nucleus Limited to help in the industrial development of the implantable receiver-stimulator and wearable speech processor for clinical trial for the US Food and Drug Administration (FDA). The implantable receiver-stimulator (Clark et al., 1983) was made more robust than the prototype developed by The University of Melbourne, and had the electronic circuit streamlined in the

light of findings with the prototype device (Figure 1). The speech processor (Clark et al., 1983) was made smaller (Figure 1), and it implemented the inaugural FO/F2 strategy. We also undertook a series of additional biological studies at this time that were subsequently of value in helping to establish for the FDA that the device was safe as well as effective (Clark et al., 1987).

The FDA clinical trial was carried out initially by centres at The University of Melbourne; University of Iowa; Baylor College of Medicine, Houston; Mason Clinic, Seattle; New York University; Good Samaritan Hospital, Portland; University of Toronto; Louisiana State University, New Orleans; Medizinische Hochschule, Hannover; and the University of Sydney. This wider clinical trial confirmed the initial results obtained by The University of Melbourne. On 40 patients it showed a mean improvement in CID open-set sentence tests from 52% for lipreading alone to 87% for lipreading combined with electrical stimulation three months postoperatively (Dowell et al., 1986). For electrical stimulation alone the mean open-set CID sentence scores on 23 patients were 16% (range 0-58%) three months postoperatively, and 40% (range 0-86%) 12 months postoperatively. This trial also showed that there was considerable variation in patient performance (a finding in all subsequent trials and for all devices to this day), and that significant learning was required to effectively use the speech sounds induced by electrical stimulation.

At the time there was considerable debate as to whether multiple-channel (multiple-electrode) stimulation was really more effective than methods of single-channel stimulation. Reports occurred from different centres which did not take into account patient variation and which used a variety of assessment procedures. For this reason a controlled comparative study was undertaken by the University of Iowa on the 3M House and 3M Vienna single-channel systems, and The University of Melbourne/Nucleus and Utah multiple-channel systems (Tyler et al., 1987, Gantz et al., 1987). The 3M House system had the speech waveform modulating a 16 kHz carrier on a monopolar electrode, the 3M Vienna system a spectrally weighted waveform on one of four common-ground electrodes, and the Utah system bandpass filtered waveforms on six monopolar electrodes. The results showed significantly better speech perception performance for the multiple-channel systems, and although the performances of The University of Melbourne/Nucleus and Utah devices were comparable in quiet the Utah device performed better in noise from multi-speech babble. It was thought this performance difference in noise was due to the difficulties a preprocessing system such as The University of Melbourne/Nucleus device would have in selecting speech from noise, whereas the Utah system which presented the unprocessed output from six fired bandpass filters would have the patients brain differentiate the speech signals. While the FDA clinical trial and comparative evaluation of The University of Melbourne/Nucleus FO/F2 speech processor were being undertaken we were actively involved in psychophysical and speech research to improve our cue-extraction speech processing system.

As our previous research had shown the need to provide additional information to improve consonant perception, in particular, the next question to be answered was whether the presentation of the first and/or third formants as well as the second formant would improve intelligibility. It was also necessary to establish by research in our laboratory that improvements were possible before industry could direct its resources to producing an alternative wearable speech processor for patient evaluation.

Consequently, the first task was to determine whether the presentation of two stimuli near-simultaneously (0.5 ms separation between pulses to avoid unpredictable current summation) on a place coding basis would result in a single or two component sensation. It was argued that if we were to convey speech information containing two or more formants the brain should have the capability of perceiving additional information from a second stimulus channel.

The perceptual dissimilarities among ten two-electrode stimuli were estimated by triadic comparisons and the resulting matrices analyzed by non-metric multidimensional scaling (Tong et al., 1983b). This showed a two dimensional solution to be the best, indicating that the dissimilarities between two electrode pairs differing in both apical and basal electrodes was even greater than the sum of two component dissimilarities. These results suggested that two electrode stimulation was perceived as a sensation with two components, and could therefore be used to present speech information with two components, such as the first and second or the second and third formants.

To assess the value of using a speech processing strategy presenting two formants we developed an acoustic model of electrical stimulation on normal hearing subjects (Blamey et al., 1984a). The acoustic model was especially necessary as all our implanted patients were involved in the clinical trial for the FDA. Multiple-channel electrical stimulation was modelled with a set of bandpass filtered noise stimuli and gave similar psychophysical results on normal hearing subjects as pulsed electrical stimulation had given on the first two implant patients (Blamey et al., 1984a). The speech perception of the two multiple-channel implant patients when using the FO/F2 strategy was then compared with that of the normally hearing listeners when using an acoustic model of the FO/F2 strategy (Blamey et al., 1984b). 22 different speech perception tests were used and very good agreement found for the two groups of subjects indicating that the acoustic model could be a useful tool for the development and evaluation of alternative speech processing strategies.

It was considered preferable to initially compare the FO/F2 strategy with one that presented the first rather than the third formant in addition to the second formant as there was good evidence from the acoustic literature that the lowest two formants were at least the most important features for recognizing a vowel (Delattre et al., 1952). The study (Blamey et al., 1985) in fact showed a significant difference between the FO/F2 and FO/F1/F2 strategies for the vowel but not the consonant test. However, with consonant confusions there were improvements in information transmission for voicing (34% to 50%), nasality (84% to 95%) and affrication (32% to 40%) but not for place (28% to 28%). We also discovered that the consonants could be classified according to the amplitude envelopes of the F2 filter output, and this also improved for the FO/F1/F2 strategy. The improved vowel perception and increased information transmission for consonant features probably accounted for the significant improvement obtained in understanding running speech as shown with the speech Tracking test.

Having shown with the acoustic model that improved speech perception scores occurred for the FO/F1/F2 compared to

the FO/F2 strategy a decision was made by industry to implement it as a wearable speech processor and this is known commercially as WSPIII. When this was clinically trialed on implant patients (Clark, 1986; Dowell et al., 1987) speech perception results were similar to those obtained with the acoustic model. This showed the benefits of coding first and second formants on a place basis, and the predictive value of the model. In quiet the closed-set medial vowel scores improved from 51% to 58% and the consonant scores from 54% to 67%. The mean open-set CID sentence scores three months postoperatively increased from 16% to 35% (Dowell et al., 1987). Improved speech results in noise were also achieved by the addition of the first formant, and they were comparable to those for the fixed filter strategy of the Utah multiple-electrode system (Dowell et al., 1987).

To further improve speech perception, especially for consonants, it was next considered appropriate to code additional information in the high frequencies on a place basis. Rather than use the third formant as we had originally proposed, we presented the filter outputs on a place coding basis in the ranges 2000-2800Hz, 2800-4000Hz and 4000-6000Hz for unvoiced sounds and two of these for voiced sounds. As initial results with this strategy showed improvements over the FO/F1/F2 strategy (Dowell et al., 1990), it was implemented by Cochlear Pty. Limited as a wearable unit and an independent comparison made for the FDA by Skinner et al., (1991). This strategy is known commercially as Multipeak-MSP.

An independent comparative study (Cohen et al., 1993) has recently reported speech perception results for the single-channel 3M Vienna, the multiple-channel Ineraid device (previously referred to as the Utah strategy), the multiple-channel Nucleus WSPIII (previously referred to as the FO/F1/F2 strategy), and the multiple-channel Multipeak-MSP device. The results confirmed the previous findings of Gantz et al., (1987) and Tyler et al., (1987) that the multiple-channel devices were superior to the single-channel system. Furthermore, open-set speech perception scores were significantly better for the Multipeak-MSP device than WSPIII. The mean open-set sentence score for WSPIII was 32% (range 2 to 76%) and for Multipeak-MSP 58% (range 9 to 97%). The open-set scores for Multipeak-MSP patients were also significantly better than for the Ineraid device (Cohen et al., 1993).

Space does not permit a summary of the research that has been undertaken to apply the FO/F1/F2 and Multipeak strategies to children, and those who are prelinguistically deaf, (born deaf or lost hearing before developing language) on the biological research to ensure the device is safe for children under two years of age. There is also not space to summarize our recent psychophysics and speech research which is showing that a strategy which extracts the six spectral maxima and presents these on a place coding basis, with amplitude modulation of a constant rate to convey voicing, is giving better results than the Multipeak-MSP device. Psychophysical and speech research in adults with cochlear implants in both ears is providing useful basic information and showing benefits in hearing speech in the presence of noise. Research to combine information presented by electrical stimulation to a cochlear implant in one ear and sound to a speech processing hearing aid in the other ear is also looking promising.

## ACKNOWLEDGEMENTS

This research was a team effort and I would like to acknowledge support from my colleagues. This research would not have been possible without funding from a number of bodies which include the National Health & Medical Research Council of Australia, the Australian Research Council the Commonwealth Department of Employment Education and Training, the Commonwealth Department of Science and Technology, the US National Institute of Health, the Channel 0 (10) Telethon and Nerve Deafness Appeal and Lions International.

## REFERENCES

Black, R. C. & Clark, G. M. (1977) "Electrical transmission line properties in the cat cochlea" *Proceedings of the Australian Physiological and Pharmacological Society*, **8,** 137.

Black, R. C. & Clark, G. M. (1978) "Electrical network properties and distribution of potentials in the cat cochlea" *Proceedings of the Australian Physiological and Pharmacological Society*. **9**, 71.

Blamey, P. J., Dowell, R. C., Tong, Y. C., Brown, A. M., Luscombe, S. M. & Clark, G. M. (1984a) "Speech processing studies using an acoustic model of a multiple-channel Cochlear Implant" *Journal of the Acoustical Society of America*. **76**(1), 104-110.

Blamey, P. J., Dowell, R.C., Tong, Y. C. & Clark, G. M. (1984b) "An acoustic model of a multiple-channel cochlear implant" *Journal of the Acoustical Society of America*. **76**(1), 97-103.

Blamey, P. J., Martin, L. F. A. & Clark, G. M. (1985) "A comparison of three speech coding strategies using an acoustic model of a cochlear implant" *Journal of the Acoustical Society of America*. **77**(1), 209-217.

Clark, G. M. (1969) "Responses of cells in the superior olivary complex of the cat to electrical stimulation of the auditory nerve" *Experimental Neurology*. **24**(1), 124-136.

Clark, G. M., Nathar, J. M., Kranz, H. G. & Maritz, J. S. (1972) "A behavioral study on electrical stimulation of the cochlea and central auditory pathways of the cat" *Experimental Neurology*. **36**, 350-361.

Clark, G. M. (1973) "A hearing prosthesis for severe perceptive deafness - experimental studies" *Journal of Laryngology and Otology*. **87**(10), 929-945.

Clark, G. M., Kranz, H. G. & Minas, H. J. (1973) "Behavioural thresholds in the cat to frequency modulated sound and electrical stimulation of the auditory nerve" *Experimental Neurology*. **41,** 190-200.

Clark, G. M., Kranz, H. G., Minas, H. & Nather, J. M. (1975) "Histopathological findings in cochlear implants in cats" *Journal of Laryngology and Otology*. **89**(5), 495-504.

Clark, G. M., Tong, Y. C., Martin, L. F. A. & Busby, P. A. (1981a) "A multiple- channel cochlear implant. An evaluation using an open-set word test" *Acta Otolaryngologica (Stockh)*. **91,** 173-175.

Clark, G. M., Tong, Y. C. & Martin, L. F. A. (1981b) "A multiple-channel cochlear implant. An Evaluation Using Open-Set CID Sentences" *Laryngoscope*. **91**(4), 628-634.

Clark, G. M., Tong, Y. C. & Martin, L. F. A. (1981c) "A multiple-channel cochlear implant. An evaluation using closed-set spondaic words" *Journal of Laryngology and Otology*. **95**(4), 461-464.

Clark, G. M., Crosby, P. A., Dowell, R. C., Kuzma, J. A., Money, D. K., Patrick, J. F., Seligman, P. M. & Tong, Y. C. (1983) "The preliminary clinical trial of a multi-channel cochlear implant hearing prosthesis" *Journal of the Acoustical Society of America*. **74**(6), 1911-1913.

Clark, G. M. (1986) "The University of Melbourne/Cochlear Corporation (Nucleus) program" *Otolaryngologic Clinics of North America*. **19**(2), 329-354.

Clark, G. M., Blamey, P. J., Brown, A. M., Busby, P. A., Dowell, R. C., Franz, B. K-H., Pyman, B. C., Shepherd, R. K., Tong, Y. C., Webb, R. L., Hirshorn, M. S., Kuzma, J., Mecklenburg, D. J., Money, D. K., Patrick, J. F. & Seligman, P. M. (1987) "The University of Melbourne-Nucleus multiple electrode cochlear implant" In: *Advances in Oto-Rhino-Laryngology.* ED. C.R. Pfaltz, (Karger, Basel). Vol. **38.**

Cohen, N. L., Waltzman, S. B. & Fisher, S. G. (1993) " A prospective randomized study of cochlear implants" *The New England Journal of Medicine.* **328,** 233-237.

Cooper, F. S., Delattre, P. C., Liberman, A. M., Borst, J. M & Gerstman, L. J. (1952) "Some experiments on the perception of synthetic speech sounds" *Journal of the Acoustical Society of America.* **24,** 597-606.

Delattre, P. C., Liberman, A.m., Cooper, F. S. & Gerstman, L. J. (1952) "An experimantal study of the acoustic development of vowel colour observations on one and two formant vowels synthesized from spectrographic patterns" *Word,* **8,** 195-210.

Dowell, R. C., Martin, L. F. A., Tong, Y. C., Clark, G. M., Seligman, P. M. & Patrick, J. F. (1982) "A 12-consonant confusion study on a multiple- channel cochlear implant patien" *Journal of Speech and Hearing Research.* **25,** 509-516.

Dowell, R. C., Mecklenburg, D. J. & Clark, G. M. (1986) "Speech recognition for 40 patients receiving multichannel cochlear implants" *Archives of Otolaryngology.* **112** 1054-1059.

Dowell, R. C., Seligman, P. M., Blamey, P. J. & Clark, G. M. (1987) "Speech perception using a two-formant 22-electrode cochlear prosthesis in quiet and in noise" *Acta Otolaryngologica (Stockh).* **104,** 439- 446.

Dowell, R. C., Whitford, L. A., Seligman, P. M., Franz, B. K.-h. & Clark, G. M. (1990) "Preliminary results with a miniature speech processor for the 22-electrode melbourne/cochlear hearing prosthesis." In: *Otorhinolaryngology, Head And Neck Surgery.* Proceedings of the XIV Congress of Oto-Rhino-Laryngology, Head and Neck Surgery, Madrid, Spain, September 10-15, 1989 ED. T. Sacristan, Ed. (Kugler & Ghedini, Amsterdam) 1167-1173.

Fant, G. (1956) "On the predictability of formant levels and spectrum envelopes on formant frequencies." In: *For Roman Jakobson,* Eds, M. Halle, H. Lurnt and Maclean, (Mounton, The Hague).

Fourcin, A. J., Rosen, S. M., Moore, B. C. J., Douek, E. E., Clarke, G. P., Dodson, H. & Bannister, L. H. (1979)" External electrical stimulation of the cochlea: clinical, psychophysical, speech-perceptional and histological findings" *British Journal of Audiology.* **13,** 85-107.

Gantz, B. J., Mccabe, B. F., Tyler, R. S & Preece, J. P. (1987) "Evaluation of pair cochlear implant designs." In: *International Cochlear Implant Symposium and Workshop-Melbourne, 1985. Annals of Otology, Rhinology & Laryngology,* Eds. G.M. Clark, P.A. Busby, Suppl. 128, **96**(1),145-147.

Liberman, A. M., Delattre, P. C., Cooper, F. S. & Gerstman, L. J. (1954) "The role of consonant-vowel transitions in the perception of the stop and nasal consonants" *Psychology Monographs.* **68,** 1-13.

Martin, L. F. A., Tong, Y. C. & Clark, G. M. (1981) "A multiple-channel cochlear implant. evaluation using speech tracking" *Archives of Otolaryngology.* **107,** 157-159.

Merzenich, M. M., Michelson, R. P., Pettit, C. R., Schindler, R. A. & Reid, M. (1973) " Neural coding of sound sensation evoded by electrical stimulation of the acoustic nerve" *Annals of Otology, Rhinology and Laryngology.* **82,** 486-504.

Merzenich, M. M. (1974) "Studies on electrical stimulation of the auditory nerve in animals and man: cochlear implants." In: Tower, *The Nervous System,* (Raven Press, New York). **3,** 337-548.

Plomp, R. (1976) "Aspects of tone sensation." In: *Series in Cognition and Perception,* eds. E.C. Carterette, & M.P. Friedman, ( Academic Press, London) 109.

Potter, R. K., Kopp, G. A. & Green, H. C. (1947) *Visible speech.*( van Nostrand, New York).

Skinner, M. W., Holden, L. K., Holden, T. A., Dowell, R. C., Seligman, P. M., Brimacombe, J. A. & Beiter, A. L. (1991) "Performance of postlinguistically deaf adults with wearable speech processor (WSPIII) and mini speech processor (MSP) of the Nucleus multi-electrode cochlear implant" *Ear and Hearing.* **12,** 3-22.

Tong, Y. C., Black, R. C., Clark, G. M., Forster, I. C., Millar, J. B., O'Loughlin, B. J. & Patrick, J. F. (1979) "A Preliminary report on a multiple-channel cochlear implant operation" *Journal of Laryngology and Otology.* **93**(7), 679-695.

Tong, Y. C., Clark, G. M., Seligman, P. M. & Patrick, J. F. (1980a) "Speech Processing for a multiple-electrode cochlear implant hearing prosthesis" *Journal of the Acoustical Society of America.* **68**(6), 1897-1899.

Tong, Y. C., Millar, J. B., Clark, G. M., Martin, L. F. A., Busby, P. A. & Patrick, J. F. (1980b) "Psychophysical and speech perception studies on two multiple-channel cochlear implant patients" *Journal of Laryngology and Otology.* **94**(11), 1241-1256.

Tong, Y. C., Clark, G. M., Blamey, P. J., Busby, P. A. & Dowell, R. C. (1982) "Psychophysical studies for two multiple-channel cochlear implant patients" *Journal of the Acoustical Society of America.* **71**(1), 153-160.

Tong, Y. C., Blamey, P. J., Dowell, R. C. & Clark, G. M. (1983a) "Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant" *Journal of the Acoustical Society of America.* **74,** 73-80.

Tong, Y. C., Dowell, R. C., Blamey, P. J. & Clark, G. M. (1983b) "Two component hearing sensations produced by two-electrode stimulation in the cochlea of a totally deaf patient" *Science.* **219,** 993-994.

Tyler, R. S., Tye-Murray, N., Preece, J. P., Gantz, B. J. & McCabe, B. F. (1987) "Vowel and consonant confusions array cochlear implant patients: do different implants make a difference?" In: *International Cochlear Implant Symposium and Workshop-Melbourne,* 1985. *Annals of Otology, Rhinology & Laryngology,* eds. G.M. Clark, & P.A. Busby, Suppl. 128, **96**(1), 141-144.

University Library

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Clark, Graeme M.

Title:
The University of Melbourne/nucleus multiple-channel cochlera implant

Date:
1993

Citation:
Clark, G. M. (1993). The University of Melbourne/nucleus multiple-channel cochlera implant. Acoustics Australia, 21(3), 91-97.

Persistent Link:
http://hdl.handle.net/11343/27321