

Frequency-place compression and expansion in cochlear implant listeners

Deniz Başkent^{a)}

Department of Biomedical Engineering, University of Southern California, Los Angeles, California 90089 and Department of Auditory Implants and Perception, House Ear Institute, 2100 W. Third Street, Los Angeles, California 90057

Robert V. Shannon

Department of Biomedical Engineering, University of Southern California, Los Angeles, California 90089 and Department of Auditory Implants and Perception, House Ear Institute, 2100 W. Third Street, Los Angeles, California 90057

(Received 28 September 2003; revised 18 August 2004; accepted 19 August 2004)

In multichannel cochlear implants, low frequency information is delivered to apical cochlear locations while high frequency information is delivered to more basal locations, mimicking the normal acoustic tonotopic organization of the auditory nerves. In clinical practice, little attention has been paid to the distribution of acoustic input across the electrodes of an individual patient that might vary in terms of spacing and absolute tonotopic location. In normal-hearing listeners, Başkent and Shannon (J. Acoust. Soc. Am. **113**, 2003) simulated implant signal processing conditions in which the frequency range assigned to the array was systematically made wider or narrower than the simulated stimulation range in the cochlea, resulting in frequency-place compression or expansion, respectively. In general, the best speech recognition was obtained when the input acoustic information was delivered to the matching tonotopic place in the cochlea with least frequency-place distortion. The present study measured phoneme and sentence recognition scores with similar frequency-place manipulations in six Med-El Combi 40+ implant subjects. Stimulation locations were estimated using the Greenwood mapping function based on the estimated electrode insertion depth. Results from frequency-place compression and expansion with implants were similar to simulation results, especially for postlingually deafened subjects, despite the uncertainty in the actual stimulation sites of the auditory nerves. The present study shows that frequency-place mapping is an important factor in implant performance and an individual implant patient's map could be optimized with functional tests using frequency-place manipulations. © 2004 Acoustical Society of America. [DOI: 10.1121/1.1804627]

PACS numbers: 43.71.Es, 43.71.Pc, 43.66.Ts [PFA]

Pages: 3130–3140

I. INTRODUCTION

Cochlear implants (CI) partially restore hearing to deaf people by electrical stimulation of the auditory nerves. The electrodes are organized in an array to deliver the spectral information roughly consistent with the tonotopic organization of the cochlea, with lower frequency information delivered in the apical region and higher frequency information delivered in the basal region. However, little attention has been paid during the standard fitting process to the spacing or absolute tonotopic location of the electrodes.

The implant processor controls how the spectral content of the acoustic input is assigned onto the electrodes. Despite the uncertainty in the current spread and the stimulation location in nerves, the tonotopic range of stimulation in the cochlea is primarily determined by the active length of the electrode array and its insertion depth. When the array is fully inserted, the most apical contact is usually 20–30 mm from the round window depending on electrode type. The active stimulation range of the electrode array is typically 13.5 or 16.5 mm in length for Clarion I and Clarion II, 16.5

mm for Nucleus 22 and Nucleus 24, and 26.4 mm for Med-El Combi 40+. Greenwood's frequency-place function (1990) describes the characteristic frequency along the organ of Corti as a function of cochlear place. Assuming an average length of 35 mm for human cochlea, an array of 16 mm would stimulate a cochlear region that corresponds to an acoustic frequency range of 1–12 kHz for a 20-mm-insertion depth, and an acoustic frequency range of 500–6000 Hz for a 25-mm-insertion depth. Similarly, a 26-mm-long array inserted to a depth of 30 mm would cover a tonotopic range of 185–11 800 Hz. While the electrode array position may vary from patient to patient, many of the present cochlear implant clinical fitting programs offer only a limited choice for the overall spectral range of analysis filters as well as partitioning of individual bandwidths; common ranges are 350–6800 Hz for Clarion II, 150 Hz–10 kHz (SPEAK strategy Table 9) for Nucleus 22, and from 200–300 to 5500–8000 Hz for Med-El Combi 40+. As a result, the acoustic frequency range assigned to the stimulation region in the cochlea can be wider or narrower than the acoustic characteristic frequency range of that region, resulting in compression or expansion of the frequency-to-place mapping, respectively. Often there is also a tonotopic shift due to the discrepancy

^{a)}Electronic mail: dbaskent@hei.org

between the actual electrode location and the acoustic information assigned. In some cases such a shift might be the result of an electrode that is not fully inserted, while in other cases the shift may be due to the assignment of the default signal processing parameters.

Başkent and Shannon (2003) measured the effects of compression and expansion in frequency-place mapping on speech recognition by normal-hearing (NH) subjects. Implant electrode arrays with different insertion depths and different number of electrodes were simulated using a noise-band vocoder (e.g., Shannon *et al.*, 1995). In the vocoder the cochlear tonotopic range of stimulation was represented by noise carrier bands while the input acoustic frequency range was determined by the frequency range of the analysis bands. The stimulation range was held constant by employing the same noise carrier bands for each condition while the analysis frequency range was made wider (compressed map) or narrower (expanded map) relative to the carrier frequency range. Speech recognition was generally better when the analysis range matched the carrier range than for any frequency-place expansion and compression condition, even when the matched condition eliminated a considerable amount of acoustic information. This result suggests that speech recognition, at least without training, is dependent on the mapping of acoustic frequency information onto the appropriate cochlear place. Fu and Shannon (1999) found a similar result with frequency-place shift. Vowel recognition by NH subjects was significantly reduced if the frequency information was presented to a simulated cochlear location more than 3 mm from its normal tonotopic location. They also observed a similar drop in performance with implant users when the input frequency range was shifted prior to implant processing. Whitford *et al.* (1993) attempted to modify the processor maps of implant users by matching the acoustic input range to the characteristic frequency of the stimulation range and observed some improvement in the open-set sentence recognition scores in low levels of noise.

A theoretical issue addressed by the present study is the flexibility of perceptual pattern recognition for altered speech tonotopic patterns. When NH subjects were tested with frequency-place maps different than the normal acoustic map, such as shifted (Fu and Shannon, 1999), compressed, or expanded maps (Başkent and Shannon, 2003), there was a reduction in phoneme and sentence recognition performance. Prior to the experiments NH subjects had experience with only the normal acoustic frequency-place map, which is perceptually “burned in” over a lifetime of hearing. In contrast, CI users had experience with two frequency-place maps: the normal acoustic map (from the previous history of normal acoustic hearing), and the frequency-place map of the implant processor, which is likely to be different than the acoustic map. It is not clear whether or how speech pattern recognition would be able to adapt to the new implant frequency-place map. The present study does not address the time course of such long-term adaptation, but rather looks at the instantaneous effects of spectral distortions in the mapping. If CI users are affected by spectral alterations in a manner similar to NH subjects, fitting the frequency-place map for an individual implant user would be beneficial. A

map optimized from the beginning might make the overall adaptation faster and the final asymptotic performance higher.

In cochlear implants there are several factors that affect the frequency-place mapping, but cannot be estimated with certainty. These factors include physical quantities such as the exact insertion depth of the electrode array, the proximity of the electrodes to the spiral ganglia where the actual stimulation occurs, and the actual length of the cochlea, physiological factors such as nerve survival pattern and the temporal and spatial pattern of stimulation in the auditory nerves, and anatomical factors such as possible structural abnormalities of the cochlea. It may be possible to obtain more detailed information about the position of the implanted electrodes using sophisticated images from radiographs (Marsh *et al.*, 1993; Cohen *et al.*, 1996) or CT scans (Ketten *et al.*, 1998; Skinner *et al.*, 2003). However, even if the relative insertion depth can be determined in individual implant patients the medial-lateral location of the electrode in the scala tympani might still not be accurately determined. Moreover, because many factors affecting the current flow cannot be imaged, it would still not be possible to know the absolute location and characteristic frequency of the neurons activated by each electrode. Because of these uncertainties, the best assessment of electrode location and frequency-place matching might be accomplished functionally rather than by imaging.

In the present study we used the nominal value for insertion depth as reported by Med-El for fully inserted electrode arrays. All participating subjects had full insertions. In reality, however, there is probably a large variation in the actual electrode locations across subjects due to individual differences in cochlear length, medial lateral electrode location, and nerve survival. One purpose of the study was to assess whether an optimum map could be obtained despite the unknown factors by starting with a map based on estimated values and fine-tuning it with behavioral tests.

In Experiment 1, we used the fact that implant users are sensitive to spectral shifts in frequency-place maps (Fu and Shannon, 1999) to find an estimate for the electrode insertion depth behaviorally. The array location producing the best performance (12–24 mm, with electrodes 4–9) was used in Experiments 2 and 3.

A typical value used in applying the Greenwood mapping function for the average cochlear length of human is 35 mm. However, measurements by Ulehlova *et al.* (1987) showed a range from 28 to 40 mm, with an average length of 34.2 mm for 28 men. When Ketten *et al.* (1998) estimated cochlear lengths of implant subjects from CT scans, they found an average length of 33 mm for 20 subjects (range 29–37.5 mm). In Experiment 2 we varied the assumed cochlear length used in the calculations of the frequency map and explored the effects on speech recognition. Subjects had peak performance at different values, but performance did not change significantly over a range of a few mm. Therefore we used the typical value of 35 mm for the average cochlear length for all subjects in Experiments 3 and 4.

The behaviorally measured parameter values for electrode array location and cochlear length from Experiments 1

TABLE I. Information about subjects, all users of Med-El Combi 40+. Reasonable scores for sentence recognition with subjects S5 and S6 (as shown with asterisks) could be obtained with simpler sentences (HINT) only, where the subjects were also allowed to listen to each sentence as many times as needed. Baseline scores were collected using subject's clinical map.

Subject	Age	Duration of profound deafness (years)-etiology	Experience with CI (years)	Baseline vowel score (corrected for chance)	Baseline consonant score (corrected for chance)	Baseline sentence score (IEEE or *HINT)	Overall acoustic input frequency range of the original map
S1	39	30-high fever	2.5	60.00	55.26	38.22	300–5500 Hz, 6 or 12 electrodes later: 200–8500 Hz, 10, 11, or 12 electrodes
S2	62	12-noise exposure	1	68.18	70.18	92.81	300–5500 Hz, all 12 electrodes
S3	46	26-unknown	2	82.50	86.67	93.94	Map 1 and 2: 300–5500 Hz Map 3: 300–7000 Hz 9 electrodes
S4	25	from birth-unknown	5	70.00	85.91	84.52	300–7000 Hz, 9 or 12 electrodes
S5	36	from birth-pregnancy rubella	3 total, 1 year with replacement	42.73	30.71	17.5* (HINT)	300–5500 Hz, all 12 electrodes
S6	40	from birth-unknown	4.5	44.55	52.63	12.8* (HINT)	300–7000 Hz, Map 1 and 2: 6 electrodes Map 3: all 12 electrodes

and 2 were used as baseline estimates for Experiments 3 and 4. From these baseline values, frequency-place compression and expansion conditions were produced with a 6-electrode processor in Experiment 3, and expansion conditions with a 12-electrode processor in Experiment 4.

II. EXPERIMENTAL METHOD

A. Subjects

Six Med-El Combi 40+ users (S1–S6), aged 25–62, participated in experiments. All were reported to have full electrode insertions at surgery. Detailed information about subjects is summarized in Table I. Center frequencies are shown in Table II for the analysis bands assigned to electrodes in the clinical maps of subjects.

Three subjects (S1, S2, and S3) were postlingually, and three (S4, S5, and S6) were prelingually deafened. All subjects were born into hearing families and therefore have used oral communication as their main communication mode, and some had been provided with speech correction therapies for long periods of time.

The baseline sentence scores given in Table I are for IEEE sentences, which were too difficult for subjects S5 and S6. They were retested with simpler HINT sentences and were allowed to listen to each sentence as many times as needed. Even with the simpler materials their open-set scores were too low to fully observe the effects of spectral manipulations. These subjects participated in Experiment 3 only. Due to time constraints S4 did not participate in Experiments 1 and 2, and S3 did not participate in Experiment 4.

TABLE II. Center frequencies of clinical maps used by the subjects.

Subject	Band-pass filter center frequencies for 12 electrodes (Hz)											
	1	2	3	4	5	6	7	8	9	10	11	12
S1, S2, S5	338	430	549	701	894	1137	1444	1845	2349	2987	3889	4918
S3	352	487	off	672	930	off	1273	1771	off	2420	3456	4544
S4	358	507	off	722	1017	off	1445	2057	off	2890	4225	6013
S6	390	off	658	off	1114	off	1867	off	3147	off	5319	off

B. Speech stimuli

Vowel stimuli (selected from recordings by Hillenbrand *et al.*, 1995) consisted of ten presentations (five male and five female talkers) of twelve medial vowels in CVC syllables, including ten monophthongs and two diphthongs, presented in a /h/-vowel-/d/ context (heed, hid, head, had, hod, hawed, hood, who'd, hud, heard, hayed, hoed). Chance level on this test was 8.33% correct, and the single-tailed 95% confidence level was 12.48% correct based on a binomial distribution.

Consonants (selected from recordings by Shannon *et al.*, 1999) consisted of six presentations (three male and three female talkers) of 20 medial consonants (b, tʃ, d, ð, f, g, ʒ, k, l, m, n, p, r, s, ʃ, t, v, w, j, z) in CVC syllables, presented in an /a/-consonant-/a/ context. Chance performance level for this test was 5% correct, and the single-tailed 95% confidence level was 8.27% correct based on a binomial distribution.

Two different sets of stimuli were used for sentence recognition tests: IEEE sentences (IEEE, 1969) spoken by a single male talker and HINT sentences (Nilsson *et al.*, 1994) spoken by multiple talkers. IEEE sentences are phonetically balanced across lists and the predictability of the words is relatively low. Subjects S1 and S2 were also retested with HINT sentences. HINT sentences are contextually easier, have fewer key words, and are more similar to natural speech in daily life compared to IEEE sentences. Subjects had not heard any of the test sentences prior to the study. Each list consisted of ten sentences and two lists were presented for each condition. The presentation order of lists was randomized across subjects and conditions.

C. Procedures

Speech stimuli were presented via a loudspeaker in a sound field at 70 dB on an A-weighted scale in a sound-treated room. In the vowel and consonant identification tests, subjects were asked to choose the phoneme they heard from a menu displayed on the computer screen. All stimuli were presented in random order via custom software (Robert, 1998). In the sentence recognition test, they were asked to repeat or type the words they heard in sentences, again presented in random order via custom software (Tiger Speech Recognition System developed by Qian-Jie Fu). The map in the experimental processor was changed before every test. The conditions of a specific experiment were presented in random order to minimize learning effects.

D. Med-EI Combi 40+ implant system

The electrode array of the Combi 40+ consists of 12 electrodes spaced 2.4 mm apart covering a total length of 26.4 mm in cochlea. The electrodes are numbered 1–12 from apex to base. The array is designed for insertions as deep as 31 mm inside the round window. Stimuli were delivered to the implant via a research TEMPO+processor, which is worn behind-the-ear (BTE) and can process acoustic frequencies from 200 Hz to 8.5 kHz.

E. Device parameters

For every subject, thresholds and maximum loudness levels were determined using the standard clinical method to customize the experimental processor for an individual subject. The input dynamic range was automatically adjusted by the processor. The TEMPO+processor uses the CIS strategy and the electrodes are stimulated in the monopolar mode.

F. Experimental conditions

In Experiments 1–3 we used a set of six electrodes that had a narrower stimulation range than the default 12 electrodes of the device. A narrower stimulation range was selected to facilitate flexibility in manipulating the spectral content of speech over a range of electrode array configurations. In the “matched” condition the input acoustic range was equal to the range of characteristic frequencies of the array, calculated by Greenwood equation. If an insertion depth of 31 mm is assumed, the six middle electrodes of the implant cover a range of 12 mm, from 12 to 24 mm from the round window. With the additional assumption of 35 mm for an average cochlear length, the normal acoustic frequencies corresponding to this range would be between 611 Hz and 3.82 kHz.

In Experiment 1, the speech was processed with an analysis frequency range of 611 Hz–3.82 kHz in each condition. A spectral shift was created by activating a different set of six electrodes in each condition, located at different distances from the round window, as shown schematically at the top of Fig. 1 and summarized in Table III. As a result, from condition 1 to 7 the cochlear location of the array of six electrodes was shifted basally by 2.4 mm per condition.

Experiment 2 simulated the effects of variation in individual cochlear length on the frequency-place mapping, by changing the assumed cochlear length used in the Greenwood equation from 31 to 39 mm in 1 mm steps. The set of six middle electrodes (4–9), which gave the peak performance in Experiment 1, were activated in all conditions. The assumed location of the most apical electrode was fixed at a constant insertion angle ($\sim 470^\circ$), equivalent to 24 mm insertion for a cochlea of 35 mm. As a result, the acoustic frequency corresponding to the apical electrode location was constant for all experimental conditions, since the Greenwood function scales the frequency-to-place map with cochlear length. As the assumed cochlear length increased the proportion of the cochlea occupied by the electrode array decreased, resulting in smaller analysis frequency range for longer cochleae, as shown in the top portion of Fig. 2. The conditions are summarized in Table IV.

In Experiment 3, frequency-place compression conditions were generated with the same set of six electrodes as the previous experiment (4–9). The acoustic input range in mm was made wider than the stimulation range (assumed to be from 12 to 24 mm from the round window) at each end by +1 mm, +2 mm, +3 mm, and +4 mm. The frequency range of the acoustic input was calculated by converting the range in mm with the Greenwood equation. Similarly, for expansion conditions the acoustic input range was made narrower than the stimulation range at each end by –1 mm, –2 mm,

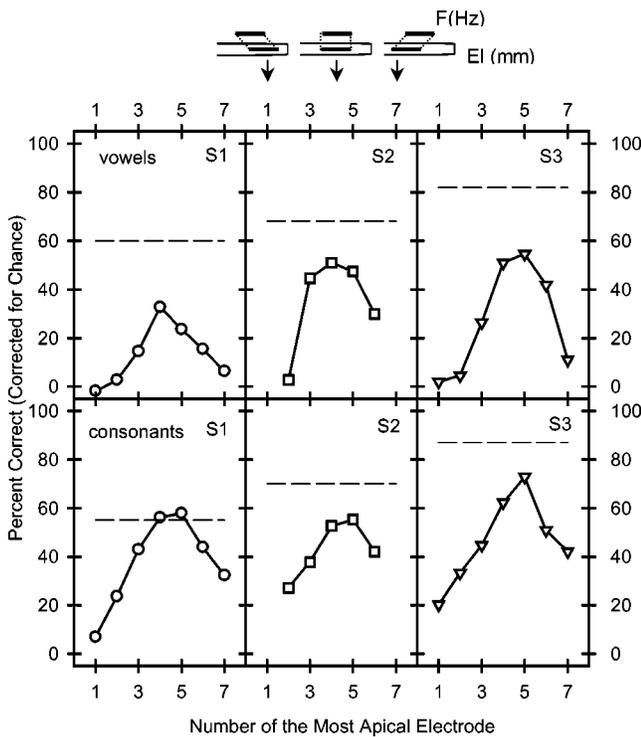


FIG. 1. Individual percent correct scores of subjects S1, S2, and S3 from Experiment 1 as a function of the shifted electrode array position. The top and bottom rows show the vowel and consonant recognition scores, corrected for chance, respectively. The dashed lines show the scores with the processor map that the patient uses in daily life. The experimental maps are schematically shown above the figure and linked to corresponding conditions in the figure with arrows. The open-ended tube represents the cochlea where the open end shows the base and the closed end shows the apex. The line in the cochlea shows the assumed position of the electrode array and the line above the cochlea shows the estimated location of the acoustic input range calculated from Greenwood's (1990) function. Different sets of electrodes from 1–6 to 7–12 were activated while the same center frequency range of 611–3.82 kHz was assigned to electrodes in every condition.

–3 mm, and –4 mm. This manipulation also resulted in narrower frequency bands that were assigned to each electrode. The compression and expansion conditions are schematically shown on top of Fig. 3 and more details can be found in Başkent and Shannon (2003), where similar conditions were simulated. The corresponding frequencies for these conditions are given in Table V.

Based on an assumed insertion depth of 31 mm, the whole array of 12 electrodes lies between 5 and 31 mm from the round window. The widest range of frequencies that Tempo+ can process is 200 Hz–8.5 kHz. When this range is translated into cochlear distance with the Greenwood equation, the corresponding range in mm is narrower than the stimulation range of 26.4 mm, the overall length of 12 elec-

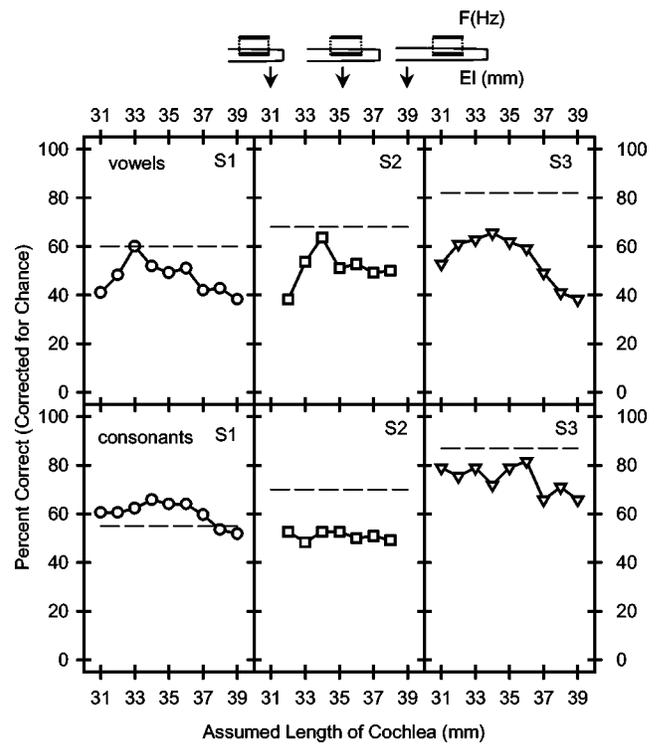


FIG. 2. Individual percent correct scores of subjects S1, S2, and S3 from Experiment 2 as a function of assumed cochlear length. The assumed cochlear length was used in the Greenwood equation to calculate the frequency range matching the stimulation range. The top and bottom rows show the vowel and consonant recognition scores, corrected for chance, respectively. Similar to Fig. 1 dashed lines show patient's performance with the processor map and the conditions are schematized above the figure. In all conditions electrodes 4–9 were activated with an array length of 12 mm and the most apical electrode was fixed at the same insertion angle of 470°. As a result, even though the array length and the insertion depth are the same, the proportion of the cochlear range covered by the array becomes wider or narrower relative to the cochlear length.

trodes. Thus, the normal map used in the implant processors may result in frequency-place expansion. Experiment 4 produced similar expansion conditions to Experiment 3. The analysis range was made narrower than the stimulation range by –3 mm, –4 mm, –5 mm, –6 mm, –7 mm, and –8 mm on each end. The experimental conditions are summarized in Table VI.

III. RESULTS

A. Experiment 1: Shifted electrode array

Subjects S1, S2, and S3 participated in Experiment 1. The individual percent correct scores from vowel and consonant recognition tests as a function of the shift condition are

TABLE III. Basal shift conditions for Experiment 1, shown as a function of the most apical electrode number in the array of six electrodes activated.

Shift condition (most apical electrode)	1	2	3	4	5	6	7
Activated electrodes in the electrode array	1–6	2–7	3–8	4–9	5–10	6–11	7–12
Center frequency range (Hz)	611–3.82 k						

TABLE IV. Center frequency ranges of the analysis bands in Experiment 2, calculated with Greenwood equation using the assumed cochlear length. Electrodes 4–9 were activated in all conditions. The most apical electrode of the array (4) was at the same insertion angle (470°) for all conditions.

Assumed cochlear length (mm)	31	32	33	34	35	36	37	38	39
Modified electrode array location (mm)	9.3–21.3	9.9–21.9	10.6–22.6	11.3–23.3	12.0–24.0	12.7–24.7	13.4–25.4	14.1–26.1	14.7–26.7
Center frequency range (Hz)	611–4.77 k	611–4.49 k	611–4.24 k	611–4.02 k	611–3.82 k	611–3.64 k	611–3.48 k	611–3.35 k	611–3.20 k

shown in Fig. 1. The scores are corrected for chance. Vowel recognition scores are presented in the top row while the consonant recognition scores are presented in the bottom row. Different symbols show scores from different subjects. The same symbol is used for the same subject in all following figures to facilitate comparison of the results across experiments. The dashed lines show the performance of subjects listening to the same stimuli with a map they use in daily life. Generally, experimental conditions resulted in lower performance levels compared to the processor map. This difference may be due to variations in experience as well as to the fact that the experimental maps only used six electrodes, and had much narrower stimulation and acoustic input ranges.

Figure 1 shows that subjects had peak performance around conditions 4 and 5, with electrodes 4–9 and 5–10 activated, respectively. Electrodes 4–9 were selected to be

used for all subjects in Experiments 2 and 3.

B. Experiment 2: Effect of assumed individual cochlear length

Figure 2 shows vowel and consonant recognition scores, corrected for chance, for the same subjects as in Experiment 1 (S1, S2, and S3) as a function of varying assumed values for cochlear length. The peaks in vowel recognition scores of S1 and S3 suggest that these subjects might have cochleae that are only 33–34 mm long. However, generally there was only a small effect on vowel recognition over a wide range of assumed cochlear lengths, and an even smaller effect on consonants. These results show that an inaccuracy in the estimate for cochlear length does not change the results significantly. Therefore an assumed length of 35 mm, which has typically been used in Greenwood mapping calculations, was selected for use in Experiments 3 and 4.

TABLE V. Frequency-place mismatch conditions of Experiment 3. Six electrodes (4–9, assumed to be located from 12 to 24 mm from the round window) were activated. For each condition the table lists the acoustic input range in cochlear distance, center frequencies of bandpass filters, and the overall frequency range of analysis bands in Hz.

Frequency-place mismatch condition	Range of acoustic input (mm)	Band-pass filter center frequencies for six channels (Hz)						Overall frequency range of analysis bands (Hz)
–4 mm (expansion)	20–16	1168	1322	1493	1684	1899	2137	1025–2367
–3 mm (expansion)	21–15	998	1200	1443	1735	2075	2475	887–2762
–2 mm (expansion)	22–14	850	1096	1404	1788	2266	2863	752–3196
–1 mm (expansion)	23–13	721	997	1361	1841	2475	3309	611–3847
0 mm (matching)	24–12	611	906	1320	1896	2700	3820	493–4522
+1 mm (compression)	25–11	509	821	1279	1953	2945	4407	394–5467
+2 mm (compression)	26–10	423	743	1239	2011	3212	5082	314–6305
+3 mm (compression)	27–9	348	670	1200	2070	3502	5855	247–7482
+4 mm (compression)	28–8	281	604	1162	2131	3816	6744	207–8082

TABLE VI. Frequency-place mismatch conditions of Experiment 4. All 12 electrodes (1–12, assumed to be located from 5 to 31 mm from the round window) were activated. For each condition the table lists the acoustic input range in cochlear distance, center frequencies of bandpass filters, and the overall frequency range of analysis bands in Hz.

Frequency-place mismatch condition	Range of acoustic input (mm)	Band-pass filter center frequency range (Hz)	Overall frequency range of analysis bands (Hz)
–8 mm (expansion)	23.2–12.8	699–3407	645–3647
–7 mm (expansion)	24.2–11.8	590–3933	535–4264
–6 mm (expansion)	25.2–10.8	495–4538	440–4981
–5 mm (expansion)	26.2–9.8	412–5231	358–5815
–4 mm (expansion)	27.2–8.8	340–6028	287–6784
–3 mm (expansion)	28.2–7.8	278–6943	227–7911

C. Experiment 3: Frequency-place compression and expansion with six electrodes

Figure 3 shows the individual percent correct scores of all subjects as a function of compression and expansion with six middle electrodes (4–9). Vowel and consonant recognition scores, corrected for chance, are plotted in the top and middle rows, respectively. S1 was tested twice with vowels and consonants. Both curves are presented here (shown with open circles in the left top and left middle panels) to demonstrate the test–retest reliability. The scores of S5 and S6,

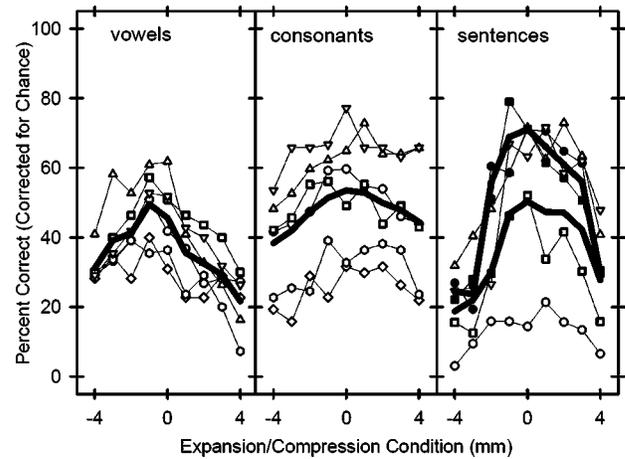


FIG. 4. Average percent correct scores of all subjects, shown with thick lines, superimposed on the individual scores from Fig. 3, shown with open symbols. In sentence recognition scores, the upper thick line shows the average scores of S1 and S2 with HINT sentences, and the lower thick line shows the average scores of subjects S1, S2, S3, and S4 with IEEE sentences.

who were both prelingually deaf, were very low, almost at the level of a single-channel processor performance for consonants. Percent correct scores for sentences are shown in the bottom row. S1 and S2 were tested with IEEE sentences (open symbols) as well as HINT sentences (filled symbols). Even with simpler sentences (HINT) S5 and S6 could not achieve significant sentence recognition.

The average scores from all subjects (thick lines) are presented in Fig. 4 superimposed on the individual scores

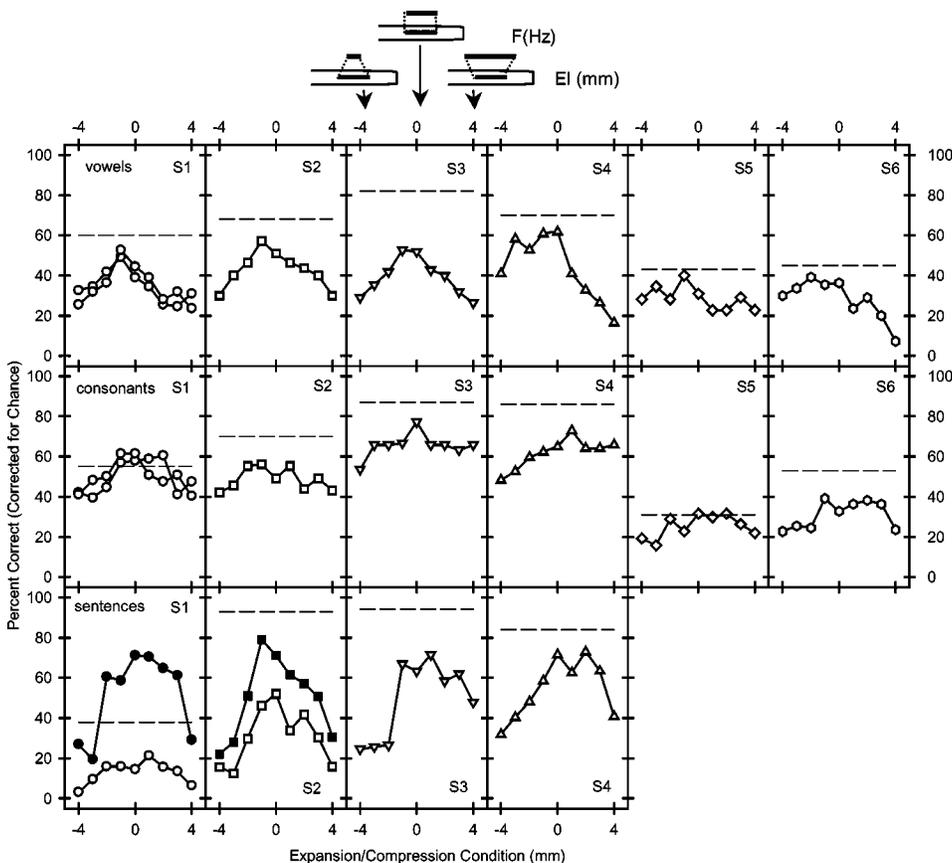


FIG. 3. Individual percent correct scores of all subjects from Experiment 3 as a function of compression and expansion in frequency-place mapping. The top and middle rows show the vowel and consonant recognition scores, corrected for chance, respectively, and the bottom row shows the sentence recognition scores. S1 was tested twice with vowels and consonants to show the test reliability. S1 and S2 were tested with both IEEE (open symbols) and HINT sentences (filled symbols). Similar to Fig. 1 dashed lines show patient's performance with the processor map and the experimental maps are schematically shown above the figure. In all conditions electrodes 4–9 were activated with an array length of 12 mm while the frequency range assigned to the electrodes was made narrower or wider than the stimulation range of the electrode array.

TABLE VII. F and p values calculated with one-way repeated-measures ANOVA for expansion and compression conditions of Experiment 3. Due to the small number of subjects HINT sentences were not included.

Expansion	F	p	Compression	F	p
Vowels, $n=6$ $F(4,20)$	13.65	<0.001	Vowels $n=6$ $F(4,20)$	13.41	<0.001
Consonants $n=6$ $F(4,20)$	15.89	<0.001	Consonants $n=6$ $F(4,20)$	5.11	<0.01
IEEE sentences $n=4$ $F(4,12)$	10.63	<0.001	IEEE sentences $n=4$ $F(4,12)$	7.61	<0.01

(symbols connected with thin lines). The left and middle panels show the average scores from vowel and consonant recognition tests, respectively. The lower thick line in the right panel shows the average score from subjects S1, S2, S3, and S4 with IEEE sentences, and the upper thick line shows the average score from subjects S1 and S2 with HINT sentences.

A one-way repeated-measures ANOVA showed that there was a significant effect of compression and expansion on performance for all stimuli (see Table VII for corresponding F and p values). In general, performance was best for the matched condition (0 mm) and poorer for both frequency-place expansion and compression. For vowels, a post hoc Tukey test revealed that -3 mm, -2 mm, and -1 mm expansion scores were not significantly different than the 0 mm condition (where input frequency theoretically matched the stimulation range). In simulations, frequency-place expansion produced a larger effect on vowel recognition than compression (Başkent and Shannon, 2003), whereas with implant users a larger performance drop was observed with compression than with expansion. As in the simulations, the effects of both expansion and compression were smaller on consonants, which are generally more robust to spectral distortions compared to vowels (Shannon *et al.*, 1998; Friesen *et al.*, 2001). A Tukey test showed no significant difference in results from -2 mm expansion to $+3$ mm compression with consonants, which is a much wider range than observed for vowels. Similar to simulations, consonant recognition scores of CI subjects dropped significantly only with high degrees of compression ($+4$ mm) and expansion (-3 mm). With IEEE sentences, the best performance was obtained around 0 mm matching condition with a tolerance of a few mm of

compression; a Tukey test showed that scores from -1 mm expansion to $+3$ mm compression were not significantly different. The performance dropped significantly with further mismatch. HINT sentences display a sharper peak around the 0 mm matched condition with a larger drop compared to IEEE sentences with increasing mismatch. The number of subjects who listened to HINT sentences (two) was not sufficient to run a statistical test.

1. Similarity of the experimental map to implant processor map

One key question in the study is whether each subject's "reference map" was determined by the normal acoustic tonotopic map or the map implemented in the clinical speech processor. Postlingually deafened patients had extensive experience with normal acoustic mapping prior to deafness, whereas prelingually deafened patients had little to no experience with the normal acoustic map, so their reference map might be determined by the implant map of the everyday processor. Başkent (2003, Fig. 4.20) showed that the reduction of speech recognition under conditions of compression and expansion could be modeled by the sum of squared differences in band center frequencies between the acoustic and experimental maps. This result was observed in implant simulations with normal-hearing listeners, and might also apply to implant users, especially if postlingually deafened.

To assess the potential influence of the normal acoustic tonotopic map and the implant processor map on the results of Experiment 3 a similarity metric was calculated:

$$\text{similarity_factor} = \frac{1}{\text{error_rms}}, \quad (1)$$

with error_rms defined as

$$\text{error_rms} = 10 \sqrt{\frac{\sum_{i=1}^N \log_{10}^2(f_{c_exp}(i)/f_c(i))}{N}}, \quad (2)$$

where N =number of electrodes, f_{c_exp} =center frequencies of the experimental map, and f_c =center frequencies of the comparison map (i.e., either the normal acoustic map or the implant processor map). The similarity factor quantifies the similarity between two maps by comparing center frequencies of the analysis bands between experimental processors and either the normal tonotopic map or the implant processor map. The value of the index ranges from one to an asymptotic zero: When the maps that are compared are identical the index is unity and as the maps differ the index decreases. The values of similarity factors for the experimen-

TABLE VIII. Similarity of experimental conditions to normal acoustic map (0 mm, tonotopically matched map) and implant processor maps (shown in Table II for each subject).

Compression/ expansion conditions					normal					
	exp -4 mm	exp -3 mm	exp -2 mm	exp -1 mm	map (0 mm)	comp +1 mm	comp +2 mm	comp +3 mm	comp +4 mm	
Similarity to normal acoustic map					1.00	0.90	0.80	0.72	0.64	
Similarity to implant processor map	S1, S2, S5	0.75	0.79	0.81	0.79	0.75	0.69	0.63	0.57	0.51
	S3	0.70	0.75	0.78	0.78	0.75	0.69	0.64	0.57	0.51
	S4	0.75	0.82	0.86	0.86	0.81	0.74	0.66	0.59	0.52
	S6	0.78	0.86	0.94	0.93	0.85	0.77	0.69	0.63	0.57

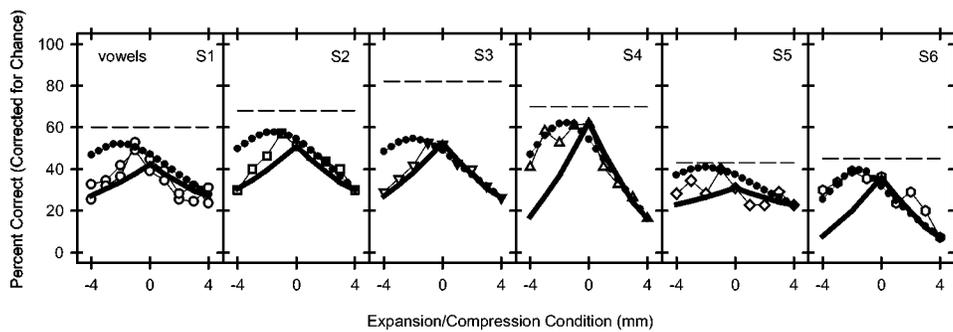


FIG. 5. Vowel recognition percent scores from Experiment 3, reproduced from Fig. 3. The solid line shows the performance predicted from the similarity of the experimental map to patient's implant processor map while the dotted line shows the prediction from the similarity of the experimental map to the normal acoustic map.

tal maps are given in Table VIII when compared to normal acoustic and implant processor maps.

Figure 5 duplicates vowel recognition scores from Fig. 3 with two similarity indices added. The similarity index function for the implant processor map (dotted line) was scaled to fit the data at the lowest and highest values. The similarity function for the acoustic tonotopic map (thick solid line) was scaled to fit the lowest point and the value at 0 mm (matching map). Figure 5 shows that the vowel recognition results of the postlingually deafened subjects S1, S2, and S3 were similar to the pattern predicted by the similarity to the normal acoustic map. The pattern of performance of the prelingually deafened subjects S4, S5, and S6 was similar to that predicted by the similarity to their implant processor map. Postlingually deaf subjects' speech recognition appears to be sensitive to the spectral mismatch relative to the normal acoustic map, similar to NH listeners. Prelingually deafened subjects, however, might not have had sufficient acoustic input during the critical period that normally establishes the acoustic tonotopic map, and so their performance appears to be determined more by similarity to the processor map.

D. Experiment 4: Expansion with all 12 electrodes

Experiments 1–3 used six electrodes covering 12 mm in the cochlea. The subjects' experience with their implant, which was as long as 5 years, was with the full stimulation range of the default 12-electrode array (26.4 mm). To confirm that similar effects occur with the entire array, we repeated the expansion conditions using all 12 electrodes of the device.

S1, S2, and S4 participated in this experiment. Figure 6 shows individual (small open symbols) and average scores (thick line) with the 12-electrode processor expansion conditions combined with average scores of the same subjects with six-electrode processor from Experiment 3 (dotted line).

The 12-electrode processor has better spectral resolution, covers a much wider stimulation range and acoustic input range, and employs additional apical electrodes, compared to the six-electrode processor. For example, for the same -3 mm expansion condition, the 12-electrode processor has a stimulation range of 26.4 mm and an analysis range of 20.4 mm, while the six-electrode processor has a stimulation range of 12 mm and an analysis range of 6 mm. As a result scores were higher with the 12-electrode processor compared to the six-electrode processor for the same expansion conditions. Yet, similar to the six-electrode processor, performance decreased as the expansion increased. The per-

formance around -3 mm expansion resulted in performance similar to that of the clinical processor. It is not clear, however, if higher performance levels could be achieved with a better matched condition because device limitations did not allow such mapping.

Figure 7 shows a comparison of the similarity metric functions to the individual vowel recognition scores of the subjects with 12-channel processor (open symbols). As a reference, vowel recognition scores with the six-channel processor are also included (filled symbols). Similar to Fig. 5, the prediction referenced to the similarity to the normal acoustic map is shown by the thick solid line and the prediction by the similarity to the subjects' processor map is shown by the dotted line. The dashed lines show subject performance with the everyday processor map. The dotted lines imply that the -4 mm and -5 mm expansion conditions with 12 electrodes are most similar to the mapping used in subjects' implant processors, except that the analysis bands of the experimental maps were partitioned in equal cochlear distances instead of logarithmic steps. Similar to Experiment 3, the performance by S1 and S2, who were postlingually deafened, followed the prediction based on similarity to the

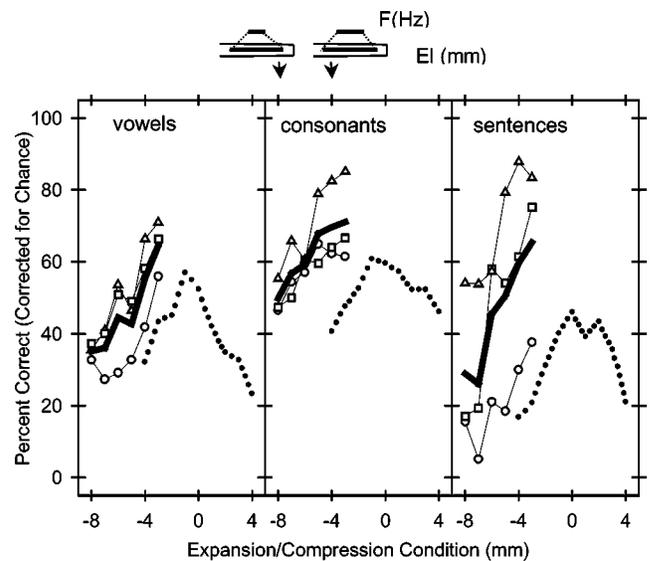


FIG. 6. Percent correct scores of subjects S1, S2, and S4 from Experiment 4. The open symbols show the individual scores for 12-electrode expansion conditions. The superimposed thick lines are the average scores of subjects with 12 electrodes. The dotted lines are the average scores of the same subjects from expansion and compression conditions with six middle electrodes, taken from Experiment 3. The expansion conditions are schematically shown above the figure.

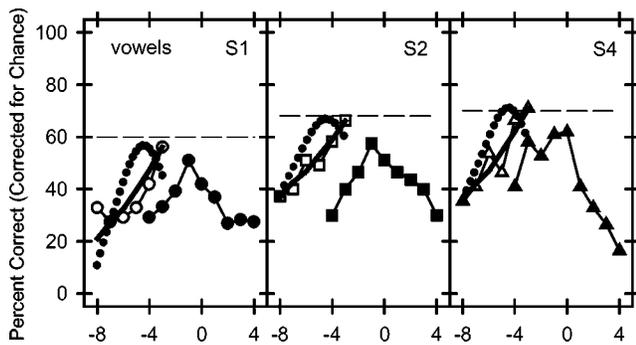


FIG. 7. Individual vowel recognition percent scores with 12-electrode processor from Experiment 4, shown by open symbols, and scores with six-electrode processor from Experiment 3, shown by filled symbols. The solid line shows the performance predicted from the similarity of the experimental map to patient's implant processor map while the dotted line shows the prediction from the similarity of the experimental map to the normal acoustic map. The dashed lines show the performance of subjects with the implant processor map.

normal acoustic map. Prelingually deafened subject S4 also showed a pattern similar to that predicted by the normal acoustic map, although the two predicted patterns were not sufficiently different to make a clear distinction, given the variability in the data in this case.

IV. DISCUSSION AND CONCLUSIONS

Speech recognition in cochlear implant listeners was significantly affected by alterations in the frequency-to-electrode mapping. Previous work in both acoustic simulations and in implant listeners had shown that speech recognition decreases with spectral distortions in the mapping, including apical-basal shift (Fu and Shannon, 1999; Dorman *et al.*, 1997), nonlinear warping (Shannon *et al.*, 1998), or compression–expansion (Başkent and Shannon, 2003). The present study expands the general pattern of frequency-place mismatch results to include frequency-place compression and expansion in cochlear implant listeners. The results have theoretical value as they quantify how speech pattern recognition is affected by alterations in the cochlear representation of speech. In addition, the present results have practical value, as they show the inherent tradeoffs between electrode array insertion depth, number of active electrodes, and input frequency range, to provide basic guidelines for optimal fitting of implant patients.

The implant subjects displayed a similar pattern of results despite the large variation in their speech recognition skills. Consistent with the acoustic simulations by Başkent and Shannon (2003) best speech recognition was obtained with frequency-place maps with the least spectral distortion. Both compression and expansion reduced recognition, especially with vowels, which are more sensitive to spectral manipulations. Yet, there was a significant difference between NH and implant subjects in their exposure to frequency-place maps, and the implant results showed two distinct patterns that might be determined by the individual subject's reference map. A simple model of frequency-place distortion fit the data from NH subjects (Başkent, 2003) and postlingually deafened subjects by weighting frequency-place mismatch relative to the normal acoustic tonotopic map. The

same model fit the data from prelingually deafened subjects based on the similarity of the experimental map to the map in their individual processor. In Experiment 4, where the expansion conditions were applied with all 12 electrodes, the performance by the postlingually deafened subjects was again closer to the prediction by the similarity of the experimental condition to the normal acoustic map.

The changes in performance observed in the present study are acute effects observed immediately after the subject was given a new map, without any time to adapt to the new processor. It is not clear how much implant patients would be able to adapt to such compressed or expanded maps over time. The results of this study show that the choice of the frequency-place map has a significant effect on speech perception and choosing a better fitting map might instantly increase the performance of an implant patient. If a patient is initially given the best-fit map, any further adaptation could then start from this high performance level. Complementary to the findings of the present study with fully inserted arrays, Başkent and Shannon (2004) showed that it is particularly important to fit patients with partial insertions, who generally do not perform as well as patients with full insertions, with an optimum map. Although studies have shown improved performance with experience over the first few months of implant use (Tyler *et al.*, 1997; Fu *et al.*, 2002; and Fu and Galvin, 2003) it has not been demonstrated that such learning has a differential effect for different processor parameters. It seems likely that experience will increase performance for any processor setting, so it may be important to start with the setting that produces the best speech recognition to optimize long-term as well as short term outcomes.

One potential difficulty in such experiments with cochlear implants is the uncertainty inherent in several key parameters such as cochlear length, electrode array insertion depth, and its lateral distance from modiolus, and the best frequencies of the nerves actually stimulated by each electrode. The experiments in the present study demonstrated that, even though the physical values of these key parameters cannot be known with certainty, optimal values for a frequency-to-electrode map can be functionally estimated starting with approximate initial values. Phoneme recognition in Experiments 1–4 was always a simple function of the underlying manipulation, showing peak performance at a certain parameter value and a drop in performance as the value of that parameter was increased or decreased. Vowels and sentences were the speech materials most sensitive to the manipulations. Given the simplicity of these functions, a clinical procedure could be developed to rapidly converge on the optimal set of parameters controlling the frequency-to-electrode mapping for an individual patient. A subset of vowels could be selected that are most sensitive to frequency-place distortion. A simple optimizing algorithm could be developed to converge on the frequency-place mapping that maximizes vowel recognition for that reduced set in each individual patient, without the costs and risks of x rays and CT scans.

ACKNOWLEDGMENTS

The authors would like to thank Med-El Corp., especially Amy Barco, for providing the equipment and software, help recruiting subjects and their travel expenses, Peter Nopp, for valuable suggestions, Rachel Cruz, Dawna Mills, Pam Fiebig, and Michelle Colburn, for help recruiting subjects, and the subjects for their valuable efforts. Also the comments by Peter Assmann, two anonymous reviewers, and Anastasios Sarampalis are greatly appreciated. Funding for the study was provided in part by NIDCD Grant No. R01-DC-01526 and Contract No. N01-DC-92100.

- Başkent, D. (2003). "Speech recognition under conditions of frequency-place compression and expansion," Ph.D thesis, University of Southern California, CA.
- Başkent, D., and Shannon, R. V. (2003). "Speech recognition under conditions of frequency-place compression and expansion," *J. Acoust. Soc. Am.* **113**, 2064–2076.
- Başkent, D., and Shannon, R. V. (2004). "Interactions between cochlear implant electrode insertion depth and frequency-place mapping," *J. Acoust. Soc. Am.* (accepted for publication).
- Cohen, L. T., Xu, J., Xu, S. A., and Clark, G. M. (1996). "Improved and simplified methods for specifying positions of the electrode bands of a cochlear implant array," *Am. J. Otol.* **17**, 859–865.
- Dorman, M. F., Loizou, P. C., and Rainey, D. (1997). "Simulating the effect of cochlear-implant electrode insertion depth on speech understanding," *J. Acoust. Soc. Am.* **102**, 2993–2996.
- Friesen, L., Shannon, R. V., Başkent, D., and Wang, X. (2001). "Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants," *J. Acoust. Soc. Am.* **110**, 1150–1163.
- Fu, Q.-J., and Shannon, R. V. (1999). "Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing," *J. Acoust. Soc. Am.* **105**, 1889–1900.
- Fu, Q.-J., Shannon, R. V., and Galvin, J. (2002). "Perceptual learning following changes in the frequency-to-electrode assignment with the Nucleus-22 cochlear implant," *J. Acoust. Soc. Am.* **112**, 1664–1674.
- Fu, Q.-J., and Galvin, J. (2003). "The effects of short-term training for spectrally mismatched noise-band speech," *J. Acoust. Soc. Am.* **113**, 1065–1072.
- Greenwood, D. D. (1990). "A cochlear frequency-position function for several species: 29 years later," *J. Acoust. Soc. Am.* **87**, 2592–2605.
- Hillenbrand, J., Getty, L., Clark, M., and Wheeler, K. (1995). "Acoustic characteristics of American English vowels," *J. Acoust. Soc. Am.* **97**, 3099–3111.
- Institute of Electrical and Electronics Engineers (1969). IEEE recommended practice for speech quality measurements.
- Ketten, D. R., Skinner, M. W., Wang, G., Vannier, M. W., Gates, G. A., and Neely, J. G. (1998). "*In vivo* measures of cochlear length and insertion depth of Nucleus cochlear implant electrode arrays," *Ann. Otol. Rhinol. Laryngol. Suppl.* **175**, 1–16.
- Marsh, M. A., Xu, J., Blamey, P. J., Whitford, L. A., Xu, S. A., Silverman, J. M., and Clark, G. M. (1993). "Radiologic evaluation of multichannel intracochlear implant insertion depth," *Am. J. Otol.* **14**, 386–391.
- Nilsson, M., Soli, S., and Sullivan, J. (1994). "Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and noise," *J. Acoust. Soc. Am.* **95**, 1085–1099.
- Robert, M. E. (1998). CONDOR: Documentation for Identification Test Program. Los Angeles. House Ear Institute, CA.
- Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). "Speech recognition with primarily temporal cues," *Science* **270**, 303–304.
- Shannon, R. V., Zeng, F.-G., and Wygonski, J. (1998). "Speech recognition with altered spectral distribution of envelope cues," *J. Acoust. Soc. Am.* **104**, 2467–2476.
- Shannon, R. V., Jensvold, A., Padilla, M., Robert, M. E., and Wang, X. (1999). "Consonant recordings for speech testing," *J. Acoust. Soc. Am.* **106**, L71–L74.
- Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., Neely, J. G., Kletzkner, G. R., Brunnsden, B., and Blocker, B. (2003). "CT-Derived estimation of cochlear morphology and electrode array position in relation to word recognition in Nucleus 22 recipients," *J. Assoc. Res. Otolaryngol.* **3**, 332–350.
- Tyler, R. S., Parkinson, A. J., Woodworth, G. G., Lowder, M. W., and Gantz, B. J. (1997). "Performance over time of adult patients using the Ineraid or nucleus cochlear implant," *J. Acoust. Soc. Am.* **102**, 508–522.
- Ulehlova, L., Voldrich, L., and Janisch, R. (1987). "Correlative study of sensory cell density and cochlear length in humans," *Hear. Res.* **28**, 149–151.
- Van Tasell, D. J., Soli, S. D., Kirby, V. M., and Widin, G. P. (1987). "Speech waveform envelope cues for consonant recognition," *J. Acoust. Soc. Am.* **82**, 1152–1161.
- Whitford, L. A., Seligman, P. M., Blamey, P. J., McDermott, H. J., and Patrick, J. F. (1993). "Comparison of current speech coding strategies," *Adv. Oto-Rhino-Laryngol.* **48**, 85–90.