Chapter 2

GAP DETECTION AS A MEASURE OF ELECTRODE INTERACTION IN COCHLEAR IMPLANTS

The results in this chapter have previously been published: Hanekom, J.J. & Shannon, R.V. 1998, "Gap detection as a measure of electrode interaction in cochlear implants", *Journal of the Acoustical Society of America*, vol. 10 no. 4, pp. 2372-2384.

1 INTRODUCTION

The individual electrodes in a modern multi-electrode cochlear implant are intended to selectively stimulate discrete neural populations. However, the assumption that discrete neural populations can be activated is not always true. It is widely assumed that stimuli applied between closely-spaced or adjacent bipolar electrode pairs lead to the localized activation of neurons, whereas widely spaced bipolar electrode pairs (including monopolar stimulation) will lead to broad electrical fields and wide areas of neural activation (van den Honert and Stypulkowski, 1987b; Busby et al., 1994). Even for a closely spaced electrode pair, a broad region can be activated at high stimulation current levels (van den Honert and Stypulkowski, 1987b). The consequence is that when two sets of bipolar electrode pairs are stimulated, and these two sets are closely spaced, overlap can occur in the neural populations excited by the stimulation currents. This overlap of neural populations can occur regardless of whether the stimuli are non-simultaneous or simultaneous. Simultaneous stimuli give rise to direct electrical field interactions, which pose additional problems for electrical stimulation, but even non-simultaneous stimuli may produce activation of overlapping neural regions.

If two electrode pairs stimulate the same neural population or overlapping neural populations, the implication is that sound sensations elicited by the two stimuli might be confused or might even be indistinguishable. This may reduce the number of independent channels of information that can be conveyed to the cochlear implant user's auditory system, presumably resulting in a deterioration of speech recognition ability. If two electrode pairs stimulate the same population of neurons and are perceptually indistinguishable, they probably cannot convey two separate channels of information.

1.1 The number of information channels in an implant

Studies by Fishman et al. (1997) and Lawson et al. (1993, 1996) indicate that increasing the number of electrodes does not necessarily lead to better speech recognition. In fact, very slight or no improvement was evident when the number of electrodes used was increased from 7 to 20. For some speech recognition tasks, no improvement in performance was found when the number of electrodes used increased from 4 to 20. In these experiments, all spectral information that is usually presented across all 20 electrodes, was applied to a limited number of electrodes, i.e. no spectral information was discarded. In a more recent study, Friesen et al. (2001) tested speech recognition as a function of the number of electrodes used in noisy conditions. Nineteen implant users of two different implants (Nucleus-22 and Clarion) participated. A general finding was that, for all noise levels, consonant and vowel recognition scores improved up to seven electrodes, and that speech recognition improved up to ten electrodes, irrespective of the implant used. A reduction in the number of channels was equivalent to a reduction in the signal-to-noise ratio at low signal to noise ratios.

These results suggest that the actual number of information channels available to these patients was not a function of the number of electrodes, and that the actual number of information channels might be limited to somewhere between 4 and 7. Interestingly, in a study with normal-hearing listeners, Shannon et al. (1995) used 4 channel processors and found that listeners achieved near-perfect speech recognition, implying that 4 information channels might be adequate, at least in quiet listening conditions. The study of Friesen et al. (2001) included normal-hearing listeners that listened to a noise-band simulation of a CIS-like processor. (CIS is a stimulation strategy used in the Clarion implant. See Wilson et al., 1991). It was found that speech recognition continued to improve up to 20 channels under similar noise conditions than used for the cochlear implant listeners.

In another study, on patients with the Nucleus cochlear implant device, Hanekom and Shannon (1996) showed that for several different seven-electrode speech processors, speech recognition performance was a function of which set of 7 electrodes were used. This indicates that different choices of which electrodes are used in a processor might lead to different numbers of information channels. A reduced number of electrodes, including only discriminable electrodes, were also used in the speech processors of eleven Nucleus cochlear implant users who participated in a study by Zwolan et al. (1997). While some subjects showed significant improvement in specific speech recognition tasks, others showed a decline in speech recognition performance. Although no strong relationship between electrode discrimination performance and speech recognition was observed, this study again indicates that the choice of electrodes in a reduced electrode processor influences speech recognition ability in some implant users. This supports the suggestion that the number of information channels is a function of the choice of electrodes in a reduced electrode processor. Lawson et al. (1996) measured a larger difference in performance between two different selections of six electrodes than between six and 20 electrodes. This suggests that there should be a way to maximize the number of information channels used for a specific subject by judicious choice of electrodes. Further maximization may be possible using electrical field focussing (Townshend et al., 1987), or by compensating for a missing patch of nerve, or by shifting the speech analysis filters to better match the electrode location (Fu and Shannon, 1999). No maximization of this sort is presently done in implant programming strategies, partly because measurement tools are not yet available and partly because the relation between the electrode interaction and information channel capacity is not well understood.

1.2 Physical factors affecting electrode interaction

To fully account for the effects of electrode interaction we must (1) identify the factors in the patterns of speech that are most important for speech recognition (Shannon et al., 1995), (2) be able to measure the electrode interaction pattern in an individual implant patient, and (3) use the information from both (1) and (2) to optimize the reception of the most important speech pattern information for an individual patient.

A number of variables can influence the interaction of electrodes in a cochlear implant user. These include the electrode placement within the cochlea and nerve survival at the cochlear level and also at the central auditory level.

1.2.1 Electrode placement.

The proximity of the electrode to the surviving neurons, as well as the impedance and paths of current flow between the electrode and the neural population, will determine the spatial selectivity of the stimulation. The impedance and the current pathways could be influenced by new bone formation in the implanted cochlea and encapsulation tissue around the electrode (Grill and Mortimer, 1994). Broad spread of activation will occur if the electrode is physically distant from the excitable neurons (along the lateral wall of the cochlea for example, rather than next to the modiolus) or if nerve survival is poor immediately adjacent to the electrode. Although not routinely used, techniques such as spiral tomography (Wang et al., 1996) are available to measure the exact placement of the electrodes inside the scala tympani. The absolute electrode location can then be used to deduce which nerve fibers will be activated. Finley et al. (1990) modeled nerve fiber activation in a finite element model with idealized electrode placement, but no work has been reported using real electrode placement data.

Although it is clear that placement of electrodes further from the modiolus requires higher stimulus levels to reach threshold and consequently leads to larger current spread, the influence of electrode placement is not yet quantified regarding the interaction or independence of information channels. It is generally assumed that placement of electrodes close to the modiolus is preferable because more focused stimulation can be achieved (Rebscher et al., 1994). Unfortunately, very few tools are available for perceptually assessing and quantifying electrode absolute location and spatial selectivity and their exact influence on speech recognition.

1.2.2 Nerve survival.

A second factor that should affect electrode interaction is the nerve survival pattern in an individual patient. Several anatomical post-mortem studies (Zimmermann et al., 1995; Linthicum et al., 1991; Fayad et al., 1991) have shown from human temporal bones that nerve survival patterns vary greatly among subjects, even for the same disease. It is not clear how the amount and pattern of neuron survival affects implant performance. However, with fewer neurons, the distance between the stimulating electrode and neurons might be larger. Certainly, the further the neurons are distant from the electrode, the larger the current required for activation and the broader the spread of activation. This, in turn, may reduce the number of independent information channels.

It is clear that it is necessary to quantify the available auditory abilities and to optimize the use of the available information channels, i.e. to optimize information transfer in the current generation of implants. Tools are needed to establish the number, the location, and the characteristics of information channels available in each individual cochlear implant user. In this chapter, gap detection is proposed as one such tool.

1.3 Gap detection as a measure of tonotopic spread

Gap detection has traditionally been used as a measure of temporal processing (Plomp, 1969). At moderate levels and higher, normal-hearing listeners can detect 3-5 ms gaps in a stimulus when identical stimuli are bounding the gap, irrespective of the frequency of the stimuli (Penner, 1976; Fitzgibbons, 1983; Florentine and Buus, 1984; Hall et al., 1996; Shailer and Moore, 1983). This results is characterized as the "within-channel" temporal resolution. However, when the frequencies or levels of the two stimuli that bound the gap are different, gap detection thresholds increase about an order of magnitude - to 30-50 ms (Divenyi and Danner, 1978; Divenyi and Sachs, 1979; Formby and Forrest, 1991, Formby et al., 1992). In this case, even the standard stimulus with no gap is perceived as having a discontinuity. The discontinuity that identifies the actual gap must be long enough to be distinctive from this no-gap, standard condition. This temporal comparison must be done centrally "across channels" in that the two stimuli bounding the gap are processed through largely independent neural pathways. A simple model of peripheral frequency resolution can largely explain these results, indicating that gap detection can indicate the degree of neural population overlap between two stimuli (Heinz et al., 1996).

In cochlear implant users, Chatterjee et al. (1998) observed that "within-channel" gap detection thresholds increase when the stimuli marking the gap were of unequal amplitude or unequal pulse rate. They concluded that the perceptual discontinuity caused by dissimilar markers complicated the gap detection task, and suggested that under these conditions gap detection thresholds may be a function both of limitations caused by peripheral mechanisms and a central perceptual distance detector. Their results also emphasize the importance of loudness balancing the stimuli marking the gap.

Shannon (1989) measured gap detection thresholds in cochlear implant users as a function of stimulus level, for both closely spaced (bipolar) and widely spaced (monopolar) electrode configurations, using sinusoids and pulsatile stimuli. He found that gap detection thresholds were a strong function of stimulus level, with the shortest gap thresholds in the order of 1.5 to 3.1 ms regardless of the separation between the active and reference electrodes. He concluded that the temporal resolution for implant subjects was as good as or better than for normal-hearing listeners. However, all measures were made with the stimuli marking the gap on a single electrode pair, i.e., no cross-channel gap detection was done.

The present study measures gap detection thresholds as an indicator of the characteristics of the available neural channels, i.e. the number of channels available, the position of these channels (which electrodes provide independent channels) and the width of the channels. A simple conceptual model is hypothesized which relates gap detection thresholds to neural excitation. When the two stimuli that bound the gap are presented on different electrode pairs, it is expected that gap thresholds will be short if the two electrode pairs stimulate the same neural population. Gap detection in this case is presumably determined by a "within-channel" temporal mechanism and so is determined by the time constant of the peripheral auditory system. As the electrode pairs are separated and the amount of neural overlap decreases, temporal information is carried in separate neural pathways, the stimuli sound more dissimilar and the gap thresholds are expected to increase. Gap detection in this case is presumably limited by the time constant of the centrally located auditory integration because the comparison is made "across-channels". As the two electrodes defining each of the electrode pairs are separated, moving from BP stimulation mode (bipolar between adjacent electrodes) toward BP+3 stimulation mode (bipolar between nonadjacent electrodes with three electrodes separating the stimulation pair), the amount of neural overlap between the two electrode pairs is also expected to increase, resulting in reduced gap thresholds. Using the same argument, the gap thresholds should presumably also be higher for lower levels of stimulation, as there would be less spread of excitation. According to this model, gap detection thresholds can be used to infer the amount of overlap in neural populations stimulated by two pairs of electrodes.

2 METHODS

2.1 Subjects

Three users of the Nucleus cochlear implant participated in this study. All were users of the Nucleus Spectra speech processor, which implements the SPEAK speech processing strategy (McDermott, 1989; McDermott et al., 1991). They were highly trained in various psychoacoustic experiments, having participated in many similar experiments over a period of months. Table 2.1 contains detailed demographic information on the three subjects.

2.2 Electrode parameters

All three subjects used the Nucleus 22 electrode array (Clark et al., 1990), implanted into the scala tympani. The electrodes are numbered from 1 at the basal end to 22 at the apical end. Adjacent electrodes were separated by 0.75 mm. Electrode pairs are referenced by their basal-most member (the *active* electrode); the *reference* electrode is the apical-most member of an electrode pair.

Table 2.1.

Subject information for the three subjects who participated in this study. Insertion depth refers to the number of electrode bands inside the cochlea. The first twenty-two electrodes are active, but eight additional inactive electrode rings aid in placing the electrode and measuring the insertion depth. Speech recognition scores for these subjects were obtained in a previous study (Fishman et al., 1997). Recognition of words from sentences was measured with the CUNY everyday sentences. For consonant and vowel recognition tests, sixteen medial consonants in a v/C/v context and eight vowels in a h/V/d context were used.

Subject	Age	Gender	Age of onset of profound hearing loss	Time of implant use	Processor type	Insertion depth	Cause of deafness	Sentence recognition	Vowel recognition	Consonant recognition
N3	55	Male	45	6 years	SPEAK	27	trauma	61	58	46
N4	39	Male	35	4 years	SPEAK	26	trauma	95	92	95
N7	54	Male	47	6 months	SPEAK	22	unknown; progressive hearing loss	71	98	75

The Nucleus speech processor allows different stimulation modes. Stimuli were presented either in bipolar mode between adjacent electrodes (BP); bipolar between nonadjacent electrodes for electrode separations up to 3 mm (BP+1: 1.5 mm separation; BP+2: 2.25 mm separation; BP+3: 3mm separation); or in pseudo-monopolar mode, using the apical-most electrode as reference electrode. Pseudo-monopolar mode is not a true monopolar mode, as the reference electrode is not located remotely, but inside the scala. In this mode, which will be called AR (apical reference) mode for simplicity, the actual mode of stimulation varies with the active electrode position, so that, for example, when electrode 20 is used as active electrode, the mode is BP+1. The spread of the current field should be larger for larger spacing between the active and reference electrodes of the pair.

2.3 Stimulus parameters

All stimuli were charge-balanced, 200 μs/phase biphasic pulses, with anodic phase first, and were presented at a stimulation rate of 1000 pulses per second. Stimuli were presented at a comfortable level of stimulation. The stimuli were loudness balanced across electrodes before the start of the experiment, using a bracketing loudness balance procedure. First, thresholds and upper loudness levels were obtained in each stimulation mode. Then the subjects were asked to choose a comfortable level of stimulation on electrode 10. All subjects chose comfort levels somewhere between 50% and 85% of their dynamic ranges in the various stimulation modes. All other electrodes were then loudness balanced to this electrode by instructing the subject to adjust the loudness of an adjustable stimulus to be just louder than, then just softer than and finally equal to the reference stimulus. Loudness was adjusted by adjusting pulse amplitude. This was repeated as many times as was necessary to obtain consistent decisions about the relative loudnesses. Loudness balancing was repeated for all conditions (each level of stimulation in each stimulation mode).

Gaps were presented between two 200 ms stimuli. These two stimuli were presented on the same electrodes in the baseline condition and on different electrodes otherwise. Gap thresholds were measured as a function of the separation of the two electrodes. In a single run, the first electrode position was held constant, and gap thresholds were measured for different positions of the second electrode. The experiment was performed in BP, BP+1, BP+2, BP+3 and a pseudo monopolar mode as described above.

A computer program generated the appropriate stimuli and recorded the subject responses. The stimuli were encoded in the correct format to enable presentation directly to the internal receiver of the Nucleus device (without using the subjects' processors), via a custom interface (Shannon et al., 1990).

2.4 Psychophysical procedure

Gap thresholds were collected using an adaptive, two-interval, forced-choice procedure. The gap was initially 100 ms and two consecutive correct decisions led to a decrease in gap size, and one error increased gap size. This procedure estimates the gap size required for 70.7% correct responses (Levitt,1971). Initially the increase or decrease was by a factor of two, but after four reversals this factor was 1.3. Data collection was for twelve reversals and the mean of the last eight reversals was used to estimate the gap threshold.

Gap detection thresholds were obtained in BP+1 mode for all three subjects using all the even-numbered electrodes as standard (the first stimulus). For each standard, the gap thresholds were measured as a function of probe electrode (just even numbered or both even and odd numbered) separation from the standard. Three repetitions were made for each measurement, which resulted in six measurements of gap threshold for each combination of stimulation electrodes when using both orderings of electrodes. (That is, when electrode i was used as standard, three measurements were obtained for probe electrode j, and when j was the standard, another three measurements were obtained with i as probe). Also, gap detection thresholds were obtained in BP, BP+2, BP+3 and AR modes for all three subjects using electrodes 6, 10 and 14 as standard. Again, gap thresholds were measured on even numbered electrodes as a function of probe electrode separation from the standard. In this task two to six measures were taken at each probe electrode location.

3 RESULTS

3.1 Gap threshold as a function of electrode separation

Figure 2.1 compares gap threshold data from Formby et al. (1996) (in normal hearing listeners) to gap threshold data from cochlear implant patients. Formby et al. (1996) measured gap thresholds as a function of marker frequency separation and the study described in this chapter measured gap thresholds as function of electrode separation in cochlear implant patients. The electrode axis in figure 2.1 is scaled to match the approximate location of the linearly spaced electrodes to the cochlear frequency-position function of Greenwood (1990). There is good agreement between the two sets of data in the shape of the gap threshold curves and in the absolute values of gap thresholds for implant subjects N3 and N7. Gap thresholds for implant listener N4 were consistently lower than those from Formby et al. at every comparison point.

Gap thresholds were measured as a function of electrode separation for ten standard electrodes (all the even numbered electrodes) for each of the three subjects (figures 2.2 to 2.4). The lowest gap thresholds were always achieved when the two stimuli that bound the gap were presented on the same electrode. The minimum values of gap threshold were near 1 ms for most electrodes for N4 and 3 to 4 ms for the other two subjects. This is consistent with the range of gap thresholds reported by Shannon (1989).

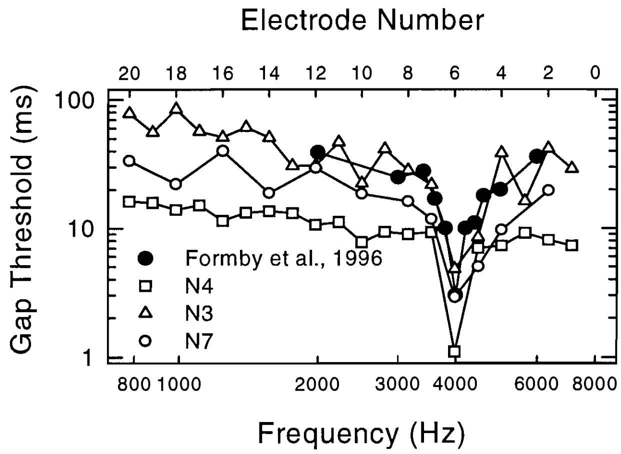

Figure 2.1.

Comparison of gap detection as a function of electrode separation with comparable results from Formby et al. (1996) on gap detection as a function of frequency separation between marker stimuli. The electrode number axis (top) has been reversed and scaled to match the approximate location and extent of the electrode according to Greenwood's (1990) formula. Data for N3, N4 and N7 was obtained in BP+1 stimulation mode.

Gap thresholds increased considerably as electrode separation increased. In general, gap thresholds increased by almost a factor of 10 as the two electrodes were separated. The absolute values and ranges of gap thresholds varied considerably among the subjects, particularly when electrodes were widely separated. Subject N4 had gap thresholds of between 10 and 20 ms for widely separated electrodes, while subject N7 had maximum gap thresholds of 20-70 ms, and N3 had maximum gap thresholds of 100-200 ms.

For most electrodes towards the basal end of the array, the spatial selectivity of the gap threshold curves was sharpest for N4, while N7 had broader selectivity, and N3 had broad "spatial tuning" that covered most of the length of the electrode array. For simplicity, the gap threshold curves will be referred to as "tuning curves". Many of the gap detection tuning curves have two portions: a sharply tuned "tip" region in the vicinity of the standard electrode, and a shallow, bowl-shaped portion for electrodes distant from the standard. These two sections may reflect two different mechanisms relating electrode similarity to gap detection.

Figures 2.2 to 2.4 show a general tendency for the slopes of the bowl-shaped portion to become steeper on the apical side of the gap threshold tuning curves (i.e. towards electrode 20) and shallower on the basal side as the standard moved from base to apex. For electrodes near the base, asymmetry was towards the apex (slopes were shallower on the apical side). This is consistent with measurements of electrode interaction in the same three subjects, using forward masking (Chatterjee and Shannon, 1998). The shallower slopes towards the base (for apical electrodes) suggest larger current flow towards the basal region, but the shallower slopes towards the apex (for basal electrodes) suggest larger current flow towards the apex. Previous measures of electrode interaction using forward masking (Lim et al., 1989) suggested larger current flow towards the basal region for stimulated electrodes at all cochlear locations, but the data presented here does not confirm this observation.

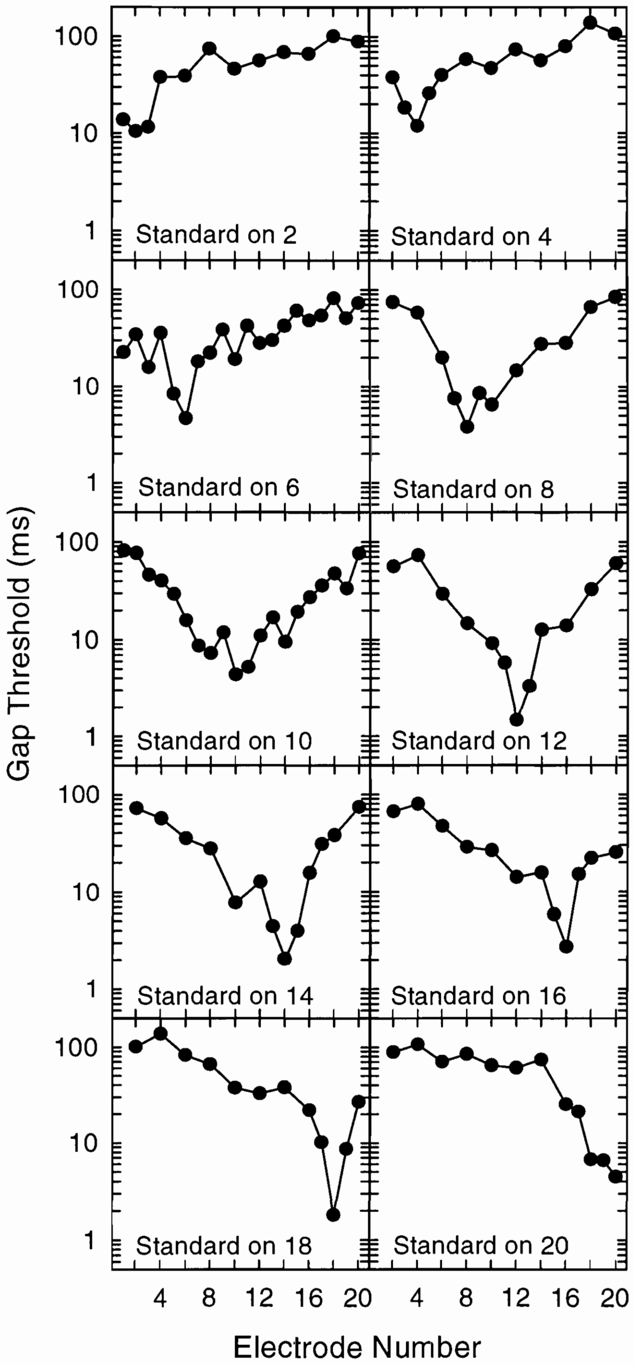

Figure 2.2.

Gap detection for N3 in BP+1 stimulation mode as a function of separation between electrodes. One of the marker bursts was presented to the standard electrode pair and the other to another electrode pair. Gap detection "tuning curves" are shown for all even numbered electrodes as standard. Gap detection thresholds were measured on all even-numbered electrodes.

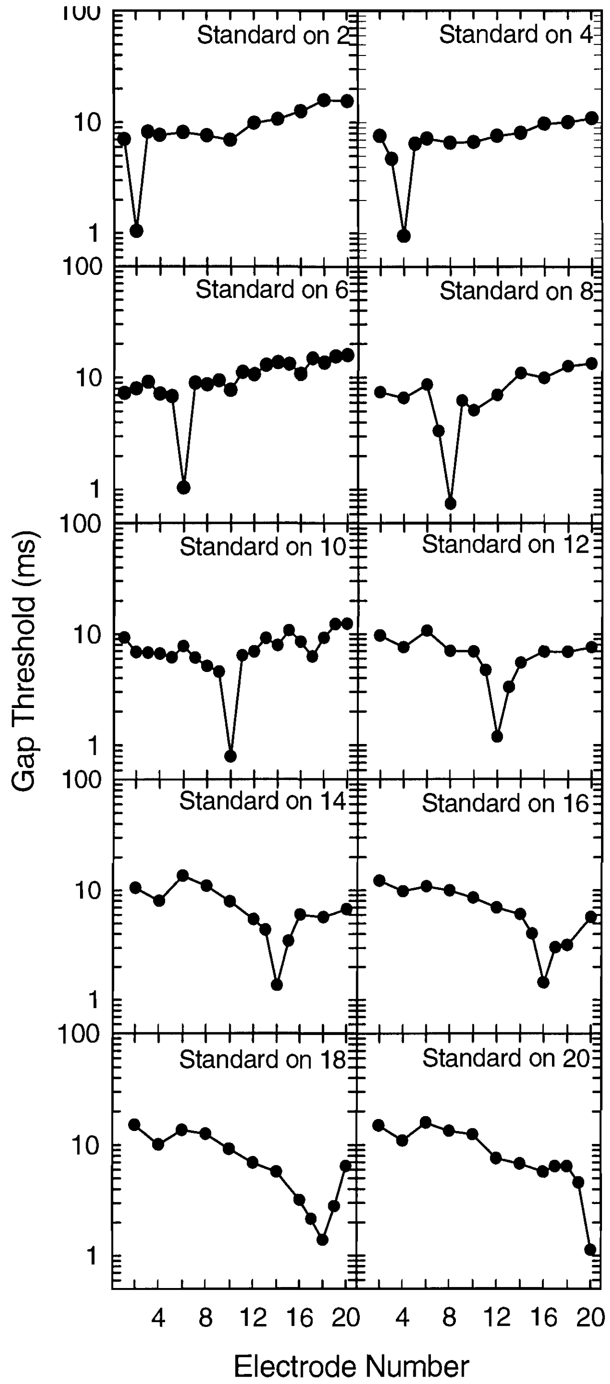

Figure 2.3.

Same as figure 2.2, but for subject N4, using BP+1 stimulation mode.

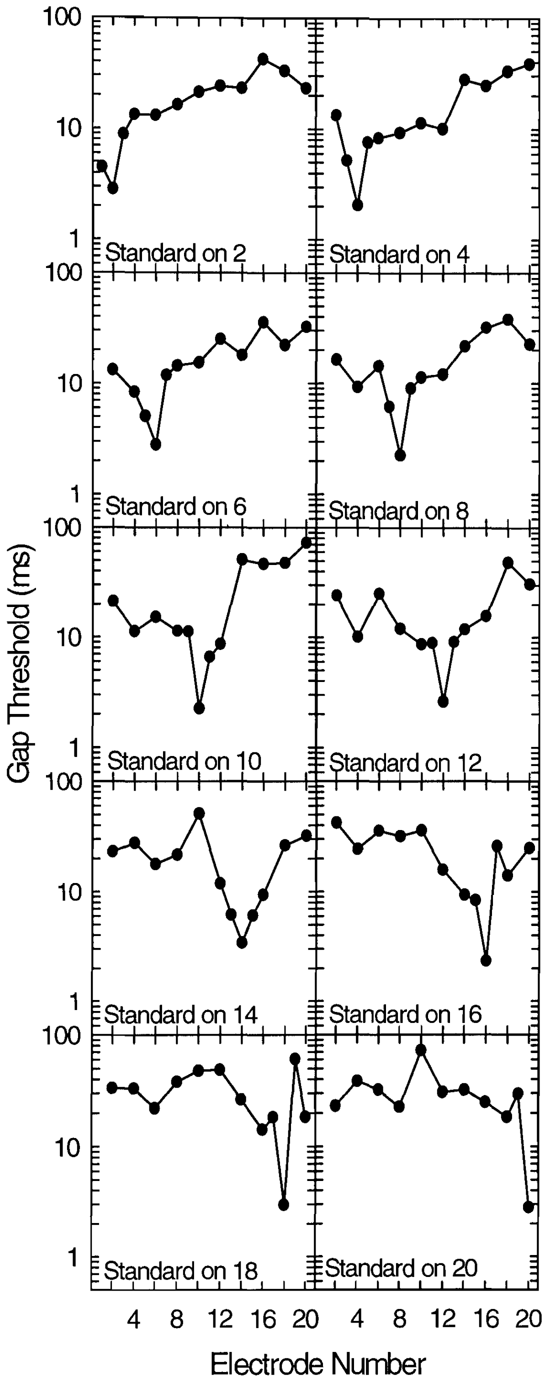

Figure 2.4.

Same as figure 2.2, but for subject N7, using BP+1 stimulation mode.

3.2 Gap thresholds as a function of level of stimulation

Figures 2.5 to 2.7 show the gap thresholds for two levels of stimulation for each of the three subjects, for three standard electrodes. Stimulation was either at a relatively loud level (comfort level was at 81% of the dynamic range for N3, 84% for N4 and 69% for N7) or a softer level (around 30% of the dynamic range for all subjects). BP+1 stimulation mode was used throughout. The most striking difference between the gap threshold tuning curves at high

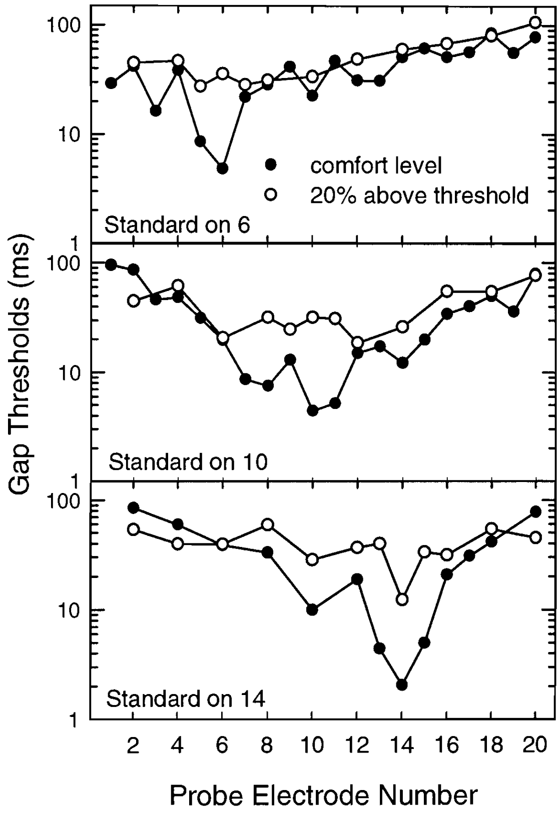

Figure 2.5.

Gap detection tuning curves for N3 for two stimulation levels. The comfortable stimulation level was at 81% of the dynamic range. Each panel represents a different standard electrode location.

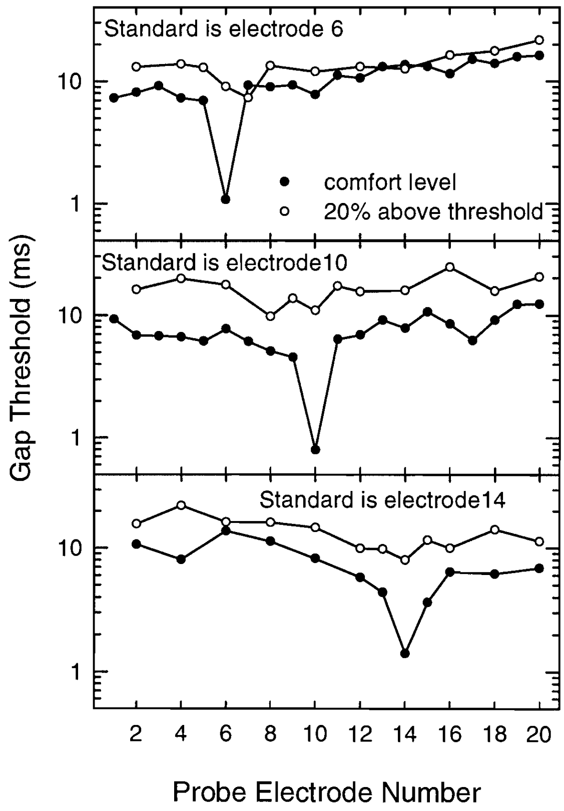

Figure 2.6.

Gap detection tuning curves for N4 for two stimulation levels at three standard electrode locations. The comfortable stimulation level was at 84% of the dynamic range.

and low levels is the absence of the sharp tip region in most cases at low levels. In the region of the tuning curve tips gap thresholds increased at the softer levels in every case. In addition to the loss of the sharp tips, for N4 there was also a 5 to 10 ms increase in gap thresholds across the whole pattern at the lower level. Subjects N3 and N7 did not show a clear shift in gap thresholds with level away from the tip region.

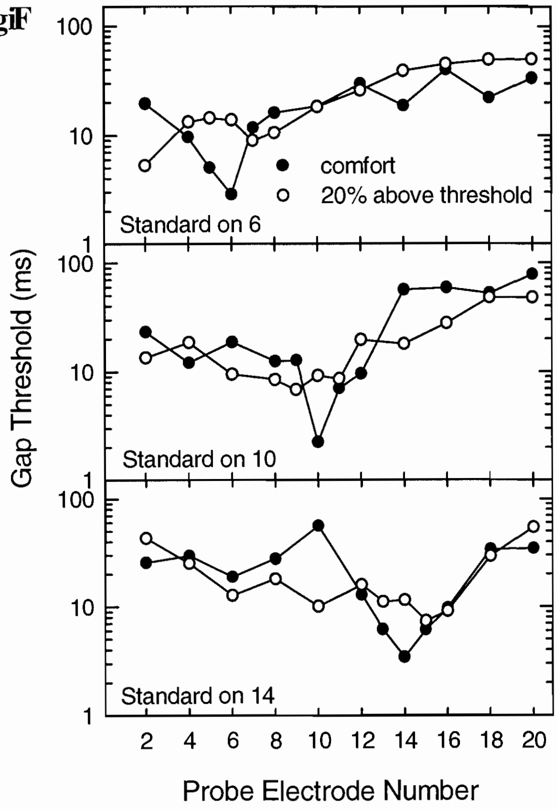

Figure 2.7.

Gap detection tuning curves for N7 for two stimulation levels at three standard electrode locations. The comfortable stimulation level was at 69% of the dynamic range.

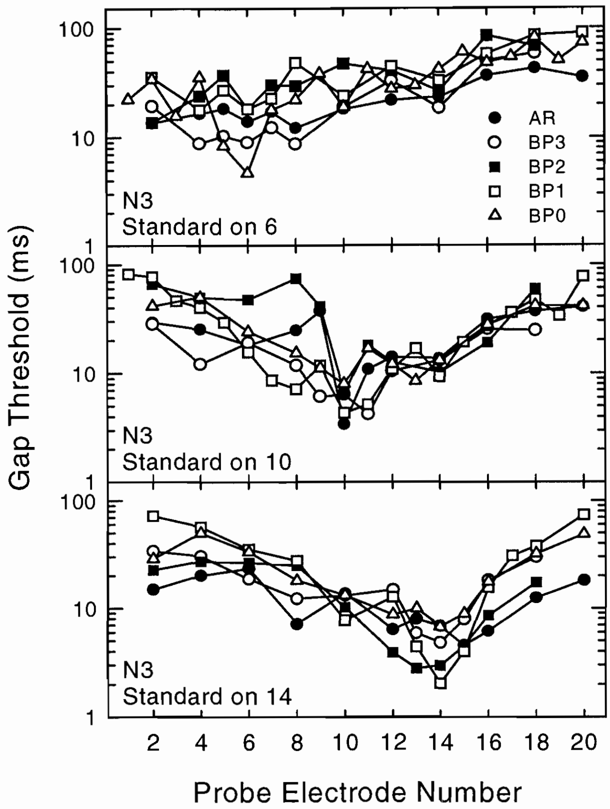

Figure 2.8.

The five curves in each panel are gap detection tuning curves for N3 for five different spacings between the two electrodes of a bipolar pair (i.e., five stimulation modes). The three panels represent measurements at three standard electrode locations.

3.3 Gap threshold as a function of mode of stimulation

N3 (figures 2.8 and 2.11) exhibited the sharpest tuning in BP+1 mode and poorer tuning in all other modes, with the exception of sharp tuning in AR mode when the standard was on electrode 6. Surprisingly, his poorest tuning occurred in BP mode, which should produce the most localized current field. Tuning curves were so flat in BP+2 mode and in AR mode when the standard was on electrode 6 that tuning curve widths could not be calculated. Gap thresholds were greater than 10 ms even at the tip of the tuning curve. There was a 10 ms difference between the AR and the BP gap threshold curves near the tip and a difference as large as 50 ms across stimulation modes away from the standard electrode.

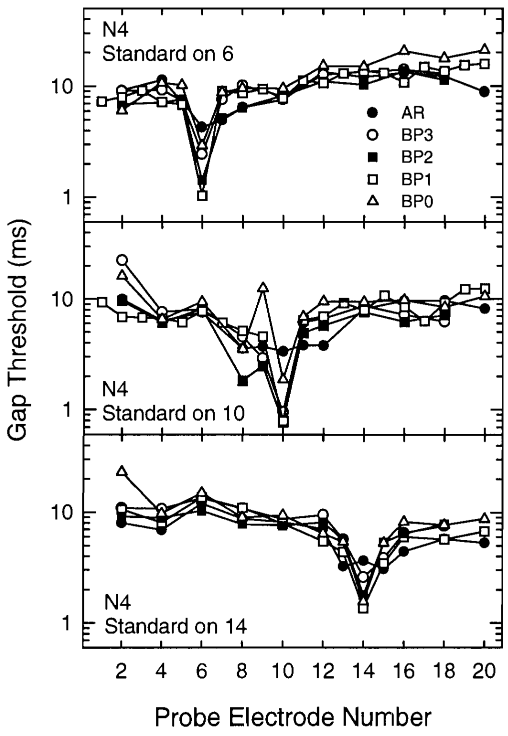

Figure 2.9.

Same as figure 2.8, but for subject N4. Gap detection tuning curves for five stimulation modes at three standard electrode locations.

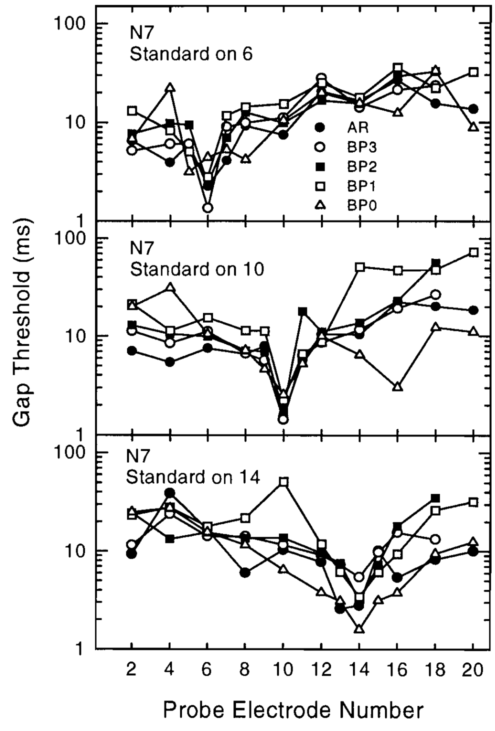

Figure 2.10.

Same as figure 2.8, but for subject N7. Gap detection tuning curves for five stimulation modes at three standard electrode locations.

N4 (figures 2.9 and 2.11) demonstrated almost the same sharp tuning in most stimulation modes, with decidedly broader tuning in AR mode. The difference between the stimulation modes was most pronounced for electrode 6, with BP having a significantly steeper slope than AR. At the tip of the tuning curve, BP+1 produced the lowest gap threshold among the stimulation modes.

For N7 (figure 2.10), tuning did not change dramatically across all conditions. Overall, BP exhibited the widest tuning - even wider than AR mode. Lowest gap threshold also did not change for N7 across all stimulation modes and standard electrode locations.

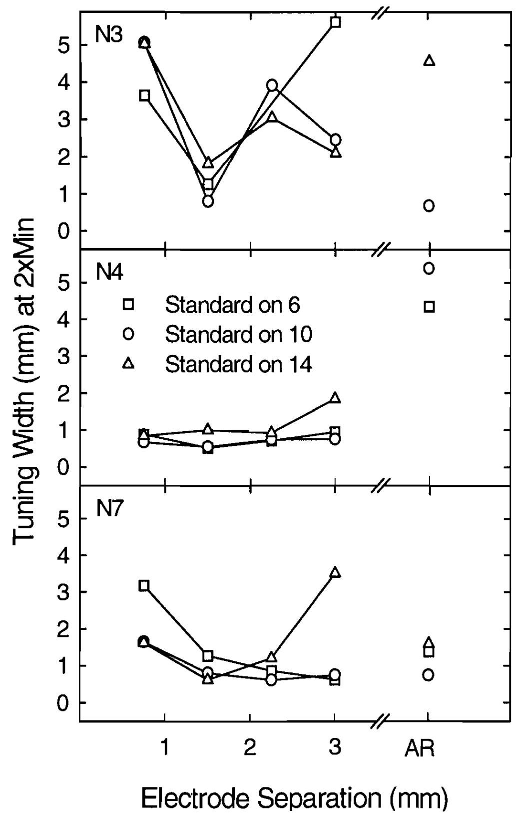

Figure 2.11.

Width of the gap detection tuning curves at twice the gap threshold value at the tip as a function of the separation of active and reference electrodes. The three curves in each panel represent the tuning width measures for three standard electrode locations.

Overall, these three implant listeners showed quite different patterns of tuning as a function of stimulation mode. Based simply on electrical field spread it would have been expected that tuning curve widths would have broadened as the separation between active and reference electrodes in each pair increased. It was expected that AR mode would produce the poorest tuning. No subject showed this expected pattern, although N4 at least showed the poorest tuning in AR mode. N3 showed the lowest gap thresholds and the sharpest tuning in BP+1 mode, which was the stimulation mode used in his normal speech processor. There was no clear change in the pattern of tuning for different standard electrode locations - similar tuning was generally observed for a given listener whether the standard electrode was 6, 10 or 14.

4 DISCUSSION

4.1 Relation between spatial selectivity and gap detection thresholds

The hypothesized relationship between gap detection thresholds and neural activation is reflected in the graphs of gap detection as a function of electrode separation (figures 2.2 to 2.4). Short gap detection thresholds were found where neural interaction was assumed to be large (zero or small electrode separation) and larger gap detection thresholds were found as the separation between the two electrode pairs increased. The conclusion is that a narrow "tuning" in the gap detection thresholds is an indication of good neural selectivity.

The data in figure 2.11 was used to calculate a two-factor ANOVA, to test the statistical relationship between tuning width and two factors: stimulation mode and subject. The ANOVA indicated a statistically significant difference between the tuning widths obtained for the three subjects ($F(2,30)=4.74$, $p=0.016$). For these data, the sentence recognition scores (table 2.1) were higher in subjects with smaller tuning widths, although this cannot be stated as a general rule as the statistical sample was too small.

The shortest gap thresholds observed at the tip of the tuning curves were similar across subjects, 1-4 ms. This is consistent with the results of Shannon (1989) who saw similar gap thresholds across patients at the highest stimulation levels. The two studies thus indicate that there is little relation between the best gap thresholds and speech recognition performance. Both studies found similar gap thresholds across subjects and included subjects with a wide range of speech recognition performance.

However, the longest gap threshold, generally observed for widely separated electrodes, may be related to speech recognition performance. Gap thresholds in this case differed by an order of magnitude between the best and poorest implant user. It is not clear what factor might underlie such a large difference in gap thresholds. The two curve segments (sharp tip and shallow bowl) may indicate the selectivity of two mechanisms of similarity between electrodes (figure 2.12). The sharp tip may reflect a peripheral/neural process that indicates the amount of overlap in the neural populations excited by the two electrodes. The "shoulder" of the gap threshold tuning curves may indicate a point of transition to a condition where electrodes do not stimulate overlapping neural populations. When electrodes are moved even further apart, a further increase in gap detection thresholds would not be expected according to our conceptual model. However, the shallow bowl portion of the function may indicate a weak effect of perceptual similarity for two electrodes that do not activate overlapping neural populations. Although the gap detection must be performed centrally in this case the shallow bowl-shaped function may indicate that there is also a mild effect of overall perceptual similarity on gap detection. The transition that is heard between electrodes that are highly distinctive complicates the gap detection task. Electrodes that are highly distinctive require longer gaps for detection than two electrodes that are less distinctive, even if no neural populations are in common in either case.

The models developed in chapters 3 and 4 shows that the gap detection tuning curves can also be predicted by peripheral mechanisms alone. In the model in chapter 3 (for gap detection in acoustic hearing), the shoulder of the gap detection tuning curve (where the sharp tip transforms into the shallow bowl) is speculated to be a transition from a gap detection task to a gap discrimination task. In the model in chapter 4 it is shown that the sharp tip is obtained when entrainment is close to 100% (i.e. spike occur on every cycle of the stimulus waveform), while the shallow bowl results when entrainment is less than 100%. This notion is supported by the data in figures 2.5 to 2.7, where the sharp tips disappear at lower stimulation levels where entrainment levels are expected to be below 100%.

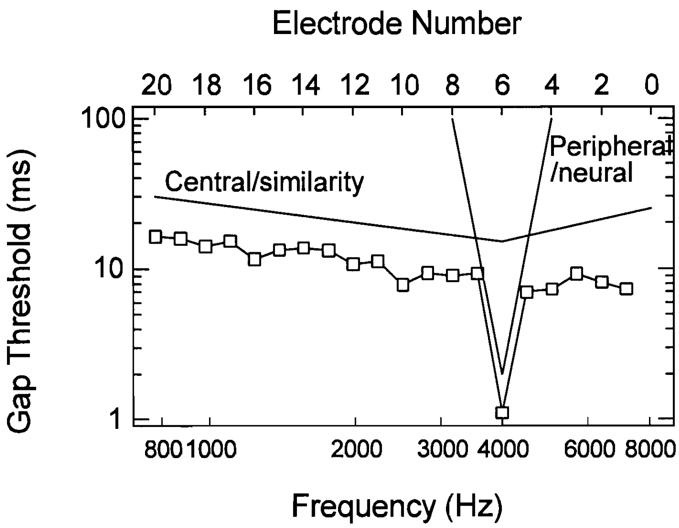

Figure 2.12.

Schematic representation of a conceptual model of gap detection. When the two markers defining the gap excite overlapping neural populations gap thresholds are lowest, but this mechanism is not useful if the neural populations do not overlap. When the neural populations of the two markers do not overlap the gap threshold is determined by a slower, central mechanism. This hypothesized mechanism is only broadly tuned in that markers that are more similar produce slightly lower gap thresholds than markers that are highly dissimilar.

Our original hypothesis was that the longer gap detection time was indicative of the time constant of a central mechanism comparing outputs from different peripheral neural channels. If this is the case, then the long gap thresholds exhibited by N3 may be long enough to interfere with recognition of speech transitions across channels. Long gap thresholds can presumably lead to the misinterpretation or missing of important temporal cues for the identification of consonants (e.g. voice onset time; Divenyi and Sachs, 1978), resulting in poorer speech recognition performance. In the case of N4, poorest gap detection was still 10-20 ms, which may be rapid enough to allow processing of all relevant cross-channel speech transitions. It may be that within-channel gap detection is not a limiting factor for speech recognition, but that long, central cross-channel comparison times can interfere with speech temporal distinctions. It is not clear what might cause the large difference observed in these gap thresholds between different subjects. It appears that there is a larger range of individual differences in the implant results than in similar conditions with acoustic hearing in normal listeners (Formby et al., 1991).

4.2 Relation between stimulation level and spatial selectivity

Gap thresholds are a strong function of stimulus level in cochlear implant users (Shannon, 1989). All of Shannon's measurements were made with identical markers before and after the gap, a condition similar to the tips of the tuning curves in the present data. The data in figures 2.5 to 2.7 also show longer gap thresholds at softer stimulus levels around the tip region. However, gap thresholds did not change as much with level away from the tip region, resulting in the loss of the sharp tips of the tuning curves at lower levels. Less interaction would be expected at lower stimulus levels, because the region of neurons activated should be smaller. Our conceptual model would predict narrower tips at low stimulus levels rather than no tips. However, the interpretation of the present data are confounded by the strong change in gap thresholds with level. In conditions where the markers before and after the gap were on different electrodes (away from the tip region), little change in gap thresholds with level was observed in two of the three subjects. According to our conceptual model this suggests that central mechanisms of gap detection are less dependent on stimulus level than peripheral mechanisms of gap detection. Clearly, more data is needed to validate this suggestion.

4.3 Relation between stimulation mode and spatial selectivity

As the active and reference electrodes in a bipolar pair are separated the electric field becomes more diffuse and spatial selectivity decreases. In general, broader tuning in the gap detection threshold curves was not found as the active and reference were separated. What was not expected was the inconsistency across subjects - gap thresholds varied much more across subjects than across stimulation modes. The ANOVA on the data in figure 2.11 indicated a statistically significant difference between the tuning widths obtained for the three subjects, but no statistical difference between tuning widths for different stimulation modes.

Three possible explanations are proposed. The first is that current spread is already so large that the effect of using stimulation modes with widely separated electrodes has little additional effect (see Lim, Tong and Clark, 1985). A second explanation is that stimulation modes with larger electrode separation do not increase current spread as much as expected. The electrical field model of Finley et al. (1990) predicted a broadly spreading field around banded electrodes, such as those used in the Nucleus implant. Other electrode designs with more localized current distributions might produce more significant and consistent variations in gap detection thresholds in different stimulation modes. A third explanation is that the effects of changing stimulation mode were confounded by the fact that stimulation current level also changed with stimulation mode. Lower currents were used in the stimulation modes with larger spacing between the active and reference electrodes, because these modes produce lower thresholds and uncomfortable loudness levels. The original goal was to change the extent of the neural population activated by using more widely spaced electrodes in a stimulation pair. However, widely spaced electrodes in a stimulation pair result in lower stimulation currents, presumably exhibiting less current spread, and so may partially offset the increased neural extent due to the electrode separation. The net observed result was that sharpness of the gap threshold tuning curves remained effectively unchanged. So it is possible that the effect of stimulation modes with anticipated larger current spread was offset by using lower currents in these stimulation modes.

An additional unanticipated result is that in AR stimulation mode gap thresholds were not the very low values that should correspond to wide spread of stimulation. For wide spread of stimulation, all neural channels receive similar input and there are more channels to aid in gap detection, so a flat tuning curve with very low values of gap thresholds was expected for all separations of electrode pairs. In fact, although AR mode had a slightly flattened curve, the curve was not entirely flat and in most instances the lowest gap threshold values were higher than the tip region of the other stimulation modes. Because true monopolar stimulation was not used, the current paths were directed towards the apical region (where the reference electrode was situated). When detecting the gap between, for example, stimuli on electrode pair (10,22) and electrode pair (19,22), gap detection is performed for an electrode pair with wide current spread activating a large neural population and an electrode pair activating a subset of the same neural population. This complete neural overlap would have been expected to produce low values of gap thresholds. This observation again suggests that the amount of neural overlap is not the only factor in determining the gap detection thresholds. Perceptual dissimilarity between the two stimuli may have confounded the peripheral gap detection mechanism. Our simple conceptual model about this peripheral mechanism cannot explain all the data presented here, and some of the trends might be ascribed to central mechanisms. This is the interpretation in Chatterjee et al. (1998), who found short gap thresholds only when the two stimuli bounding the gap were identical; gap thresholds were long if the two stimuli were perceptually different in any way (pitch or loudness). This result suggests that even if the two stimuli marking the gap activate mostly overlapping neuron populations, the differences in neuron activation may produce a sufficiently different percept that the gap detection decision is primarily central. However, the relative importance of central and peripheral processing mechanisms is unknown.

4.4 Channel characteristics

Gap detection threshold was employed as a tool to provide more insight into the channel characteristics, i.e. the number of channels, the location of channels, the width of channels and the factors that determine channel characteristics. The present results show that the gap detection tuning curves are wider for some choices of electrodes and stimulation modes, and also vary widely across subjects. However, it is not clear at this time how to interpret the gap detection tuning curves in terms of information channels. The following two assumptions are proposed to assist in defining channel width:

(1) Each subject has a minimum gap threshold when the two stimuli are presented to the

same electrode and a large gap threshold when the two electrodes are widely separated. *It was initially assumed that the value of the gap threshold corresponds to the relative amount of neural interaction.* Thus, for each electrode for each subject, there exists a value of gap threshold relative to the minimum such that larger values of gap threshold correspond to negligible channel interaction.

(2) *It is then necessary to make an assumption as to how much neural interaction is negligible*, i.e. how much interaction can be tolerated between two neural channels for them to still be distinct channels. This value is unknown, but it might correspond to the "shoulder" of the gap threshold "tuning curves". It is speculated that this shoulder indicates the point of transition from a peripherally-limited task to a centrally limited task (figure 2.12). As a first approximation a fixed gap threshold value at 40% between the lowest and highest gap threshold can be used, which is in the general vicinity of the shoulder of the gap threshold tuning curves. Another candidate measure for deciding whether two electrode pairs correspond to two different channels, is electrode discriminability. As the electrode pairs are separated and the amount of neural overlap decreases, the stimuli become easier to discriminate. The electrode separation at a chosen level of electrode discrimination (say 75% correct) may be used to find the corresponding gap threshold (from the gap threshold tuning curves). This defines a minimum gap threshold value for electrodes to be discriminable. Gap thresholds larger than this value correspond to two electrodes constituting two different channels.

Electrode discriminability was measured in these same subjects (Hanekom and Shannon, 1996) as described in Appendix 2.A. The 40% measure does not match very well with the electrode discriminability measure. The electrode discrimination measure may be too strict. Two electrodes might be discriminable if there is any difference in the neural populations that they stimulate, but that might not be enough difference to allow them to be independent information channels. Better choices than the 40% measure may be available, but this gives an example of how channels may be defined and this measure was used in the discussion that follows.

Using these assumptions, the following deductions emerge from the results.

(1) Number of channels

The number of distinct channels may be small. Estimating the number of channels from the above assumptions, it is found that N3 (a relatively poor user) may have only a few information channels available, and the upper limit in the number of available channels may be around six or seven (for N4, the best user in the group). Clearly, the number of channels will generally be less than the number of electrodes. One key question raised by this observation is whether improvements in speech recognition can be achieved by selecting electrodes appropriately (Zwolan et al., 1997; Henry et al., 1997; Hanekom and Shannon, 1996). Specifically, can better speech recognition be achieved with a smaller number of independent electrodes or a larger number of interacting electrodes?

When speech processors are programmed with a subset of the total number of available electrodes, many combinations of electrodes are available from which to select. Because of the pattern of interactions in an individual subject, processors with the same number of electrodes can be selected that will have quite different numbers of independent channels (Hanekom and Shannon, 1996). Hanekom and Shannon (1997), using fourteen different seven-electrode processors, made a very simple estimation of the number of channels using gap detection thresholds and the assumptions above and found significant correlation between vowel recognition and the estimated number of channels. Presumably, this relation was due to a clearer formant structure when a larger number of distinct channels were available.

(2) Width and location of channels

The shapes of the gap threshold curves suggest that the tuning may be broad or that channels are relatively wide and typically span many electrodes. Channels become only slightly wider when using stimulation modes with widely spaced active and reference electrodes. Channels are in general wider for the poorest user (N3) and narrower for the best user (N4) in the group. Using the assumptions above, channel width may be between 2 electrodes for N4 (1.5 mm) and 14 electrodes for N3 (10.5 mm). The location of the best (most selective) channels may be deduced directly from the gap detection curves.

(3) Factors determining the characteristics of channels

At least four physical factors determine channel characteristics: (1) electrode placement relative to the remaining nerve fibers; (2) electrode design, which determines the electrical field distribution (Finley et al., 1990), e.g. the Nucleus has a banded electrode design while the Clarion device (Schindler and Kessler, 1993) has a radial electrode placement; (3) nerve survival (Zimmerman et al., 1995); and (4) the current pathway that the stimulation current follows between the active and reference electrodes. Apart from these physical factors, channel characteristics may also be influenced by central auditory nervous system processing.

In existing implants we can only control the current pathway and current spread to some degree by choice of the stimulation electrode pair. Results reported here for the Nucleus device indicate very little difference between the gap threshold tuning curves for different stimulation electrode pair separations and larger variations across subjects. Electrode placement and nerve survival are fixed for an individual implant patient and so cannot be modified after surgery to achieve a larger number of channels.

Although much research has focused on the physical factors influencing cochlear implant user performance, the important question is how these effect the information actually *received*. It is proposed that more research needs to be concentrated on how the *channel capacity* depends on or is related to the physical aspects of cochlear stimulation (electrode design, electrode placement, electrical fields patterns and nerve survival patterns).

4.5 Implications for cochlear implants

4.5.1 Comparison of electrode designs

Different electrode designs exhibit different current spread characteristics (Finley et al., 1990). The results suggest that electrode designs cannot be compared by simply calculating which design produces the most localized current field. All the subjects in our study had the same electrode design, but large differences in selectivity were observed. Selectivity is not a linear function of either current spread or the spacing between the stimulation electrode pair.

4.5.2 Reduced electrode processors

In any nonsimultaneous delivery of biphasic pulses to a number of electrodes there is an inherent trade-off between the number of electrodes and the overall pulse rate. As the number of electrodes decreases a higher pulse rate can be maintained on each electrode. However, the trade-off between pulse rate and number of electrodes is not well understood in terms of their importance to speech recognition. Several recent studies (Fishman et al., 1997; Lawson et al., 1993, 1996) suggest that implant patients are not making full use of all electrodes. It is possible that better speech performance could be achieved with a smaller number of electrodes, selected to be maximally independent channels, that are stimulated at a higher pulse rate. Techniques such as gap detection, forward masking (Shannon, 1983, Lim et al., 1989; Chatterjee and Shannon, 1998), electrode discrimination (Hanekom and Shannon, 1996; Kileny et al., 1997; Zwolan et al., 1997; Henry et al.,1997), or loudness summation (Fu et al., 1996) could be used to help select electrodes for inclusion or exclusion in a processor that uses only a subset of all available electrodes.

4.5.3 Choice of electrodes for a reduced electrode processor

Gap detection thresholds may also be used to compare different choices of speech processors regarding the number of channels, using the assumptions mentioned earlier. Although the actual number of channels is unknown, this method could be used to find the speech processor that maximizes the number of calculated channels. As discussed earlier, Hanekom and Shannon (1997) found that seven electrodes in a processor can lead to a quite different number of distinct channels depending on which electrodes are chosen. Thus, a simple relationship between number of electrodes and quality of speech recognition cannot be assumed.

4.5.4 Choice of electrode pair separation (stimulation mode).

It has been widely assumed that closely spaced bipolar electrodes are necessary for achieving good spatial selectivity in a cochlear implant. However, the present gap detection tuning curves show only minor differences in spatial selectivity as a function of the separation of the bipolar pair. Indeed, in a recent study electrode discrimination and speech recognition were each similar for monopolar and bipolar stimulation (Zwolan, Kileny et al., 1997). To the extent that stimulation mode does effect channel interaction, the optimal configuration may change from one end of the electrode array to the other. This may result in the use of multi-mode speech processors, with each information channel optimized by choosing the electrodes and stimulation modes that result in the best selectivity. Present clinical speech processor fitting software for the Nucleus device allows mixed-mode processor designs, but this feature is not generally used.

5 CONCLUSIONS

(1) Gap detection thresholds are a function of the physical separation of the electrode pairs used for the two stimuli that bound the gap. Gap thresholds increase from a minimum when the two stimuli are presented on the same electrode pair to a maximum when the two stimuli are presented on widely separated electrode pairs. This change may be due to a change-over from a peripheral, within-channel gap detection process for closely spaced electrode pairs to a central cross-channel process for widely spaced electrode pairs.

(2) When the two marker bursts are presented to the same electrode, gap detection thresholds are similar across subjects at 1-4 ms. Gap thresholds for widely separated electrodes vary considerably among subjects and may be related to speech recognition performance, with better implant users having lower gap thresholds in this condition.

(3) The area of neural activation by each electrode (as inferred from the width of the tip region of the gap detection tuning curves as a function of electrode pair separation) varies across subjects and across electrodes. For the three subjects in the present study, the better implant users exhibit sharper tuning, i.e., a smaller area of neural activation around each stimulation pair.

(4) Using stimulation modes with larger separation between active and reference electrodes has limited effect on spatial selectivity. AR stimulation mode, although presumably having larger current spread, has better neural selectivity than BP mode for some subjects. This implies that there is no fixed optimal stimulation mode, but that the optimal stimulation mode may vary across subjects and from one end of the electrode array to the other.

The research presented in this chapter was supported in part by the NIDCD (National Institute on Deafness and other Communication Disorders). The research was made possible by travel grants towards the first author by the University of Pretoria, South Africa and by the House Ear Institute, Los Angeles.

APPENDIX 2.A
SUMMARY OF ELECTRODE DISCRIMINATION STUDY

Electrode discriminability was measured in the same subjects that participated in the study described in this chapter. Details can be found in Hanekom and Shannon (1996), but as this publication may not be easily accessible, a brief description of the study is given here.

This electrode discrimination study determined the amount of electrode confusion with a pitch discrimination experiment. A place pitch ranking matrix (or electrode discrimination matrix) was compiled by using a very simple psychophysical procedure. Consecutive stimuli of 500 ms, separated by a brief quiet interval of 200 ms, were presented on two of the subject's electrodes. The subject's task was to judge which stimulus was higher pitched.

Each stimulus pair consisted of stimuli on two different electrodes stimulated in BP+1 mode. All stimuli were current-balanced biphasic pulses, positive phase first. Stimulation rate was

1000 pulses per second and the pulse phase duration was 200 microseconds. Stimuli were presented at a comfortable level of stimulation above 50%, but below 80% of the dynamic range of the subject. The stimuli were balanced for loudness to minimize confusions between loudness and pitch.

A computer program generated the appropriate stimuli and recorded the subject responses. The electrodes used for the stimulation pairs during the pitch discrimination experiment were completely randomized for each run. One run consisted of the presentation of all possible combinations of electrodes, in both orders of presentation, excluding comparisons of electrodes with themselves. Twenty runs were completed in BP+1 mode for each of the subjects, which gave a total of forty comparisons of each electrode with every other electrode.

Subject reaction, indicating which stimulus was judged to be higher-pitched, was recorded for each stimulation pair and compiled into a response matrix. The response matrix tabulated the number of times that the more basal electrode of a stimulation pair was judged to be higher pitched than the more apical (which would be the expected order based on the tonotopic organization of the cochlea). This matrix was then converted to a percentage correct matrix, under the assumption that a judgement of the more apical electrode to be higher pitched than the more basal in a specific stimulation pair, was an incorrect decision. This resulted in a lower-triangular matrix. This lower-triangular matrix was then converted to a matrix of d' values. The d's gave an indication of the perceptual distance between the stimuli.

Figures 4.34 and 4.36 summarize some of the results. It was found the three subjects required different distances between electrodes for two electrodes to be discriminable ($d'>1$). This distance (ΔE, the number of inter-electrode distances) was also found to be variable across the electrode array for all three subjects. N3 required electrodes to be far apart to be discriminable. Figure 4.36 suggests that electrodes 1 to 16 were not discriminable in N3's case. N4 generally required a ΔE of 2 (see figure 4.34), while N7 generally required a ΔE of 4 or more.