A SPEECH PROCESSING STRATEGY FOR MULTIPLE-ELECTRODE COCHLEAR IMPLANT PROSTHESSES
Y.C. Tong and G.M. Clark, Department of Otalaryngology, University of Melbourne, Parkville, Vic. 3052, Australia

Speech studies in a number of research centres have shown that useful speech information could be presented to deaf patients using single or multiple electrode cochlear implant prostheses (Parkins & Anderson, 1983). In our laboratory, speech processing strategies were formulated on the basis of psychophysical results. This paper examines the psychophysical characteristics of the hearing sensations produced by electrical stimulation using scala tympani electrodes in postlingually deaf patients; a speech processing strategy is then discussed on the basis of these characteristics.

Sensations related to electrode position

One of the most important observations made from the electric stimulation of the human cochlea was the orderly variation of perceptual characteristics with the position of the intra-cochlear electrodes. At a fixed repetition rate, pitch varied from low to high and sharpness from dull to sharp for individual electrodes ordered in the apical to basal direction (Tong et al., 1982). In addition, vowel labels could be assigned to the sensations produced by individual electrodes (at a fixed rate), and the second formant frequency of the vowel was consistent with the acoustic characteristic frequency of the location of the activated electrode (Tong et al., 1979). This orderliness of perceptual characteristics was also reflected in the perceptual organisation of the hearing sensations produced by stimulation using pairs of electrodes. In a study employing triadic comparisons and analysis by multi-dimensional scaling, it was shown that two-electrode stimulation was perceived as a two-component sensation, and the two-components corresponded in an orderly fashion to the position of the apical and basal electrode of the two-electrode pair (Tong et al., 1983a).

The most likely mechanism for this orderly variation is the activation of different groups of residual auditory nerve fibres by electric current, and subsequent processing of this 'place' information in the more central nuclei of the auditory system. This is supported by a number of electrophysiological findings. Firstly, selective activation of groups of inferior collicular neurons by intra-cochlear electrodes has been demonstrated by Schindler et al. (1977). Secondly, these same authors have found a correspondence between the characteristic frequency of the electrically activated inferior collicular neurons and that of the cochlear place where the electrode was located, and therefore, established that the information about the 'place' of the electrode is preserved in an orderly fashion in the more central nuclei of the auditory system. Finally, the perception of the variation of pitch and sharpness with electrode position did not appear to be due to temporal information as this variation can be perceived for stimuli at a fixed repetition rate and therefore producing constant temporal discharge characteristics. The temporal discharge patterns for first order auditory neurons and collicular neurons have been studied by Kiang and Moxon (1972) and Merzenich et al. (1973), and the results showed that the discharges were synchronous with the periodic electrical stimulus, and no additional temporal information other than that corresponding to the frequency of electric stimulation was observed.

Although a 'place' mechanism appears to be attractive in explaining some of the psychophysical results in electric stimulation, it should be noted that in acoustic stimulation the temporal pattern of discharge in auditory neurons is also considered to be an important factor in the encoding of frequency information and therefore pitch perception. The temporal pattern (to be described later) is particularly important for stimuli at high sound pressure levels where a purely 'place' mechanism does not account for the frequency resolving acuity of the auditory system because of saturation (Sachs & Young, 1979). However, the central mechanism by which this temporal information is processed is not known. One suggestion is that the fine time structure of discharge is merely a vehicle for the transmission of place information to the higher centers of the auditory system (Evans, 1978). Our argument that place information with fixed temporal characteristics can be conveyed to and utilised by the central auditory system in electric stimulation does not contradict...
this suggestion; it merely highlights the possibility that the mediating fine time structure can be bypassed provided that a spatially restricted group of fibers can be independently activated.

Sensations related to repetition rate

In addition to electrode position, repetition rate is another electrical parameter influencing the pitch percept produced by an electric stimulus. The major psychophysical results are summarized as follows: a) the discrimination performance for electric repetition rate was much poorer than that for acoustic signals such as pure tones and multi-component tone complexes in normal hearing subjects (Tong et al., 1982); b) the discrimination performance was poorer for repetition rates above 250 pps in comparison to that for lower repetition rates (Redington et al., 1978); Vingenich et al., 1973); c) the increase in (scaled) pitch with repetition rate was less pronounced above 250 pps (Tong et al., 1979; Tong et al., 1983b); d) the pitch sensation produced by a single-electrode stimulus with repetition rate below 250 pps could be matched to that for an acoustic pure tone with the same frequency in a cochlear implant patient with residual hearing in the unimplanted ear (Bilger et al., 1977); and e) the discrimination performance of repetition rate for single-electrode stimulation was found to be as good or better than that for single-electrode stimulation (Tong & Clark, 1983).

The difference in discrimination performance between electric stimuli differing in repetition rate and acoustic tones differing in frequency is not surprising because there are significant differences in the representation of the first order neurons of the auditory system for these two types of stimuli. Taking acoustic pure tones as an example, the information about the intensity and frequency is transmitted in the form of a spatial-temporal discharge pattern governed by the travelling wave motion of the basilar membrane.

The discharge pattern for pure tones can be summarized by the following characteristics: a) the distribution of the effective discharge rate (with respect to the spontaneous rate) of the array of neurons varies in the same fashion as the displacement envelope of the basilar membrane vibration, and is asymmetric about the place of maximum displacement; b) the stochastic discharge pattern is phase-locked to the input tone for frequencies below 5 kHz, and there is a gradual phase variation for neurons innervating successive locations along the cochlear partition; c) period histograms of the discharge pattern demonstrate temporal spread of discharges over the half-period of the vibration towards the scala vestibuli and over the whole period for very low frequency tones; and d) information concerning the period of the tone can also be seen in interspike interval histograms where peaks at integral multiples of the stimulus period were observed (Rose et al., 1967).

This complex discharge pattern for acoustic tones is not simulated, however, for electrical stimulation. The spatial distribution of the neural discharge rate depends on the current distribution in the scala tympani and cochlear partition complex (Black & Clark, 1980), and for bipolar electrodes and electrodes with a common ground only residual nerve fibers in a restricted region in the proximity of the electrode are activated (Vingenich et al., 1979; Tong et al., 1979). Over this restricted region, electrophysiological evidence (Hodgkin & Huxley, 1952; Kiang & Moxon, 1972) on electric stimulation of neural tissue suggests that the most probable instant of excitation of the neurons is synchronous with a particular phase of an above threshold stimulus, and the phase variation amongst the neural discharges, observed in acoustic stimulation, is therefore not reproduced. Another discrepancy is the more restricted temporal spread of the discharges with electric stimulation, as shown in period histograms (Kiang & Moxon, 1972).

From these descriptions of the discharge patterns, it is apparent that the information concerning the difference in repetition rate for electric stimuli is reflected in the period of the neural discharge. On the other hand, for acoustic tones and tone complexes differing in fundamental frequency and harmonic structure, a combination of place, period and phase information is available to the auditory system. The mechanism that is responsible for the processing of this combination of information for acoustic signals.
unfortunately, is not clear (Kiang, 1968; Evans, 1978; Young &Suchs, 1980). However, the difference in perceptual performance between electric and acoustic stimulation does highlight the limitation of the performance of an auditory process receiving only 'period' information as its input.

The poorer scaling and discrimination performance at higher repetition rates (above 250 pps) is consistent with the findings in the inferior colliculus, by Merzenich et al., (1973) who have shown that period information in the form of peaks in interspike interval histograms was only preserved for electrical stimulation frequencies up to about 400 Hz. Furthermore, there was a gradual degradation of the distinctness of the peaks from about 300 Hz.

The pitch sensation produced by electric stimulation at repetition rates below 250 pps is similar to 'rattle pitch' (Flanagan & Gutman, 1960; Plomp, 1976), a component of the pitch percept produced by acoustic pulse trains. The similarities are reflected in the following observations: a) both types of pitch sensation can be matched to acoustic pure tones with the same frequency (Flanagan & Gutman, 1960; Bilger et al., 1977); b) the upper limit of acoustic repetition rates for which 'rattle pitch' is most prominent is about 200 pps (Flanagan & Gutman, 1960); and c) both types of sensation exhibit buzz-like quality at low repetition rates. In fact, the mechanism responsible for the perception of these two types of sensations may be the same. In this regard, the masking experiments of Rosenberg (1965) showed that 'rattle pitch' could be most effectively masked by a band of noise centered around 5000 Hz. This result indicated that rattle pitch is derived from the time interval (period) between the neural discharges phase-locked to the amplitude envelope of the combined wave form of the unresolved high frequency harmonics. In relation to these masking experiments, it is worth noting that, in addition to 'rattle pitch', there is a competing pitch component related to the resolvable harmonics for acoustic stimulation using pulse trains, and this component can be masked by a band of noise centered around the frequencies of the dominant harmonics responsible for its perception (Rosenberg, 1965).

Interaction of electrode position and pulse rate

From the above discussions, it is apparent that the psychophysical observations relating to the position of the electrode may be accounted for by a 'place' code, while those related to the repetition rate may be accounted for by a 'period' or 'temporal' code.

Another observation was that the interaction between these two codes appeared to be small. Two psychophysical studies were conducted to investigate the possible effects of the interaction (Tong et al., 1983).

The first of these showed that the information provided by rising or falling repetition rate trajectories superimposed on individual electrodes or electrode trajectories could be used as an indicator of the direction of intonation variation. These results obtained for electric stimulation are comparable with those for acoustic stimulations with fundamental frequency trajectories superimposed on the syllable 'oh'. This correspondence between acoustic and electric results suggests that the perception of pitch contours related to electric repetition rate in the presence of a variation in electrode position is similar to the perception of pitch contours related to fundamental frequency superimposed on a variation in spectral envelope of an acoustic signal.

The second of those studies showed that the dissimilarities amongst the hearing sensations produced by steady state stimuli differing in electrode position and repetition rate were characterized by two perceptual components, relating to the two electric parameters respectively. In addition, comparison with acoustic data (Plomp, 1976) showed that the two perceptual components in electric stimulation closely parallel those related to spectral envelope and fundamental frequency in acoustic stimulation.

Sensations related to time-varying electrode position and repetition rate

For signals with time-varying electrode position at a fixed repetition rate, psychophysical results (Tong et al., 1982) showed that the 'place' code is capable of conveying time-varying frequency information with rapid frequency shifts. The discriminability of electrode trajectories differing in the
direction and extent of the electrode shift over a 100 ms time span has been demonstrated in Tong et al., (1982), and these results showed that this performance was maintained for electrode shifts with a much shorter time span (down to 25 ms). Furthermore, for the electrodes implanted in the basal coil of our patients, the hearing sensations produced by these short-duration electrode shifts were reported to be similar to those produced by acoustic speech signals characterized by second formant frequency transitions.

On the other hand, for single-electrode signals with time-varying repetition rate below 250 pps, the results in Tong et al., (1982) show that the discrimination performance related to the 'period' code deteriorated when the duration of the shift in repetition rate was shortened. A further difference between the 'place' and 'period' codes for time-varying signals is that the sensations produced by shifts in repetition rate were described as 'the same word with different pitches', while those produced by electrode shifts were described as 'different words'.

A speech processing strategy

In the light of the above discussions a speech processing strategy, which converts the frequencies of the spectral emphases of a speech signal to electrode positions, and the fundamental (voicing) frequency to repetition rate may be proposed. This strategy makes use of the ability of the auditory system in discriminating rapid electrode shifts to convey segmental speech information which is characterized by rapid frequency shifts. The suprasegmental information contained in the fundamental frequency of a speech signal, on the other hand, can be mediated by the 'period code' related to pulse rate as this is better discriminated in slowly time-varying signals. Furthermore, the phenomenological reports of patients also showed that the hearing sensations produced by electrode shifts were associated with those produced by acoustic shifts of spectral emphasis, while repetition rate variations were associated with changes in fundamental frequency.

In addition to the encoding of the spectral emphases and the fundamental frequency, the amplitude of the speech signal may be converted to the current level of the electrical stimulus in this speech processing strategy. From the results obtained in a previous study (Tong et al., 1979) and results from other research centers, it has been shown that a full range of loudnesses, from threshold to discomfort, can be evoked by the variation of the current level. The loudness growth due to increases in current level, however, was much steeper than the growth for acoustic stimulation in normal hearing subjects. It is therefore apparent that, for the effective utilization of the current level for the encoding of signal amplitude, an appropriate compression algorithm should be used. This is, of course, feasible with current electronic and computing technology.


