Place Pitch Discrimination and Speech Recognition in Cochlear Implant Users

Johan J Hanekom

Department of Electrical and Electronic Engineering
University of Pretoria

Robert V Shannon

Head: Department of Auditory Implant and Perception
House Ear Institute
Los Angeles, California, United States of America

ABSTRACT

The considerable variability in speech perception performance among cochlear implant patients makes it difficult to compare the effectiveness of different speech processing strategies. One result is that optimal individualized processor parameter setting is not always achieved. This paper investigates the relationship between place pitch discrimination ability and speech perception to establish whether pitch ranking could be used as an aid in better patient-specific fitting of processors. Three subjects participated in this study. Place pitch discrimination ability was measured and this information was used to design new channel to electrode allocations for each subject. Several allocations were evaluated with speech tests with consonant, vowel and sentence material. It is shown that there is correlation between the perceptual pitch distance between electrodes and speech perception performance. The results indicate that pitch ranking ability might be used both as an indicator of the speech perception potential of an implant user and in the choice of better electrode configurations.

OPSOMMING

Die beduidende verskille in spraakherkenningsvermoë van kogleêre-inplantpasiënte bemoeilik die vergelyking van die effektiwiteit van verskillende spraakverwerkingsstrategieë. 'n Gevolg is dat die individuele instelling van spraakverwerkerparameters vir pasiënte nie altyd optimaal gedoen word nie. Hierdie artikel ondersoek die verband tussen plek-toonhoogtediskriminasie en spraakherkenning om te bepaal of toonhoogte-rangskikking nuttig is as hulpmiddel vir beter gebruiker-spesifieke passing van spraakverwerkers. Drie proefpersone het aan hierdie ondersoek deelgeneem. Plek-toonhoogtediskriminasie is gemeet en die inligting hieruit is gebruik vir die ontwerp van nuwe afbeeldings van elektrodes op kanale. Verskeie afbeeldings is evalueer met spraaktoetse met vokaal-, konsonant- en sinsmateriaal. Daar word aangetoon dat daar korrelasie bestaan tussen toonhoogtediskriminasie en spraakherkenningsvermoë. Die resultate wys dat toonhoogte-rangskikkingsvermoë gebruik kan word as beide 'n indikator vir die spraakherkenningspotensiaal van 'n kogleêre-inplantgebruiker en vir die beter keuse van elektrodekonfigurasies.

KEY WORDS: cochlear implants, multi-electrode stimulation, pitch discrimination, speech recognition, neural selectivity, perceptual distance.

INTRODUCTION

Two parameters, which influence the speech perception abilities of cochlear implant users, are the quality of spectral information and the quality of temporal information received by their electrically activated auditory systems. This paper focuses on the importance of spectral information. Multiple-electrode stimulation is preferred in cochlear implants, because it is generally accepted that the tonotopic organization found along the length of the cochlea in the healthy auditory system is retained to some degree for electrical hearing. Many research studies, including earlier work by Eddington (1980) and a recent study by Nelson, Van Tasell, Schroder & Soli (1995), have shown that electrodes stimulating the more basal areas of the cochlea result in higher perceived pitches (or sharper tonal quality) and stimulation closer to the apex results in lower perceived pitches.

It would be natural to assume that tonotopic organization in electrical hearing is retained by multiple electrodes selectively stimulating discrete neural populations. However, the assumption that discrete neural populations can be activated is not always true. When electrodes are closely spaced, considerable overlap occurs in the neural populations excited by the stimulation current, a problem which was addressed by Townshend & White (1987). This

is the result of spread of electrical current in the biological medium of the cochlea.

The implication is that if two electrodes stimulate the same neural population or overlapping neural populations, sound sensations elicited by the two stimuli might be confused or might even be indistinguishable. This reduces the number of independent channels of information that can be conveyed to the cochlear implant user's auditory system. It is important to realize that the number of independent channels of information is not equal to the number of electrodes.

The question which presents itself then, is how much overlap in stimulation of auditory neurons does occur in the cochlea, and how important is this in the patient's ability to understand speech? Although it is generally assumed to be true, is it really beneficial to have excitation of distinct neural populations? In other words, on the one hand it is assumed that multi-electrode implants perform well because they utilize the tonotopical organization of the cochlea, but on the other hand current spread inside the cochlea might defeat this purpose by having different electrodes stimulating identical neural populations. If many electrodes stimulate the same neural population, why do some patients perform so well on open set speech recognition tasks? A secondary question which follows is whether using simpler implants with a smaller number of electrodes (which in turn could potentially be cheaper and more reliable) would not perform as well as implants with many electrodes.

The questions formulated above and related questions have been investigated by a number of researchers. The effect of the number of electrodes on speech perception performance has been addressed in Lawson, Wilson, Zerbi & Finley (1996). They found that by increasing the number of electrodes, speech perception performance is improved, but for as few as four to seven electrodes, speech performance levels are close to what can be achieved by using ten or twenty electrodes.

Busby, Whitford, Blamey, Richardson & Clark (1994) studied patients' abilities to rank electrode pitch as a function of stimulation mode of the Nucleus[1] implant device. (Clark, Tong & Patrick (1990) give detailed descriptions of this device). As will be explained in the text to follow, the Nucleus device can utilize different stimulation modes, which produce differences in the amount of current spread from the electrodes. Busby *et al.* (1994) found that the ability to rank pitch for stimulation on a specific electrode (in other words, place pitch) was related to the mode of stimulation used (and thus the amount of current spread).

Nelson *et al.* (1995) studied the relationship between pitch ranking ability (or electrode ranking ability) and consonant perception in ten subjects using the Nucleus cochlear implant. They found correlation between the consonant perception and pitch discrimination, but they found little correlation between recognition of consonants based on recognition of place cues and place pitch perception. However, they concluded that this might be related to the speech processing strategy not taking full advantage of the user's ability to do place pitch ranking.

This paper addresses some of the questions mentioned above. Specifically, we investigate the question of the re-

lationship between speech perception performance and the stimulation of overlapping neural populations in the cochlea. The approach used is to find a measure of the amount of overlap among stimulated neural populations. This is accomplished by determining the amount of electrode confusion with a pitch discrimination experiment. An additional question, which we also address, is whether it is possible to improve speech discrimination with the correct design of electrode configuration.

EXPERIMENTAL APPROACH

One way to quantify the amount of overlap among neural populations stimulated by different electrodes, is to measure the amount of confusion between electrodes. We measure the pitch discrimination between electrodes as a measure of the amount of electrode confusion, which is then also a measure of the amount of overlap in the neural populations stimulated. This is in turn a measure of the amount of current spread resulting from electrical stimulation. The procedure we use is to compile a place pitch ranking matrix (or electrode discrimination matrix) which is transformed to perceptual distance values, as explained below. Based on the pitch discrimination information, we then design various maps and evaluate the speech perception performance for these.

We focus primarily on the relationship between pitch discrimination data and speech perception performance for various different maps in a single subject, while repeating only some of the tests in other subjects. In this respect our approach differs from that used by Nelson *et al.* (1995). They compared speech perception abilities in ten subjects using their everyday maps[2], and related this to the subjects' place pitch ranking abilities. However, they did not study the effect of using different maps in the same patient. The rationale for this would be to investigate to which *dimensions* of place pitch discrimination speech perception is related. The term *dimensions* refers to the fact that place pitch discrimination ability might be related to various physical electrode parameters, for example electrode spatial separation and current spread from the electrodes, but also to perceptual distance between electrodes. Our study evaluates nine different maps in the same subject, to establish whether speech recognition performance is related to the perceptual distance between electrodes or to other dimensions of pitch discrimination ability.

It will become clear in the description of the properties of the pitch discrimination abilities of each of the subjects why more maps were evaluated in one specific subject. The three subjects who participated included a good user,

---

[1] One of the most widely used cochlear implant devices is the Nucleus, manufactured by Cochlear Pty Limited in Australia and by their United States subsidiary, Cochlear Corporation.

[2] The term *map* used in this context, usually refers to the patient-specific settings that are made to the Nucleus processor. A map is a table of values with the threshold and uncomfortable loudness stimulation current values for each electrode. The map also holds information on the specific *frequency allocation table* used. The frequency allocation table is a table specifying the filter cutoff frequencies used in the twenty channels of the Nucleus speech processor. Different frequency allocation tables are available. The software used in this study to program the Nucleus device allows the user an extra option, namely to allocate multiple filter channels to the same electrode. When we refer to *map* in this article, we are actually referring to this filter-to-electrode allocation. The other parameters we used for the maps were from the subjects' everyday maps and remained unchanged throughout the experiments, unless noted otherwise. *Everyday map* refers to the regular maps that the subjects used daily.

a fair user and a relatively poor user. We relate our findings to this observation.

Perceptual distance between stimuli is quantified by d', a measure often used in psychophysical studies to express perceptual distances in forced choice experiments. This measure was used in the context of pitch discrimination studies by Townshend et al. (1987) and more recently by Nelson et al. (1995). Smaller values of d' indicate more confusion between stimuli. Negative values of d' in the pitch discrimination experiment indicate pitch reversal. For no confusion between stimuli, the largest value of d' (3.29) is achieved. The d's can be calculated from signal detection theory as described in Green & Swets (1966). Hacker & Ratcliff (1979) tabulated the values of d' for a two alternative forced choice experiment such as is described here.

We used a vowel test, a consonant test and a sentence test in the evaluation of speech perception ability with each map. In previous studies (for example Nelson et al., 1995), consonant perception was used to assess speech perception ability. We included a vowel test, because vowels are recognized primarily by their formant structure (Dubno & Dorman, 1987) and as such their recognition should be dependent on the ability to activate discrete neural populations selectively. Sentence material was included in the speech perception tests to evaluate open set speech understanding for the various different maps.

The number of electrodes used in the maps in our experiments was six or seven throughout. The reason for this choice was that it was found that seven-electrode maps gave speech perception performance levels that allowed some play for the speech perception scores to improve or deteriorate. A study by Lawson et al. (1996) showed that speech recognition demonstrated a rapid decline when the number of electrodes was lowered from seven to four to two and to one, and that with a seven-electrode map it is possible to get speech discrimination close to what can be achieved by ten or twenty-electrode maps. An advantage of using a reduced number of electrodes is that the stimulation rate of the Nucleus processor increases. Higher stimulation rates have proved to result in better speech recognition performance (Wilson et al., 1991).

A corollary of this study is to establish a procedure which could optimize an electrode map when only a small number of electrodes are used. Fewer electrodes than the twenty available in the Nucleus might be used for several reasons. Other implants that use fewer electrodes are available. It is also possible that only a small number of stimulation sites are available as a result of poor neural survival, or that a reduced number of electrodes are available because of electrode damage. Another possible application for fewer electrodes would be a future lower cost device.

The rest of this paper is discussed in two sections. In the first section, the pitch discrimination experiment is described and in the second section the results are used in the design of maps which are evaluated with speech perception experiments.

PITCH DISCRIMINATION EXPERIMENT

METHODOLOGY

Subjects

Three users of the Nucleus cochlear implant participated in this study. All of them were users of the Nucleus Spectra speech processor. This processor implements the SPEAK speech processing strategy, which is described in Skinner et al. (1994). Table 1 contains detailed information on the three subjects.

Electrode parameters

All three subjects used the Nucleus 22 electrode array (described in Clark et al., 1990, p. 114), implanted into the scala tympani. The electrode bands in this array are separated by 0.75 mm. We refer to the two electrodes of a stimulation pair as the stimulation electrode and the return electrode. Stimulation was always with current-balanced biphasic pulsatile waveforms, with the positive pulse preceding the negative pulse. The stimulation electrode was the electrode on which a positive first biphasic pulse could be measured if the other electrode was used as reference. The return electrode was always the more apical of a stimulation pair. The electrode numbering convention used in this paper is as follows: an electrode pair is referred to by the stimulation electrode number, and with the stimulation mode known (explained below), the return electrode is implicitly known. Electrode 1 was the most basal electrode and electrode 20 was the most apical electrode used as stimulation electrode.

The Nucleus speech processor allows different stimulation modes. In BP mode[3], the stimulation electrode and

TABLE 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.

Subject	Age	Gender	Age of onset of profound hearing loss	Time of implant use	Processor type	Insertion depth	Cause of deafness
EWB	55	Male	45	6 years	Spectra	27	trauma
JEM	39	Male	35	4 years	Spectra	26	trauma
REK	54	Male	47	6 months	Spectra	22	progressive hearing loss

[3] The abbreviation BP, for bipolar, is used throughout this paper. This is the standard abbreviation used in texts by Cochlear Pty Limited to indicate the Nucleus device's stimulation modes.

Johan J Hanekom en Robert V Shannon

the return electrode are adjoining. In BP+1 mode one electrode separates the stimulation electrode and the return electrode. In BP+2 mode two electrodes separate the stimulation electrode and return electrode. Stimulation modes up to BP+3 were used in this study. The symbol ΔE is used for electrode spatial separation.

Stimulus parameters

Only the BP+1 stimulation mode was used for subjects EWB and REK, because their everyday maps utilized this mode of stimulation. For subject JEM, who also used a BP+1 everyday map, we used stimulation in BP, BP+2 and BP+3 modes in addition to BP+1 mode. The reasons for this will become clear in the discussions to follow.

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. The length of a single presentation on a specific electrode was 500 ms and no ramping was used. Stimuli were presented at a comfortable level of stimulation above 50%, but below 80% of the dynamic range of the subject. 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. The coded stimuli were presented directly to the internal receiver coils of the subjects (the subjects' processors were not used), via a custom interface (described in Shannon, Adams, Ferrel, Palumbo & Grandgenett, 1990).

Psychophysical procedure

A place pitch discrimination matrix was measured for each of the three subjects. BP+1 stimulation mode was used for all three subjects, and in addition we also measured pitch discrimination matrices for BP, BP+2 and BP+3 mode for subject JEM.

The pitch 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. The procedure was controlled by a computer program. The subject had to indicate his choice by depressing either of two buttons, with one button corresponding to the first sound and the other button to the second sound. The subsequent pair of stimuli was only presented after a delay of 200 ms after the patient had made his choice. There was no option for repetition of the pair of stimuli.

Each stimulus pair consisted of stimuli on two different electrodes. The stimuli were balanced for loudness to minimize confusions between loudness and pitch. This was done by a loudness balancing procedure which preceded the initial pitch discrimination experiments. A reference electrode was chosen, usually electrode 10, which is in the middle of the electrode array. The subject was asked to indicate a comfortable listening level, and for this purpose the stimulation levels were varied somewhere between 50% and 80% of the dynamic range of the subject. Then the loudness balancing was conducted. This consisted of the reference stimulus being presented, followed by a stimulus on any of the other electrodes. This stimulus pair was repeated as many times as necessary in order to find a level setting for the second stimulus which had the same

apparent loudness as the first stimulus. To be more accurate in pinpointing the stimulation level which was equal in loudness to the reference, subjects were also asked to find a level that was just noticeably softer than the reference and also to find a level just noticeably louder than the reference. In this way three datapoints were established for each electrode's stimulation level and we were able to make a good estimation of the stimulus current necessary to have all the electrode stimuli of equal loudness.

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. For BP+1 mode, this amounted to twenty electrodes times nineteen comparisons, equalling a total of three hundred and eighty comparisons per run. Each run took between fifteen and twenty minutes to complete. To be able to calculate a measure for the amount of confusion between electrodes, a considerable number of comparisons were needed. 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. The larger the number in the percentage correct matrix, the higher the d′ value. The maximum value of d′ is 3.29 for 100% consistency in judgement.

Although the value of d′ was also an indication that there was a definite pitch difference between two stimuli, the value of d′ is not a direct indication of how much that pitch difference was. Two closely spaced electrodes might activate two discrete neural populations and might have a very distinct pitch difference, resulting in a large d′, without the pitch difference being very large. d′ is more probably an indication of whether overlapping or discrete neural populations are stimulated.

If perfect tonotopicity was present, electrodes with lower numbers (located more basally) would be expected to be consistently judged higher in pitch than electrodes with higher numbers (the more apical electrodes) in a paired comparison. Also, if completely separate neural populations (non-overlapping) were stimulated by each electrode, then even for close electrode spacing, no confusion between electrodes was expected. On the other hand, if current spread from the electrodes was large enough so that two electrodes stimulated the same neural population, confusion is expected between the pitches of these two electrodes. Thus, the data in the pitch discrimination matrix is an indication of the amount of current spread in the cochlea.

RESULTS

The stimulus-response matrices for the three subjects are shown in figures 1 to 4 after they have been converted

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	0.36																		
3	0.62	-0.07																	
4	1.00	0.70	0.91																
5	0.18	0.51	0.25	0.00															
6	0.36	-0.36	0.62	-0.62	0.07														
7	0.07	-0.25	0.62	-0.18	-0.18	0.00													
8	1.53	0.36	1.53	0.43	0.62	1.00	0.51												
9	1.66	1.81	2.09	1.53	1.53	1.35	1.14	0.25											
10	1.66	2.33	2.33	2.33	2.33	2.33	1.24	1.66	0.36										
11	2.90	1.81	3.29	3.29	1.81	1.66	2.09	1.66	1.24	0.91									
12	2.90	1.81	2.90	3.29	2.90	2.90	3.29	2.33	1.81	1.53	0.51								
13	2.90	3.29	3.29	3.29	3.29	2.90	2.90	2.09	3.29	1.81	1.66	1.14							
14	3.29	2.90	2.90	3.29	2.33	2.90	3.29	3.29	1.81	1.66	1.81	1.24	0.91						
15	3.29	3.29	3.29	3.29	2.90	3.29	3.29	2.90	2.90	2.33	1.66	2.33	1.66	1.14					
16	2.90	3.29	3.29	3.29	3.29	3.29	3.29	2.90	2.90	3.29	2.90	2.09	1.53	1.35	1.14				
17	3.29	3.29	3.29	3.29	3.29	3.29	3.29	2.90	3.29	2.90	1.53	2.33	2.33	1.81	1.81	0.18			
18	3.29	3.29	3.29	3.29	3.29	2.90	2.90	3.29	2.90	2.09	2.90	3.29	1.66	1.53	1.66	1.00	0.91		
19	3.29	3.29	3.29	3.29	3.29	2.33	3.29	2.33	3.29	2.90	3.29	3.29	2.09	2.90	1.66	1.24	1.14	0.62	
20	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	2.90	2.90	2.33	2.33	2.09	2.09	2.09	1.14	0.78	0.70	0.36

FIGURE 1. Pitch discrimination matrix for subject JEM for BP+1 mode of stimulation. Positive values of d′ in the table indicate that the electrode numbered at the top was judged to have a higher pitch than the electrode numbered at the left. d′ values of 3.29 indicate 100% consistency in pitch judgement and values higher than 1.5 indicate an 85% consistency. JEM had a large area of good pitch discrimination.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
2	0.18																
3	1.19	0.74															
4	0.00	0.18	-0.74														
5	0.18	0.74	-0.55	0.54													
6	0.54	0.18	-0.74	0.18	0.18												
7	1.19	0.18	0.18	0.74	0.95	0.36											
8	1.81	1.47	1.47	2.33	1.81	1.47	0.95										
9	2.33	3.29	1.81	2.33	3.29	1.81	2.33	1.47									
10	3.29	2.33	2.33	3.29	3.29	2.33	2.33	2.33	0.74								
11	3.29	3.29	3.29	3.29	3.29	2.33	3.29	1.81	1.81	0.95							
12	2.33	3.29	3.29	3.29	3.29	3.29	3.29	3.29	2.33	3.29	1.81						
13	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	1.19	1.47					
14	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	1.47	0.18				
15	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	2.33	2.33	1.81	1.47	0.18			
16	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	1.81	0.74	0.36		
17	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	1.19	1.19	0.18	0.36	
18	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	3.29	2.33	1.47	3.29	2.33	0.95	0.95	0.54	0.18

FIGURE 2. Pitch discrimination matrix for subject JEM for BP+3 stimulation mode. Note that the number of available electrodes was only 18 instead of 20. That is because the return electrode, for electrode 18 as stimulation electrode in BP+3 mode, is electrode band 22, which is the most basal electrode band. Also, note that the pitch discrimination pattern differs from that in figure 1. Some electrodes which did not exhibit good discrimination in BP+1 mode, had good discrimination in BP+3 mode. Compare the d′ between electrodes 11 and 12 in these stimulation modes. The inverse was also true for some electrodes.

to d′ values. Two aspects are immediately evident from these matrices. First, the distribution of areas where electrode discrimination was better, varied considerably for the three subjects. This might have been dependent on electrode placement, with areas of good electrode discrimination being where the electrodes were situated closer to the (surviving) nerves. This result underlines the fact that there are two important but uncontrollable factors in perception of sound in cochlear implants: (1) placement of electrode and (2) nerve survival.

If a d′ of 1.5 was (arbitrarily) taken as a criterion for largely independent neural populations (this corresponds

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	0.36																		
3	0.36	0.36																	
4	1.47	0.95	0.54																
5	0.74	0.74	-0.36	0.00															
6	1.19	1.19	1.81	0.18	0.18														
7	0.95	1.47	1.81	0.74	-0.36	0.18													
8	1.19	0.95	0.74	0.00	-0.55	-0.74	0.18												
9	0.54	-0.18	0.36	0.00	-1.47	-0.95	-1.19	-0.74											
10	1.47	1.19	1.47	0.54	0.00	0.74	0.00	0.54	1.47										
11	1.81	2.33	2.33	1.19	1.47	0.95	0.18	0.54	1.47	0.54									
12	2.33	1.81	1.19	1.81	0.74	0.54	1.19	1.19	1.19	0.74	-0.36								
13	3.29	3.29	1.81	1.19	1.47	1.19	0.95	0.74	3.29	0.74	0.18	-0.18							
14	2.33	3.29	2.33	3.29	0.54	1.19	0.74	0.54	1.47	0.18	0.18	0.54	0.00						
15	2.33	3.29	1.81	1.81	0.95	1.19	1.19	1.47	1.81	0.54	0.36	0.54	0.36	-0.36					
16	3.29	2.33	1.81	1.81	1.81	0.54	0.18	1.81	1.47	0.74	0.36	0.36	0.00	0.18	0.00				
17	3.29	0.95	1.81	2.33	0.36	0.36	0.54	0.74	1.47	1.19	0.36	0.18	0.36	-0.18	-0.36	0.74			
18	1.47	0.00	0.36	0.36	-0.36	-0.18	-0.18	0.00	0.74	-0.55	-0.95	1.47	-0.74	-0.18	-0.74	-1.19	-0.95		
19	0.18	-0.74	-0.36	-0.55	-0.18	1.47	1.47	1.47	0.95	-2.33	1.81	1.47	1.19	1.81	1.47	1.81	0.00	-0.55	
20	1.19	-0.55	0.36	-0.18	-0.18	-1.19	0.74	-0.36	0.55	-0.95	-1.19	1.42	1.19	1.81	1.19	1.47	1.81	-0.36	1.47

FIGURE 3. Pitch discrimination matrix for subject REK. The values in the table are the calculated d′ values. Positive values of d′ indicate that the electrode numbered at the top was judged higher in pitch than the electrode numbered at the left more than half of the time. For normal tonotopic organization, the d′ values are expected to be positive, as electrode 1 is the most basal electrode and electrode 20 the most apical. REK had a region of good pitch discrimination toward the left side of the matrix.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
2	-0.07																		
3	-0.70	0.00																	
4	0.14	-0.14	0.00																
5	0.54	-0.07	0.62	0.54															
6	-0.21	0.07	-0.21	0.21	0.14														
7	-0.43	-0.14	-0.07	-0.07	-0.43	0.36													
8	0.29	-0.07	-0.21	-0.21	-0.21	-0.21	-0.14												
9	0.14	-0.29	0.62	0.21	-0.07	0.00	0.07	-0.14											
10	0.87	-0.14	0.36	-0.07	0.43	0.62	0.07	0.29	0.36										
11	0.54	0.21	0.87	0.07	0.14	0.00	0.21	0.43	0.43	-0.14									
12	0.78	0.29	0.70	0.14	-0.14	0.70	0.14	0.21	0.29	0.36	-0.07								
13	0.36	0.00	0.54	0.07	0.21	0.54	0.36	0.54	0.54	-0.14	-0.36	0.00							
14	0.29	-0.21	0.21	-0.21	0.00	0.07	0.21	0.21	0.14	-0.43	0.00	-0.62	-0.29						
15	0.43	-0.07	0.07	0.14	-0.36	0.14	0.07	0.00	0.14	0.07	-0.78	-0.43	0.14	0.21					
16	0.54	-0.21	0.00	-0.36	0.00	-0.43	-0.14	0.07	-0.70	-0.70	-0.36	-0.87	-0.95	-0.36	0.07				
17	0.95	0.54	1.14	0.62	0.70	0.70	0.87	0.87	0.54	0.21	0.70	0.07	0.29	1.05	0.78	1.05			
18	1.05	0.78	0.87	1.24	0.87	0.87	0.62	1.35	0.95	0.70	0.70	0.70	0.29	0.78	1.14	1.24	0.43		
19	1.66	0.78	1.14	1.14	1.14	1.35	0.95	1.99	1.35	0.70	1.05	0.95	0.78	1.24	1.05	1.81	1.05	-0.07	
20	1.14	1.47	1.05	1.05	1.05	1.66	1.35	1.47	0.95	1.24	1.24	0.87	1.05	1.66	1.24	1.47	1.14	1.14	0.78

FIGURE 4. Pitch discrimination matrix for subject EWB. This subject exhibited poor pitch discrimination throughout the electrode array, although he did a little bit better toward the apical side.

to 85% consistency in pitch ranking), then subject EWB had almost no discrimination between any of his electrodes, although he did a little better toward the apical side. Subject REK had a region of good pitch discrimination, but primarily for large electrode separations. Subject JEM had very good electrode discrimination throughout his electrode array when the electrode spatial separation was two or more. His best region was near to the middle of his electrode array.

More details concerning the analysis of the properties of the subjects' pitch discrimination matrices are discussed when the maps used in the speech perception experiment are described below.

SPEECH PERCEPTION EXPERIMENT

METHODOLOGY

Objective

The speech perception experiment was used to investigate whether there was correlation between subjects' ability to discriminate among electrodes based on place pitch (as reflected by the information in the pitch discrimination matrices), and speech perception performance. Two parameters (which could be deduced from the pitch discrimination matrices) were identified as being important in speech perception: the perceptual distance between pitch sensations elicited by electrodes as reflected by the d′ values and ΔE, the electrode spatial separation. The maps used where chosen to reflect the effect on speech perception of varying these two parameters.

Speech materials

We used vowel, consonant and sentence material to evaluate speech perception performance. All the speech material was available on laser disk.

The vowel test used a set of eight vowels in a /hVd/ context uttered by a male talker. Each vowel syllable was repeated three times in randomized order. The consonant test used sixteen consonants in a /aCa/ context. The consonant syllables were repeated five times in randomized order during a single consonant test.

The sentence test consisted of a set of thirty-six sentences uttered by a female speaker. No repetitions of the same sentences were used and a new set of sentences was used for the evaluation of every map. None of the subjects had been tested with the specific sentence material before.

Procedure

The software used for the creation of the subject maps allowed the programmer to allocate the output of any of the Nucleus processor's filters to any electrode. Multiple filter outputs could be allocated to the same electrodes. This enabled us to use maps utilizing a reduced set of electrodes, while still presenting all the spectral information from the twenty filters to the electrodes. The operation of the Nucleus implant is such that reducing the number of electrodes increased the stimulation rate on the electrodes that were used (Shannon et al., 1990).

Subjects were tested with their regular BP+1 maps and frequency allocation tables which they had been using daily

for several months. The only changes to these maps were that the regular filter-to-electrode allocations were replaced by maps with a reduced number of electrodes of which multiple filter channels were allocated to each electrode. The specific choices of electrodes which were used in the various different maps are explained below. The subjects wore these experimental maps for a period of two full days before evaluation with the speech material.

Tests were conducted in a sound-isolated booth. Speech stimuli were presented at 60 dB SPL. Speech stimuli were played from a laser disc player through high quality audio loudspeakers. A computer program controlled the presentation of the speech material to the subjects. The subjects responded by indicating their choice on a computer keyboard. The computer program recorded the subject reaction for the vowel and consonant tests, and compiled stimulus-response matrices for these. The computer program also controlled the laser disc player for presentation of the sentences. The subject had to repeat as many words as he could understand from the sentence material, which was then recorded by the experimenter.

Map parameters

Details of the maps that were used for each patient are explained below. As explained earlier, all maps were seven-electrode maps, except two of JEM's maps, which were six-electrode BP+3 maps. Maps were chosen to give either large or small cumulative values of d′, and maps with similar cumulative d′ were evaluated using different electrode spatial separations. Both orderly and disorderly electrode spatial separations were used. Table 2 summarizes the maps used. The motivation for the choice of each of the maps is given in the descriptions below.

We evaluated more maps for JEM than for the other two subjects. This was primarily dictated by the fact that subject JEM's pitch discrimination matrix provided more degrees of freedom in the choice of various maps. This statement will become clear in the explanation of the procedure used to choose electrodes to be used in the maps.

This procedure was very simple and was as follows: A list of all possible combinations forming seven-electrode maps was compiled (for example electrodes 1, 2, 3, 4, 5, 6, 7 or electrodes 1, 3, 5, 7, 9, 11, 13 and so on). Then, for each subject, the cumulative d′ was calculated for all these electrode sets by simply adding the corresponding six d′ values. For example for subject REK, for the electrode set consisting of electrodes 1, 2, 3, 4, 5, 6 and 7, the six d′ values to be added were 0.36 (d′ between electrode 1 and 2), 0.36 (the d′ between electrode 2 and 3), -0.36, 0, 0.18 and 0.18. The cumulative d′ was thus calculated as 0.36.

It was assumed that electrodes were distinguishable if d′ was larger than 1.5. From subject REK's pitch discrimination matrix in figure 2, it can be seen that only a small region in the lower left corner of this matrix produced d′ values in the region of 1.5 or larger, and then for an electrode spacing of four or larger. For smaller a criterion level of d′>1, a larger area in the d′ matrix was identified. Between electrodes 9 and 10 a d′ of 1.47 was produced, but for all other electrodes the electrode spatial separation had to be at least three to find a d′ larger than 1. It can also be seen that electrodes could not be discriminated at the apical end of the electrode array. In this area all electrodes were confused and d′ values were generally very low. This means that to find a seven-electrode set with a large value

of cumulative d′, the electrodes had to be chosen more toward the basal side of the array. Even then we would not be able to choose an electrode set with all d′ values greater than 1. For example, if we started with electrode 1, the next electrode had to be electrode 5 for d′ greater than 1. Then, the third electrode had to be electrode 11 to have a d′ greater than 1 between electrode 5 and the third electrode to be used. From electrode 11 toward the apical side of the electrode array, all electrodes were confused.

Comparing this to JEM's pitch discrimination matrix, it becomes clear that for JEM there was much more flexibility in possible choices of electrodes for good cumulative d′. A d′ of greater than one was found for all elec-

trodes if electrode spatial separation was eight, and for most electrodes if the spatial separation was only two or larger, excluding electrodes to the more basal side of the array. This gave us many different possibilities of choice of seven-electrode maps with good cumulative d′. Confusion between electrodes occurred primarily on the basal electrodes.

There seemed to be total confusion of electrode pitch for all electrodes in subject EWB's case, except for a few electrodes near the basal end of the array. To find a d′ of larger than 1, electrode spatial separation had to be at least 14 in the more apical half of the electrode array, and no d′s of 2 or larger were found. This gave very little flex-

TABLE 2. Results for the speech perception experiment. The number D is a simple measure for the disorder in electrode spacing as calculated by equation 1. ΔE_a is the average spacing between electrodes. d'_c is the cumulative d′. No of sites refer to the number of discrete stimulation sites when d′ > 1.5 is used as criterion. Cons and Sent are abbreviations of Consonant and Sentence, respectively. The table lists the stimulation modes used for each electrode for the mixed stimulation mode map for JEM (map 8). The abbreviations BP1 and BP2 are used for BP+1 and BP+2 stimulation modes, respectively.

SUBJECT: JEM									
Map no	Electrodes used	Stim mode	d'_c	ΔE_a	D	No of sites	Vowel test	Cons test	Sent test
1	3,4,9,13,15,17,19	BP+1	10.3	2.7	2.2	5	92	88	95
2	6,7,8,9,10,11,12	BP+1	2.54	1	1	1	54	76	20
3	2,4,6,9,11,17,2	BP+1	4.98	3	2	2	88	88	88
4	4,5,10,11,16,17,19	BP+1	7.46	2.5	4	3	79	83	85
5	3,4,9,13,15,17	BP+3	6.53	2.8	3	3	83	89	92
6	2,4,6,9,11,17	BP+3	7.27	3	2.3	4	79	85	82
7	12,13,14,15,16,17,18	BP+1	5.42	1	1	1	62	81	66
8	2,5,8,11,14,17,20 (BP,BP,BP2,BP2,BP2, BP2,BP1)	mixed	13+	3	1	5	92	90	97
9	2,4,7,10,13,16,19	BP+1	6.64	2.8	1.1	3	88	85	98

SUBJECT: REK									
Map no	Electrodes used	Stim mode	d'_c	ΔE_a	D	No of sites	Vowel test	Cons test	Sent test
1	1,3,6,10,12,14,16	BP+1	4.37	2.5	1.8	2	46	56	38
2	1,2,5,12,13,14,19	BP+1	0.60	3	5	1	58	65	74
3	2,4,6,8,10,12,14	BP+1	2.00	2	1	1	50	55	38
4	2,5,8,11,14,17,20	BP+1	-0.87	3	1	1	46	88	94
5	1,4,7,10,13,16,19	BP+1	0.41	3	1	1	67	74	87

SUBJECT: EWB									
Map no	Electrodes used	Stim mode	d'_c	ΔE_a	D	No of sites	Vowel test	Cons test	Sent test
1	3,5,8,16,17,18,20	BP+1	3.72	2.8	3.25	1	38	49	30
2	1,4,7,9,12,14,16	BP+1	-0.55	2.5	1.5	1	62	53	33
3	1,4,7,10,13,16,19	BP+1	0.86	3	1	1	67	49	57

ibility in choice of electrodes if a seven-electrode map was to be constructed, because although a range of different cumulative values for d´ could be found, the electrode pitch sensations could not be discriminated for many of the possible electrode pairs. The hypothesis we wanted to test was whether electrodes that were clearly distinguishable were better choices to achieve good speech perception performance. At least we wanted to be able to compare good and poor electrode discrimination.

Consequently, only three different maps were tested for subject EWB. Two choices of cumulative d´ were contrasted: a map with EWB's best possible cumulative d´ (3.72), and a map with a very poor cumulative d´ of -0.55 were used. Map 2 (with the poorer cumulative d´) had the electrodes spaced more orderly than map 1, with ΔE at least 2. ΔE was 2 toward the apex and 3 toward the base. Map 1 had a very disorderly electrode spacing with a big gap between electrodes 8 and 16 and some electrodes having spatial separation of only 1. The third map had an intermediate value of cumulative d´ and all ΔE were 3 throughout.

Five maps were tested for subject REK. Map 1 utilized both a good cumulative d´ and a reasonably good electrode spacing. Although the electrode spacing was irregular, ΔE was equal to or larger than 2. The cumulative d´ of 4.37 was close to the best cumulative d´ of 5.12 achievable for REK. The second map had a small cumulative d´ of only 0.6, but with the interesting property that the d´ values between every second electrode were relatively large. This map possibly stimulated only four distinct sites in the cochlea. Map 3 had a cumulative d´ of 2, which was near 50 % of the maximum achievable d´, with a very orderly electrode spacing of 2. The electrodes used were situated in the middle of the electrode array. Maps 4 and 5 also used very orderly electrode spacings, but this time the spatial separations used were 3. The cumulative d´s were the small value of -0.87 and 0.41 respectively.

Nine maps were tested for subject JEM. These maps tested speech perception performance for a spectrum of cumulative d´s, from very low to the highest achievable values for this subject. Some of these maps had very orderly electrode spacing, and some of them had very irregular electrode spacing. Six of the maps used BP+1 mode (which was the stimulation mode used in JEM's everyday map). Two of the maps used BP+3 mode and one used different stimulation modes on different electrodes. Map 1 was chosen for maximum d´ (10.34) in BP+1 mode. The electrode spacing was irregular with the largest electrode spatial separation of 5 between electrodes 4 and 9. Map 2 was chosen for low d´ (2.54) and had an electrode spacing of 1. Map 3 had an irregular spacing with a large separation gap between electrodes 11 and 17. The map was chosen for its cumulative d´ of 4.98, in the middle of the d´ value of maps 1 and 2. Map 4 was chosen to be similar to map 2 of subject REK. Although this map had a large cumulative d´ of 7.46, electrodes were grouped into pairs, with the spacing between the pairs being only 1, and these electrodes were essentially stimulating the same areas within the cochlea. Map 5 en map 6 used exactly the same electrodes as maps 1 and 2 respectively, but BP+3 stimulation mode was used. These maps were chosen to test whether the wider spread of stimulation current in BP+3 mode would influence the speech results negatively. Thus, if d´ was a good indication of current spread, then larger current spread would result in smaller d´ values and poorer

speech perception performance. Map 7, a BP+1 map, tested another ΔE = 1 map, this time closer to the base, where JEM had better electrode discrimination. Map 8 was an attempt to create a map with the maximum achievable d´, with the best possible electrode spacing. Different stimulation modes were used on different electrodes. The stimulation modes were chosen to give the best achievable d´s. To obtain the d´s for the other stimulation modes, pitch discrimination matrices were compiled for these. The cumulative d´ for this multi-stimulation mode map was larger than 13. The exact value is not known, because the d´s between electrodes 5 and 8, and also between electrodes 17 and 20, were unknown because the electrodes in these pairs used different stimulation modes for each electrode. Map 9 was a ΔE = 3 map with an intermediate value of d´.

RESULTS

Although it was clear that electrode spatial separation was a parameter in determining speech perception performance, we had to find a way to quantify this parameter. Various simple measures of how well the electrodes were spaced, were used. We calculated a value for the average ΔE, $ΔE_a$, by simply adding all the inter-electrode separations and dividing by the number of inter-electrode separations (six for a seven-electrode map).

We also calculated a number, D, for the amount of electrode disorder caused by using irregular electrode spacing. The rationale for using this measure of disorder was that when we used these seven-electrode maps, we always added the same filter outputs, but for each choice of map these filter outputs were relayed to a different electrode. In some map we connected two neighbouring filters to electrodes that were spaced far apart. Thus, we were actually introducing spectral distortion, and the number D calculated for electrode disorder gave some indication of how much this distortion was.

A third measure of the quality of our choice of electrode spacing was to find a measure for the number of discrete stimulation sites activated. This could be found from the d´s under the assumption that the perceptual distance between two electrodes was an indication of how much overlap there was in neural excitation. We assumed that a d´ of 1.5 reflected reasonably little overlap, because this value of d´ translated to an 85% consistency in pitch judgements.

Table 2 summarizes the results for all the maps tested for the three subjects. The electrodes used in each map, the cumulative d´s and the three measures reflecting the electrode spatial separations are included in the table. The results are given as percentage correct scores for the vowel, consonant and sentence tests. Several interesting observations can be made from these results when we investigate the relative contribution of the two parameters (d´ and electrode spacing) to speech perception performance.

Simple linear regression analysis was used to relate various map parameters to the results of the speech perception tests. Regression lines and correlation coefficients are given in the figures. Significance of correlation was tested by the t-test and a 5% level of significance was used.

Speech perception as a function of cumulative d´

As explained earlier, perceptual pitch distance as measured by d´ was used as measure of the amount of overlap

in neural population stimulated by electrodes. For a specific seven-electrode map, cumulative d′ is also simply a measure of a subject's ability to pitch rank this set of electrodes. Figures 5 to 7 relate speech perception results to this measure for each of the subjects. The three figures for each subject give results for vowels, consonants and sentences, respectively.

As was also demonstrated by Nelson et al. (1995), significant correlation between pitch ranking ability and speech perception performance was evident. There was, however, considerable variation in how pitch ranking ability was correlated with speech perception. The regression lines and correlation coefficients for linear regression are given in the figures.

In general, for subject JEM, for whom the most data were available, higher cumulative values of d′ were correlated to better performance on the three speech perception measures. There was a correlation of around 70% for

all three speech perception measures to cumulative d′. There was a significant increase in vowel perception with larger values of d′, the total range being nearly 40%. Consonant perception increased over a smaller range of 14%.

Far too few data points were available for subject EWB to be conclusive. No correlation between vowel perception and cumulative d′ was evident (r=0.03), but there was correlation with consonant perception (r=0.95) and sentence perception (r=0.68).

The vowel test did not have significant correlation to the cumulative d′ for subject REK. The weakest vowel performance was achieved for the smallest and the largest values of d′. This subject showed an inverse relationship between cumulative d′ and consonant scores. The best consonant recognition scores were achieved for small values of d′; increasing the cumulative d′ decreased consonant perception. The correlation coefficient of 0.84 is significant. The range of consonant performance scores was

FIGURE 5. Speech perception as a function of cumulative d′ for subject JEM. Regression lines for linear regression are fitted through the data. The correlation coefficients are shown above each panel. Although the correlation coefficients are significant for all three speech perception performance measures, the slope was shallow for the consonant test.
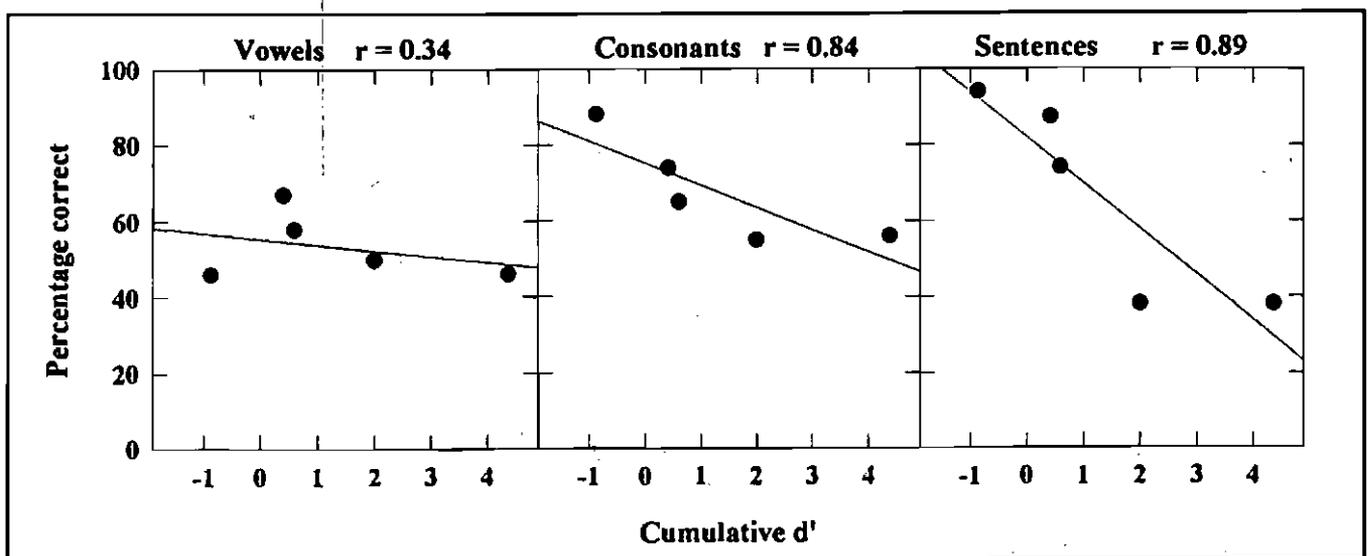

FIGURE 6. Speech perception as a function of cumulative d′ for subject REK. Regression lines for linear regression are fitted through the data. The correlation coefficients are shown above each panel. This subject demonstrated a decline in performance levels for increasing cumulative d′.

more than 30%. There was even stronger correlation with the perception scores for sentence material, the relationship also being inverse.

The general tendency in all three subjects was that increase from very small cumulative d's to intermediate values influenced speech perception scores more than increases from intermediate d' values to large d' values. The variability among subjects and also the strong inverse relationship between d' and speech perception scores for REK, prompted us to investigate which other parameters also influenced speech perception.

Speech perception as a function of average electrode spacing

As described earlier, three measures were used to quantify the influence of electrode spacing on speech perception. Figures 8 to 10 show speech perception performance as function of the average electrode spacing. The number E_a is an indication of the general spacing of the electrodes

used in a specific map. The larger the number, the more widely spaced the electrodes are.

In general, there was significant correlation between electrode spacing and speech perception for the three subjects. The correlation coefficients are given in figures 8 to 10. Although not enough data were available for subject EWB, the tendencies shown were that consonant and sentence perception was correlated to electrode spacing, but that vowel perception had little correlation with electrode spacing. REK also showed little correlation between electrode spacing and vowel perception, but JEM showed significant correlation between electrode spacing and vowel perception. His worst vowel perception performance was for the two maps with $\Delta E = 1$. Good vowel performance was achieved for all maps when ΔE was larger than 2 for most electrodes. JEM also showed strong correlation between consonant performance and ΔE, but as mentioned previously, the range of consonant scores was small and improvement with larger ΔE was minimal. REK achieved his best consonant scores for $\Delta E = 3$. Also, REK had his
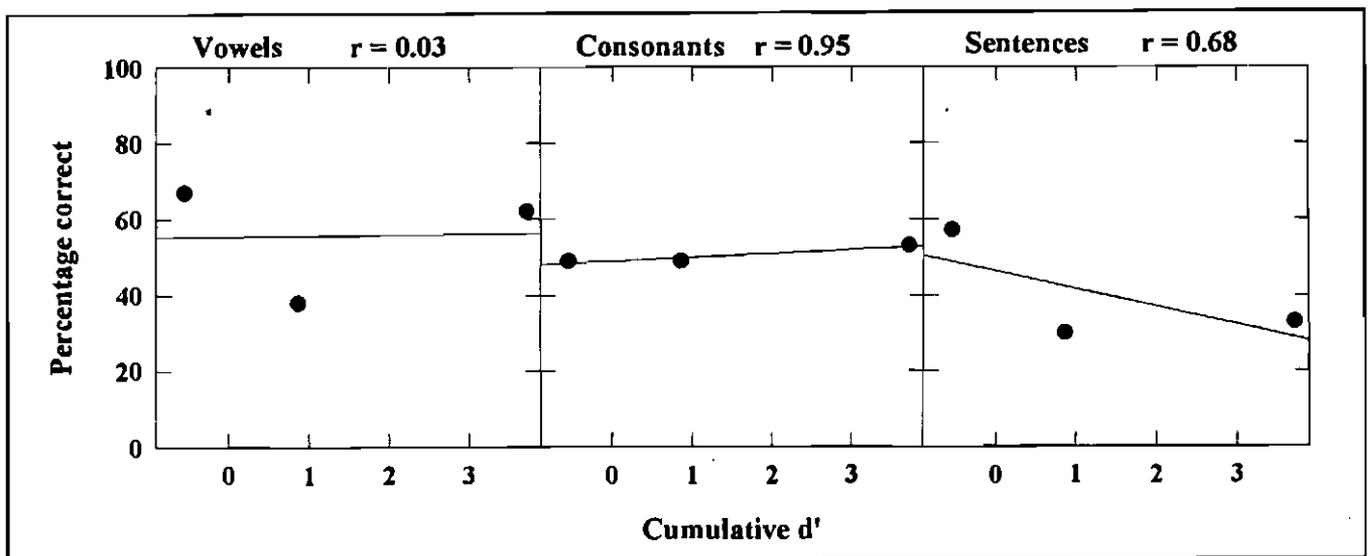

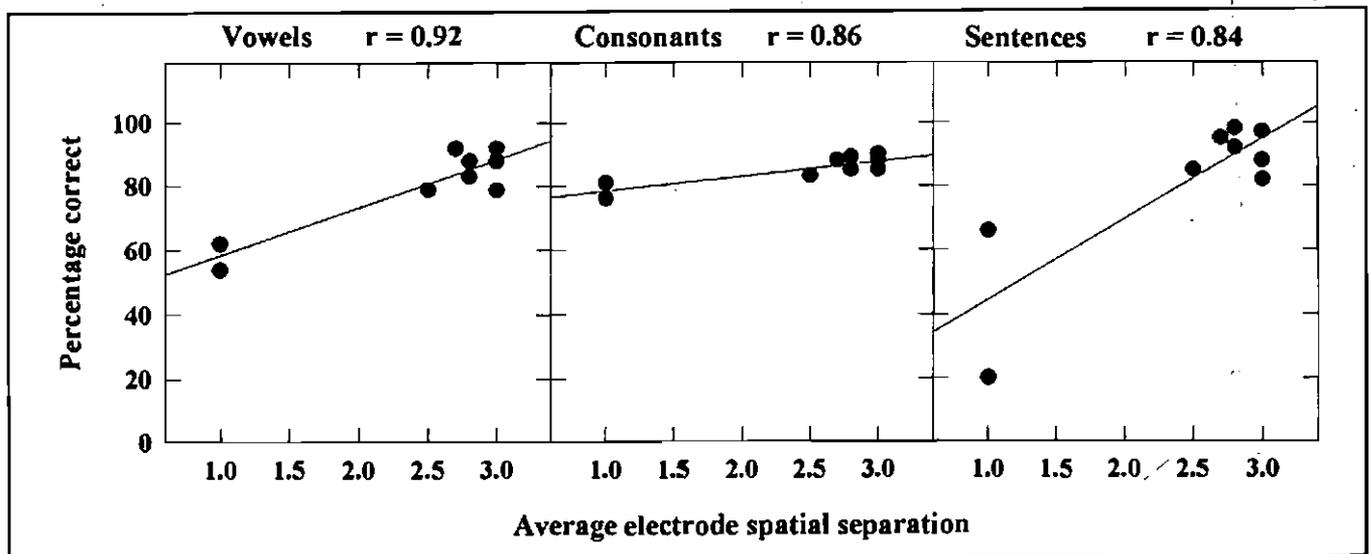**FIGURE 7.** Speech perception as a function of cumulative d' for subject EWB.

FIGURE 8. Speech perception as a function of average electrode spatial separation for subject JEM. Linear regression lines fitted through the data indicate correlation between electrode separation and speech perception performance. The correlation coefficients are given above each panel.

best sentence scores with ΔE = 3. JEM achieved good sentence scores for all ΔE above 2.5, and his worst sentence scores were for ΔE = 1.

Speech perception as a function of disorder of electrode spacing

The second measure used to quantify electrode spacing was the amount of disorder in the electrode choice. Looking at the results for average electrode spacing, one gets the impression that these might somehow be correlated to how regular (or how disordered) the chosen electrode spacing was. A simple measure of the amount of disorder of electrode spacing (a number to which we assign the symbol D) was calculated with this equation:

$$D = \frac{(Sum\ of\ the\ three\ largest\ electrode\ spacings)}{(Sum\ of\ the\ three\ smallest\ electrode\ spacings)} \quad (1)$$

For the BP+3 maps with only six electrodes, we used only the two largest and two smallest electrode spacings in this equation. The values obtained for D are reflected in table 2.

If the spatial separations among all electrodes are equal (1, 2 or 3), the number D equals 1. Larger numbers correspond to more disorder in electrode spacing. For JEM, map 3 had one large gap, map 1 had two large gaps and map 4 had three groups of electrodes with spacing of 1, separated by two large gaps. The increasing numbers calculated with the above equation correspond to this increasing amount of disorder in electrode spacing.

Interestingly, disorder in spacing was either unimportant in the speech perception tests, or inverse relationships were found between the amount of disorder and speech perception performance. JEM showed very little correlation between amount of disorder and any of the speech perception tests (correlation coefficients of 0.25, 0.23 and 0.32 for the vowels, consonants and sentences respectively). EWB had significant correlation between disorder and vowel performance (r=0.99), the two being
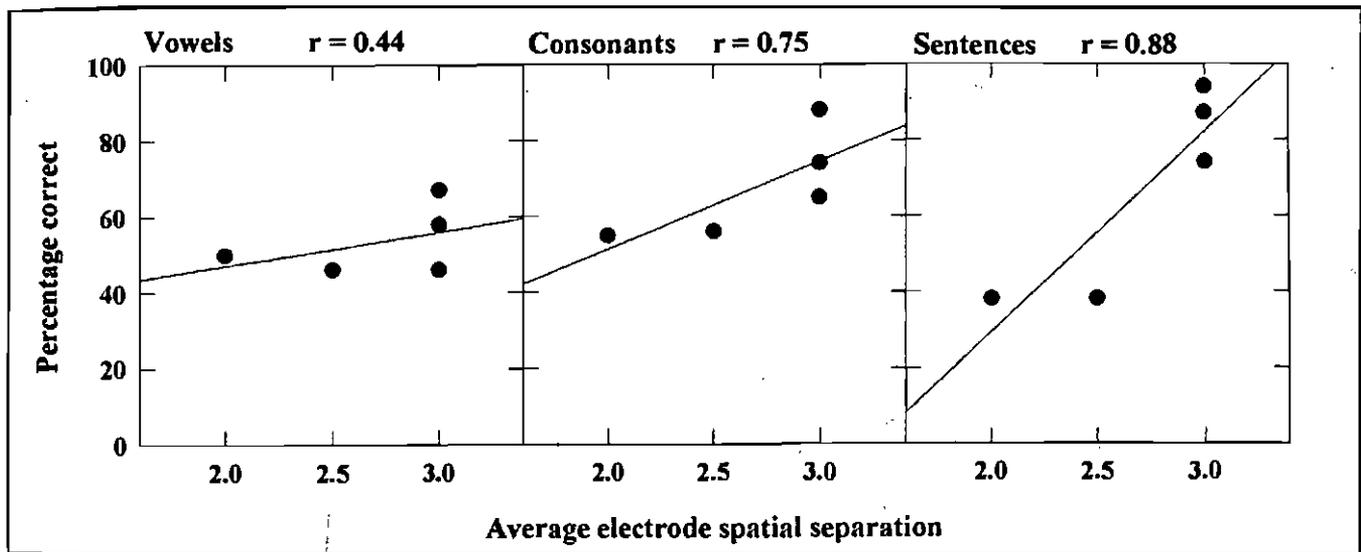

FIGURE 9. Speech perception as a function of average electrode spatial separation for subject REK. Linear regression lines are fitted through the data and correlation coefficients are given above each panel.
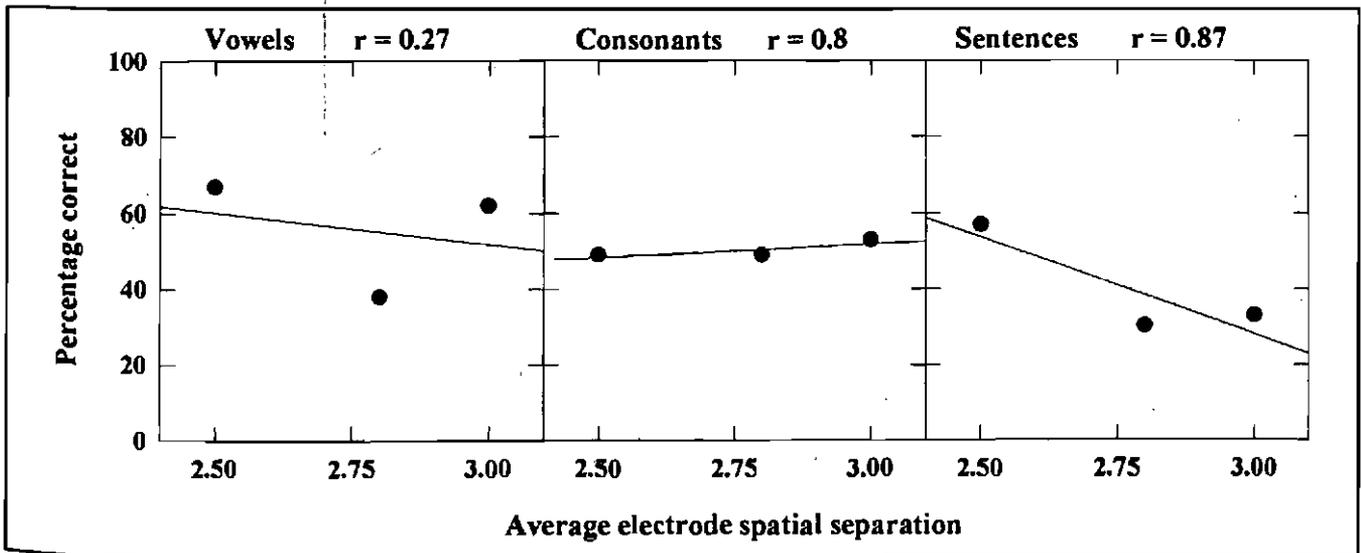

FIGURE 10. Speech perception as a function of average electrode spatial separation for subject EWB. Linear regression lines are fitted through the data and correlation coefficients are given above each panel.

inversely related, and he also had correlation between sentence performance and electrode disorder (r=0.75). REK had no correlation between D and speech perception, except for vowels, where he also demonstrated an inverse relationship (r=0.77 for vowels, r=0.17 for consonants and r=0.33 for sentences).

Speech perception as a function of the number of discrete stimulation sites

As explained earlier, a measure for the number of discrete stimulation sites was the number of electrodes in a specific map with $d' > 1.5$. Based on this criterion, both REK and EWB had no or very little discrimination between different electrodes in the maps used. Figure 11 shows the speech perception performance as a function of the number of electrode sites for JEM. All the measures of speech performance showed significant correlation (r=0.7 or larger) with the number of discrete stimulation sites.

Speech perception as a function of position of electrodes

Most of the maps for subject JEM were spread over the entire range of electrodes, except his map 2 and map 7. Map 2 was in the middle of the electrode array, and map 7 was on the apical end of the electrode array. Vowel and consonant scores improved slightly with the electrodes in the more apical position, but the sentence score improved dramatically from 20% to 66%. No similar data were measured for the other two subjects.

Speech perception as a function of the stimulation mode

Subject JEM was the only subject for whom speech perception tests were conducted as a function of stimulation mode. Maps 1 and 3 in BP+1 mode were repeated in BP+3 mode. The goal was to evaluate the effect of larger current distribution on speech perception. Current spread is

larger in BP+3 mode than in BP+1 mode (Busby *et al.*, 1994), therefore it was assumed that neural selectivity decreased and it was expected that speech perception performance would decrease.

The BP+3 maps utilized the same electrodes as the corresponding BP+1 maps, except that the most basal electrode in each map was omitted. This decreased the number of discrete stimulation sites for map 5 (a BP+3 map) in comparison to map 1 (a BP+1 map), but, contrary to expectation, the number of discrete stimulation sites *increased* for map 6, the BP+3 counterpart for map 3 (a BP+1 map). For both maps significant (10%) decreases in vowel perception scores were observed. This might be explained by the BP+3 maps using only six instead of seven electrodes. An alternative explanation might be the fact that vowels are primarily recognized by their formant structure, and for the BP+3 mode the formant structure was less pronounced. However, there is no significant decrease in performance for the consonant and sentence tests.

DISCUSSION

THE RELATIONSHIP BETWEEN SPEECH PERCEPTION AND PITCH DISCRIMINATION ABILITY OF SUBJECTS

The experiments clearly identified both physical electrode spacing and perceptual electrode distance (as reflected by the cumulative d') as parameters determining speech perception performance. (Note that pitch ranking ability as measured by cumulative d' is dependent on the specific set of electrodes used in a map.) These two parameters are related: the physical electrode spacing influences the amount of current distribution from electrodes. This in turn determines the neural selectivity that can be achieved. The actual number of discrete neural sites that are activated is also a function of other factors, including electrode distance from neurons, neural survival (Zimmermann, Burgess & Nadol, 1995) and the conductive properties of the biological medium. Thus, electrode spacing is
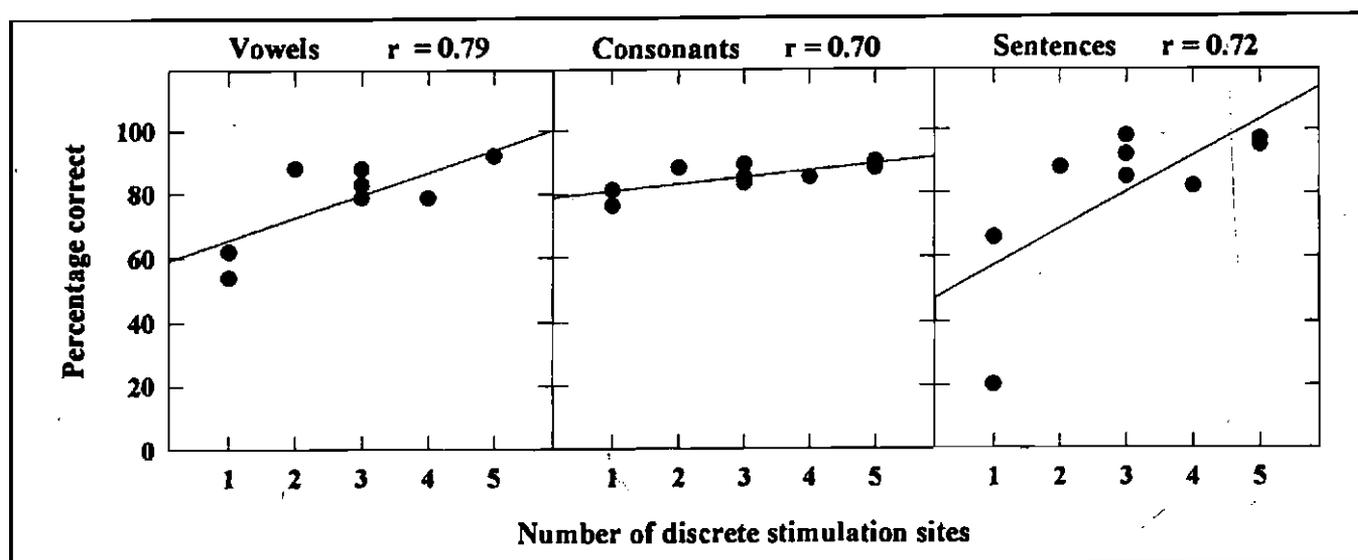

FIGURE 11. Speech perception as a function of the number of discrete stimulation sites for subject JEM. Linear regression lines are fitted through the data and correlation coefficients are given above each panel. The number of discrete sites were calculated for each map by taking a d' value of larger than 1.5 as criterion for two electrodes stimulating discrete neural populations.

a physical parameter determining electrical current spread. On the other hand, perceptual distance is a psychophysical parameter, meaning that this is a derived parameter, dependent on neural selectivity as well as experimental and subject-dependent variables. Thus, although there is correlation between current distribution and neural selectivity, the relationship is subject-dependent. This is demonstrated by comparing the neural selectivity for BP+3 and BP+1 modes for JEM, as reflected by the d′ matrices in figures 1 and 2.

The results also indicate that better *overall* pitch discrimination ability (measured over the set of all twenty electrodes) is related to better speech perception ability. This is demonstrated by comparing overall pitch discrimination ability to speech perception abilities. One measure of overall place pitch discrimination ability of a subject is the maximum cumulative d′ achievable for a specific number of electrodes. In this study, JEM achieved a maximum seven-electrode cumulative d′ of 10.34 in BP+1 mode. This can be compared to a maximum achievable cumulative d′ of just more than 5 for REK and 3.72 for EWB. Relating this to speech perception abilities as reflected by this study, JEM was the best implant user in the group and EWB derived the least benefit from his implant.

The relationship between the amount of benefit that the user derives from the implant and the pitch discrimination ability can also be quantified with two other measures of overall pitch discrimination ability. These are (1) the area of the d′ matrix with large d′ values (say d′>1.5) and (2) the minimum ΔE required for electrode discrimination. If d′>1.5 is used as the criterion for electrode discriminability, it is seen that JEM generally required a ΔE of 2, while REK required a ΔE of 4 or more and EWB required large ΔE values to have d′>1.5.

SPEECH PERCEPTION AND SPECTRAL DISTORTION

Disorder in the choice of electrode spacing did not seem to be an important parameter determining speech perception performance. However, the results obtained by Shannon *et al.* (1995) suggested that spectral distortion might be an important factor determining cochlear implant users' speech perception abilities. An explanation for this apparent anomaly might be that different current distributions from each electrode, overlap in neural excitation and neural survival patterns might already introduce so much spectral distortion that the effect of poor electrode allocation is simply swamped.

SPEECH PERCEPTION AND THE NUMBER OF DISCRETE NEURAL SITES

Although the number of discrete neural excitation sites in the cochlea (therefore, the number of discrete spectral locations) is correlated to speech perception performance, other information in the speech signal may also be utilized to understand electrical speech. This was evident from results achieved by EWB and REK. Although being poorer users than JEM, both demonstrated significant speech perception even with the excitation of so few discrete neural sites. These subjects might rely on temporal information for speech recognition. This paper emphasised the importance of cochlear place information (spectral information), but other studies have demonstrated the im-

portance of temporal information in speech recognition (see, for example, Shannon, Zeng, Kamath, Wygonski & Ekelid (1990) or Dorman *et al.* (1990)).

USING PITCH RANKING DATA TO DETERMINE BETTER SUBJECT-SPECIFIC MAPS

It is possible to make better choices of electrodes to be used in a reduced electrode map by basing the electrode choices on pitch discrimination data. The choice of a smaller number of electrodes leads to improved stimulation rates in the Nucleus processor (Shannon *et al.*, 1990). It has been shown that higher stimulation rates lead to better speech perception performance (Wilson *et al.*, 1991).

Also, the careful choice of electrodes can be used to choose the set of electrodes with the best neural selectivity. The best neural selectivity is achieved by paying attention to two parameters: the physical electrode spacing, which determines current spread, and perceptual distance as measured with the d′s, which is related to neural selectivity. Furthermore, current distribution can be controlled by choice of stimulation mode. Different stimulation modes result in different pitch ranking ability, as demonstrated in subject JEM and also by Busby *et al.* (1994). Stimulation modes with larger current spread do not necessarily lead to decreased neural selectivity, as was demonstrated by map 6 of subject JEM, where more discrete neural channels were achieved than in the BP+1 equivalent of this map. It might be possible to achieve better neural selectivity by using multi-stimulation mode maps.

By varying both electrode spacing and stimulation mode, current distribution patterns might be obtained which give larger perceptual distances between electrodes. This was accomplished in map 8 for subject JEM.

This technique for individualized fitting is more useful when subjects have good pitch discrimination abilities. For subjects in which pitch discrimination ability is not very good in a specific stimulation mode, other stimulation modes might be used. For subjects with poor pitch discrimination ability it might be advantageous to use fewer electrodes in order to increase the stimulation rate. The rationale is that if discrete excitation sites are few, place pitch abilities will be poor regardless of the number of electrodes used, and place pitch resolution may be substituted with better temporal resolution by increasing the stimulation rate.

CONCLUSIONS

(1) The ability to discriminate electrodes based on place pitch, varies considerably among subjects. Also, if it is assumed that perceptual distance between electrodes is related to current spread in the cochlea, the patterns of current spread vary considerably among subjects.

(2) The variability in performance among subjects makes it difficult to compare the effectiveness of different speech processing strategies. Also, there is a need for alternative procedures for better individualized fitting of processors. Our results indicate that pitch ranking ability might be used both to assess implant user potential and to choose better electrode configurations.

(3) Two parameters that can be related to subjects' ability to rank pitch according to place of stimulation influence speech perception performance: electrode spatial

separation and perceptual distance between electrodes.

(4) It is possible to make better choices of electrodes to be used in a reduced electrode map by basing the electrode choices on pitch discrimination data.

(5) It might be possible to achieve better neural selectivity by using multi-stimulation mode maps.

ACKNOWLEDGEMENTS

REFERENCES

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. *Journal of the Acoustical Society of America*, 95, 5, 2658-2669.

Clark, G.M., Tong, Y.C. & Patrick, J.F. (1990). *Cochlear Prostheses*. Melbourne: Churchill Livingstone Incorporated.

Dorman, M.F., Soli, S., Dankowski, K., Smith, L.M., McCandless, G. & Parkin, J. (1990). Acoustic cues for consonant identification by patients who use the Ineraid cochlear implant. *Journal of the Acoustical Society of America*, 88, 5, 2074-2079.

Dubno, J.R. & Dorman, M.F. (1987). Effects of spectral flattening on vowel identification. *Journal of the Acoustical Society of America*, 82, 5, 1503-1511

Eddington, D. (1980). Speech discrimination in deaf subjects with cochlear implants. *Journal of the Acoustical Society of America*,

68, 885-891.

Green, D.M. & Swets, J.A. (1966). *Signal Detection Theory and Psychophysics*. New York: John Wiley and Sons Incorporated.

Hacker, M.J. & Ratcliff, R. (1979). A revised table of d' for M-alternative forced choice. *Perception and Psychophysics*, 26, 2, 168-170.

Lawson, D.T., Wilson, B.S., Zerbi, M. & Finley, C.C. (1996). Speech processors for auditory prostheses. *Third Quarterly Progress Report*, February through April 1996, NIH Contract N01-DC-5-2103.

Nelson, D.A., Van Tasell, D.J., Schroder, A.C. & Soli, S. (1995). Electrode ranking of "place pitch" and speech recognition in electrical hearing. *Journal of the Acoustical Society of America*, 98, 4, 1987-1999.

Shannon, R.V., Zeng, F-G., Kamath, V., Wygonski, J. & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science*, 270, 303-304.

Shannon, R.V., Adams, D.D., Ferrel, R.L., Palumbo, R.L. & Grandgenett, M. (1990). A computer interface for psychophysical and speech research with the Nucleus cochlear implant. *Journal of the Acoustical Society of America*, 87, 2, 905-907.

Skinner, M.W., Clark, G.M., Whitford, L.A., Seligman, P.M., Staller, S.J., Shipp, D.B., Shallop, J.K., Everingham, C., Menapace, C.M., Arndt, P.L., Antogenelli, T., Brimacombe, J.A., Pijl, S., Daniels, P., George, C.R., McDermott, H.J. & Beiter, A.L. (1994). Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. *The American Journal of Otology*, 15, supplement 2, 15-27.

Townshend, B. & White, R.L. (1987). Reduction of electrical interaction in auditory prostheses. *IEEE Transactions on Biomedical Engineering*, 34, 11, 891-897.

Wilson, B.S., Finley, C.C., Lawson, D.T., Wolford, R.D., Eddington, D.K. & Rabinowitz, W.M. (1991). New levels of speech recognition with cochlear implants. *Nature*, 352, 236-238.

Zimmermann, C.E., Burgess, B.J. & Nadol, J.B. (1995). Patterns of degeneration in the human cochlear nerve. *Hearing Research*, 90, 192-201.