PITCH AND VOWEL PERCEPTION IN COCHLEAR IMPLANT USERS

Peter J Blamey and Elvira S Parisi

Department of Otolaryngology
University of Melbourne

ABSTRACT - Two methods of determining the pitch or timbre of electrical stimuli in comparison with acoustic stimuli are described. In the first experiment, the pitch of pure tones and electrical stimuli were compared directly by implant users who have residual hearing in the non-implanted ear. This resulted in a relationship between frequency in the non-implanted ear and position of the best-matched electrode in the implanted ear. In the second experiment, one- and two-formant synthetic vowels, with formant frequencies covering the range from 200 to 4000 Hz, were presented to the same implant users through their implant or through their hearing aid. The listeners categorised each stimulus according to the closest vowel from a set of eleven possibilities, and a vowel centre was calculated for each response category for each ear. Assuming that stimuli at the vowel centres in each ear sound alike, a second relationship between frequency and electrode position was derived. Both experiments showed that electrically-evoked pitch is much lower than that produced by pure tones at the corresponding cochlear location in normally-hearing listeners. This helps to explain why cochlear implants with electrode arrays that rarely extend beyond the basal turn of the cochlea have achieved high levels of speech recognition in postlinguistically deafened adults without major retraining or adaptation by the users. The techniques described also have potential for optimising speech recognition for individual implant users.

INTRODUCTION

Cochlear implants produce hearing sensations in profoundly deaf implantees by electrically stimulating residual nerve fibres using electrodes surgically inserted into the cochlea. In the electrical coding of speech, acoustic frequency ranges are mapped onto electrodes. The question arises as to whether the pitch of the electrically evoked hearing sensations are similar to those evoked by the corresponding acoustic stimuli in normally-hearing listeners. The sensations are similar enough for many postlinguistically deaf implant users to understand speech with a minimum of retraining, but it is unlikely that the electric signals sound identical to the acoustic ones. There will also be differences between implant users arising from the variable insertion depth of the electrode array, the number of electrodes in use, and the frequency-to-electrode mapping. The most direct method of determining pitch is to ask implant users to compare electric and acoustic stimuli, but studies of this sort have been hampered by the fact that very few implant users have useable hearing for acoustic signals. In 1978, Eddington *et al* reported pitch matching results for one unilaterally deaf volunteer. They concluded that pitch matching was "roughly consistent with electrode position and tonotopic maps of the cochlea derived from basilar membrane motion and hearing loss measurements". Two other pitch matching studies based on data from one single-channel House implant user (Bilger *et al* 1977) and one Ineraid multichannel implant user (Dorman *et al* 1994) have investigated the matching of electrical pulse rate to pure tone frequency up to 400 Hz. Several other studies (Tong *et al* 1982, Shannon 1983, Tong & Clark 1985, Townshend *et al* 1987, Busby *et al* 1994) have investigated the relative pitch of electric signals using identification, scaling, and discrimination paradigms. The latter studies have established that electrode placement, electrode configuration, and rate of stimulation all affect the perceived pitch, and that the pitch decreases tonotopically from basal to apical electrode positions. They have not determined the pitch of electric stimuli in an absolute fashion that can be compared with acoustic stimuli, however. A knowledge of the absolute pitch of electric stimuli for individuals, or as a function of position in the cochlea would be very useful in optimising the frequency mapping for cochlear implants.

The present paper describes two methods of comparing acoustic and electric stimuli to study the relationship between perceived pitch and electrode position. The first method directly compared the pitch of acoustic pure tones in one ear with electric signals in the other ear. The second method compared the categorisation of synthetic vowels presented acoustically to one ear and electrically to the other ear. The main questions addressed were: Whether the two methods would give comparable relationships between frequency and electrode position, and whether the electrode used in the matched stimulus would correspond in position to the place of maximum basilar membrane motion produced by the acoustic signal in a normal cochlea.

METHOD

Subjects

The eight subjects were postlinguistically deafened adults with some residual hearing in the non-implanted ear. Average audiometric thresholds in the non-implanted ear at octave frequencies from 250 Hz to 4 kHz were 81 (7), 98 (8), 102 (6), 100 (6), and 90 dB HL (4) respectively. The numbers in brackets are the numbers of subjects with measurable thresholds at each frequency. All subjects were implanted with the 22-electrode cochlear implant (Clark *et al* 1987). All subjects were regular cochlear implant or hearing aid users and the majority wore both devices most of the time.

Determination of electrode positions

A modified Stenver's view X-ray, with the intracochlear portion of the electrode array in the plane of the film (Marsh *et al* 1993) was used to provide reliable electrode position information. Angular positions for electrodes were measured about the centre of the cochlear spiral, with the hook region of the basilar membrane at 0°. Angles were used because the position of the electrode array within the scala tympani is close to the outer wall of the cochlear spiral and distances along the array are not equal to distances along the basilar membrane. Angular positions eliminate this discrepancy. The angles of the most apical electrodes for the patients ranged from 201 to 394°. Bredberg (1968) provides anatomical data from humans that indicate the relationship between angles and proportional distances along the Organ of Corti. These proportions were used in conjunction with the formula of Greenwood (1961) to calculate characteristic frequencies corresponding to angular positions in the cochleae of normally-hearing listeners.

Pitch comparison experiment

Acoustic and electric stimuli were produced by computer control of a "bimodal" aid which combined a cochlear implant speech processor and a hearing aid processor. The electric signals were pulse trains with duration 500 ms at a fixed rate on a bipolar electrode pair. The acoustic stimuli were pure tones of duration 500 ms. The tones were presented via an Oticon AN1000 hearing aid receiver and ear mould at levels that were comfortable for the listener. The levels varied between subjects and between frequencies because of differences in thresholds and dynamic ranges. Before starting a block of pitch comparisons, the intensity of each electric stimulus was adjusted to match the loudness of the acoustic stimulus. Electric and acoustic stimuli were presented in pairs. The subject was asked which stimulus was higher, and the judgements were repeated to test their reliability. The subject listened to the alternating signals for as long as was necessary to make a decision (usually about three alternations). There was sometimes a broad range in which the subject seemed uncertain which stimulus was higher. To cope with this situation, the subject was allowed to choose from three responses: "acoustic higher", "electric higher", or "both the same". A set of pure tone frequencies was chosen for each subject, spanning the frequency range of useable hearing in octave or half-octave steps. Within a block of trials, the frequency of the acoustic signal was kept fixed, the pulse rate of the electric signal was kept fixed at 100, 250 or 800 pps, and the position of the stimulated electrode was varied.

Vowel perception experiment

Two sets of vowels in an /hVd/ frame were synthesised using the parallel branch of the Klatt (1980) synthesiser. The single-formant set had formant frequencies (SF) ranging from 200 to 1000 Hz in 100 Hz steps, and from 1000 to 4000 Hz in 200 Hz steps. The two-formant set had first formant (F1) frequencies ranging from 300 to 900 Hz in 100 Hz steps and second formant (F2) from 600 to 2400 Hz in 200 Hz steps. The resulting stimuli were normalised to the same level in dBA. Each stimulus was presented ten times to the subject, listening through one ear only (implant or hearing aid). The listener responded with one of the eleven steady state vowels: /i,ɪ,ɛ,æ,a,ʌ,ɔ,ɒ,ʊ,u,ɜ/. For each vowel response, the formant frequencies (SF, F1, and F2 separately) were averaged over all the stimuli in the category to produce a "vowel centre". For the implant data, the frequencies of the vowel centres were converted to electrode numbers using the "frequency-to-electrode map" in the subject's speech processor. The angular position of the electrode was then determined from the X-ray data as indicated above.

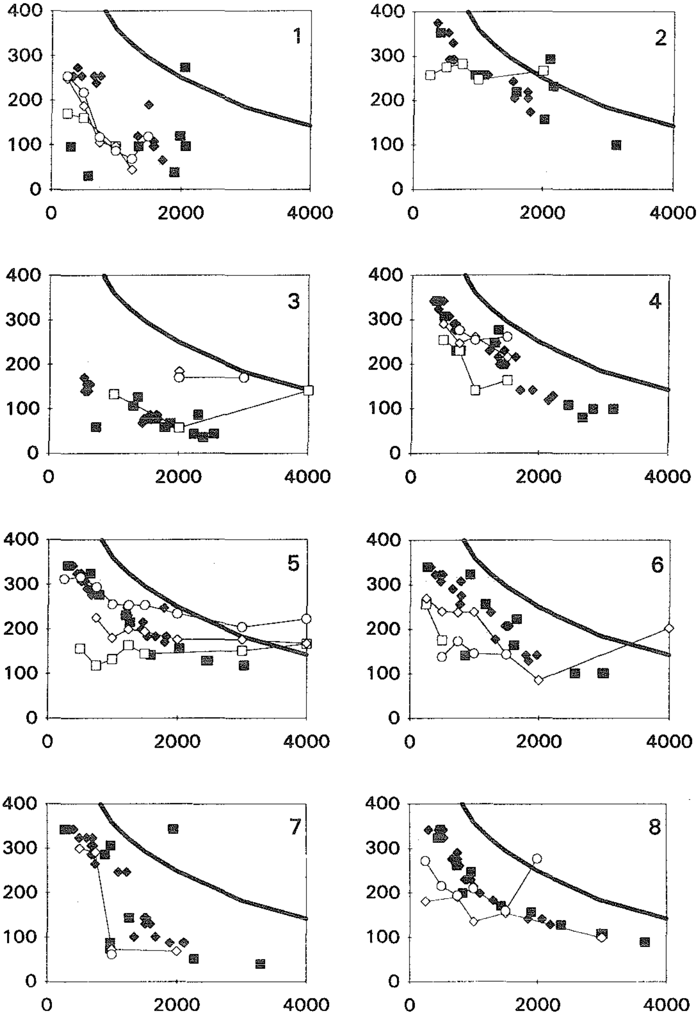

Figure 1. Angular positions of electrodes (vertical axis) corresponding to acoustic frequencies (horizontal axis) for each subject. Open symbols represent pitch matching results for electric pulse rates of 100 pps (□) 250 pps (◊) or 800 pps (o). Filled symbols represent results from the vowel perception experiment using SF (■) or F1 and F2 (◆). The darker line shows the frequency variation in a normal cochlea.

RESULTS

Pitch comparison experiment

Fig. 1 shows the results from the two experiments in a separate plot for each of the eight subjects. The open symbols linked by lines represent the best-matching electrodes for different frequencies from the first experiment. Squares represent points for an electric pulse rate of 100 pps, diamonds for 250 pps, and circles for 800 pps. The symbols represent the electrode positions at which responses changed from "implant higher" to "hearing aid higher" as the electric stimulus was shifted in a basal to apical direction in the cochlea. Because of time limitations and differences in the subjects' residual hearing in the non-implanted ear, it was not possible to test every subject at every pulse rate and every acoustic frequency. In a small number of cases, the subjects reported that the electric and acoustic stimuli sounded identical, but this was generally not the case. Many of the subjects found the pitch comparisons difficult to do, and this also indicates that the electric and acoustic sounds differed in timbre qualities such as roughness, fullness, etc as well as pitch. It is also possible that electric stimuli evoke sensations that are more like tone complexes or noises than pure tones, so that they may have more than one pitch component. Despite the difficulty of the task, most of the responses followed the expected pattern of increasing pitch with more basal electrode position. This is consistent with previous psychophysical investigations of pitch referred to in the introduction. In most cases, the results for different pulse rates are not identical, indicating that pulse rate as well as electrode position affects the perceived pitch of the electric signals. For most subjects, the points for higher rates lie above and to the right of points for lower rates, indicating that electric stimuli at higher rates match higher acoustic frequencies, as expected.

Vowel perception experiment

In this experiment, vowel centres were calculated for each response category by calculating a weighted mean of the formant frequencies for stimuli eliciting that response. Assuming that these vowel centres for the hearing aid sound the same as the vowel centres for the implant, we find a set of matched acoustic and electric stimuli and plot a function relating frequency to electrode position. In Fig. 1, the filled squares correspond to single-formant vowel centres and filled diamonds correspond to the F1 and F2 values from the two-formant vowel centres. The points tend to cluster along a curve that falls from left to right, indicating that more basal electrodes (at lower angles) are perceived similarly to higher frequency acoustic formants, as expected. For some subjects there are a few points for single-formant vowels that fall a long way away from the main group of points. This may have been caused by the fact that some subjects used some vowel responses very infrequently for the single formant vowels, so that the vowel centres were determined by only a few trials. This would make these points less reliable than those based on a larger number of responses. Apart from these isolated points, the one- and two-formant vowel data are very consistent with each other.

DISCUSSION

Comparison of pitch matching and vowel perception results

The two sets of data occupy overlapping ranges of frequency and electrode position, and show qualitatively similar trends. They are not identical, and this is not surprising considering the differences in the stimuli and the different perceptual processes involved in categorising vowels and comparing the pitch of sounds presented to the two ears. For most subjects, the pitch matching data have a gentler slope from left to right than the vowel data, although the slopes are very similar in those regions where the pitch matching data are close to the vowel data. At low frequencies (especially for 100 pps), the matching data tend to lie to the left of the vowel data, indicating that the psychophysical stimuli were perceived as lower in pitch than the vowel formants presented at the same angular positions. At high frequencies, the matching data tend to lie to the right of the vowel data. The differences between the matching and the vowel data may stem from the influence of the electric pulse rate. It has been demonstrated by Tong *et al* (1983) that rate and electrode position have orthogonal effects on the hearing sensation. It is likely that the listeners based their pitch comparisons on an overall percept that included both rate and place components. In the vowel experiment, rate plays a similar role to voice fundamental frequency so that it would not be expected to have a large effect on the judgements of vowel category.

Comparison with normal hearing

The solid lines in Fig 1 represent the angular positions of points on the basilar membrane that correspond to pure tone frequencies in normally-hearing listeners, based on the work of Greenwood (1961) and Bredberg (1968). All but seven of the experimental points for pitch matching, and two for the vowel study fall to the left of the solid lines. In other words, the pitch percept evoked by electrical stimulation with an electrode at a given position in the cochlea usually corresponds to a lower frequency tone than would normally be perceived at that position. The vowel data tend to lie on curves that are roughly parallel to the normal curves, and some of the pitch matching data intersect the normal curves at frequencies above 2000 Hz. The magnitude of the downward shift is one to three octaves, depending on the subject and the angular position considered.

There are several factors that may lower the "matched" pitch of electrical stimulation in the present experiments. The first is the actual site of excitation of the nerves. It is likely that there are few surviving dendrites, and that the site of excitation is in Rosenthal's Canal. It is known that the dendrites do not travel radially outward from Rosenthal's Canal, and that the ganglion cells in the modiolus extend around 1.75 turns compared to 2.75 turns for the Organ of Corti (L Seldon, personal communication). To a first approximation, one would expect the angle corresponding to a given frequency to be reduced by a factor of 1.75/2.75. Applying this correction to the lines in Fig. 1 would bring them much closer to the data. Another factor is the effect of hearing impairment in the non-implanted ear. The effect of this factor is not easily quantified, but it is unlikely that it could account for frequency shifts of the magnitude observed here. Finally, one should consider the effect of learning or plasticity. It is possible that the patients became accustomed to the sounds produced by the implant to such an extent that the matching of psychophysical and vowel stimuli was influenced by their daily experience. If this process was complete, the acoustic frequency matched to each electrode would correspond to the frequency that was usually mapped onto that electrode. This learning effect would account quite well for the vowel data because the vowel centre frequencies for the implant and hearing aid were usually within half an octave of each other for all vowels. It should be noted, however, that this explanation implies that the listener is judging pitch or vowel category according to different criteria depending on whether the sound is heard in the implanted ear or the non-implanted ear.

IMPLICATIONS FOR FREQUENCY MAPPING

In the implant speech processor, frequency bands are "mapped" to particular electrodes, and the fundamental frequency of the voice, F0, is represented by the pulse rate or amplitude modulation within frequency bands. Until recently, it was assumed that the pitch of an electric stimulus would approximate the pitch of a pure tone which produces a maximum of excitation on the basilar membrane at the corresponding position in a normally-hearing listener as shown in Fig. 1. Because the electrode array cannot be inserted into the full length of the cochlear spiral, the mapping of frequencies from 300 Hz to 6 kHz onto electrodes in the basal turn of the cochlea (0-360° in Fig. 1) would have implied a considerable upward pitch shift. The results of Fig. 1 imply that the perceived pitch range for many implant users is actually quite close to the normal range covered by speech stimuli. Patients with a relatively shallow electrode insertion will still experience an upward pitch shift, but usually less than an octave. The size of this perceptual shift would be similar to the difference between adult and child voices, and thus within the normal range of variability experienced by normally-hearing listeners. This result helps to explain why most postlinguistically deafened adults can recognise speech at reasonably high levels of accuracy within a short time of the implant operation.

The successful use of the vowel perception method for determining perceived pitch of electrical signals suggests that it may be used in optimising the frequency-to-electrode mapping used for cochlear implant patients. The vowel category of the sound produced by each electrode would be determined first. Results from normally-hearing listeners would be used to determine the formant frequencies for the vowel centres, and these frequency values would be mapped onto the electrodes corresponding to the appropriate vowels. This technique may reduce the need for the patient to adapt to the electrically evoked hearing sensations.

ACKNOWLEDGMENTS

The authors wish to acknowledge the help of Jon Shallop and Patti Arndt in carrying out part of the study in Denver, Laurie Cohen in the analysis of the X-rays, and Gary Dooley in data collection and writing the programs for experimental control and data analysis. The research was funded by the National Health and Medical Research Council and the Cooperative Research Centre for Cochlear Implant, Speech, and Hearing Research.

REFERENCES

Bilger, R., Black, F., Hopkinson, N., Myers, E., Payne, J., & Stenson, N. (1977) *Evaluation of subjects presently fitted with implanted auditory prostheses*, Ann. Oto. Rhino. Laryngol. 86, (Suppl 38) 1-176.

Bredberg, G. (1968) *Cellular pattern and nerve supply of the human organ of Corti*, Acta Oto-Laryngologica Suppl. 236, 1-135.

Busby, P.A., Whitford, L.A., Blamey, P.J., Richardson, L.M., & Clark, G.M. (1994) *Pitch perception for different modes of stimulation using the Cochlear multiple-electrode prosthesis*, J. Acoust. Soc. Am. 95, 2658-2669.

Clark, G.M., Blamey, P.J., Busby, P.A., Dowell, R.C., Franz, B.K., Musgrave, G.N., Nienhuys, T.G., Pyman, B.C., Roberts, S.A., Tong, T.C., Webb, R.L., Kuzma, J.A., Money, D.K., Patrick, J.F., & Seligman, P.M. (1987) *A multiple-electrode cochlear implant for children*, Arch. Otolaryngol. 113, 825-828.

Dorman, M.F., Smith, M., Smith, L., & Parkin, J.L. (1994) *The pitch of electrically presented sinusoids*, J. Acoust. Soc. Am. 95, 1677-1679.

Eddington, D.K., Dobelle, Brackman, Mladejovsky, & Parkin, C. (1978) *Auditory prostheses research with multiple channel intracochlear stimulation in man*, Ann. Otol. Rhinol. Laryngol. Suppl. 53, 5-39.

Greenwood, D.D. (1961) *Critical bandwidth and the frequency coordinates of the basilar membrane*, J. Acoust. Soc. Am. 33, 1344-1356.

Marsh, M.A., Xu, J., Blamey, P.J., Whitford, L.A., Xu, S.A., Silverman, J.M., & Clark, G.M. (1993) *Radiological evaluation of multiple-channel intracochlear implant insertion depth*, Am. J. Otol. 14, 386-391.

Shannon, R.V. (1983) *Multichannel electrical stimulation of the auditory nerve in man. I. Basic Psychophysics*, Hear. Res. 11, 157-189.

Tong, Y.C. & Clark, G.M. (1985) *Absolute identification of electric pulse rates and electrode positions by cochlear implant patients*, J. Acoust. Soc. Am. 77, 1881-1888.

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*, J. Acoust. Soc. Am. 71, 153-160.

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

Townshend, B., Cotter, N., Van Compernolle, D., & White, R.L. (1987) *Pitch perception by cochlear implant subjects*, J. Acoust. Soc. Am. 82, 106-115.

Author/s:
Blamey, Peter J.; Parisi, Elvira S.

Title:
Pitch and vowel perception in cochlear implant users

Date:
1994

Citation:
Blamey, P. J., & Parisi, E. S. (1994). Pitch and vowel perception in cochlear implant users. In Fifth Australian International Conference on Speech Science and Technology.

File Description:
Pitch and vowel perception in cochlear implant users