J Am Acad Audiol 17:733–746 (2006)

The Effect of Intensity on Pitch in Electric Hearing and Its Relationship to the Speech Perception Performance of Cochlear Implantees

Cila Umat*†
Hugh J. McDermott†
Colette M. McKay†‡

Abstract

This study investigated the effect of intensity on pitch in electric hearing and its relationship to the speech perception ability of cochlear implantees. Subjects were 13 adult users of the Nucleus 22 cochlear implant system, using either the Spectra22 or ESPrit22 speech processor and the SPEAK speech processing strategy. A multidimensional scaling technique was employed. Speech perception was measured using sentences and vowels. All measurements were performed in a soundfield condition, and subjects wore their own speech processors with their normally used settings. Results showed a significant correlation between the degree of deviation of the subjects' stimulus spaces from the "ideal" space and subjects' performance with the sentences, but not with the vowels. A significant correlation was found between subjects' response variability in performing the multidimensional scaling task and their speech perception measures, suggesting that spectral smearing or underlying cognitive abilities might affect implantees' speech perception performance.

Key Words: Cochlear implants, multidimensional scaling, pitch, speech perception

Abbreviations: ALSCAL = Alternating Least-Squares Scaling; CNC = consonant-nucleus-consonant; INDSCAL = Individual Differences Scaling; MDS = multidimensional scaling; NBN = narrow-band noise; RMDS = Repeated MDS; RSQ-I = individual squared correlation value from the INDSCAL analysis; RSQ-R = Overall squared correlation value from the RMDS analysis; SIT = Speech Intelligibility Test; Stress-I = individual Stress value from the INDSCAL analysis; Stress-R = overall Stress value from the RMDS analysis

Sumario Este estudio investiga el efecto de la intensidad sobre el tono en auxiliares auditivos eléctricos y su relación con la capacidad de percepción del lenguaje en individuos con implante coclear. Se incluyeron 13 adultos usuarios del sistema de implante coclear Nucleus 22, utilizando el procesador de lenguaje Spectra22 o el ESPrit22, y la estrategia de procesamiento SPEAK. Se empleó una técnica de ordenamiento en escalas multidimensionales. La

*Department of Audiology and Speech Sciences, Universiti Kebangsaan Malaysia, Kuala Lumpur, Malaysia; †Department of Otolaryngology, The University of Melbourne, Australia; ‡Neuroscience Research Institute, Aston University, Birmingham, United Kingdom

Cila Umat, Department of Audiology and Speech Sciences, Faculty of Allied Health Sciences, Universiti Kebangsaan Malaysia, Jalan Raja Muda Abdul Aziz, 50300 Kuala Lumpur, Malaysia; Phone: 603-2691-4230/603-2698-1420, ext. 216; Fax: 603-2698-6039; E-mail: cilajasp@yahoo.com

Part of this study was presented orally at the XVI National Conference of the Audiological Society of Australia, Brisbane, Queensland, Australia, May 17–20, 2004.

The first author was sponsored by a joint scholarship from the Malaysian government and the Universiti Kebangsaan Malaysia throughout her study in Australia. The second and the third authors had received funding from the Australian National Health and Medical Research Council and the Garnett Passe and Rodney Williams Memorial Foundation.

percepción del lenguaje fue medida usando frases y vocales. Todas las medidas fueron realizadas en condición de campo sonoro, y los sujetos utilizaron sus propios procesadores de lenguaje, en sus ajustes habituales. Los resultados mostraron una correlación significativa entre el grado de desviación de los espacios del estímulo para los sujetos a partir del espacio "ideal", y el desempeño de los sujetos con las frases pero no así con las vocales. Se encontró una correlación significativa entre la variabilidad de respuesta de los sujetos para realizar las tareas de la prueba de escalas multidimensionales y sus medidas de percepción del lenguaje, sugiriendo que la disposición espectral o las capacidades cognitivas subyacentes pueden afectar el desempeño en cuanto a la percepción del lenguaje de los implantados.

Palabras Clave: Implantes cocleares, escalas multidimensionales, tono, percepción del lenguaje

Abreviaturas: ALSCAL = Escalas Alternantes de cuadrados mínimos; CNC = consonante-núcleo-consonante; INDSCAL = Escala de diferencias individuales; MDS = Escalas Multidimensionales; NBN = ruido de banda estrecha; RMDS = MDS repetido; RSQ-I = valor cuadrado de correlación individual a partir del análisis INDSCAL; RSQ-R = valor cuadrado de correlación global a partir del análisis RMDS; SIT = Prueba de Inteligibilidad del Lenguaje; Stress-I = valor de tensión individual a partir del análisis INDSCAL; Stress-R = valor de tensión individual a partir del análisis RMDS

There is a wide range of speech perception performance among cochlear implantees. Many factors contribute to this variability, including device-related factors such as speech processors and speech-processing strategies, as well as individual-related factors such as subjects' ability to discriminate electrodes or process auditory signals. The present study examined one of the individual factors that may affect the speech perception performance of adult cochlear implantees. Specifically, this study investigated the effects of intensity on pitch perception in electric hearing and its relationship to the speech perception performance of individual implantees.

The impetus for the present study came from an earlier work reported by Henry et al (2000). In that study, the ability to discriminate electrodes when level was randomly varied was significantly correlated with the ability to perceive speech information in the low- to mid-frequency regions (170–2680 Hz). One possible explanation for the difficulty in discriminating electrodes when level varies is that the perceived pitch may change with level. Several studies using direct electrical stimulation have noted that the pitch evoked by a single electrode can shift with level in some subjects and that the degree and direction of such pitch shifts differ among individual subjects (Shannon, 1983; Townshend et al, 1987; Pijl, 1997). For example, Townshend et al (1987) found that one subject perceived a higher pitch as the level increased, whereas another perceived a lower pitch as the level increased.

The effect of intensity on pitch perception could be explained by changes in the neural excitation pattern with level. The differing distribution pattern of the surviving nerve cells and the state of the cochlea may explain differences in severity and direction of the changes in pitch with level. For example, the current flow in the cochlea may be altered by the presence of new bone growth. Depending on the state of the cochlea, as the level increases, the current may flow more horizontally, stimulating adjacent neurons, or more vertically, activating the neurons in more apical turns.

The present study used a multidimensional scaling (MDS) technique (Schiffman et al, 1981) to examine the percepts evoked by changes in level and frequency of narrow-band noise (NBN) stimuli. An MDS technique is useful to show graphically the independent perceptual dimensions evoked by a set of stimuli. The output of the MDS analysis is a "stimulus space" in which rank order of the Euclidean distances among the stimuli corresponds with least error to the rank order of perceptual distances. The MDS technique has been used in several studies to examine the dimensionality of the pitch percept associated with activation of different electrode places (Tong et al, 1983; McKay et al, 1996; Collins and Throckmorton, 2000; Henshall and McKay, 2001; McKay and Henshall, 2002). In these studies, it was found that the pitch evoked by single electrodes of varying place is a single-dimension percept, which is best represented as a U-shaped curve in a two-dimensional space. Collins and Throckmorton (2000) showed that the curved nature of the stimulus space is probably due to subjects having difficulty ranking large perceptual distances.

It is well known that loudness and pitch percepts in hearing are perceptually independent. In this study we varied the frequency and level of acoustic stimuli presented via the subjects' own speech processors. Both frequency changes and level changes can produce changes in both loudness and pitch percepts. In this study, the stimuli differing only in frequency were loudness balanced with the intention that frequency changes would produce only pitch changes. Therefore, if level changes do not evoke changes in pitch, it was predicted that the MDS stimulus space would show two orthogonal dimensions corresponding to the frequency and level of the stimuli, and these dimensions would be associated with the percepts of pitch and loudness, respectively. If, on the other hand, level changes evoked both pitch and loudness perceptual changes, the frequency and level aspects of the stimuli would not evoke orthogonal perceptual dimensions. Thus, in this study, the degree of deviation from an "ideal" MDS stimulus space (in which frequency and level produced orthogonal dimensions of pitch and loudness) was used to quantify the amount of influence that level changes have on the pitch percept for each subject.

We examined the relationship between each individual subject's deviation from the ideal stimulus space and their speech perception performance, with the hypothesis that such deviations would particularly affect vowel perception. Two further hypotheses were tested: that speech perception performance would be correlated with the relative importance of the pitch and loudness cues in the MDS task, and that speech perception performance would be correlated with the variability in each subject's responses on the MDS task.

METHOD

Subjects

Subjects were 13 postlingually deafened adult users of the Nucleus 22 cochlear implant system, using either the Spectra22 or ESPrit22 speech processor with the SPEAK speech processing strategy (Skinner et al, 1994). They were selected at random depending on their availability to participate in a research project at the time the study commenced, and were recruited from the Cochlear Implant Clinic of the Royal Victorian Eye and Ear Hospital in East Melbourne, Australia. Subjects' demographic details are given in Table 1.

Table 1. Subjects' Demographic Details

Subject	Age (years)	Years of profound deafness	Year of implantation	Etiology of deafness
S1	48	8	1992	Unknown
S2	68	1	1992	Hereditary
S3	66	16	1988	Otosclerosis
S4	67	16	1992	Hereditary
S5	78	22	1985	Unknown
S7	70	49	1988	Multiple ear infections and surgeries
S8	73	33	1989	Unknown
S9	63	22	1987	Unknown
S10	57	6	1993	Hereditary
S11	59	Right (R) - 44 Left (L) - 17	1988	R - since birth L - head injury
S12	66	1	1986	Meningitis
S13	64	21	1988	Otosclerosis
S14	57	R - 12 L - 7	1993	Otosclerosis

Subjects were tested using their own speech processors with their normally used processor settings throughout the experiment, with the exception of S2 and S14. These subjects changed their processors from ESPrit22 to Spectra22 for the vowel perception test. Table 2 shows a summary of each subject's map.

Stimuli

Stimuli for the MDS experiment were NBNs, generated using the Cool Edit Pro software program, with three different center frequencies (500, 1000, and 2000 Hz) and three different equal-loudness levels (*soft* [S], *medium* [M], and *loud* [L]). In the following, the stimuli are denoted by "NX," where "N" is the center frequency of the NBN, and "X" is the loudness level (for example, 500L is the 500 Hz NBN at the loud level). A different stimulus set was created for each subject to ensure that equal loudness levels across frequencies were achieved. The loudness-balanced levels were determined by performing a loudness category task prior to the MDS experiment.

In the loudness category task, NBN stimuli with four different center frequencies (250, 500, 1000, and 2000 Hz) and six different levels over a 25 dB range were generated and presented to each subject in a soundfield condition. A seven-category loudness scale was employed with descriptors ranging from 1 *(uncomfortably loud)* to 7 *(inaudible)*. Subjects were asked to indicate verbally which category on the scale best described the loudness of each stimulus. Four blocks of trials, with each block consisting of 72 trials (24 stimuli x 3 presentations), were presented to each subject, with the first block serving as a practice list. The resultant loudness growth functions (mean levels in dB SPL versus loudness category) were used to determine the three equal loudness levels across three frequencies for each individual subject.

Calibration was performed in the audio test booth to adjust the level of each of the nine NBN stimuli (i.e., 500L, 500M, 500S, 1kL, 1kM, 1kS, 2kL, 2kM, and 2kS) to that required. The stimuli were then paired with each other, with a 700 msec silent interval separating the stimuli comprising each pair. All possible pairs were generated (9 stimuli x 8), resulting in a total of 72 pairs of stimuli per MDS block of trials.

Procedures

MDS Dissimilarity Judgments

For the MDS dissimilarity judgments, the subject was seated in the test booth at the calibrated position, approximately one meter in front of the loudspeaker. The stimulus pairs were presented in a random order. The subject's task was to judge how different the pairs of sounds were in each presentation. Their responses were recorded on an

Table 2. Summary of Subjects' Maps Used in This Study

Subject	Speech processor	Sensitivity setting	Number of channels	Stimulation mode	Frequency range (Hz)
S1	ESPrit22	3.5	18	BP+1	150–7885
S2	ESPrit22	Fixed – 24 dB	20	BP+1	120–8658
S3	ESPrit22	3.5	16	BP+3	150–5744
S4	Spectra22	2.5	15	Variable BP (+1, +2, +4)	240–7844
S5	ESPrit22	3.5	17	BP+2	150–6730
S7	ESPrit22	3.5	20	CG	120–8658
S8	ESPrit22	Fixed – 21 dB	16	BP+1	150–5744
S9	ESPrit22	2	19	BP+2	120–7390
S10	Spectra22	5	17	BP+3	120–5384
S11	Spectra22	3	16	CG	150–5744
S12	ESPrit22	4	16	Mixed (CG, BP+1, BP+2)	150–5744
S13	Spectra22	2	14	BP+2	171–5603
S14	ESPrit22	3.5	17	BP+1	150–6730

Note: Fixed sensitivity setting on the ESPrit22 means that the clinician disabled the manual control on the processor, which can be programmed either as a sensitivity control or as a volume control. The value set in dB is a relative gain value, in which 21 dB approximates sensitivity setting 8 on the Nucleus 24 SPrint body-worn processor. BP + n (n = 1, 2, 3, 4) stimulation mode is the bipolar electrode configuration with a spacing of n inactive electrodes between the active and indifferent electrodes; CG is common-ground mode.

undifferentiated 10 cm line scale marked with “Exactly the same” at one end and “The most different” at the other end. They were instructed to listen to any differences at all that they were able to perceive and to place a mark on the line scale to indicate the magnitude of the perceived differences between the stimuli.

Prior to commencing the dissimilarity judgments test session, subjects were familiarized with all of the nine stimuli. A training session then followed in which the pairs that were considered likely to be most different were presented first to every subject; for example, 500L-2kS or vice versa, as well as the more similar pairs such as 1kL-1kM or vice versa. They were then guided to mark their responses on the line scales. Subsequently, a practice list was carried out comprising between 20 and 30 trials using the MACarena software program (see acknowledgments). Test sessions were started only when subjects indicated that they were confident to perform the MDS trials. Two subjects (S5 and S8) required additional training to become confident in their ability to do the task. Testing involved six blocks of trials per subject over at least two separate sessions.

Subjects’ responses on the line scales were transformed into numbers between 0 and 100, and entered into six data matrices (i.e., with one data matrix per test block). The data were then submitted to the ALSCAL (Alternating Least-Squares Scaling) software program (Young and Lewyckyj, 1979) for analyses using the Individual Differences Scaling (INDSCAL) and Repeated MDS (RMDS) procedures.

The output stimulus space from INDSCAL analysis is the best representation of perceptual dimensions that are common to all subjects (i.e., the “ideal” stimulus space). For this analysis, the input comprised of 13 data matrices representing the results from the 13 subjects in the study (each matrix containing the sum of the six test-block matrices for each subject). In this experiment, if level did not affect pitch systematically across the 13 subjects, we expected that the INDSCAL stimulus space would depict the frequency and level parameters of the stimuli as associated with orthogonal perceptual dimensions (i.e., pitch and loudness, as discussed earlier in the introduction). INDSCAL also allows analysis of individual subject differences in assigning the common dimensions to the stimuli. Results of the analysis known as “weights” provide information about the relative importance of each of the common dimensions for each subject (i.e., the relative importance of the pitch and loudness cues to individual subjects), and the individual data can be assessed to see how closely they fit the overall “ideal” space.

Individual subject data were examined using RMDS analysis. The six data matrices for the subject served as the ALSCAL input data. The resulting stimulus space from the RMDS analysis represented the perceptual distances between the stimuli for that subject.

Speech Perception Measures

The *Speech Intelligibility Test* (SIT) sentences (Magner, 1972) were administered in quiet at an average level of approximately 60 dBA. The SIT test consists of 40 lists of digitally recorded sentences on a compact disc (CD), spoken by a native Australian male speaker. Each list comprises 15 sentences with each sentence containing three to eight key words. In total, there are 80 key words per list. Two lists were selected at random for the test, and the same lists were used for all subjects. Subjects were asked to repeat the sentences, and the mean percent correct key words identified from the two lists (mean SIT scores) were calculated for each subject.

The Australian version of the consonant-nucleus-consonant (CNC) monosyllabic words test (Henry et al, 1998) was used to measure vowel perception. Three lists of CNC words, each consisting of 50 meaningful words spoken by a female talker having an average Australian accent, were randomly selected, and the same lists were used for all subjects. The average peak speech level for presentation was 61 dBA. The vowel scores from the three lists were averaged to obtain the mean CNC vowels score.

In both speech tests, the stimuli were presented via a loudspeaker, routed from a MADSEN Itera Diagnostic Audiometer (used to set the presentation levels of the speech signals), which was connected to the CD player. Subjects were seated at the calibrated position in the test booth for the speech measures, using the same speech-processor settings as they used in the MDS experiment.

RESULTS

MDS

INDSCAL Analysis

The first step in MDS analysis is to establish the appropriate number of dimensions in which to represent the stimulus space solution. One factor used in this assessment is the way that Stress and R^2 (RSQ) values change as a function of the number of dimensions used. Stress is an index that assesses the mismatch of data and corresponding distances in the stimulus space (i.e., the lower the Stress values, the better). RSQ shows the proportion of variance in the data accounted for by the MDS model (i.e., the higher the RSQ values, the better). Analysis of the changes in Stress and RSQ values revealed three as the maximum number of dimensions appropriate for analysis of the INDSCAL group configuration. Generally, the number of dimensions where a knee-point occurs (i.e., where increasing the number of dimensions further does not greatly improve the Stress and RSQ values) indicates the most appropriate number of dimensions for analysis.

The derived three-dimensional INDSCAL group stimulus space is shown in Figure 1. The configuration shows that the plane consisting of dimensions 1 and 3 (left panel) separates the stimuli into three groups differing in frequency. In contrast, the dimension orthogonal to this plane (dimension 2; the vertical direction in the right panel) separates the stimuli systematically according to level. Note that the dimension 1/3 plane shows the frequencies arranged in the typical U-shaped pattern. As explained in the introduction, the curved nature of the relationship between the stimuli does not mean that there are in fact two pitch dimensions, but is just an artifact produced by using stimuli that are very different in frequency. Since the group space obtained by INDSCAL shows the level and frequency parameters in the stimulus to produce orthogonal perceptual dimensions, we can be fairly confident that the perceptual dimensions are pitch and loudness. Thus, the dimension 1/3 plane in Figure 1 will be referred to as the "pitch plane," and dimension 2 as the "loudness dimension." The group space will be referred to as the "ideal" space in the further analysis, since it represents well the ideal situation in which level does not cause systematic significant pitch changes. The value describing how well each subject's data conforms to this ideal space can then be thought of as a measure of the interaction of pitch with level for individual subjects.

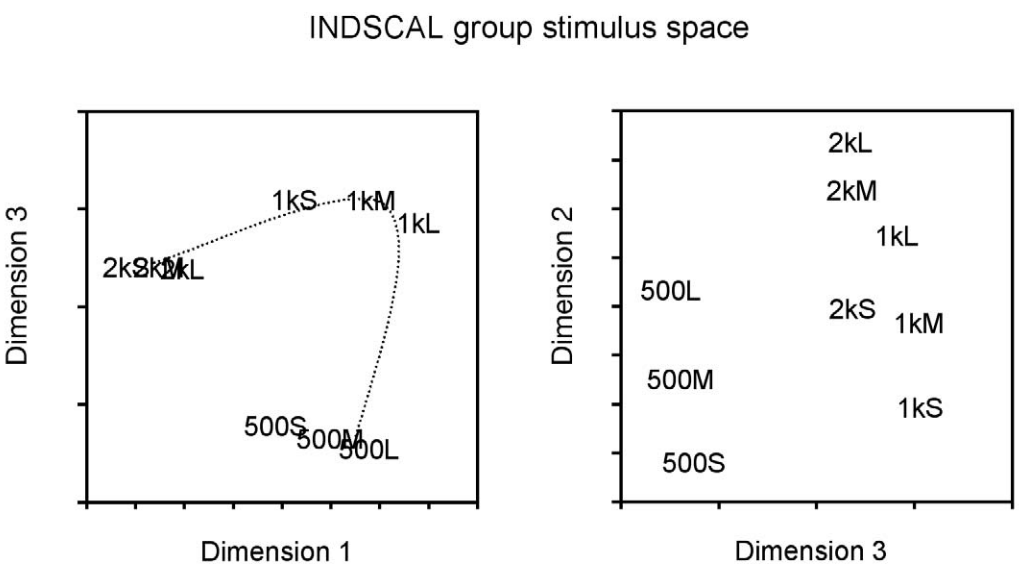

Figure 1. The three-dimensional INDSCAL group stimulus space. The U-shaped pattern observed in the dimension 1/3 plane (left panel) separates the stimuli of different frequencies, whereas dimension 2 (the vertical direction in the right panel) separates the stimuli on the basis of level (soft [S], medium [M], and loud [L]).

How closely the individual subjects' data fitted this group configuration was indicated by the individual RSQ values (hereafter denoted as "RSQ-I") and Stress values (hereafter denoted as "Stress-I"). The RSQ-I value shows how well an individual subject's data fit this ideal stimulus configuration, whereas the Stress-I value reflects the error in fitting the INDSCAL configuration. The RSQ-I and Stress-I values for each subject are shown in the second and third columns of Table 3. From this table, it can be seen that data from S14 fitted the INDSCAL configuration best (i.e., had the highest RSQ-I among the subjects), while S5's data were the poorest in fitting the ideal MDS stimulus space (i.e., had the lowest RSQ-I).

The subjects' "weight space" consists of vectors: one vector for each subject. The length of a weight vector reflects the amount of variation in the subject's data that is accounted for by the group stimulus space, and the angle of the vector indicates the relative importance of the dimensions for that subject when the dissimilarity judgments were performed. The length of the vector for subject k (V_k) is the square root of the sum of squares of the weights (ϖ_{ka}) attached by subject k to each dimension a:

$$V_k = \sqrt{\sum_{a=1}^{3} \omega_{ka}^{2}} = \sqrt{(\text{RSQ-I})} \quad \ldots\ldots\ldots\ldots(1)$$

The angle, θ_k, is the angle of the weight vector relative to the pitch plane and, hence, is a measure of the relative weight the subject placed on loudness cues relative to pitch cues. This angle is calculated according to the following formula:

$$\text{Sin } \theta_k = \frac{\omega_{kD2}}{V_k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

Thus,

$$\theta_k = \text{Arcsine } \left(\frac{\omega_{kD2}}{V_k} \right) \quad \ldots\ldots\ldots\ldots\ldots\ldots(3)$$

An angle of 45° would indicate that the subject weighted the pitch and loudness dimensions equally when differentiating the stimuli in the dissimilarity judgments. The calculated θ_k for each subject is shown in Table 3 (column 4). S11 had the smallest angle, suggesting that this subject put more weight on the pitch cue than the loudness cue when differentiating the stimuli in the MDS experiment.

RMDS Analysis

The RMDS analysis was performed to obtain individual subjects' stimulus spaces. Examination of the Stress and RSQ values for individual subjects' results suggested that three was the maximum number of dimensions appropriate for analysis of the stimulus spaces. For ease of viewing and comparing the data across subjects, the three-dimensional stimulus spaces were orientated in such a manner that the 1000 Hz stimuli (1kL, 1kM, and 1kS) overlapped each other as much as possible. In general, orientation of the data points in such a way resulted in dimensions 1 and 2 demonstrating the typical U-shaped pattern while stimuli in the third

Table 3. The Results of INDSCAL and RMDS Analysis for Each Subject

Subjects	RSQ-I (INDSCAL)	Stress-I (INDSCAL)	Angles θ_k (°)	RSQ-R (RMDS)	Stress-R (RMDS)
S1	0.882	0.106	27.8	0.946	0.084
S2	0.972	0.051	29.8	0.930	0.088
S3	0.937	0.075	46.9	0.830	0.111
S4	0.839	0.120	35.7	0.912	0.105
S5	0.835	0.126	36.7	0.921	0.109
S7	0.915	0.099	21.4	0.965	0.074
S8	0.873	0.107	43.9	0.761	0.141
S9	0.936	0.077	28.5	0.915	0.102
S10	0.959	0.061	30.9	0.941	0.085
S11	0.943	0.077	17.9	0.973	0.058
S12	0.958	0.061	47.9	0.877	0.104
S13	0.935	0.078	26.8	0.839	0.142
S14	0.978	0.056	26.8	0.951	0.085

Note: Shown are the individual subjects' data fit to the three-dimensional INDSCAL stimulus space (RSQ-I and Stress-I), the angles (θ_k) that indicate the relative importance of the pitch and loudness cues for individual subjects, and the RSQ-R and Stress-R values for individual subjects' three-dimensional RMDS stimulus spaces.

dimension show systematic level changes. Examples of RMDS stimulus spaces for three subjects—S2, S5, and S13—are depicted in Figure 2. The results for these three subjects were chosen to illustrate representative examples of the main configuration of stimulus spaces observed in this study. S2's stimulus space is an example showing clearly independent dimensions of pitch (dimensions 1/2) and loudness (dimension 3). S5's results showed a strong interaction of frequency and level parameters within the stimulus space, as evidenced by the overlapping data points for 500L, 500M, and 1kS (see the top panel of S5's graph), irrespective of the orientation of the stimulus space. The perceptual dimensions of S5's stimulus space are not clearly interpretable as having a pitch plane and a separate loudness dimension. For example, in the dimension 1/2 plane, the 500 Hz stimuli of differing level are not grouped together, as is the case for the other frequencies. Two other subjects (S3 and S12) also showed similar findings to that of S5, that is, significant interaction of frequency and level parameters in their stimulus spaces. Some other subjects (S1, S4, S7, and S13) did not differentiate the different intensity levels in either the 500 Hz or the 2000 Hz stimuli (i.e., the data points for 500L, 500M, and 500S, or 2kL, 2kM, and 2kS overlapped in the derived stimulus spaces for these subjects). An example is shown in Figure 2 for S13's stimulus space in which the 500L, 500M, and 500S stimuli were closely grouped, irrespective of how the space was oriented to view the data. These observations are interesting given that all subjects were able to distinguish these stimuli in the loudness category task.

The overall RSQ (RSQ-R) and Stress (Stress-R) values for the RMDS analysis are shown in Table 3 (columns 5 and 6). These indices, which represent how well each subject's data fit their own best-fitting stimulus space, partly reflect the degree of uncertainty in subjects' responses when performing the MDS dissimilarity judgments.

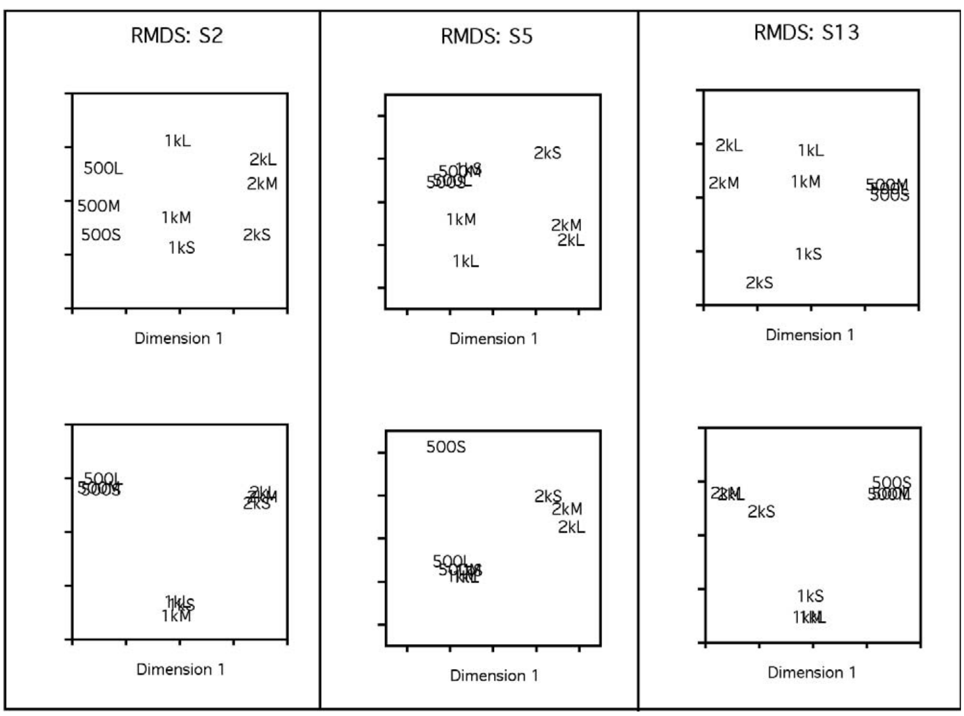

Figure 2. Examples of three-dimensional RMDS stimulus spaces for three subjects (S2, S5, and S13) that represent the main configurations observed among the subjects' results in this study. S2's graph is an example of a clearly interpretable stimulus space in which the dimension 1/2 plane could be interpreted as the pitch plane and dimension 3 could be interpreted as loudness. S5's graph is an example in which the perceptual dimensions are not as clearly interpretable as having a pitch plane and a separate loudness dimension (see text). S13 is an example of subjects who appeared not to differentiate the different intensity levels in the 500 Hz (or 2000 Hz) stimuli. The 500L, 500M, and 500S points in this graph were closely grouped in the loudness dimension (dimension 3), irrespective of how the space was oriented to view the data.

A high degree of uncertainty (leading to variability and inconsistency in the responses) will contribute to a low RSQ-R or high Stress-R value. From the table, it can be observed that S8's data was the most variable; only 76.1% of the variance was accounted for by S8's stimulus space. S11 provided the most consistent responses.

Speech Perception Performance

Individual subjects' speech perception performance on sentences and vowels are shown in Figure 3. There was a considerable variation in the implantees' speech perception performance. The vowel scores for two of the subjects, S2 and S14, were obtained with different speech processors than those used when performing the MDS task. Thus, the CNC vowel scores for these two subjects were excluded in the subsequent correlation and regression analyses with the MDS results.

Correlating MDS and Speech Perception Results

To test the hypothesis that individual subjects' perception of changes in frequency and intensity are correlated with their speech perception performance, the goodness of fit of each subject's data to the ideal stimulus space in Figure 1 (as given by the RSQ-I and Stress-I) was correlated with their speech perception scores. These indices reflect how far individual subjects' data deviated from the INDSCAL configuration, which showed independent pitch and loudness dimensions.

Both the RSQ-I and Stress-I were significantly correlated with the mean SIT scores. The correlation between the RSQ-I and mean SIT scores was $r = 0.656$ ($F[1,11] = 8.30$, $p = 0.015$). The Pearson correlation coefficient for the association between the Stress-I values and the mean SIT scores was $r = -0.610$ ($F[1,11] = 6.51$, $p = 0.027$). Significant results were not found, however, when either the RSQ-I or Stress-I values were correlated with the mean CNC vowels scores. The correlation between the RSQ-I and mean CNC vowels scores was $r = 0.217$ ($p = 0.522$), and the correlation between the Stress-I and mean CNC vowels was $r = -0.122$ ($p = 0.722$). Thus, our main hypothesis that the variation seen in pitch when intensity changes significantly correlates with speech perception performance was not fully supported, as a significant correlation was found only with the sentences and not with the vowels. It was expected that vowel perception would be more sensitive than sentence perception to the effects of pitch changes with level. The regression plots are shown in Figure 4.

We also examined the correlation between the indices from the RMDS analysis (RSQ-R and Stress-R) and the speech perception scores. As discussed earlier, these measures partly reflect the uncertainty of each subject in giving perceptual distance estimates. We found a significant correlation between the Stress-R and the mean SIT scores: $r = -0.636$ ($p = 0.019$) but not between the RSQ-R and the mean SIT ($r = 0.449$, $p = 0.123$). When correlated with the mean CNC vowels scores, a stronger correlation was observed with Stress-R ($r = -0.820$, $p = 0.002$), while the correlation between the RSQ-R and mean CNC vowels scores was $r = 0.618$ ($p = 0.043$). The regression plots are shown in Figure 5.

To test the hypothesis that the relative importance of the pitch and loudness cues to individual subjects in the MDS task significantly correlated with their speech perception performance, the angles of the weight vectors relative to the pitch plane were correlated with the speech perception scores. Speech is a complex signal having acoustic components that vary in frequency and intensity over time. It was hypothesized that implantees possibly used specific strategies to differentiate the phonemes, which could be inferred from the relative importance of the pitch and loudness cues in results of the MDS experiment. No significant correlations between the angles and either of the speech perception measures were observed.

DISCUSSION

The present study investigated the effect of intensity on pitch perception in electric hearing and its relationship to the speech perception of adult cochlear implantees. The study utilized the MDS technique and NBN stimuli that varied in frequency and intensity to vary the pitch and loudness percepts.

The MDS results revealed that three-

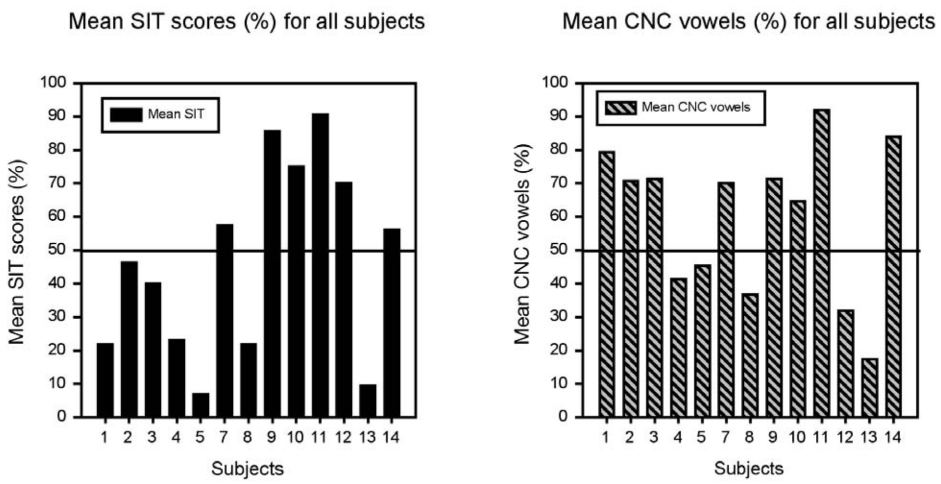

Figure 3. The Mean SIT and Mean CNC vowels scores (percent correct) across subjects. It should be noted that the vowel scores for S2 and S14 were obtained with different speech processors than those used when performing the MDS task and the SIT sentences test.

dimensional stimulus spaces were appropriate to represent the data from all subjects. In general, the frequencies could be separated in a two-dimensional plane (showing a U-shaped pattern), and the levels could be separated in an orthogonal third dimension. Thus, the perceptual dimensions could be interpreted as pitch and loudness. The present result in which a U-shaped pattern was observed in the pitch plane is consistent with findings from earlier studies related to electrode-place pitch using the MDS technique (Tong et al, 1983; McKay et al, 1996; Collins and Throckmorton, 2000; Henshall and McKay, 2001; McKay and Henshall, 2002).

The present results confirm that pitch and loudness are two independent percepts in electric hearing as evidenced by the orthogonal dimensions in the INDSCAL stimulus configuration, which could be interpreted as pitch and loudness dimensions. It was hypothesized that individual implantees differed in the effect of intensity changes on pitch. Visual inspection of the individual stimulus spaces obtained from the RMDS analysis supported this hypothesis. For example, although S2 showed no interaction between pitch and level, S5 showed an unusual stimulus configuration that could not be interpreted as having distinct pitch and loudness dimensions (see Figure 2). This individual variability was also consistent with the RSQ-I and Stress-I values from the INDSCAL analysis. As discussed above, these measures reflect how well the individual's data fitted the "ideal" group space. It can be seen that S2 had the best fit (lowest Stress-I) and S5 the worst fit (highest Stress-I). The finding that pitch changed with level in some implantees was consistent with the reports of previous studies using direct electrical stimulation on single electrodes (Shannon, 1983; Townshend et al, 1987; Pijl, 1997). The individual variation in the direction of the pitch shifts with level was also consistent with the findings of Townshend et al (1987).

Our main hypothesis, however, was not fully supported by the experimental results. We hypothesized that the variation of pitch with intensity changes would correlate significantly with the speech perception performance of the implantees, in particular with vowel perception. We did find significant correlations between the RSQ-I (and Stress-I) and the mean SIT scores ($p = 0.015$ and $p = 0.027$, respectively), but not with the mean CNC vowel scores. It would be expected that, since vowel perception depends on accurate perception of formant frequencies, variations in pitch with level might affect vowel recognition more than sentence perception. In explaining the significant

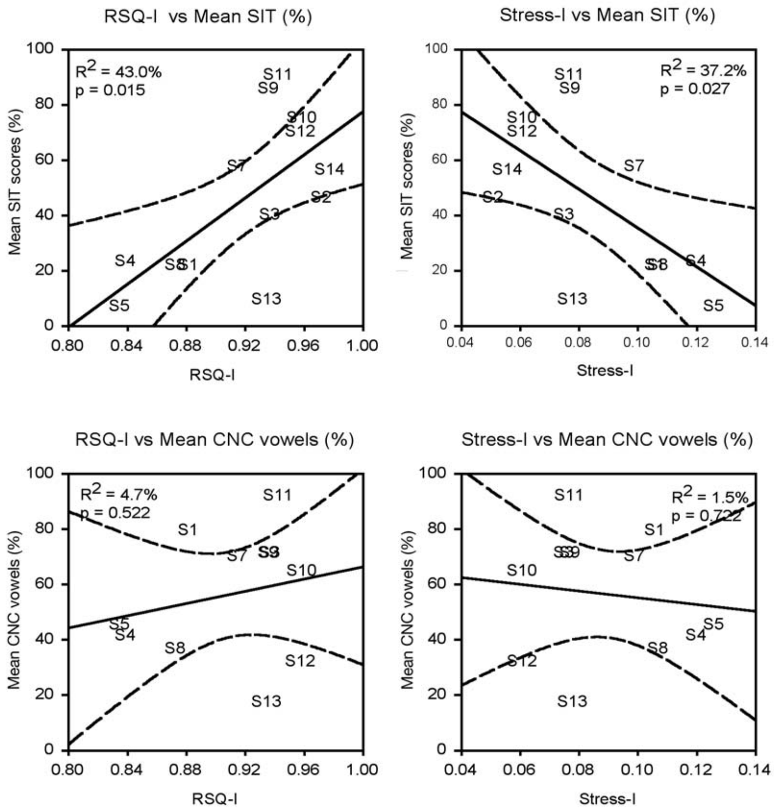

Figure 4. Regression plots for RSQ-I versus Mean SIT scores (top left) and Stress-I versus Mean SIT scores (top right). Regression plots for RSQ-I versus Mean CNC vowels scores (bottom left) and Stress-I versus Mean CNC vowels scores (bottom right). For all graphs, the solid lines are the regression lines while the dashed lines represent 95% confidence intervals.

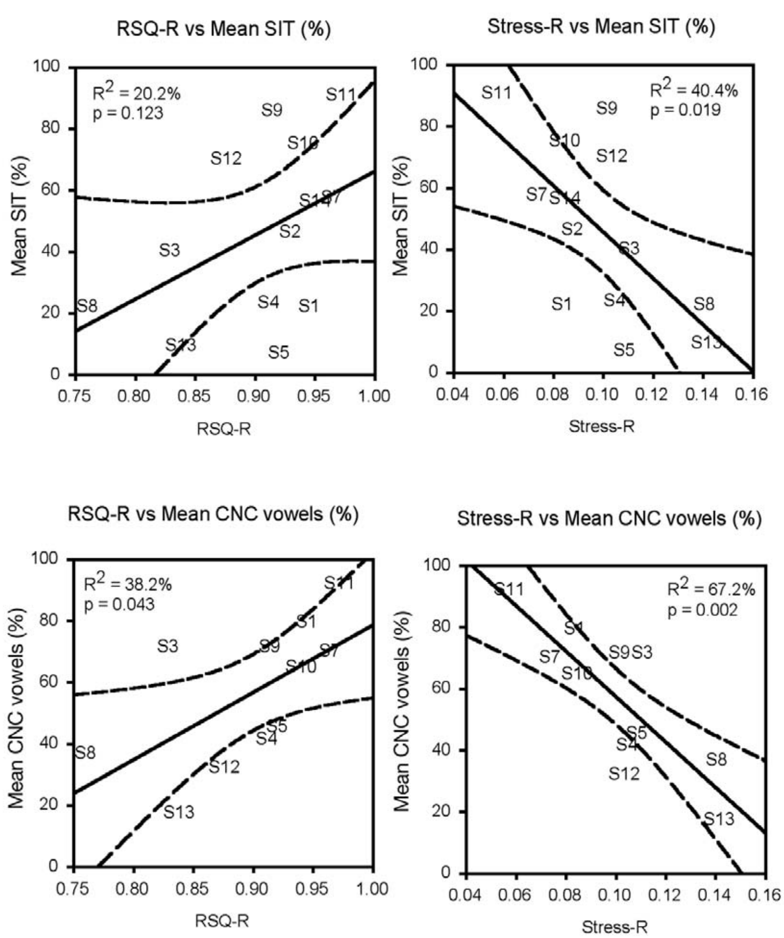

Figure 5. Regression plots for RSQ-R versus Mean SIT scores (top left) and Stress-R versus Mean SIT scores (top right). Regression plots for RSQ-R versus Mean CNC vowels scores (bottom left) and Stress-R versus Mean CNC vowels scores (bottom right). For all graphs, the solid lines are the regression lines while the dashed lines represent 95% confidence intervals.

relationship between the RSQ-I and the mean SIT scores, it is possible that other cues in sentences such as suprasegmental information or consonant perception were affected by the pitch and intensity interaction, thus contributing to the significant result. It might also be possible that the natural intensity variation within the sentences (in contrast to the CNC words, which were presented at a constant intensity) led to more performance variability being due to pitch/level interactions. Finally, it should be pointed out that the correlation with vowel scores used results from only 11 subjects (in contrast to 13 for the sentences).

We found significant correlations between the speech perception measures and the indices from the RMDS analysis (Stress-R and RSQ-R), which in part quantify the degree of uncertainty or difficulty that subjects had when performing the MDS dissimilarity judgments. That is, the higher the variability or inconsistency in subject responses, the lower the RSQ-R (or higher Stress-R). It was found that subjects with more variability in their responses had poorer speech perception performance, as measured using both sentences and vowels (see Figure 5). There are two broad reasons why this correlation may occur: the subjects with more inconsistent responses may suffer from an inability to perceive differences among electrode positions (due perhaps to such factors as current spread in the cochlea), or these subjects may suffer from a more central or cognitive disadvantage that would affect both speech and MDS results.

It is well known that implantees vary greatly in their ability to discriminate electrode positions (Nelson et al, 1995; Hanekom and Shannon, 1996; Henry et al, 1997; Throckmorton and Collins, 1999; Donaldson and Nelson, 2000; Henry et al, 2000). Also, it has been shown that frequency resolution ability is highly correlated with speech understanding (Hawks et al, 1997; Henry and Turner, 2003; Laback et al, 2004). The experiment presented here was not designed to measure electrode discrimination or frequency resolution, and the NBNs used were separated by relatively large frequency differences. Nevertheless, subjects would have differed in the absolute perceptual distances they experienced between the different stimuli. It is plausible that subjects for whom the stimuli all sounded very similar would have responded in a more variable way in the MDS task and would have experienced more difficulty distinguishing different speech sounds as well.

Subjects performing an MDS dissimilarity judgment need to be attentive to the first stimulus and hold information about it in their memories for the subsequent comparison with the second stimulus. Two separate processes may be involved in performing this task: one is to perceive the differences between the stimulus pairs as discussed above, and the other is to make a judgment about how different the stimuli are in each pair (Nosofsky, 1985). In Nosofsky's study, semicircle stimuli varying in size and a radial line with varying angle of orientation were used, and subjects were asked to identify the stimuli. The results were computed in an identification confusion matrix. Nosofsky fitted his data to a version of the MDS model and found that the results supported the notion that there is a distinction between "perceptual" independence and "decisional" independence. Both factors may contribute to variable responses, and hence lower RSQ-R (or higher Stress-R) in the RMDS analysis. For example, if a subject perceived most of the stimuli as the same, this might contribute to his or her uncertainty in judging how different the stimuli were. Alternatively, a subject may perceive the differences but have difficulty in assigning a relative magnitude to them.

According to the information processing (IP) model proposed by researchers in the area of cognitive psychology (see, for example, Pisoni, 2000; Pisoni and Cleary, 2004), perception involves attention and memory. Poor attention and limitation in the working memory of an individual subject might affect what the subject perceives. The working memory is a cognitive system that holds the incoming information for a brief period (short-term storage) while at the same time applying a cross-check (recall and retrieve) between the incoming information and that held in long-term storage. Nosofsky (1991), for example, generated 34 schematic faces varying in eye height, eye separation, nose length, and mouth height, and performed an MDS study with 138 undergraduates from Indiana University. Subjects rated the similarity of the faces on a scale from most dissimilar to most similar. The data were consistent with the theory that recognition decisions are based on similarity comparisons with stored information in the memory. If the information were not stored adequately in memory, subjects would not be able to take full advantage of comparisons with stored information when asked to make recognition judgments.

These various possible sources of uncertainty may affect both the MDS results and speech perception, and might thereby contribute to the significant correlation found in the present study between variability in responses in the dissimilarity judgments and speech understanding. Some implantees may perceive different speech phonemes as very similar, or have difficulty in deciding the differences between phonemes due to limitations in their information-processing capability, hence contributing to their difficulty in perceiving the speech sounds. The results imply that differences in frequency selectivity of the electrical stimulation or differences in underlying cognitive abilities might underlie differences in speech perception.

The role of cognitive function in the speech perception of adult cochlear implantees has not been extensively studied. In one study, Lyxell et al (2003) found that a certain level of working memory capacity was needed to allow implant users to follow an oral conversation with a relatively low level of effort. However, the findings from several other researchers do not support the proposition that the speech perception of adult cochlear implantees may be affected by underlying cognitive abilities (Parkin et al, 1989; Gantz et al, 1993; Collison et al, 2004). Collison et al (2004), for example, conducted the *Test of Nonverbal Intelligence-3* (TONI-3) and the *Woodcock-Johnson III Tests of Cognitive Abilities* (Verbal Comprehension Section [WJ-III VCS]) with implantees. In these tests, participants identified relationships among abstract figures and solved problems following manipulation of these figures. The word recognition test involved the recognition of consonant-vowel-consonant (CVC) words. No significant relationships were found between the cognitive and speech perception measures. Research with children, however, has revealed that nonverbal intelligence is one of the significant predictive factors of good speech perception performance postimplantation (Geers et al, 2003; Moog and Geers, 2003; Willstedt-Svensson et al, 2004). Further investigation is needed to examine the role of cognitive function in speech perception performance of adult implantees.

CONCLUSIONS

In conclusion, we have confirmed that pitch and loudness are two independent percepts in electric hearing and can best be observed using a three-dimensional MDS stimulus space. It was found that, in some subjects, pitch changed with level of stimulation. How far subjects' perception deviated from the ideal stimulus configuration (in which level changes affected only the stimulus dimension orthogonal to pitch) was found to be correlated significantly with their speech perception as measured using sentences, but not with vowels. The relative importance of pitch and loudness cues in performing the MDS task did not correlate with the speech perception measures. The indices that reflect the degree of uncertainty or inconsistency in subjects' responses when performing the MDS dissimilarity judgments were significantly correlated with both the speech perception measures, suggesting that difficulty in discriminating the stimuli or poor underlying cognitive abilities might contribute to subjects' poor speech perception performance.

Acknowledgments. The study was undertaken at the Department of Otolaryngology, University of Melbourne, and the Bionic Ear Institute, Melbourne, Australia. The authors would like to thank all the research subjects for their time and enthusiastic participation in the study. Dr. Wai Kong Lai from the Department of Otorhinolaryngology Head and Neck Surgery, University Hospital, in Zürich, Switzerland, is acknowledged for providing the MACarena software program used in the experiment.

REFERENCES

Collins LM, Throckmorton CS. (2000) Investigating perceptual features of electrode stimulation via a multidimensional scaling paradigm. *J Acoust Soc Am* 108:2353–2365.

Collison EA, Munson B, Carney AE. (2004) Relations among linguistic and cognitive skills and spoken word recognition in adults with cochlear implants. *J Speech Lang Hear Res* 47:496–508.

Donaldson GS, Nelson DA. (2000) Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and

SPEAK speech processing strategies. *J Acoust Soc Am* 107:1645–1658.

Gantz BJ, Woodworth GG, Abbas PJ, Knutson JF, Tyler RS. (1993) Multivariate predictors of audiological success with multichannel cochlear implants. *Ann ORL* 102:909–916.

Geers A, Brenner C, Davidson L. (2003) Factors associated with development of speech perception skills in children implanted by age five. *Ear Hear* 24:24S–35S.

Hanekom JJ, Shannon RV. (1996) Place pitch discrimination and speech recognition in cochlear implant users. *S Afr J Commun Disord* 43:27–40.

Hawks JW, Fourakis MS, Skinner MW, Holden TA, Holden LK. (1997) Effects of formant bandwidth on the identification of synthetic vowels by cochlear implant recipients. *Ear Hear* 18:479–487.

Henry BA, McDermott HJ, McKay CM, James CJ, Clark GM. (1998) A frequency importance function for a new monosyllabic word test. *Aust J Audiol* 20:79–86.

Henry BA, McKay CM, McDermott HJ, Clark GM. (1997) Speech cues for cochlear implantees: spectral discrimination. In: Clark GM, ed. *XVI World Congress of Otorhinolaryngology Head and Neck Surgery: Cochlear Implants.* Sydney: Monduzzi Editore - International Proceedings Division, 89–93.

Henry BA, McKay CM, McDermott HJ, Clark GM. (2000) The relationship between speech perception and electrode discrimination in cochlear implantees. *J Acoust Soc Am* 108:1269–1280.

Henry BA, Turner CW. (2003) The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *J Acoust Soc Am* 113:2861–2873.

Henshall KR, McKay CM. (2001) Optimizing electrode and filter selection in cochlear implant speech processor maps. *J Am Acad Audiol* 12:478–489.

Laback B, Deutsch WA, Baumgartner WD. (2004) Coding of vowellike signals in cochlear implant listeners. *J Acoust Soc Am* 116:1208–1223.

Lyxell B, Andersson U, Borg E, Ohlsson I-S. (2003) Working-memory capacity and phonological processing in deafened adults and individuals with a severe hearing impairment. *Int J Audiol* 42:S86–S89.

Magner ME. (1972) *A Speech Intelligibility Test for Deaf Children.* Northampton, MA: Clarke School for the Deaf.

McKay CM, Henshall KR. (2002) Frequency-to-electrode allocation and speech perception with cochlear implants. *J Acoust Soc Am* 111:1036–1044.

McKay CM, McDermott HJ, Clark GM. (1996) The perceptual dimensions of single-electrode and nonsimultaneous dual-electrode stimuli in cochlear implantees. *J Acoust Soc Am* 99:1079–1090.

Moog JS, Geers AE. (2003) Epilogue: major findings, conclusions and implications for deaf education. *Ear Hear* 24:121S–125S.

Nelson DA, Tasell DJV, Schroder AC, Soli S, Levine S. (1995) Electrode ranking of "place pitch" and speech recognition in electrical hearing. *J Acoust Soc Am* 98:1987–1999.

Nosofsky RM. (1985) Overall similarity and the identification of separable-dimension stimuli: a choice model analysis. *Percept Psychophysics* 38:415–432.

Nosofsky RM. (1991) Tests of an exemplar model for relating perceptual classification and recognition memory. *J Exp Psych Hum Percept Perform* 17:3–27.

Parkin JL, Stewart BE, Dankowski K, Haas LJ. (1989) Prognosticating speech performance in multichannel cochlear implant patients. *Otolaryngol Head Neck Surg* 101:314–319.

Pijl S. (1997) Pulse rate matching by cochlear implant patients: effects of loudness randomization and electrode position. *Ear Hear* 18:316–325.

Pisoni DB. (2000) Cognitive factors and cochlear implants: some thoughts on perception, learning, and memory in speech perception. *Ear Hear* 21:70–78.

Pisoni DB, Cleary M. (2004) Learning, memory, and cognitive processes in deaf children following cochlear implantation. In: Zeng FG, Popper AN, Fay RR, eds. Vol. 20 of *Cochlear Implants: Auditory Prostheses and Electric Hearing.* New York: Springer-Verlag, 377–426.

Schiffman SS, Reynolds ML, Young FW. (1981) *Introduction to Multidimensional Scaling—Theory, Methods and Applications.* New York: Academic Press.

Shannon RV. (1983) Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear Res* 11:157–189.

Skinner MW, Clark GM, Whitford LA, Seligman PM, Staller SJ, Shipp DB, Shallop JK, Everingham C, Menapace CM, Arndt PL, Antogenelli T, Brimacombe JA, Pijl S, Daniels P, George CR, McDermott HJ, Beiter AL. (1994) Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. *Am J Otol* 15:15–27.

Throckmorton CS, Collins LM. (1999) Investigation of the effects of temporal and spatial interactions on speech recognition skills in cochlear-implant subjects. *J Acoust Soc Am* 105:861–873.

Tong YC, Blamey PJ, Dowell RC, Clark GM. (1983) Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant. *J Acoust Soc Am* 74:73–80.

Townshend B, Cotter N, Compernolle DV, White RL. (1987) Pitch perception by cochlear implant subjects. *J Acoust Soc Am* 82:106–115.

Willstedt-Svensson U, Lofqvist A, Almqvist B, Sahlen B. (2004) Is age at implant the only factor that counts? The influence of working memory on lexical and grammatical development in children with cochlear implants. *Int J Audiol* 43:506–515.

Young FW, Lewyckyj R. (1979) *ALSCAL User's Guide.* Chapel Hill: Data Analysis and Theory Associates.