JARO 02: 087–103 (2001)
DOI: 10.1007/s101620010065

JARO
Journal of the Association for Research in Otolaryngology

Effects of Electrode Configuration and Place of Stimulation on Speech Perception with Cochlear Prostheses

BRYAN E. PFINGST, KEVIN H. FRANCK, LI XU, ERIK M. BAUER, and TERESA A. ZWOLAN

Department of Otolaryngology, Kresge Hearing Research Institute, University of Michigan Health System, Ann Arbor, MI 48109-0506, USA

Received: 25 July 2000; Accepted: 7 December 2000; Online publication: 14 May 2001

ABSTRACT

Recent research and clinical experience with cochlear implants suggest that subjects' speech recognition with monopolar or broad bipolar stimulation might be equal to or better than that obtained with narrow bipolar stimulation or other spatially restricted electrode configurations. Furthermore, subjects often prefer the monopolar configurations. The mechanisms underlying these effects are not clear. Two hypotheses are (a) that broader configurations excite more neurons resulting in a more detailed and robust neural representation of the signal and (b) that broader configurations achieve a better spatial distribution of the excited neurons. In this study we compared the effects of electrode configuration and the effects of longitudinal placement and spacing of the active electrodes on speech recognition in human subjects. We used experimental processor maps consisting of 11 active electrodes in a 22-electrode scala tympani array. Narrow bipolar (BP), wide bipolar (BP+6), and monopolar (MP2) configurations were tested with various locations of active electrodes. We tested basal, centered, and apical locations (with adjacent active electrodes) and spatially distributed locations (with every other electrode active) with electrode configuration held constant. Ten postlingually deafened adult human subjects with Nucleus® prostheses were tested using the SPEAK processing strategy. The effects of electrode configuration and longitudinal place of stimulation on recognition of CNC phonemes and words in quiet and CUNY sentences in noise (+10 dB S/N) were similar. Both independent variables had large effects on speech recognition and there were interactions between these variables. These results suggest that the effects of electrode configuration on speech recognition might be due, in part, to differences among the various configurations in the spatial location of stimulation. Correlations of subjective judgments of sound quality with speech-recognition ability were moderate, suggesting that the mechanisms contributing to subjective quality and speech-recognition ability do not completely overlap.

Keywords: cochlear prosthesis, electrode configuration, place of stimulation, speech perception, human

INTRODUCTION

It has long been assumed that the optimal configuration of stimulated electrodes on each channel of a cochlear prosthesis would be one that restricted the longitudinal spread of current, thus maximizing channel separation. However, in the last eight years a few studies of the effects of electrode configuration on speech recognition have reached the surprising conclusion that broad configurations work just as well if not better than narrow configurations in a multichannel prosthesis and that patients most often prefer the broader configuration (Lehnhardt et al. 1992; von Wallenberg et al. 1995; Zwolan et al. 1996; Pfingst et al. 1997; Kileny et al. 1998). These studies all used various versions of Nucleus® prostheses (Cochlear Corporation, Englewood, CO) in which the scala tympani electrode array consisted of 22 band-shaped electrodes

---

Correspondence to: Bryan E. Pfingst, Ph.D. • Kresge Hearing Research Institute • 1301 East Ann Street • Ann Arbor, MI 48109-0506. Telephone: (734) 763-2292; fax: (734) 764-0014; email: bpfingst@umich.edu

spaced at 0.75 mm center to center. Bipolar electrode configurations were longitudinal bipoles with separations of 0.75 mm (designated BP), 1.5 mm (designated BP+1), or 5.25 mm (designated BP+6). In the Nucleus implant, when bipolar stimulation is used, the term "active electrode" is used arbitrarily to refer to the more basally located member of the electrode pair, and the term "return electrode" is used to refer to the more apical member of the pair. Monopolar configurations consisted of stimulation between one intracochlear electrode (the active electrode) and an extracochlear electrode (the return electrode). Two monopolar configurations were available. For the Nucleus 20+2 implant (Zwolan et al. 1996), the return electrode was a ball electrode placed under the temporalis muscle (MP1) or an electrode on the casing of the implanted receiver (MP2). For the Nucleus 20+2L implant (Kileny et al. 1998), the MP1 configuration differed from that in the Nucleus 20+2 implant in that the return electrode was placed into the lateral wall of the cochlea near the apex. The narrow bipolar configurations (BP and BP+1) are assumed to produce narrow (i.e., spatially restricted) patterns of neural excitation, and the broad bipole (BP+6) and the monopoles are assumed to produce broad excitation patterns.

Lehnhardt et al. (1992) were among the first to compare narrow and broad electrode configurations in human subjects. They compared monopolar and bipolar electrode configurations in five subjects with Nucleus 20+2 cochlear implant systems using the MPEAK speech processing strategy. Subjects were tested in an ABA design, where A was the monopolar configuration and B was the bipolar (BP+1) configuration. The subjects had three months experience with each configuration before changing to the next. Results varied from test to test and were confounded with learning effects. With a consonant test, one of the five subjects showed steady improvements over time, three showed a decrease in performance when switched from monopolar to bipolar and then an increase when switched back to monopolar, and one showed an increase in performance when switched from monopolar to bipolar and then a decrease when switched back to monopolar.

In another study, Zwolan et al. (1996) used a balanced crossover design (ABCABC) to test six subjects with Nucleus 20+2 implants using the MPEAK speech processing strategy. The study compared performance with one bipolar (BP+1) and two monopolar (MP1 and MP2) configurations on six speech-recognition tests after subjects had two weeks of experience with each configuration. Electrode configuration did not have a large or consistent effect on speech recognition. However, five of six subjects expressed a preference for the monopolar configurations based on the quality of the sound.

A similar study examined effects of electrode configuration on speech perception in Nucleus 20+2L cochlear implants using the SPEAK speech processing strategy (Kileny et al. 1998). That study used a fivefold replication of an orthogonal Latin square design with nine subjects. Subjects were tested using five experimental processor maps after two months of experience with each map. (The term "map" refers to a set of values that define all stimulation parameters: frequency bands, number of electrodes, electrode configuration, etc.) Three maps (BP+1, MP1, and MP2) used 20 active electrodes and the other two maps (MP1-10 and MP2-10) used ten basal active electrodes. Six speech tests were used. Electrode configuration did not have a large or consistent effect on speech recognition among the maps with 20 active electrodes. Two of the subjects showed better speech-recognition performance on all six tests in one or both monopolar configurations. The remaining subjects showed no consistent differences in speech-recognition results across the three 20-electrode maps. All nine subjects preferred the monopolar 20-electrode maps.

Pfingst et al. (1997) compared a narrow and a broad bipolar electrode configuration (i.e., BP and BP+6, respectively) in subjects using the Nucleus 22 or 20+2 cochlear implant systems that use the MPEAK speech processing strategy. The comparison was done using two experimental maps in which only 11 channels were activated, and the active electrodes for both electrode configurations were in the basal region of the electrode array. Subjects received no practice with the experimental maps prior to speech-recognition testing. Of 12 subjects tested with the BP basal and the BP+6 maps, six showed significantly better speech-recognition performance with the broad (BP+6) configuration and none showed significantly better performance with the narrow (BP) configuration. The authors suggested two possible interpretations of this result. One was that the broader configuration yielded better performance in some subjects because it excited a greater number of neurons resulting in better fidelity. The alternative interpretation also assumed that a greater number of nerve fibers were stimulated by the broader configuration, but that the greater spread of excitation resulted in stimulation of more optimally *located* neurons. This latter hypothesis could take a number of forms. For example, a broader current field might serve to better span patchy areas of nerve loss along the tonotopic axis of the cochlea. Alternatively, it might be that a spread of excitation toward the apex would be advantageous, particularly since fibers apical to the electrode array, which encode the low-frequency components of the speech signal in a normal hearing ear, are often not reached by the apical electrodes that carry the low-frequency components of the electrical signal. The later interpretation was particularly applicable in the Pfingst et al. (1997) study because all of the narrow bipolar pairs of electrodes were located toward the basal end of the electrode array, whereas the broader BP+6 electrode pairs necessarily had the "return" electrode from each pair located more apically. Feasibility of this interpretation was tested in the study reported here.

The goal of the current study was to better understand the effects of the longitudinal placement of the electrodes that carry the outputs of specific channels (i.e., specific frequency bands) and the effects of the electrode configuration on speech perception. To achieve this, we examined the effects on speech recognition and sound quality of several experimental maps in which the active electrodes were located in the basal, middle, or apical region of the electrode array and of maps in which the active electrodes were distributed throughout the array. These electrode placements were studied under narrow bipolar (BP), wide bipolar (BP+6), and, where possible, monopolar (MP2) electrode configurations.

METHODS

Subjects

Data were collected from ten postlingually deaf adult subjects who had been implanted and followed at the University of Michigan. All of the subjects used cochlear implant systems supplied by Cochlear Corporation, Englewood, CO. Four had the Nucleus 22 (N22) and six had the Nucleus 24 (CI24M) cochlear implant. All subjects had at least six months of experience with the implant, native use of English, and fewer than two nonfunctional intracochlear electrodes. In addition, they all had satisfactory open-set speech-recognition performance and were mentally and physically fit for testing. Subjects were paid for time and travel expenses. The use of human subjects in this study was reviewed and approved by the University of Michigan Medical School Institutional Review Board.

The processors for all subjects using the N22 implant were programmed with the SPEAK processing strategy (Skinner et al. 1994; Whitford et al. 1995) in the devices that they used every day. Of the CI24M users, three normally used the SPEAK processing strategy and three used the ACE processing strategy. The ACE strategy is conceptually the same as the SPEAK strategy but typically uses a higher pulse rate and a few other slightly different parameters (Cochlear Corporation 1999). Additional subject details are included in Table 1.

Equipment

The N22 cochlear implant had an array of 22 electrodes that were surgically implanted into the scala tympani through a cochleaostomy. Electrodes were labeled 1 through 22, starting at the basal end of the array. The CI24M cochlear implant also had an array of 22 intracochlear electrodes but, in addition, had two extracochlear electrodes: a ball electrode placed under the temporalis muscle and an electrode on the casing of the internal receiver. These two extracochlear electrodes were used for two monopolar configurations: MP1 and MP2, respectively. The normal electrode configuration of the CI24M implant was MP1+2, where the return path was the MP1 and MP2 electrodes in parallel. In this study, MP2 was used for the monopolar configuration.

The parameters for the experimental maps used in these experiments were controlled using Cochlear Corporation's Diagnostic and Programming System (DPS) software version 6.125 and version 7 for users of the N22 and CI24M devices, respectively. The software communicated with the laboratory speech processors via an IF4 ISA card and the Dual Processor Interface (DPI) for the N22 and via an IF5 ISA card and the Processor Control Interface (PCI) for the CI24M. The speech processors controlled transmission of radio frequency (RF) pulses to the internal receiver/stimulator. The internal receiver/stimulator then decoded the RF information to the correct stimulation pulse parameters.

All testing using experimental maps was conducted using one of two laboratory speech processors and headsets. For N22 subjects we used a Spectra speech processor (serial number 346609) with an HS-6 headset and for CI24M subjects we used a SPrint speech processor (serial number 408594) with an HS-8 headset. This procedure avoided any chance of confounding effects caused by differences in the individual subjects' processors and microphones.

Speech-recognition test materials were presented using a Sony CDP-C250Z compact disk player. A GSI-1715 audiometer was used to control the speech signal amplitude. A Rane ME-60 graphic equalizer was used to flatten the frequency response. A Crown D-75 amplifier was used to increase the signal amplitude. Signals were presented through a TDC 4A loudspeaker positioned 1 m away from the subject at 0° azimuth inside an Acoustic Systems (Model RE 242S, Austin, TX) double-walled sound-attenuating booth.

The speech test level was calibrated periodically with a sound-level meter (Brüel and Kjær Type 2231, Naerum, Denmark). Subjects were not present during the calibration. The sound-level meter was positioned near where the headset microphone would be located during the test sessions. A fast time setting (i.e., time

TABLE 1

Subject information

Subject number	*Sex*	*Age (years)*	*Device*	*Normal coding strategy*	*Normal electrode configuration*	*Stiffening rings inserted*	*X-ray estimate of insertion depth*[a]	*Duration of deafness (years)*[b]	*Length of device use (years)*	*Symbol used in this article*
1	M	49	CI24M	SPEAK	MP1+2	4	5	1	0.6	◀
2	F	76	N22	SPEAK	BP+1	7	7	2	4.4	○
3	F	52	N22	SPEAK	BP+1	9	8	30	3.1	△
4	F	75	N22	SPEAK	BP+1	7	NA	11	8.6	□
5	F	32	CI24M	SPEAK	MP1+2	0	2	4	0.5	■
6	F	51	CI24M	SPEAK	MP1+2	4	6	11	1.5	●
7	F	50	CI24M	ACE	MP1+2	5	5	3	1.5	▶
8	F	39	CI24M	ACE	MP1+2	1	3	1	0.7	▲
9	F	32	CI24M	ACE	MP1+2	4	5	13	0.5	◆
10	F	48	N22	SPEAK	BP+1	8	NA	6	13.4	◇

[a] The number of electrodes that were apical to the most-superior electrode in the cochlea is reported here (see Methods and Fig. 3 for details). NA indicates that the x-rays were not available.

[b] Defined as the amount of time between patient's report of the date of onset of profound hearing loss and the date of implant activation.

TABLE 2

Independent variables

Electrode configurations	*Electrode locations*	*Overall bandwidths*
BP	Spaced	Narrow
BP+6[a]	Basal	Wide
MP2[b]	Centered	
	Apical	
	Full-array[c]	

[a] Implemented only in a basal active-electrode location.

[b] Implemented only with CI24M implant.

[c] Used in BP and MP2 configurations. Only the wide overall bandwidth was applicable.

constant of 125 ms) and an "A" frequency weighting were set in the sound-level meter during the calibration to the speech materials. A spectrum analyzer (Stanford Research Systems, model SR760 FFT, Sunnyvale, CA) was used for the frequency response calibration to both narrow and broadband signals. An equalizer (Rane ME-60 graphic equalizer, Mukilteo, WA) was adjusted to assure compliance with ANSI 3.6 specifications.

Research design

Three independent variables were tested in various combinations to address the aims of these experiments: electrode configuration, place of stimulation, and overall bandwidth. Electrode configuration and place of stimulation were the main independent variables. The overall-bandwidth variable was used to control for changes in the total bandwidth of presented frequencies imposed when the 11-electrode maps were used. Parameters of each independent variable are listed in Table 2.

Three electrode configurations were tested: two bipolar (narrow and wide) and one monopolar. For the narrow bipolar (BP) configuration, both the active and the return electrodes were located in the cochlea. The active and return electrodes were adjacent, with a center-to-center distance of approximately 0.75 mm (Fig. 1). The more basal electrode was labeled as the active electrode. The wide bipolar (BP+6) configuration (Fig. 1) was similar, but the electrode separation was 5.25 mm (six nonstimulated electrodes in between the basal active and the apical return electrodes). For the monopolar configurations (MP2), the intracochlear electrodes were active, and a single electrode located on the casing of the internal receiver served as the return.

Place of stimulation was defined in the BP and BP+6 configurations as the portion of the electrode array spanned by the active and return electrodes and defined in the MP2 configurations as the portion spanned by the active electrodes. Longitudinal positions included the entire array, the basal half of the array, the centered half of the array, and the apical half of the array (Fig. 2). All of the apical, centered, and basal longitudinal positions were achieved using 11 adjacent active electrodes. Longitudinal positions spanning the entire array were achieved using either 20 adjacent active electrodes (called the "full array") or 11 spaced active electrodes (every other electrode active). With the preceding definition of place of stimulation, BP+6 (active electrodes numbered 3–13 and return electrodes numbered 10–20) spanned most of the array.

The overall bandwidth presented to the electrode array (defined as the bandwidth from the lowest frequency presented to the most apical electrode to the highest frequency presented to the most basal electrode) depended on the number of electrodes used and the bandwidth assigned to each electrode pair.

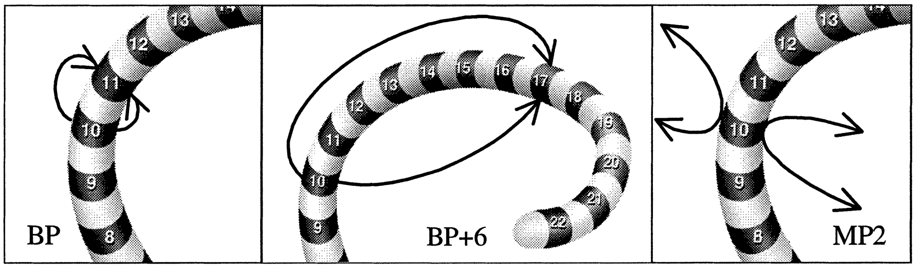

FIG. 1. Schematic illustrations of the three electrode configurations tested in the current experiments: Narrow bipolar (BP), wide bipolar (BP+6), and monopolar (MP2). In each of the three panels, the numbered dark bands represent the electrodes of a Nucleus® N22 or CI24M implant. Arrows indicate current flow from the active to the return electrode, which occurs during one phase of the biphasic pulse. In the monopolar (MP2) case, the current flow is to a remote extracochlear electrode. Activation of only one site in each electrode array is illustrated.

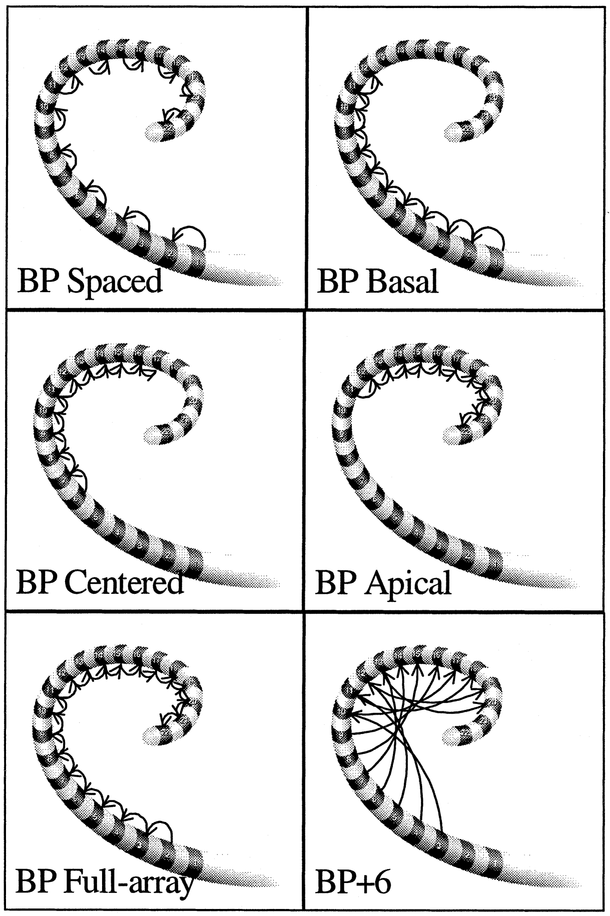

FIG. 2. Experimental maps where the independent variable was place of stimulation. The upper two rows illustrate various electrode placements for 11-electrode maps. For these illustrations, a narrow bipolar configuration was used. Arrows illustrating current flow are arbitrarily drawn to show current flowing from the "active" electrode (i.e., the more basal member of the bipolar pair) to the "return" electrode. Similar electrode locations were tested in a monopolar configuration where the active electrodes were in the same location but the return electrodes were extracochlear. The lower-left panel illustrates the full array, which utilized 20 active electrodes. The lower-right panel illustrates the BP+6 map.

When 11-electrode maps were used and a single processor channel was assigned to each stimulation site, the bandwidth assigned to each electrode pair was slightly larger than the bandwidth assigned to the same electrode pair when the full array was used. However, the overall bandwidth was still smaller in the 11-electrode map than in the full-array map. Wider overall bandwidths, comparable to the overall bandwidth of the full-array map, were achieved with the 11-electrode maps by pairing two processing channels to each electrode site (double mapping). The resulting total bandwidth of the 11-electrode maps was similar to the total bandwidth used in the subject's normal processor. Note that double mapping may result in doubling the stimulation rate for certain electrode pairs when the two adjacent analysis channels assigned to that electrode pair contain spectral peaks in a particular analysis cycle.

The internal filter settings for the Spectra and SPrint processors (which drive the N22 and CI24M cochlear implants, respectively) were slightly different, so the frequency range allocated to each electrode differed slightly between these two processors. For subjects using the N22 cochlear implant, frequency allocation table number 6 (frequency range of 109–7,871 Hz) was used for the full-array and the 11-electrode double-mapped wide-overall-bandwidth maps. Frequency allocation table number 12 (frequency range of 240–4,288 Hz) was used for 11-electrode single-mapped narrow-overall-bandwidth maps. For subjects using the CI24M cochlear implant, frequency allocation table number 6 (frequency range of 116–7,871 Hz) was used for the full-array and 11-electrode wide-overall-bandwidth maps, and frequency allocation table number 13 (frequency range of 160–4,666 Hz) was used for 11-electrode narrow-overall-bandwidth maps.

In total, there were nine bipolar maps (spaced, basal, centered, and apical maps with both narrow and wide overall bandwidth; and full-array), two BP+6 maps (narrow and wide overall bandwidth), and nine MP2 maps (analogous to the nine bipolar maps). Users of the N22 system were tested with a maximum of 11 maps because monopolar configurations were not available with their implants. Users of the CI24M system could be tested with a maximum of 20 maps. Subject 1 had incomplete data (2 tests with the wide-bandwidth maps missing), as did subjects 2, 3, and 5 (no wide-bandwidth maps) because they became unavailable for testing during that part of the experiment.

Procedures

The first step in the experiment was to determine the subject's threshold (T level) and maximum comfortable loudness level (C level) for each electrode in the array for each of the two (BP and BP+6 for N22) or three (BP, BP+6, and MP2 for CI24M) electrode configurations. Procedures were similar to those used in fitting implants clinically. T and C levels were established using the method of adjustment with a control knob for adjusting the level of the current. The stimuli consisted of 200-μms/phase symmetric-biphasic pulses presented at a rate of 250 pulses/s with a 500-ms on/off duty cycle. The processors were set in "current level" mode so that pulse duration would remain constant. First in the bipolar electrode configuration, T levels were determined from the apical electrodes to the basal electrodes. C levels were then determined in the same direction. In order to check for the effects of adaptation to the high-level stimuli, the C levels for the apical four electrodes were rechecked. Variation by more than three programming units necessitated repetition. After setting T and C levels on each electrode, an apical-to-basal sweep of all electrodes at the threshold level was presented. The subject was asked to be sure that each presentation was heard and that all presentations were of equal loudness. Adjustments to T levels were made accordingly and then an apical-to-basal sweep of all electrodes at the maximum comfortable loudness level was presented. The subject was asked to be sure that none of the presentations were uncomfortably loud and that all presentations were of equal loudness. Adjustments to C levels were made accordingly. The cochlear implant fitting software combined all the stimulation parameters into a "map". The map was informally tested to be sure that stimulation was comfortable when using the processor volume and sensitivity used in the experiments. Modifications to global C levels in percentage of the dynamic range were made if the map resulted in excessive loudness. The maps were then saved and the mapping procedure was repeated for the BP+6 and MP2 configurations.

After all maps were completed, speech-recognition testing was performed. The order of testing the experimental maps was randomized and each experimental map was tested three times. For each of the 33 (N22) or 60 (CI24M) test sessions, the experimental processor was programmed with the corresponding map. Additionally, three tests were performed using the subject's everyday map with the subject's personal processor unit. All speech materials were presented at a level averaging 64 dB(A) measured during individual word or syllable presentations using a fast time setting (time constant of 125 ms). During speech-recognition testing, the speech processor was set at a normal volume and sensitivity level (sensitivity = 3 for Spectra; sensitivity = 7 and volume = 10 for SPrint). The subjects were not permitted to adjust the processor.

The Spectral Peak (SPEAK) processing strategy was used for all conditions in this experiment. In this strategy the acoustic signal was passed through a bank of bandpass filters. Filter settings were distributed linearly at lower frequencies and logarithmically at higher frequencies (lin–log frequency spacing). The output of each filter was associated with one channel, with low-frequency outputs sent to more apically located electrodes (higher electrode numbers) and high-frequency outputs sent to more basally located electrodes (lower electrode numbers). A subset of channels was stimulated during a given cycle and the pulses were interleaved so that no two electrode sites were stimulated at the same time. Pulse rate for the SPEAK processing strategy was about 250 pulses/s on any given stimulated electrode. The subset of channels to be stimulated on a given cycle was determined by the largest peaks in the outputs of the filters. The relative location of the stimulated channels in the electrode array, and thus on the longitudinal (tonotopic) axis of the scala tympani, was determined by the frequencies at which the spectral peaks occurred. The amplitude of the pulses was modulated in proportion to the amplitudes of the spectral peaks.

Materials for speech-perception testing were CNC words, CNC phonemes, and CUNY sentences in noise. CNC materials were presented from the Minimum Speech Test Battery for Adult Cochlear Implant Users on a compact disk (House Ear Institute, Los Angeles, CA). One list containing 50 CNC words in quiet was presented during each of the three test sessions. In seven subjects, one list of 12 CUNY sentences with a 10-dB signal-to-noise ratio (multitalker babble noise) was also presented for each of the test sessions. CUNY sentences were presented from the Cochlear Corporation's Investigational Test Battery compact disk.

Data analysis

Binomial-variable analysis (Thornton and Raffin 1978) was used to determine statistical significance of differences in speech-recognition scores between maps for individual subjects. Statistical significance testing on group means was performed using standard *t* statistics.

Two intervening variables that must be considered in interpreting the effects of place of stimulation are (1) the proximity of the experimental map to the user map (i.e., the map that the subject used every day) to which the subject has adapted for some period of time, and (2) the proximity of the experimental map to the tonotopic map of the normal cochlea. In order to determine if these variables might have contributed to the speech-recognition performance of the subjects, we compared each experimental map with the user map and the tonotopic map. Two metrics, referred to as map differences, were computed as follows to represent the frequency differences between the experimental map and the user map [Eq. (1)] or the normal tonotopic map [Eq. (2)]:

$$\frac{\sum_{i=1}^{n} \left| \log \frac{F_{e(i)}}{F_{u(i)}} \right|}{n} \tag{1}$$

where F_e is the center frequency of an active electrode in an experimental map, F_u is the center frequency of the corresponding active electrode of the user map, and n is the number of active electrodes used in the experimental map; and

$$\frac{\sum_{i=1}^{n} \left| \log \frac{F_{e(i)}}{F_{t(i)}} \right|}{n} \tag{2}$$

where F_t is the normal tonotopic frequency on the basilar membrane of the cochlea at the position of the active electrode for monopolar stimulation or at the position halfway between the active and return electrodes for bipolar stimulation. The tonotopic frequency was defined by Greenwood's (1990) formula, $F_t = 165.4(10^{0.06x} - 0.88)$, where x is the distance (in mm) of the electrode from the apex. It was assumed that the basilar membrane length is ~35 mm and the audible frequency range is 20–20,000 Hz.

On each Nucleus electrode array, there are 10 stiffening rings basal to the 22 electrodes. The 10 stiffening rings and the 22 electrodes are spaced equally at an interval of ~0.75 mm (center to center). In the ten subjects reported here, the number of stiffening rings (n) inside of the cochlea, based on the surgeon's report, ranged from 0 to 9 (Table 1 and Fig. 3). The position (x) of an active electrode for monopolar stimulation (in mm from the apex) was then calculated using the formula: $x = 35 - [0.75(n + \text{active electrode number} - 1) + 0.375]$. The formula was slightly adjusted for calculation of the position halfway between the active and return electrodes for BP and BP+6 stimulation: $x = 35 - [0.75(n + \text{active electrode number} - 1 + 0.5m) + 0.375]$, where m is 1 for BP and 7 for BP+6 stimulation.

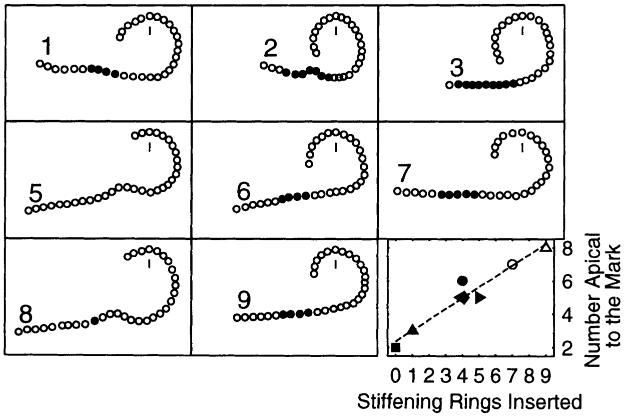

FIG. 3. Assessment of insertion depth based on x-rays for eight of the ten subjects. Each panel (except for the lower-right panel) shows the electrode and stiffening ring locations obtained from digitized photographs of transorbital x-ray films for one subject. The subject number is indicated on the left. The images of electrode arrays implanted to the right ears are flipped 180° to the left so that all of the images are shown in the same orientation. The vertical bar represents the vertical plane and a length of 1 mm. The upper tip of the bar points to the most superior electrode in the cochlea. There are ten stiffening rings, and those reported inserted into the cochlea by the surgeon are filled black. The lower-right panel plots the relationship of the stiffening rings inserted (abscissa) and the number of electrodes that were apical to the most superior electrode (ordinate). Each symbol represents an individual subject, as indicated in Table 1. The dashed line is the linear fit to the data. The correlation coefficient $r = 0.969$.

Plain-film x-rays were available for eight of the ten subjects. These were used to check the insertion depth estimates that were based on the surgeon's reports of the number of rings inserted. The x-rays were taken at the time of implant activation, approximately one month following the implant surgery, using a standard posterior–anterior transorbital orientation. Depth of implant insertion was estimated based on the following procedure: The films were digitized using a digital camera (Sony model DKC-CM30) attached to a surgical microscope. For each digitized x-ray, the coordinates of the 22 implanted electrodes plus 10 stiffening rings and a ruler lying in the vertical axis were recorded. All of the images were aligned relative to the vertical (midline) axis. The electrode located at the most superior location was identified (see vertical mark in Fig. 3). The number of electrodes that were apical to this most-superior electrode was counted and reported as the relative insertion depth (Table 1). Figure 3 shows the locations of the electrodes and the results of the quantitative assessment of the insertion depth. It is noteworthy that the surgeon's reports of the number of rings inserted and the relative insertion depths that were estimated based on the above procedure were highly correlated (Fig. 3, lower-right panel, $r = 0.969$).

Following speech-recognition testing with each map, subjects completed a questionnaire in which they rated the subjective sound quality of the map and their ability to recognize words by circling one of the following adjectives: terrible, very poor, poor, fair, good, very good, or excellent. In our data analysis, these adjectives were represented by the numbers 1–7. Spearman's rank correlation analyses were used to compare the subjects' subjective quality ratings to their subjective ability ratings and to compare quality and ability ratings to speech-recognition performance. The sign test was used to determine the statistical significance of electrode-configuration effects on the subjective ratings.

RESULTS

Relationships among speech tests

Speech-recognition results obtained in the present study for CNC words, CNC phonemes, and CUNY sentences in noise were highly correlated with one another. The CNC word and CNC phoneme results from 360 tests in ten subjects showed a correlation coefficient of 0.950. The CNC word and CUNY sentence results from 184 tests in seven subjects showed a correlation coefficient of 0.873. The CNC phoneme and CUNY sentence results from 184 tests in seven subjects showed a correlation coefficient of 0.885. All correlation coefficients were statistically significant ($p < 0.01$). Therefore, in the following sections, we report only results from CNC phoneme tests as the scores for speech-recognition performance.

Effects of electrode configuration

Figure 4 shows the speech-recognition scores for CNC phonemes for the six CI24M subjects with MP2 and BP configurations at various locations of the active electrodes. Each score is the mean percent correct across three CNC phoneme tests. In the spaced locations (left panel), speech-recognition performance was better in the monopolar configuration than in the bipolar configuration for all but one (S8: upright triangles) of the six subjects. The group-mean speech-recognition scores for the MP2 configuration were 9.3 percentage points higher than those for the BP configuration. This difference was statistically significant (paired t test, $p < 0.01$). When electrode locations were restricted to basal, centered, or apical halves of the electrode array (middle panels), only one subject showed consistently better performance with the MP2 configuration (S5, filled squares). Note that this subject also showed the largest difference (~30%) in performance between the MP2 and BP configurations in the spaced locations. For the group of six subjects, no statistically significant difference was found between the average performance for the two configurations using the three compressed (basal, centered, and apical) maps.

In four of the six CI24M subjects, we also compared the speech-recognition scores for CNC phonemes using MP2 and BP configurations and a full array of closely spaced electrode sites (Fig. 4, right panel). All four subjects performed better with the MP2 configuration than with the BP configuration. The group-mean performance of the four subjects was 8.7 percentage points higher for the MP2 configuration than for the BP configuration and this difference was statistically significant (paired t test, $p < 0.05$).

Speech-recognition performance for the wide bipolar configuration (BP+6) was compared with that for the narrow bipolar (BP) configuration and the monopolar (MP2) configuration at the basal electrode locations. Comparisons were made with basal locations because the active electrode locations were similar for these three maps. This comparison is analogous to the Pfingst et al. (1997) study. Comparisons were also made with the centered maps because the region spanned by the current path between active and return electrode locations for the BP+6 configuration was comparable to that for the BP centered and MP2 centered maps. Figure 5 (upper panel) shows the speech-recognition performance for CNC phonemes for all ten subjects using BP basal, BP+6, and MP2 basal configurations. Consistent with the previous report (Pfingst et al. 1997), most of the subjects in this study showed better speech-recognition performance with the wide bipolar map (BP+6) than with the narrow bipolar basal map (BP basal). Statistical analysis (binomial-variable analyses) revealed that eight of the ten subjects did significantly better with the BP+6 map than with the BP basal map. Only one subject (S1: left-pointing triangles) did better with the BP basal map. The remaining subject (S8: upright triangles) showed no statistically significant differences in performance between the two configurations. The group mean of speech-recognition scores for the BP+6 map was 10.5 percentage points higher than that for the BP basal map and this difference was statistically significant (paired t test, $p < 0.01$).

In five out of six subjects in whom a monopolar configuration could be tested, speech-recognition performance was better with the BP+6 map than with the MP2 basal map. These differences were statistically significant in two of the subjects (binomial-variable analyses, $p < 0.05$). One subject (S1: left-pointing triangles) did better with the MP2 basal map. The difference between the group-mean scores for the BP+6 and the MP2 basal maps was not statistically significant.

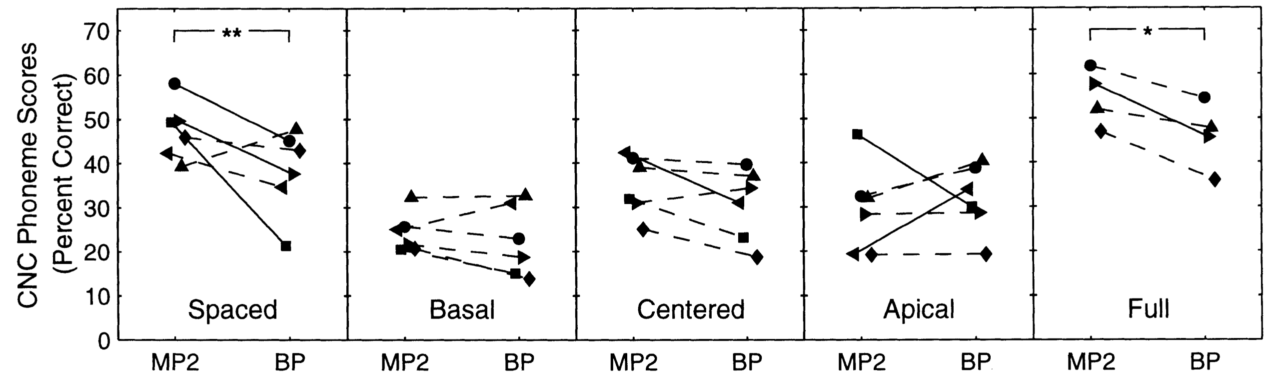

FIG. 4. Speech-recognition scores with MP2 and BP configurations at various electrode locations for six CI24M subjects. Panels from left to right represent spaced, basal, centered, and apical locations for the 11-electrode maps, and the map using the full array of electrodes. Each line represents speech-recognition performance of one of the six subjects that used CI24M implants, which have both monopolar and bipolar stimulation modes available. Only four of these subjects were tested with the full array. Each score is the mean percent correct for phonemes across three lists of CNC words. The solid lines indicate that the differences between the scores with MP2 and BP configurations are statistically significant (binomial-variable analysis, $p < 0.05$). The dashed lines indicate that the differences are not statistically significant. Each symbol represents an individual subject, as indicated in Table 1. Locations of data points on the abscissa are jittered slightly to avoid overlap. Group-mean differences between the speech-recognition performance for monopolar (MP2) and bipolar (BP) configurations were statistically significant in the spaced and full-array conditions as indicated by the asterisks (paired t test: $^{**}p < 0.01$, $^{*}p < 0.05$).

When performance with the BP+6 map was compared with performance with the centered BP or MP2 maps (Fig. 5, lower panel), few statistically significant differences were found at either individual or group levels. However, an "atypical" subject (S1: left-pointing triangles) did remarkably worse with the BP+6 map than with the centered BP or MP2 maps. Subject S9 (filled diamonds) showed significantly higher scores with the BP+6 map than with the centered BP map.

Effects of place of stimulation

Figure 6 shows the effect of longitudinal electrode location on speech recognition. Overall, the centered electrode location yielded the highest speech-recognition scores among basal, centered, and apical locations. Results from basal and apical locations did not differ significantly from each other.

For almost every subject, centered locations produced higher speech-recognition scores than did basal electrode locations regardless of electrode configuration (BP or MP2). The differences in group-mean scores for both the BP and the MP2 centered vs. basal locations were 11.2 percentage points. These differences were statistically significant (paired t test, $p < 0.05$).

Place of stimulation seemed to interact with electrode configuration and with the subject's previous stimulation history when we compared the centered and apical locations. For the MP2 configuration (Fig. 6, upper panel), five of six subjects tested showed higher speech-recognition scores with the centered location than with the apical location, although this difference was statistically significant for only one subject (S1: left-pointing triangles) (binomial-variable analysis). One subject (S5: filled squares) did remarkably better with the apical location. It should be noted that this subject had the shallowest insertion depth of the electrode array among all subjects (Table 1). We address the effects of insertion depth on speech recognition in the next section and in the Discussion section. For the BP configuration (Fig. 6, lower panel), all six CI24M subjects (filled symbols), whose normal daily electrode configurations were monopolar, performed almost equally well with either centered or apical locations. However, all four N22 subjects (open symbols), whose normal daily electrode configurations were bipolar, performed better with centered locations than with apical locations. This difference was statistically significant when evaluated using individual (binomial-variable analysis, $p < 0.05$) and group (paired t test, $p < 0.05$) analyses.

The spaced maps tended to produce higher speech-recognition scores than the three compressed maps (basal, centered, and apical) (Fig. 6). For monopolar stimulation (upper panel), speech-recognition scores of the six CI24M subjects for the spaced maps were an average of 12.4% higher than those for the centered maps. For bipolar stimulation (lower panel), the scores of the four N22 subjects (open symbols) for the spaced maps were on average 17.8% higher than those for the centered maps. Both group-mean differences were statistically significant (paired t test, $p < 0.05$).

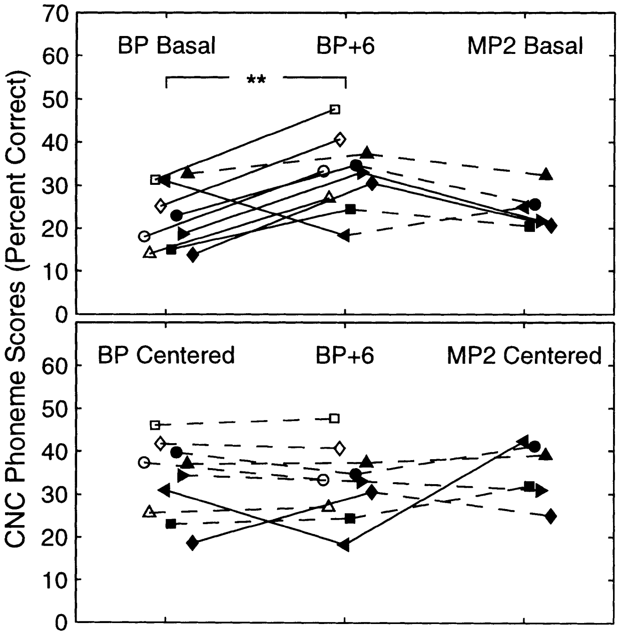

FIG. 5. Speech-recognition scores of all subjects using BP, BP+6, and MP2 configurations. Only the six CI24M subjects could be tested in the MP2 configuration. The upper panel is for the BP basal and MP2 basal locations and the lower panel is for the centered locations. The BP+6 configuration is the same in both panels. The active electrodes for this configuration were in the basal region of the array, but the overall distribution of electrodes (active to return) was centered. Each line represents one subject. The solid lines indicate that the differences between the scores with the BP and BP+6 maps or between the scores with the BP+6 and MP2 maps were statistically significant (binomial-variable analysis, $p < 0.05$). The dashed lines indicate no statistically significant difference. Each symbol represents an individual subject, as indicated in Table 1. The filled symbols represent CI24M subjects in whom all three configurations were tested. The open symbols represent N22 subjects in whom only the bipolar configuration was tested. Group-mean differences in speech-recognition performance were statistically significant only between the BP basal and BP+6 configurations, as indicated by the asterisks (paired t test: ** $p < 0.01$). Locations of data points on the abscissa are jittered slightly to avoid overlap.

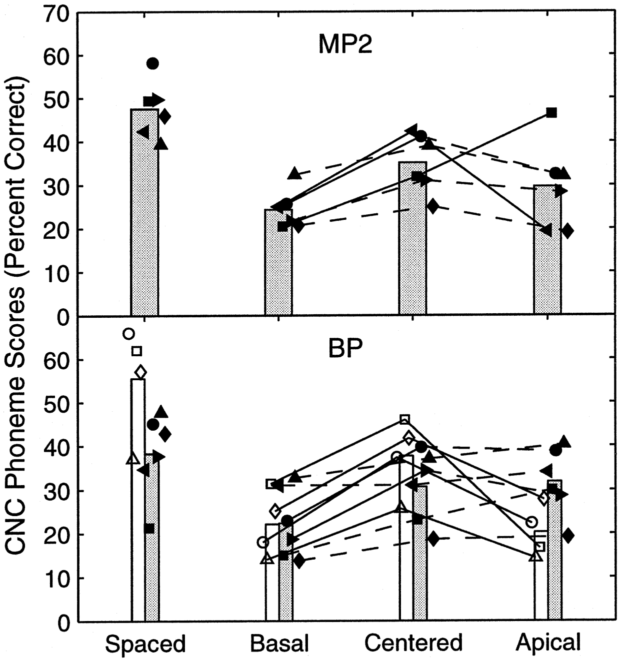

FIG. 6. Speech-recognition scores as a function of electrode location (spaced, basal, centered, and apical). **Upper panel**: Monopolar (MP2) stimulation. **Lower panel**: Bipolar (BP) stimulation. Each symbol or line represents one subject. The solid lines indicate that the differences between the scores with basal and centered or between the scores with centered and apical locations are statistically significant (binomial-variable analysis, $p < 0.05$). The dashed lines indicate no statistical significance. The symbol assigned to each subject is given in Table 1. The gray and open bars represent the means of speech-recognition scores for the CI24M and the N22 subjects, respectively. Statistical analyses of group means are given in the text. Locations of data points on the abscissa are jittered slightly to avoid overlap.

Effects of relationship of experimental maps to user map and to normal tonotopic map

Figure 7 shows the computed map differences between the experimental maps (basal, centered, apical, spaced, and BP+6) and the user map (left panel) or the tonotopic map (right panel). For the map differences relative to the user map, the centered and spaced maps tended to produce the smallest map differences, whereas for the map differences relative to the tonotopic map, the smallest map differences were found for the apical maps.

Map differences relative to the user map showed a correlation with the speech-recognition scores for CNC phonemes. Figure 8 (left panel) plots the speech-recognition scores with the experimental maps as a function of the differences between the experimental maps and the user maps. The slopes of the linear fits to such functions, which indicate the strength of the dependence of speech-recognition scores on the map differences relative to the user map, tended to increase in magnitude with the duration of use of the cochlear implants (Fig. 9, solid line). On the other hand, differences between the experimental maps and the tonotopic map did not correlate well with speech-recognition performance (Fig. 8, right panel). The slopes of the linear fits to the function of speech-recognition scores and the map differences relative to the tonotopic map (Fig. 8, right panel) were shallower than those relative to the user map (Fig. 8, left panel) and they tended not to depend on duration of use of the cochlear implants (Fig. 9, dashed line). Statistical analysis (z statistics) of the slopes of the lines in Figure 9 indicated that the slope of the dashed line was not significantly different from 0 ($p = 0.64$) and that the slope of the solid line was significantly different from 0 ($p < 0.01$). Therefore, our data indicate that after more than six months of wearing a cochlear implant, the normal tonotopic map exerts little influence on

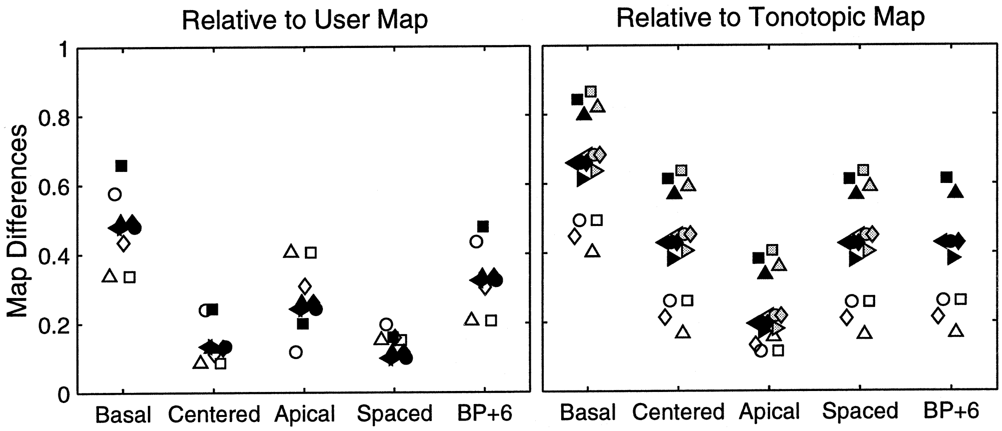

FIG. 7. Map differences relative to the user map [**left panel**; from Eq. (1)] or relative to the tonotopic map [**right panel**; from Eq. (2)] at the basal, centered, apical, spaced positions and the BP+6 configuration. In the left panel, each symbol represents an individual subject, as indicated in Table 1. In the right panel, each of the N22 subjects is represented by one symbol whereas each of the CI24M subjects is represented by two symbols, one in black for bipolar configuration and one in gray for monopolar configuration. Locations of data points on the abscissa are jittered slightly to avoid overlap.

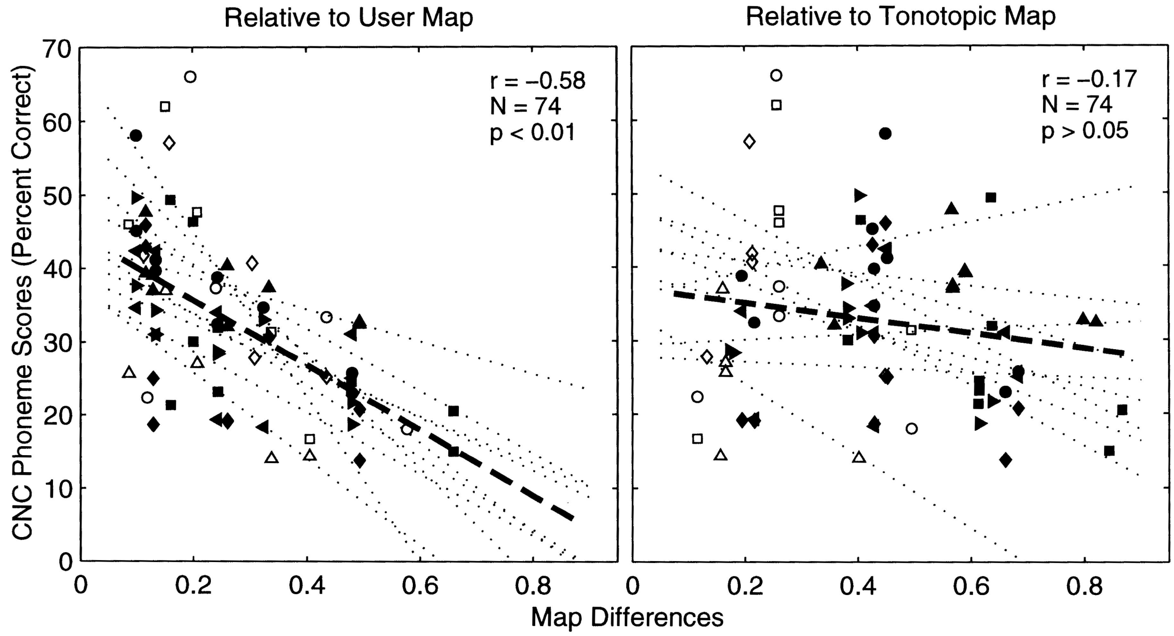

FIG. 8. Relationship between the speech-recognition scores and the map differences relative to the user maps (**left panel**) or relative to the tonotopic map (**right panel**). Each symbol type represents an individual subject, as indicated in Table 1. Each thin dotted line represents the least-squares fit of data from one subject. The thick dashed lines represent the least-squares fit of data from all subjects.

speech-recognition performance with an experimental map. The user map, however, plays an important role in determining the speech-recognition performance when an experimental map is introduced and such an effect continues to strengthen with increasing duration of implant use and experience with the user map.

Effects of bandwidth

With the 11-electrode experimental maps, two overall bandwidths were tested: one more narrow than that in the subject's user map and one similar to that in the user map. The narrow overall bandwidth was the software default, as it used one channel per electrode pair. Because the experimental maps had fewer electrodes, the overall bandwidths in the 11-electrode maps were narrower than those of the users' maps. To achieve an overall bandwidth similar to that in the user's map, each electrode site was assigned two channels (double mapping). Narrow and wide bandwidths were tested in order to determine their influence on speech-recognition performance across electrode configuration, spacing, and location.

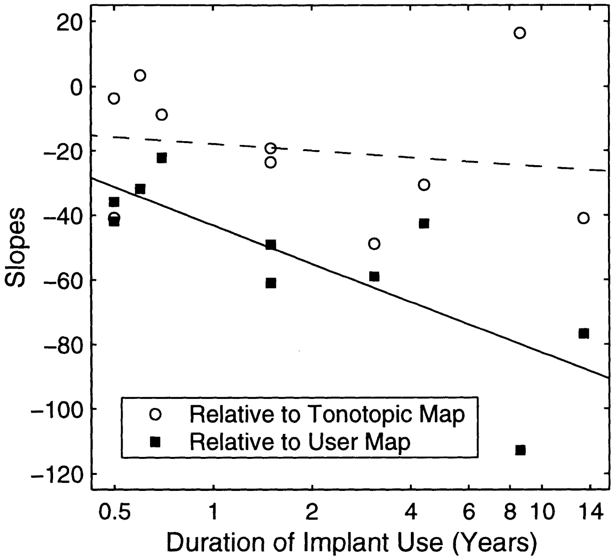

FIG. 9. Slopes of the percent correct vs. map difference functions from Figure 8 plotted as a function of duration of implant use. These slopes represent the degree of dependence of speech-recognition performance with the experimental maps on the proximity of the experimental maps to the user map or the tonotopic map. Each symbol represents one subject. The filled squares and open circles plot the slopes of the least-squares fit obtained from the left and right panels of Figure 8, respectively, as a function of the logarithm of the duration of implant use. The solid and dashed lines represent the linear fit of the filled squares and open circles, respectively. Statistics analysis indicates that the dashed line has a slope not significantly different from 0 ($p < 0.64$) and that the solid line has a slope significantly different from 0 ($p < 0.01$).

No consistent effects of bandwidth were observed in this study. Of the 74 comparisons made between narrow and wide bandwidths, only five were statistically significant (two were better in wide bandwidth, three were better in narrow bandwidth). In only one case could this be predicted by map difference scores [Eq. (1)]. In the bipolar centered map for subject S4, narrow and wide bandwidths had map difference scores relative to the user map of 0.09 and 0.22, respectively. Performance with this map was significantly higher for the narrow-bandwidth map than it was for the wide-bandwidth map.

All ten subjects that participated in the present study showed relatively high scores with their own maps in their daily used speech processors (user maps). The mean percent correct of speech recognition (CNC phonemes) was 65.5%, ranging from 42.2% to 80.9%. In general, performance with the experimental maps (where subjects had no practice) was poorer than that with the user maps. However, the experimental maps yielding the best scores (usually the spaced-electrode maps) yielded comparable scores to those obtained with the user maps in four of the ten subjects.

Subjective judgments

We found that there existed a strong correlation between the subjective judgments of the sound quality and the subjective judgments of the ability to recognize speech. In the ten subjects, rank correlation coefficients (r_s derived from Spearman's rank correlation analysis ranged from 0.84 to 1.00 (all $p < 0.01$) with a median of 0.92. Data from one representative subject (S6) are shown in Figure 10 (left panel) in which the subjective judgments of the sound quality are plotted on the abscissa and the subjective judgments of the ability to recognize words are plotted on the ordinate. Each data point represents the mean of the three subjective rankings for each map. One can see that all data points are close to a straight line parallel to the main diagonal line. Figure 10 (right panel) shows the pooled data from all ten subjects. It is evident that most of the subjective judgments are on or close to the main diagonal line. The high correlation between the subjective judgments of the sound quality and judgments of the ability to recognize words suggests that the two judgments are not independent of each other.

Subjective judgments of both sound quality and the ability to recognize words predicted the speech-recognition scores to some extent. Spearman's rank correlation coefficients for speech-recognition scores and the subjective judgments of sound quality ranged from 0.30 to 0.74 (all $p < 0.05$) with a median of 0.55 for the ten subjects. Similarly, the rank correlation coefficients between the speech-recognition scores and the subjective judgments of the ability to recognize speech ranged from 0.31 to 0.78 (all $p < 0.05$) with a median of 0.59. Figure 11 shows data from one subject, S6, whose rank correlation coefficients were representative of the medians of the population.

Previous studies have reported that the subjects tended to prefer monopolar configurations to bipolar configurations (von Wallenberg et al. 1995; Zwolan et al. 1996; Kileny et al. 1998). In our study, we compared monopolar and bipolar configurations in terms of the subjective judgments of the sound quality and subjective judgments of the ability to recognize words. Figure 12 plots the mean ranking of each subject's judgments of the sound quality in both MP2 and BP configurations at various active-electrode locations for the six subjects who could be tested in both MP2 and BP configurations. Only at the centered location did all six subjects show higher scores for MP2 than for BP. The sign test indicated that these differences were statistically significant ($p = 0.031$). At other electrode locations, however, no statistically significant differences in ranking were detected (sign test, $p > 0.05$).

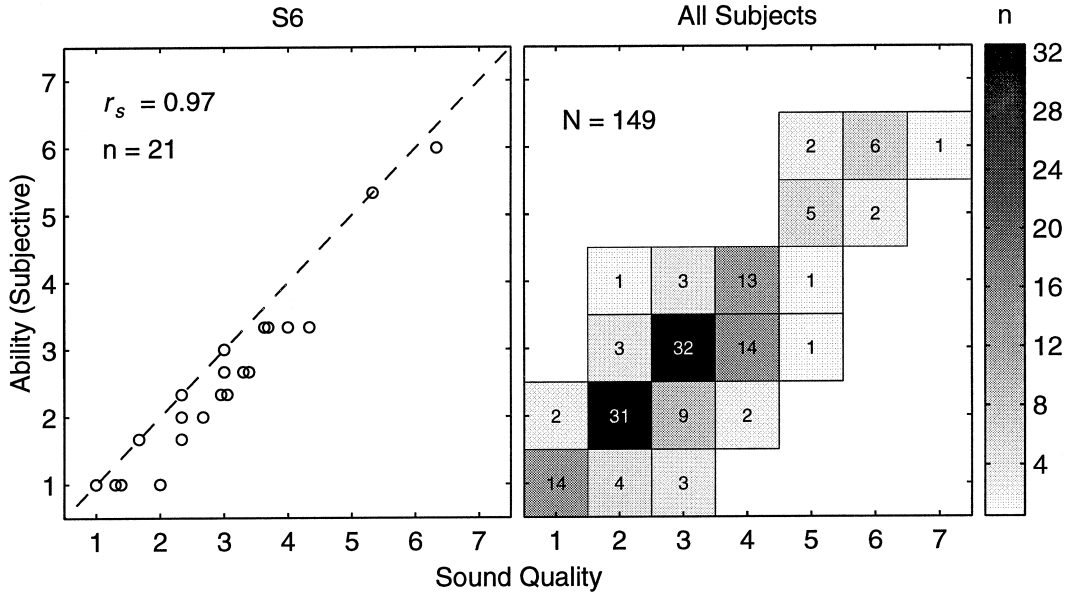

FIG. 10. Relationship between subjective judgments of the sound quality and subjective judgments of the ability to recognize words. (**Left panel**) Data from representative subject (S6). Each data point represents the mean subjective judgments of three sessions of speech tests. The subjective ability and quality scales ranged from 1 = "terrible" to 7 = "excellent." (**Right panel**) Pooled data from all ten subjects. A square is plotted when any of the data points fall into such a space. The number in the middle of each square and the associated gray scale represents the number of data points that fell into the square. It is evident that most data points fell on or near the main diagonal line.

DISCUSSION

The electrode configuration chosen for implementation in commercial cochlear prostheses has varied over the years. Decisions about which configuration or configurations to offer have been based on a combination of theoretical and practical considerations. The initial Nucleus® implants were designed only for bipolar stimulation, based on the assumption that narrow bipolar configurations would maximize the independence of neural populations stimulated by individual channels of the multichannel prosthesis (Cochlear Corporation 1993). As a contemporary of the early Nucleus designs, the Ineraid® prosthesis used monopolar stimulation, but the electrodes were much more widely separated in the scala tympani than those in Nucleus implants (Youngblood and Robinson 1988). Even with wide separation of stimulation sites, it was demonstrated that currents delivered simultaneously to adjacent electrodes did interact and, indeed, there was evidence that eliminating this current interaction by nonsimultaneous stimulation improved speech recognition (Wilson et al. 1991). Contemporary Nucleus prostheses (Cochlear Corporation 1999) primarily use monopolar stimulation, based on the speech-recognition studies described in the Introduction and on the fact that monopolar stimulation requires less current than narrow bipolar stimulation. The Clarion® prosthesis initially attempted to achieve channel independence with bipolar stimulation on radially oriented, closely-spaced electrodes, but this was not generally successful because of high current requirements. Recently, the Clarion prosthesis offered the option of monopolar or diagonally oriented bipolar stimulation, and subjects' preferences were mixed (Osberger and Fisher 1999).

The mechanisms underlying differences across patients in preference for strategy and in speech-recognition performance are poorly understood. It seems likely that electrode configuration interacts with other variables so that the effects of electrode configuration vary from case to case. As we come to recognize the mechanisms underlying the effects of electrode configuration, we will gain a better understanding of these interactions and increase our ability to predict and control the variables that affect speech recognition and quality of electrical hearing.

In the experiments reported in this article, we found an interaction between effects of electrode configuration and effects of longitudinal electrode placement. For the 11-electrode maps, we found that the effects of electrode configuration (BP vs. MP2) were significant only for the spaced electrode placement. There are several possible interpretations of this finding. First, it might be that for the three compressed maps (basal, middle, and apical), any beneficial effects of the monopolar configuration, such as excitation of more neurons, were countered by deleterious effects of channel interaction as a result of the close spacing of the electrode pairs. To test this hypothesis, we com-

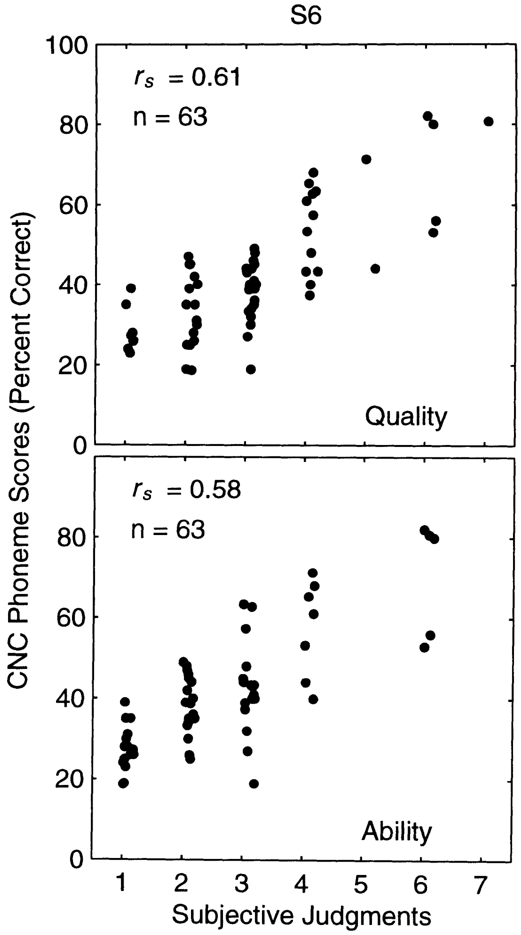

FIG. 11. **Upper panel**: Relationship between speech-recognition scores and subjective judgments of the sound quality and **lower panel**: relationship between speech-recognition scores and subjective judgments of the ability to recognize words. Data are from one representative subject (S6). Each data point was obtained from one test session. Locations of the data points on the abscissa are jittered slightly to avoid overlap. The subjective ability and quality scales ranged from 1 = "terrible" to 7 = "excellent."

pared speech-recognition performance for MP2 and BP stimulation using a full 20-electrode map where the spacing between the stimulation sites was the same as in the 11-electrode compressed maps. The difference in speech recognition with monopolar versus bipolar stimulation under these maps was found to be similar to the difference under the spaced map (Fig. 4). This result suggests that longitudinal electrode location rather than electrode spacing was the primary variable interacting with electrode configuration to affect speech recognition.

Another possible explanation of the interaction between electrode configuration and place of stimulation concerns the similarity of the experimental maps to the users' normal maps. All six of the subjects in this experiment who could be tested with monopolar stimulation used monopolar stimulation in their normal everyday maps. Familiarity and practice with a particular spatial pattern of electrical stimulation can have a large effect on speech recognition, as discussed below. The compressed maps were perhaps so far removed from the users' normal maps that neither electrode configuration made them familiar enough to have an effect.

Note that the four subjects who normally used bipolar stimulation in their everyday map (the N22 subjects) did better with the spaced map than with the compressed maps (see Fig. 6). The relative difference in percentage points between performance in the spaced vs. centered maps for the N22 subjects was equivalent to the difference between these maps for monopolar stimulation in the CI24M subjects who normally used monopolar stimulation in their everyday maps (Fig. 6). This argues in favor of the hypothesis that familiarity with the maps contributed to the observed interaction between electrode configuration and place of stimulation.

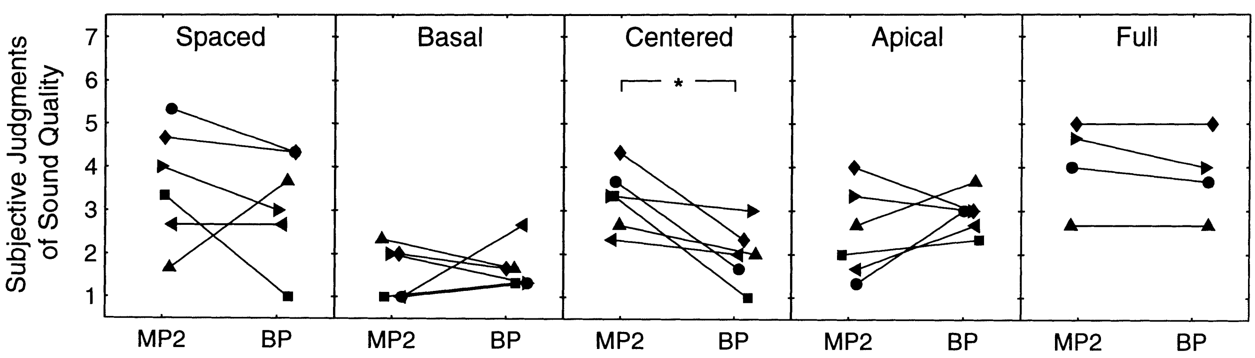

FIG. 12. Comparison of the mean subjective judgments of the sound quality between MP2 and BP configurations at various electrode locations for six CI24M subjects. As in Figure 4, the panels from left to right represent spaced, basal, centered, apical locations for the 11-electrode maps and the full array of electrodes. Each line represents the mean subjective judgments of the sound quality of one of the six subjects that had both monopolar and bipolar stimulation modes available. Each symbol represents an individual subject, as indicated in Table 1. The group-mean quality judgment scores were significantly different only for the 11-electrode centered map as indicated by the asterisk (paired *t* test: * $p < 0.05$). The subjective quality scales ranged from 1 = "terrible" to 7 = "excellent." Locations of data points on the abscissa are jittered slightly to avoid overlap.

Data from the current study suggest that the advantage found in the Pfingst et al. (1997) study for the BP+6 configuration (with more basally located active electrodes and more apically located return electrodes) might have been due to the location of the stimulation being more toward the middle of the electrode array than with the BP basal map. Data on place of stimulation suggest that stimulation centered in the electrode array was preferable to basal stimulation. On the other hand, in the current study we found no significant advantage of monopolar stimulation over bipolar stimulation for electrodes in the basal location (Fig. 4). This suggests that monopolar stimulation did not achieve the advantage that was expected to result from spread of excitation toward a more desirable location in the tonotopic axis.

It must be recognized that our estimates of the sites of stimulation along the longitudinal axis of the cochlea are only rough approximations. A number of factors can contribute to errors in the estimation of the site of neural stimulation and to variability in accuracy of the estimates from case to case. These include variations in cochlear length and other anatomical features from subject to subject (Úlehlová et al. 1987), unknown nerve-survival patterns, and imprecisely defined current pathways. Estimation of site of stimulation across subjects, of course, will be more variable than estimates of the relative location of sites within subjects, but both measures are subject to error from a number of sources.

Previous studies of the effects of longitudinal stimulus location on speech recognition have produced mixed results. Kileny et al. (1992) showed no statistically significant difference between performance obtained using a full 20-channel map and a map that used only 10 channels in the basal end of the electrode array. However, in a later study with some different conditions, Kileny et al. (1998) found poorer performance when only the basal 10 channels were stimulated compared with the full array. Fu and Shannon (1999) found that the longitudinal position of the stimulated electrodes made a significant difference in speech-recognition performance that depended on the range of stimulus frequencies assigned to each stimulation site. The importance of longitudinal electrode location has also been demonstrated using acoustic simulations of cochlear prostheses (Dorman et al. 1997).

There were a number of variables that differed among these studies that might have contributed to the differences in results. For example, Kileny et al. (1992) used an $F_0F_1F_2$ processor and allowed the subjects six months of experience with each experimental map before testing speech recognition. Kileny et al. (1998) used a SPEAK processing strategy with two months' experience, and Fu and Shannon (1999) used as a CIS processing strategy and tested subjects after no practice with each map. In the current study, we used the SPEAK processing strategy and no practice. Using acoustic simulations similar to those used by Dorman et al. (1997), Rosen et al. (1999) have shown that practice is an important variable in experiments where the location of stimulation along the tonotopic axis is shifted experimentally. Thus, it is quite reasonable to assume that the lack of significant effects observed by Kileny et al. (1992) was due to the six months of experience that the subjects had with each map before testing. Consistent with this assumption is the observation by Kileny et al. (1992) that subjects showed a decrease in performance immediately after shifting from either the full configuration to the reduced, basal configuration or vice versa. Also, Fu and Shannon (1999) have argued that the lack of correlation between the optimal frequency assignments to the various channels of the implant and implant insertion depth is evidence that subjects learn to adapt to whatever map they are given. However, the time course of this adaptation, and the variables that affect it, have not been clearly defined. In our experiment, the one subject (S5) who consistently showed her best speech-recognition performance (among the compressed maps) with electrodes in the apical position was also the subject with the shallowest insertion depth (all 10 stiffening rings outside the cochlea). Examination of the x-rays confirmed the shallow insertion (Fig. 3). This subject had been using her prosthesis for a little over six months (Table 1), suggesting that relatively large shifts from the normal tonotopic cochlear map can have effects that last for several months. The data from Figure 9 in our experiments suggest that the influence of the user's normal map on the performance with an experimental map might increase over many years.

An advantage of using the short-term procedures, where the subject is given little or no training with the experimental maps, is that the effects of the independent variables are most obvious under these conditions. It is often possible to reduce some of the effects of the independent variables by long-term training of the subjects, thus making the effects more difficult to detect. However, we currently know little about the details of these training effects. It is not known if training completely overcomes the effects of new stimulation patterns or only reduces the impact of these effects. This area requires considerable additional research.

The CNC phoneme scores for subjects in our study were similar to those obtained in an IDE-controlled multicenter clinical trial involving 62 subjects with Nucleus CI24M prostheses (Arndt et al. 1999). This suggests that our subjects were representative of a larger population.

In this experiment with cochlear prostheses, we found high correlations among scores on the various speech tests. This is in agreement with previous studies of speech recognition using acoustic hearing (Boothroyd and Nittrouer 1988; Olsen et al. 1997). It suggests that recognition of the components of sentences is highly predictive of the subject's ability to recognize whole sentences. In contrast, the correlation between speech recognition and the subjective judgments of the quality of the speech sounds was less strong. This suggests that there might be some differences in the mechanisms underlying speech recognition and mechanisms underlying quality of the perceived sound. Both subjective quality and speech-recognition ability are important to implanted subjects. In addition, subjective-quality judgments might reflect variables that are important for hearing of non-speech stimuli such as music. While improving speech recognition is an appropriate short-term focus for cochlear implant research, in the long run we must consider a broader spectrum of the auditory experience.

SUMMARY AND CONCLUSIONS

We found high correlations between various tests of speech recognition across a variety of conditions of electrical stimulation. This is in general agreement with previous studies using acoustic hearing (Boothroyd and Nittrouer 1988; Olsen et al. 1997).

There was an interaction between the effects of electrode configuration and place of stimulation. For the basal, centered, and apical 11-electrode maps, there was no significant difference between performance with monopolar (MP2) and narrow bipolar (BP) configurations. With the 11-electrode map with spaced active electrodes and with the full map, the group of subjects implanted with the CI24M device performed significantly better with the monopolar configuration (MP2) than with the bipolar (BP) configuration. This latter result might have been affected by the subjects' previous experience since all subjects that could be tested in the monopolar configuration used it in the prostheses that they wore daily.

Speech-recognition performance for the wide bipolar configuration (BP+6) was better than that for the basally located, narrow bipolar configuration (BP), which had similarly located active electrodes. This result is in agreement with the findings of Pfingst et al. (1997). Performance with the BP+6 configuration was similar to that of the BP centered location, which was more comparable to the BP+6 in the longitudinal distribution of the stimulation. This suggests that the better speech-recognition performance with BP+6 than with BP basal was due more to the place of stimulation than to the size of the stimulated neural population.

Speech-recognition performance for basally located monopolar (MP2) stimulation was not significantly better than that for basally located, narrow bipolar (BP) stimulation. This suggests that the broader electrode configuration did not overcome the disadvantage of a poor electrode location.

Of the 11-electrode maps, the centered location yielded the best speech recognition. This result was predicted by the proximity of the stimulus placement to that in the user's everyday map. The ability of proximity to the user's map to predict performance with the experimental maps increased as a function of the user's experience with their everyday map over the 14-year period of use available for study in this subject population.

Subjective judgments of the quality of the experimental maps were only moderately correlated with the subject's speech-recognition ability with these maps. This suggests that the mechanisms contributing to sound quality and speech-recognition ability do not completely overlap. Since improvements in both speech-recognition ability and subjective quality would benefit the patients, both measures should be considered in evaluating the stimulus features important for electrical hearing.

ACKNOWLEDGMENTS

This work was supported by NIH-NIDCD grants R01 DC 03808 and T32 DC 00011. We express appreciation to Alex Arts, M.D., for assistance in interpreting the subjects' x-rays and to Julie Oldfield for technical assistance.

REFERENCES

ARNDT PA, STALLER SJ, HOLDEN LK. (1998) Application of advanced speech encoding strategies in cochlear implant recipients. Paper presented at the 11th Annual Meeting of the American Academy of Audiology, Miami, FL.

BOOTHROYD A, NITTROUER S. Mathematical treatment of context effects in phoneme and word recognition. J. Acoust. Soc. Am. 84:101–114, 1988.

COCHLEAR CORPORATION. Audiologist's Handbook Cochlear Limited Englewood, CO, 1993.

COCHLEAR CORPORATION. Nucleus Technical Reference Manual Z43470, Issue 1 Cochlear Limited Lane Cove, Australia, 1999.

DORMAN MF, LOIZOU PC, RAINEY D. Simulating the effect of cochlear-implant electrode insertion depth on speech understanding. J. Acoust. Soc. Am. 102:2993–2996, 1997.

FU QJ, SHANNON RV. Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant. Ear. Hear. 20:321–331, 1999.

GREENWOOD DD. A cochlear frequency-position function for several species—29 years later. J. Acoust. Soc. Am. 87:2592–2605, 1990.

KILENY PR, ZIMMERMAN–PHILLIPS S, ZWOLAN TA, KEMINK JL. Effects of channel number and place of stimulation on performance with the Cochlear Corporation multichannel implant. Am. J. Otol. 13:117–123, 1992.

KILENY PR, ZWOLAN TA, TELIAN SA, BOERST A. Performance with the 20 + 2L lateral wall cochlear implant. Am. J. Otol. 19:313–319, 1998.

LEHNHARDT E, GNADEBERG D, BATTMER RD, VON WALLENBERG E. Experience with the cochlear miniature speech processor in adults and children together with a comparison of unipolar and bipolar modes. ORL J. Otorhinolaryngol, Relat. Spec. 54:308–313, 1992.

OLSEN WO, VAN TASELL DJ, SPEAKS CE. Phoneme and word recognition for words in isolation and in sentences. Ear. Hear. 18:175–188, 1997.

OSBERGER MJ, FISHER L. SAS–CIS preference study in postlingually deafened adults implanted with the Clarion® Cochlear Implant. Ann. Otol. Rhinol. Laryngol. 108:74–79, 1999.

PFINGST BE, ZWOLAN TA, HOLLOWAY LA. Effects of stimulus configuration on psychophysical operating levels and on speech recognition with cochlear implants. Hear. Res. 112:247–260, 1997.

ROSEN S, FAULKNER A, WILKINSON L. Adaptation by normal listeners to upward spectral shifts of speech: Implications for cochlear implants. J. Acoust. Soc. Am. 106:3629–3638, 1999.

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, BEIRER AL. Evaluation of a new Spectral Peak coding strategy for the Nucleus 22 Channel Cochlear Implant System. Am. J. Otol. 15:15–27, 1994.

THORNTON AR, RAFFIN MJ. Speech-discrimination scores modeled as a binomial variable. J. Speech. Hear. Res. 21:507–518, 1978.

ÚLEHLOVÁ L, VOLDŘICH L, JANISCH R. Correlative study of sensory cell density and cochlear length in humans. Hear. Res. 28:149–151, 1987.

VON WALLENBERG EL, BATTMER R-D, DODEN I, GNADEBERG MS, HAUTLE K, LENARZ T. Place-pitch and speech perception measures with bipolar and monopolar electrical stimulation of the cochlea. Ann. Otol. Rhinol. Laryngol. 104(suppl 166):372–375, 1995.

WHITFORD LA, SELIGMAN PM, EVERINGHAM CE, ANTOGNELLI T, SKOK MC, HOLLOW RD, PLANT KL, GERIN ES, STALLER SJ, MCDERMOTT HJ, GIBSON WR, CLARK GM. Evaluation of the Nucleus Spectra 22 processor and new speech processing strategy (SPEAK) in postlinguistically deafened adults. Acta Otolaryngol. (Stockh) 115:629–637, 1995.

WILSON BS, FINLEY CC, LAWSON DT, WOLFORD RD, EDDINGTON DK, RABINOWITZ WM. Better speech recognition with cochlear implants. Nature. 352:236–238, 1991.

YOUNGBLOOD J, ROBINSON S. Ineraid (Utah) Multichannel Cochlear Implants. Laryngoscope. 98:5–10, 1998.

ZWOLAN TA, KILENY PR, ASHBAUGH C, TELIAN SA. Patient performance with the Cochlear Corporation “20+2” implant: Bipolar versus monopolar activation. Am. J. Otol. 17:717–723. 1996.