

Effects of Stimulus Level on Speech Perception with Cochlear Protheses

KEVIN H. FRANCK,^{1,2} LI XU,^{1,3} AND BRYAN E. PFINGST¹

¹*Kresge Hearing Research Institute, Department of Otolaryngology, University of Michigan Health System, Ann Arbor, MI 48109-0506, USA*

²*Center for Childhood Communication, The Children's Hospital of Philadelphia, Philadelphia, PA 19104, USA*

³*School of Hearing, Speech and Language Sciences, Ohio University, Athens, OH 45701, USA*

Received: 16 October 2001; Accepted: 23 May 2002; Online publication: 16 July 2002

ABSTRACT

This study is one of a series that examines stimulus features important for cochlear implant function. Here, we examine effects of stimulus level. In subjects with cochlear implants, a number of psychophysical tests of temporal discrimination (pulse rate discrimination, gap detection, etc.) show marked improvement as a function of stimulus level through most or all of the dynamic range, while electrode-place discrimination can improve or degrade as a function of level. In this study, effects of these combined potential influences were studied by examining the effects of stimulus level on syllable identification. We tested two hypotheses: that syllable identification varies as a function of stimulus level and that level and electrode configuration interact in affecting syllable identification. We examined vowel and consonant identification as a function of stimulus level for bipolar and monopolar electrode configurations. We used experimental processor maps where upper and lower stimulation limits of each electrode pair were equated to eliminate confounding effects of dynamic range, which varies across subjects and electrodes. For each channel, stimulation amplitude was set to a fixed percentage of its dynamic range. Eight adult subjects with Nucleus CI24M implants were tested using the SPEAK processing strategy. With each electrode configuration, stimulus levels were tested from 0% to 90% of the dynamic range in nine steps. The effects

on consonant and vowel identification were similar. Phoneme identification was usually better for monopolar than for bipolar stimulation. In the lower half of the dynamic range, syllable identification usually increased as a function of stimulus level. In the upper half of the dynamic range, syllable identification continued to increase as a function of level to 90% of the dynamic range for some subjects, while for others there was no appreciable change or a decrease as a function of level. Decreases in performance at high levels were more common with monopolar than bipolar stimulation. These results suggest that if speech processors are programmed to optimize level for each individual, speech perception performance could be improved.

Keywords: auditory prosthesis, cochlear implant, stimulus level, electrode configuration, speech perception, human

INTRODUCTION

Speech perception performance varies considerably among cochlear implant users. The variability is assumed to be caused by many factors, ranging from variables associated with the design of the prosthesis, to those related to the individual using the device. It is crucial to understand the source of speech perception variability so that controllable variables can be adjusted to improve performance with the device. The purpose of the present study was to examine the

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

effects of one such variable, electrical stimulus level, on speech perception in implanted human subjects.

There is ample evidence that stimulus level affects performance with cochlear implants. A number of psychophysical difference limens are known to decrease (improve) as electrical stimulus level increases. This is particularly true for temporally based discrimination such as pulse-rate discrimination (Pfungst et al. 1994), frequency discrimination (Pfungst and Rai 1990), modulation-frequency discrimination (Morris and Pfingst 2000), modulation-frequency detection (Shannon 1992, 1993), and gap detection (Shannon 1989). Level difference limens in dB also improve as a function of level, being smaller at higher levels (Pfungst et al. 1983; Shannon 1983). Electrode discrimination improves as a function of level in some subjects at some electrode locations, but it deteriorates as a function of level in others and, in a third group, level has nonmonotonic effects (McKay et al. 1999; Pfingst et al. 1999).

The mechanism by which electrical stimulation level affects perception might be related to the spatial extent of neural excitation. Increasing the level of electrical stimulation causes larger activating-potential fields and thus leads to an increase in the size of the stimulated neural population. This has been shown with neurophysiological recordings at all levels of the auditory pathway from auditory nerve (Kral et al. 1998) through inferior colliculus (Rebscher et al. 2001; Snyder et al. 1990) to auditory cortex (Bierer and Middlebrooks 2002). With more neurons contributing to the representation of temporal cues at higher electrical stimulus levels, we predict increased psychophysical and speech perception performance. With the same model, one might predict degradation in the specificity of tonotopic stimulation as increased electrical stimulus level results in greater overlap of adjacent populations of stimulated neurons. However, as noted above, discrimination of one stimulation site from another degrades as a function of level only in some cases and it improves in others.

Given the assumption that effects of level are related to the degree of current spread, we suggest that other factors that presumably affect current spread might interact with current level to influence psychophysical performance and speech perception. Monopolar stimulation, where an intracochlear electrode is referenced to a distant electrode, is known to result in larger activating-current fields than does bipolar stimulation, where an intracochlear electrode is referenced to another proximal intracochlear electrode, provided the same current is delivered through each configuration (Spelman et al. 1995). However, it is not known to what degree broader configurations produce broader spread of excitation relative to that produced by narrower configurations at the currents

required to achieve comfortable listening level, because the narrower configurations usually require higher current levels. For psychophysical discrimination of temporal and level information, electrode configuration has relatively smaller effects compared with the effects of stimulus level (Morris and Pfingst 2000). The degree to which electrical stimulus level and electrode configuration interact to influence both speech perception and subjective loudness judgments has not been investigated.

Of course, we must recognize that the effects of stimulus level and electrode configuration might be mediated by mechanisms other than spread of excitation, or by mechanisms that are confounded with spread of excitation, such as the longitudinal position of excitation along the axis of the scala tympani. Evidence that the site of stimulation shifts as a function of stimulus level is that some subjects report a pitch shift when level increases (Townshend et al. 1987). Such shifts in site of stimulation probably also occur when the electrode configuration is broadened and this can affect speech perception (Fu and Shannon 1999; Pfingst et al. 2001).

A consideration when testing effects of stimulus level on speech perception is that the speech signal spans a large part of the subject's electrical dynamic range. The electrical dynamic range is the difference between the amount of current required for the threshold of detection (T level) and the amount of current that results in a comfortable but maximally loud percept (C level). The magnitude of the electrical dynamic range varies across implant users and across stimulation sites in individual implants. One can manipulate the level of the acoustic signal, but automatic gain control and other features of the auditory prosthesis processors confound interpretation of the results. A more controlled procedure is to compress the range of currents within the total dynamic range of electrical hearing so that all components of the signal fall within a restricted range. Programming the processor using artificially elevated T levels and/or artificially reduced C levels can achieve this. Thus, it is possible to reduce the dynamic range of the electrical signal and move the resulting electrical signal up or down within the dynamic range of electrical hearing.

Recent studies have investigated the consequences of reducing the dynamic range of electrical stimulation on speech perception. Zeng and Galvin (1999) showed that reducing dynamic range in cochlear implant users had no significant effect on phoneme identification in quiet. They also found only a small deleterious effect on phoneme identification in noise. Skinner et al. (1999) showed that reducing the dynamic range by increasing the threshold levels resulted in a slight improvement in speech perception, particularly for lower intensity acoustic signals (50 and

TABLE 1

Subject information					
<i>Subject number</i>	<i>Sex</i>	<i>Age (years)</i>	<i>Normal coding strategy</i>	<i>Duration of deafness (years)^a</i>	<i>Duration of device use (years)</i>
1	F	53	SPEAK	11	2.2
2	F	33	SPEAK	3	2.0
3	F	51	SPEAK	13	0.8
4	M	50	ACE	16	1.6
5	F	47	SPEAK	1	1.0
6	F	70	SPEAK	20	0.5
7	M	67	ACE	3	0.6
8	F	49	ACE	29	0.6

^aDefined as the amount of time between patient's subjective report of onset of profound hearing loss and implant activation.

60 dB SPL). Neither of these studies examined the effects of the reciprocal manipulation, namely compressing the signal into the lower region of the dynamic range of electrical hearing. Loizou et al. (2000) showed, using 6-channel acoustic simulations of cochlear implant processors, that reduced dynamic range should adversely affect speech perception abilities of cochlear implant users more for vowels than for consonants.

In a previous study (Pfungst et al. 2001) where subjective judgments of sound quality and syllable identification were quantified, we found that the two measures did not respond identically to variation in stimulus parameters. Thus, we believe that quantitative assessment of subjects' judgments of sound quality provides useful information that can supplement the understanding of implant performance that we gain through objective speech testing. In the present study we assessed the effects of stimulus level both on speech perception and on the subjects' evaluations of the quality of the speech sounds.

METHODS

Subjects

Data were collected from eight adult postlingually deaf subjects who had been implanted and followed clinically at the University of Michigan. All of the subjects used Nucleus CI24M cochlear implant systems supplied by Cochlear Corporation (Englewood, CO). All subjects had at least 6 months' experience with the implant, native use of spoken English, and no nonfunctional intracochlear electrodes. In addition, subjects were required to have open-set speech perception skills and had to be mentally and physically fit for testing. Subjects were monetarily compensated for their 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.

Five subjects used the Spectral Peak (SPEAK) processing strategy (Skinner et al. 1994; Whitford et al. 1995) and three used the ACE processing strategy (Cochlear Corporation 1999) in the devices that they used everyday. Cycling continuously, these coding strategies analyze bandpass-filtered channels of information and present biphasic amplitude-modulated pulses sequentially to a subset of electrodes whose channels contain high-intensity information. The ACE strategy is conceptually the same as the SPEAK strategy but typically cycles faster (resulting in a higher pulse rate) and uses two additional channels of stimulation. Additional subject details are provided in Table 1.

Equipment

The CI24M cochlear implant has an array of 22 intracochlear electrodes that are implanted into the scala tympani through a cochleostomy. The implant also has 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 can be used in three monopolar configurations: MP1, MP2, and MP1 + 2. For MP1 and MP2, the reference path is to one of the two electrodes. For MP1 + 2 (the default configuration in the clinical software), the reference path is to the two electrodes in parallel. The MP1 + 2 configuration was used in this study, as well as by each subject outside the laboratory.

Configuration parameters and stimulation limits were defined using Cochlear Corporation's Diagnostic and Programming System for Windows (Win-DPS) software version R116. The software programmed speech processors using an IF5 ISA card and the Processor Control Interface (PCI). The speech processor controlled transmission of radio frequency (RF) pulses to the internal receiver/stimulator. The internal receiver/stimulator then decoded the RF informa-

tion to the correct stimulation pulse parameters. All testing using experimental maps was conducted using the laboratory's SPrint speech processor (serial number 408594) and microphone. This procedure avoided any chance of confounding effects due to differences in the individual subjects' equipment.

Syllable identification test materials were presented from an IBM personal computer using custom software written in the MATLAB (MathWorks Inc., Natick, MA) programming environment. A Rane ME-60 graphic equalizer (Rane Corp., Mukilteo, WA) was used to compensate for the frequency response of the loudspeaker. A Crown D-75 amplifier (Crown International Inc., Elkhart, IN) was used to increase the signal amplitude. Signals were presented through a TDC 4A loudspeaker (TDC, Randolph, MA) positioned 1 m away from the subject at 0° azimuth inside an Acoustic Systems (Austin, TX, model RE 242S) double-walled sound attenuation booth.

The peak speech test level was calibrated periodically with a sound-level meter (Brüel & Kjær, Naerum, Denmark, type 2231). Subjects were not present during calibration. The sound-level meter microphone was positioned near where the headset microphone would be located during the test sessions. A fast time setting (i.e., time constant of 125 ms) and an "A" frequency weighting were set in the sound-level meter during calibration of the speech materials. A spectrum analyzer (Stanford Research Systems, Sunnyvale, CA, model SR760 FFT) was used for the frequency response calibration to both narrow- and broad-band signals. The graphic equalizer was adjusted to assure compliance beyond American National Standards Institute (ANSI) 3.6 specifications, which describe tolerances for speech signals used in clinical audiometry (ANSI 1996).

Experimental design

The independent variables tested in this experiment were stimulus level and electrode configuration. For the bipolar configuration, both the active and the return electrodes were located in the scala tympani. The active and return electrodes were adjacent, with a center-to-center distance of 0.75 mm. The more basal electrode was labeled as the active electrode. For the monopolar configuration (MP1 + 2), the intracochlear electrode was labeled as active and both the electrode located on the casing of the internal receiver and the electrode placed under the temporalis muscle, in parallel, served as the return. For all cases, the stimulation site was designated by the location of the active electrode.

Nine maps with different stimulus levels (see below), plus a full-dynamic-range map, were created for each of the two electrode configurations, for a total of

20 maps. Other mapping parameters such as the frequency allocation table and number of maxima presented were held constant at software default values (Cochlear Corporation 1999). The dependent variables tested in this experiment were speech syllable identification performance for vowels and consonants and subjective judgments of speech loudness and quality.

Procedures

The first step in the experiment was to determine the subject's T level and C level for each site in the electrode array for each of the two electrode configurations. Procedures were similar to those used in fitting implants clinically. For all measurements and calculations, manufacturer-defined current-level programming units were used. The 256 current-level units range from approximately 10 μA peak to approximately 1750 μA peak, increasing by approximately 2% per unit (Cochlear Corporation 1999). T and C levels were established using the method of adjustment with a control knob that allowed the subject to adjust the current level. The stimuli consisted of 200 μs /phase pulses presented at a rate of 250 pps with a 500 ms on/off duty cycle. First, in the bipolar electrode configuration, T levels were determined from the apical electrodes to the basal electrodes. Then, C levels were 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 determining T levels 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 if needed. 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 was uncomfortably loud and that all presentations were of equal loudness. Adjustments to C levels were made if needed. The values were then saved and the mapping procedure was repeated for the monopolar configuration.

The cochlear implant-fitting software combined the configuration parameters and stimulation levels into a "map." The full dynamic range between the T levels and C levels was completely compressed to a single value (100% dynamic range compressed). The level of this compressed signal was varied between threshold and comfort values in 11 equally spaced steps, omitting 10% because of time constraints and 100% because of comfort constraints (Fig. 1). The

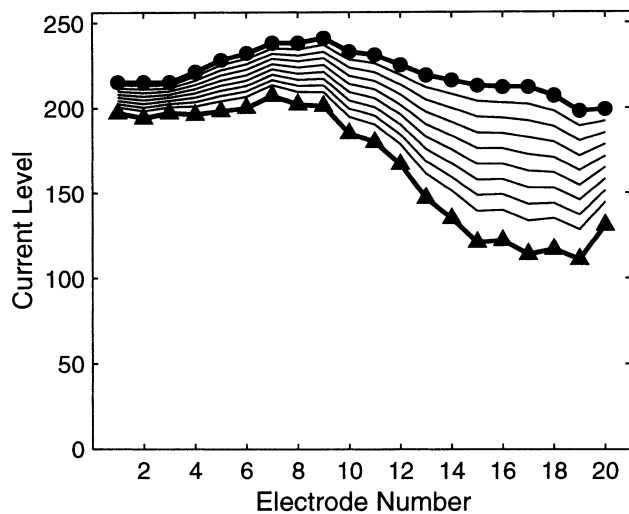


FIG. 1. Dynamic range compression: example Subject 7. This figure illustrates how the programming levels for experimental maps are determined using Subject 7's data. Circles and triangles represent C levels and T levels, respectively, measured in current-level programming units across the electrode array. Thin lines represent the mapping levels set in the compressed-dynamic-range experimental maps at 20%, 30%, 40%, 50%, 60%, 70%, 80%, and 90% of the range between the measured T levels and C levels. The level of 0% dynamic range is identical to the T levels.

sets of electrode stimulus levels associated with each of the nine conditions were not tested for equal loudness. In addition to the nine levels using the compressed dynamic range, a full-dynamic-range condition was used.

For speech perception testing, the Spectral Peak (SPEAK) processing strategy was used (Skinner et al. 1994; Whitford et al. 1995). Vowel and consonant syllable identification tests were administered. The vowel syllable identification test used 12 vowels presented in a /h/-vowel-/d/ context (had, hayed, hawed, head, heard, heed, hid, hod, hoad, hood, hud, who'd) (Hillenbrand et al. 1995). Two male (talkers 48 and 49) and two female (talkers 39 and 44) talkers were used of the multiple talkers available in these test materials. The consonant syllable identification tests used 20 syllables in a consonant-/a/ context (ba, cha, da, fa, ga, ja, ka, la, ma, na, pa, ra, sa, sha, ta, tha, va, wa, ya, za) (Shannon et al. 1999). One male (talker 3) and one female (talker 3) talker were used of the multiple talkers available for this test. Within each run, a stimulus token was chosen randomly (without replacement) from all 48 tokens in the vowel syllable identification test or from all 40 tokens in the consonant identification test. Alphabetic representations of the 12 vowel or the 20 consonant stimuli were presented in a grid on a computer screen and the subject responded by pointing the cursor to the appropriate symbol using a computer mouse and depressing the left mouse but-

ton. No feedback was provided and the subjects were instructed to guess if they were not sure.

The 20 experimental maps (9 levels with the compressed dynamic range plus 1 map with the full dynamic range, times 2 electrode configurations) were tested three times each in random order. For each test, the experimental processor was programmed with the appropriate map, and then one set of consonant and one set of vowel stimuli were presented in succession. Syllables spoken by the talkers were randomized. The speech stimuli were delivered via loudspeaker in the sound field at 64 dB(A). During speech syllable identification testing, the speech processor was set at a constant sensitivity and volume level (sensitivity = 7 and volume = 9), and the subjects were not allowed to adjust the processor.

After speech syllable identification testing for each map, subjects scaled the sound quality of the map by moving a slider bar on the computer screen in between the adjectives "excellent" and "terrible." Similarly, they rated the subjective loudness of the map by moving the slider bar between "too soft" and "too loud." Sound-quality judgments were obtained from only four subjects (Subjects 3, 4, 7, and 8) because this part of the experiment was not implemented until after the first four subjects had completed the study.

Statistical analysis

The effects of stimulus level on performance scores (0–90% of the dynamic range), electrode configuration (monopolar and bipolar), and syllable type (consonant and vowel) were analyzed using a three-way analysis of variance (ANOVA) of the data for each subject. Subject 6 was not tested on eight of the high-level stimuli because they were uncomfortably loud. Two of these missing points were in the vowel test at 80% of the dynamic range (one out of three missing in both the monopolar and bipolar configurations). In order to complete the ANOVA, the average of the two taken measures was used to fill in the missing data points. The remaining six missing data points were from the monopolar stimulation at 90% of the dynamic range condition. Data were filled in with the scores from the 80% of the dynamic range condition.

Group data were analyzed with two additional three-way ANOVAs, one considering the effects of subject, level, and speech test, and another considering the effects of subject, level, and electrode configuration.

RESULTS

Effects of stimulus level

The effects of stimulus level on vowel and consonant identification differed across subjects, electrode configuration, and test materials (Fig. 2). In some cases,

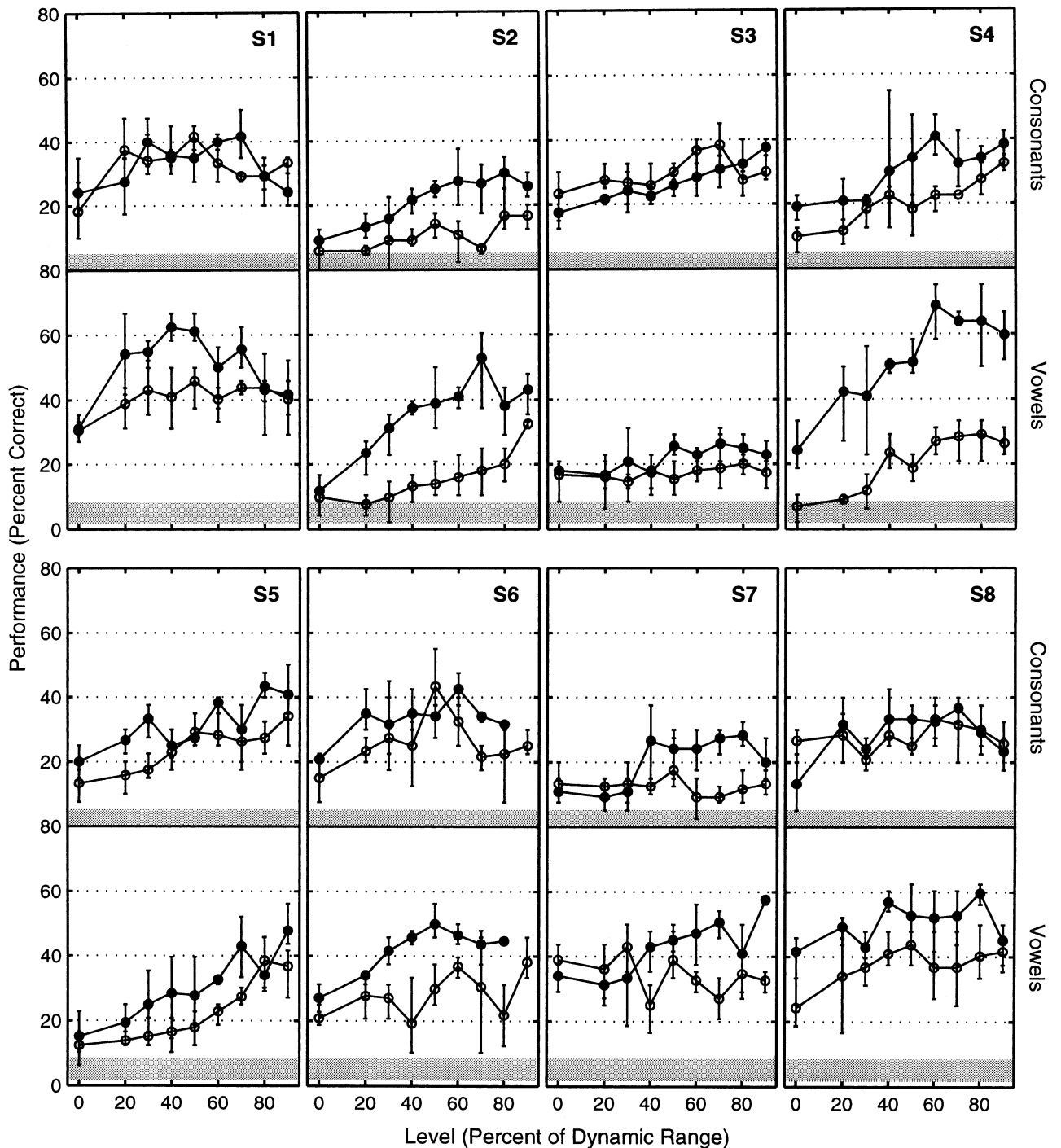


FIG. 2. Syllable identification as a function of level. Syllable identification data as a function of level are plotted for each subject in a separate panel. Subject number is shown in the upper right of each panel. Consonant and vowel data are shown in the upper and lower plot for each subject, respectively. Data for the bipolar and monopolar electrode configurations are plotted with open and filled symbols, respectively. Error bars represent the range of the results of the three speech tests. The gray areas indicate chance performance

level (i.e., 5% and 8.33% for consonant and vowel tests, respectively). For Subject 6, eight speech tests of the high-level stimuli were not tested because they were uncomfortably loud. Two of these missing points were in the vowel test at 80% of the dynamic range (one for bipolar and one for monopolar electrode configuration). The remaining six missing data points were from the monopolar stimulation at 90% of the dynamic range condition.

vowel and consonant identification scores increased as a function of stimulation level through most or all of the dynamic range (Subjects 2, 3, 4, 5, and 7). For

others, vowel and consonant identification increased to a maximum score near the middle of the dynamic range of stimulus levels, and then decreased. This

TABLE 2

ANOVA results									
Subject number	1	2	3	4	5	6	7	8	All
(1) Stimulus level (% dynamic range)	***	***	***	***	***	***	*	***	***
(2) Syllable type (consonant and vowel)	***	***	***	***		***	***	***	***
(3) Electrode configuration (monopolar and bipolar)	***	***		***	***	***	***	***	***
Interaction between 1 and 3		***					***		
Interaction between 1 and 2		*							
Interaction between 2 and 3	**	***	**	***		*		***	***
Interaction between 1, 2, and 3									

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

pattern was termed “rollover” (Subjects 1, 6, and 8). No differences were seen between these two groups of subjects in the vowel and consonant identification scores of the full-dynamic-range condition nor in the magnitude of the subjects’ dynamic ranges.

In each subject, there was a statistically significant effect of stimulus level on speech syllable identification performance ($p < 0.001$, except for Subject 7 where $p < 0.05$). When analyzed across all subjects, the effect of stimulus level on speech syllable identification performance was also significant ($p < 0.001$). Plots of consonant and vowel speech perception performance vs. stimulus level for both monopolar and bipolar electrode configuration are shown in Figure 2. The results of the statistical analysis of these data are shown in Table 2. Group results showed increased speech perception performance with increased stimulus level (Fig. 3).

Interaction between electrode configuration and stimulus level

It was hypothesized that vowel and consonant identification performance as a function of stimulus level would differ for the monopolar and bipolar electrode configurations. Specifically, if both a high stimulus level and a monopolar electrode configuration contributed to a larger spatial extent of neural activation, we might expect to find a higher incidence of rollover of scores when monopolar stimulation was used. However, the data did not support this hypothesis. When analyzed on a subject-by-subject basis, electrode configuration and stimulus level showed significant interactions in only two subjects, 2 and 7 ($p < 0.001$). Interaction was not statistically significant when the data were analyzed across subjects.

Interaction between syllable type and electrode configuration

When syllable identification was analyzed with respect to differences in the specific test used (vowel vs.

consonant) and electrode configuration, interactions were found. Significant interactions ($p < 0.001$) between these variables were found for Subjects 2, 4, and 8. Interactions were found for Subjects 1, 3, and 6 at lower confidence levels (Table 2). Averaged as a group, this effect was highly significant ($p < 0.001$). These interactions reflected a larger effect of electrode configuration on the vowel tests than on the consonant tests.

Effects of dynamic range compression

Per subject or averaged across subjects, syllable identification scores with the full-dynamic-range condition were not significantly different from the maximum score of the compressed-dynamic-range conditions (see Fig. 4) For the vowel test, the average maximum score across subjects and electrode configurations was 40% correct for the full-dynamic-range condition and 45% correct for the maximum of the compressed-dynamic-range conditions. For the consonant test, the average score across subjects and electrode configurations was 41% correct for the full-dynamic-range condition and 35% correct for the maximum of the compressed-dynamic-range conditions.

Subjective judgments

Complete sets of subjective judgments of sound quality and loudness were obtained from four of the eight subjects (Subjects 3, 4, 7, and 8). For Subjects 3, 4, and 7, sound quality increased with stimulus level. For each of these subjects, syllable identification performance increased with stimulus level and did not show rollover. Subject 8, who did show rollover in syllable identification performance, demonstrated sound-quality judgments that did not change markedly with stimulus level. For all subjects, the monopolar electrode configuration sounded better than the bipolar electrode configuration. The left panel of Figure 5 shows the mean subjective sound-quality

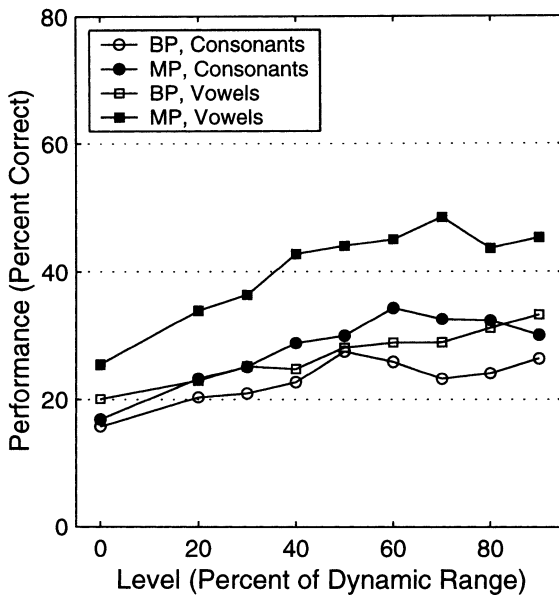


FIG. 3. Group mean of syllable identification as a function of level. Data for the bipolar and monopolar electrode configurations are plotted with open and filled symbols, respectively. Consonant and vowel test results are plotted with circles and squares, respectively. BP = bipolar electrode configuration; MP = monopolar electrode configuration.

judgments of these four subjects. For monopolar stimulation, the subjective rating of the condition with the full dynamic range intact was within the span of the subjective ratings of the 100% compressed-dynamic-range conditions. Higher stimulus-level conditions with no amplitude modulation cues received higher quality judgments than the full-dynamic-range condition with amplitude modulation cues intact.

There was little difference in loudness judgments between the monopolar and bipolar configurations (Fig. 5, right panel). With both electrode configurations, loudness increased linearly with stimulus level. None of the subjects showed rollover in loudness judgments. Loudness estimates for bipolar and monopolar configurations were approximately equal, except for the 80% and 90% conditions, where monopolar was louder than bipolar. However, the maximum difference was ~ 0.1 within the scale of 0–1. The sound-quality preference for the monopolar configuration over the bipolar configuration was therefore not due to the differences in the loudness of the two configurations. The loudness of the full-dynamic-range map was approximately equal to the compressed-map loudness at levels of 70–80% of the dynamic range.

DISCUSSION

Effects of stimulus level on syllable identification differed across subjects, stimulus conditions, and test

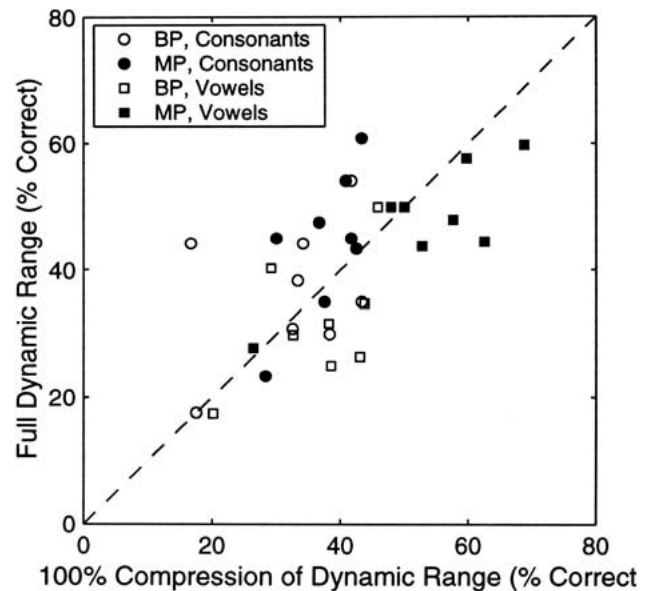


FIG. 4. Syllable identification for conditions with full dynamic range versus conditions with 100% compression of dynamic range. Syllable identification for full dynamic range is plotted as a function of the maximum score of the compressed-dynamic-range conditions. Data for the bipolar and monopolar electrode configurations are plotted with open and filled symbols, respectively. Consonant and vowel test results are plotted with circles and squares, respectively. BP = bipolar electrode configuration; MP = monopolar electrode configuration.

materials. For approximately half of the cases, syllable identification scores increased as a function of stimulus level through most or all of the tested range. For the remaining cases, there was an increase as a function of stimulus level, and then a decrease at higher levels. In this regard, the effects of stimulus level on syllable identification are more similar to the variable effects of stimulus level on electrode discrimination (McKay et al. 1999; Pfungst et al. 1999) than to the more consistent effects on temporal discriminations (Pfungst et al. 1983, 1994, 1999; Pfungst and Rai 1990; Shannon 1983, 1989, 1992, 1993). These results might be due to an interaction of positive and negative effects of stimulus level. Positive effects could include better encoding of speech signals due to increased discharge rates and increased numbers of fibers carrying the signal. Negative effects could include rate saturation and increased channel overlap. Since subjects presumably differ considerably in the condition of auditory nerve fiber population, we would expect these competing effects to interact differently across subjects.

In this study we hypothesized that there would be an interaction between electrode configuration and stimulus level because both affect the number of fibers activated by the stimulus. However, we found no statistically significant interaction between electrode configuration and stimulus level. The hypothesized

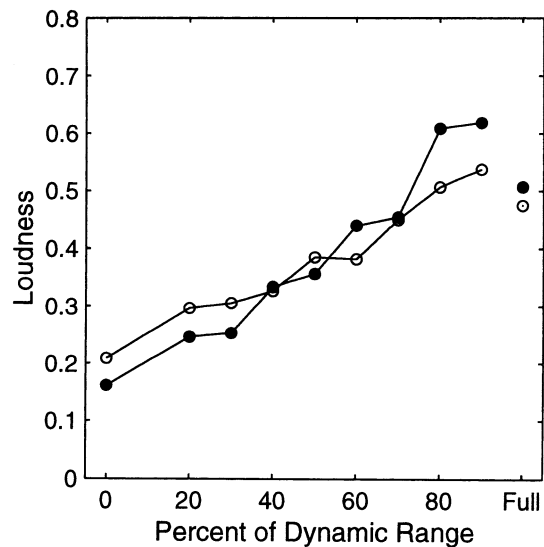
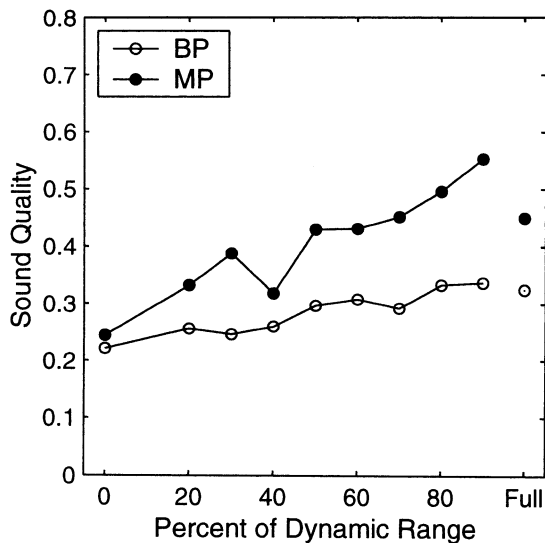


FIG. 5. Subjective ratings. Mean subjective judgments of the sound quality (left panel) and loudness (right panel) as a function of level across the four subjects with complete data (Subjects 3, 4, 7, and 8). In both panels, data for the bipolar and monopolar electrode configurations are plotted with open and filled symbols, respectively.

In the figure legend, BP = bipolar electrode configuration; MP = monopolar electrode configuration. The label of "Full" on the abscissa indicates the full-dynamic-range condition. The sound-quality scale ranged from 0 ("terrible") to 1 ("excellent"). The loudness scale ranged from 0 ("too soft") to 1 ("too loud").

increased incidence of rollover with monopolar stimulation was not seen in this population of subjects. This suggests that a more complex set of variables affect speech perception, with the extent of neural activation not being a dominant variable.

The monopolar electrode configuration resulted in statistically higher syllable identification scores than did the bipolar electrode configuration in seven out of eight subjects. In this study, a higher proportion of subjects performed better in the monopolar electrode configuration compared with the bipolar electrode configuration than in previously published studies (Kileny et al. 1998; Lehnhardt et al. 1992; Pfungst et al. 1997; Zwolan et al. 1996). This result was probably due to the fact that the subjects in this experiment all used a monopolar configuration in their everyday processors, as discussed below.

Compared with the full-dynamic-range condition, 100% dynamic range compression had no statistically significant effect on syllable identification in quiet. These results agree with those of Zeng and Galvin (1999). They showed that the resilience of phoneme identification with dynamic range compression applied to a full set of electrodes (20), as well as to a reduced set (10 and 4). They also showed that dynamic range compression had a negative effect on phoneme identification in noise. Our results differed from the predictions of Loizou et al. (2000) from acoustic simulations of a 6-channel Continuous Interleaved Sampling (CIS) processing strategy. Loizou et al. (2000) predicted a negative effect of dynamic range compression on phoneme identification,

where vowels would be more adversely affected than consonants. Differences between performance with the SPEAK strategy and predictions with the 6-channel CIS strategy could be attributed to the fact that when the dynamic range is compressed, the SPEAK coding strategy can continue to present spectral cues over the relatively larger number of channels. The CIS strategy, conversely, relies more on amplitude modulation cues to present the temporal envelope of the signal, and thus would be more affected by dynamic range compression.

For the subset of subjects with complete subjective rating data, sound quality generally paralleled syllable identification results. For the group data, sound quality was judged better with the monopolar electrode configuration than with the bipolar configuration, and this effect increased as a function of stimulus level. The subjective preference for the monopolar electrode configuration, as with the syllable identification results, might be due to adaptation since monopolar stimulation was used in the subjects' everyday processors. The sound quality of the full-dynamic-range condition was within the range of the ratings of the compressed-dynamic-range condition in the monopolar electrode configuration. This was not the case in the bipolar electrode configuration, where the maximum subjective sound-quality ratings of the compressed-dynamic-range conditions were approximately equal to the full-dynamic-range condition. In the bipolar condition, this result could be due to the small range of relatively poor sound-quality judgments reported by the sub-

jects. Due to user experience and amplitude modulation information with a full dynamic range, we expected that full-dynamic-range conditions would sound better than the compressed-dynamic-range conditions, but this was not the case at high stimulus levels.

Loudness increased with stimulus level in a similar fashion for monopolar and bipolar electrode configurations. The loudness of the full-dynamic-range condition was equal to that of the compressed-dynamic-range condition at 70–80% of the dynamic range. At a sensitivity setting of 7, 64 dB(A) is approximately 70% of the linear operating range presented by the SPrint processor (Cochlear Corporation 1999). Thus, in the full-dynamic-range condition, the perceived loudness of the stimulus is consistent with the device settings.

The results highlight stimulus level as one of the important mapping variables that differentially affects syllable identification across subjects. The results suggest that if speech processors are programmed to optimize stimulus-level input–output functions, speech perception performance could be improved. However, such optimization would need to be done on an individual basis because subjects differ in the pattern of effects of stimulus level and the pattern might depend on other variables such as electrode configuration. Thus, in some cases, subjects might benefit from maximizing stimulus level within the electrical dynamic range for most sounds while in others there might be a benefit from concentrating stimuli near the middle of the dynamic range. This can be done by optimizing each subject's Q value programming parameter. The Q value determines the percentage of the subject's electrical dynamic range that is allocated to the top 10 dB of the speech processor's channel amplitude range (Cochlear Corporation 1999). For example, Q values of 20 and 50 place 80% and 50%, respectively, of the dynamic range in the top 10 dB of the processor's range. Subjects who benefit from having more of the acoustic dynamic range in the upper electric dynamic range would be mapped with a lower Q value. Subjects who benefit from having more of the acoustic dynamic range in the middle electric dynamic range would be mapped with a higher Q value.

The study also contributes to our understanding of how patients' descriptions of sound quality are related to their syllable identification abilities, and how their loudness judgments are related to stimulus levels. Because sound-quality results did not parallel syllable identification abilities as stimulus levels were changed, programming based on what sounds best to the patient might not yield optimal performance. On the other hand, subjective loudness judgments were in agreement with stimulus level.

In most studies that compare experimental performance and preference results, a potential confounding factor is the user's experience with their everyday settings. Everyday settings might result in better performance and better quality judgments because they are most familiar. In this study, all the subjects were using a monopolar electrode configuration every day and this might have biased the performance and preference results toward the monopolar experimental configurations. Experience and familiarity with typical stimulus levels would have biased all of the subjects in the same manner because each used the same Q value. Therefore, differences that were observed in the shape of the stimulus level functions are likely due to differences among subjects, such as differences in neural survival. Another potential confounding factor in the stimulus level data is the fact that the criteria for C levels probably differ across subjects. That is, some subjects might be more conservative than others, setting their C levels to lower loudness levels. This factor was controlled to the greatest extent possible by getting instructions from the same researcher when the subject was setting stimulus levels. This bias could still affect the data, but this does not alter the clinical implications of the study, because C levels were collected in the experiment the same way that they are determined clinically.

SUMMARY AND CONCLUSIONS

1. Effects of electrical stimulus level on syllable identification differed across subjects, with approximately half showing a steady performance increase as a function of level through most or all of the electrical dynamic range, and half with maximum scores near the middle of the dynamic range.
2. The effects of stimulus level on syllable identification were not systematically different for vowels and consonants.
3. The monopolar electrode configuration resulted in better syllable identification scores and subjective quality ratings. This configuration was the subjects' everyday electrode configuration.
4. Complete (100%) dynamic range compression, at an optimum stimulus level, had no statistically significant effect on vowel or consonant identification or on subjective sound quality.
5. Effects of stimulus level on perceived loudness were similar for monopolar and bipolar electrode configurations. The loudness of the full-dynamic-range condition was equivalent to between 70% and 80% of the dynamic range in the 100% compressed-dynamic-range conditions.

ACKNOWLEDGMENTS

This work was supported by NIH-NIDCD grants R01 DC 03808, T32 DC00011, and F32 DC00470. We appreciate the technical assistance provided by Julie Oldfield.

REFERENCES

- ANSI. American National Standards Institute ANSI S3.6-1999, American National Standard Specification for Audiometers. Available from The Acoustical Society of America, Melville, NY, (1996).
- BIERER JA, MIDDLEBROOKS JC. Auditory cortical images of cochlear-implant stimuli: Dependence on electrode configuration. *J. Neurophysiol.* 87:478–492, 2002.
- Cochlear Corporation. Nucleus Technical Reference Manual. Cochlear Limited, Lane Cove, Australia, 1999.
- 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.
- HILLENBRAND J, GETTY LA, CLARK MJ, WHEELER K. Acoustic characteristics of American English vowels. *J. Acoust. Soc. Am.* 97:3099–3111, 1995.
- KILENY PR, ZWOLAN TA, TELIAN SA, BOERST A. Performance with the 20 + 2L lateral wall cochlear implant. *Am. J. Otol.* 19:313–319, 1998.
- KRAL A, HARTMANN R, MORTAZAVI D, KLINKE R. Spatial resolution of cochlear implants: The electrical field and excitation of auditory afferents. *Hear. Res.* 121:11–28, 1998.
- LEHNHARDT E, GNADEBERG D, BATTMER RD, VON WALLEMBERG 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.
- LOIZOU PC, DORMAN M, FITZKE J. The effect of reduced dynamic range on speech understanding: implications for patients with cochlear implants. *Ear Hear.* 21:25–31, 2000.
- McKAY CM, O'BRIEN A, JAMES CJ. Effect of current level on electrode discrimination in electrical stimulation. *Hear. Res.* 136:159–164, 1999.
- MORRIS DJ, PFINGST BE. Effects of electrode configuration and stimulus level on rate and level discrimination with cochlear implants. *J. Assoc. Res. Otolaryngol.* 1:211–223, 2000.
- PFINGST BE, BURNETT PA, SUTTON D. Intensity discrimination with cochlear implants. *J. Acoust. Soc. Am.* 73:1283–1292, 1983.
- PFINGST BE, FRANCK KH, XU L, BAUER EM, ZWOLAN TA. Effects of electrode configuration and place of stimulation on speech perception with cochlear prostheses. *J. Assoc. Res. Otolaryngol.* 2:87–103, 2001.
- PFINGST BE, HOLLOWAY LA, POOPAT N, SUBRAMANYA AR, WARREN MF, ZWOLAN TA. Effects of stimulus level on nonspectral frequency discrimination by human subjects. *Hear. Res.* 78:197–209, 1994.
- PFINGST BE, HOLLOWAY LA, ZWOLAN TA, COLLINS LM. Effects of stimulus level on electrode-place discrimination in human subjects with cochlear implants. *Hear. Res.* 134:105–115, 1999.
- PFINGST BE, RAI DT. Effects of level on nonspectral frequency difference limens for electrical and acoustic stimuli. *Hear. Res.* 50:43–56, 1990.
- 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.
- REBSCHER SJ, SNYDER RL, LEAKE PA. The effect of electrode configuration and duration of deafness on threshold and selectivity of responses to intracochlear electrical stimulation. *J. Acoust. Soc. Am.* 109:2035–2048, 2001.
- SHANNON RV. Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear. Res.* 11:157–189, 1983.
- SHANNON RV. Detection of gaps in sinusoids and pulse trains by patients with cochlear implants. *J. Acoust. Soc. Am.* 85:2587–2592, 1989.
- SHANNON RV. Temporal modulation transfer functions in patients with cochlear implants. *J. Acoust. Soc. Am.* 91:2156–2164, 1992.
- SHANNON RV. Reply to: "Comment on Temporal modulation transfer functions in patients with cochlear implants" [*J. Acoust. Soc. Am.* 93:1649–1650 (1993)]. *J. Acoust. Soc. Am.* 93:1651–1652, 1993.
- SHANNON RV, JENSVOLD A, PADILLA M, ROBERT ME, WANG X. Consonant recordings for speech testing [letter]. *J. Acoust. Soc. Am.* 106:71–L74, 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.
- SKINNER MW, HOLDEN LK, HOLDEN TA, DEMOREST ME. Comparison of two methods for selecting minimum stimulation levels used in programming the Nucleus 22 cochlear implant. *J. Speech Lang. Hear. Res.* 42:814–828, 1999.
- SNYDER RL, REBSCHER SJ, CAO K, LEAKE PA, KELLY K. Chronic intracochlear electrical stimulation in the neonatally deafened cat. I. Expansion of central representation. *Hear. Res.* 50:7–33, 1990.
- SPELMAN FA, PFINGST BE, CLOPTON BM, JOLLY CN, RODENHISER KL. Effects of electrical current configuration on potential fields in the electrically stimulated cochlea: Field models and measurements. *Ann. Otol. Rhinol. Laryngol. Suppl.* 166:131–136, 1995.
- TOWNSHEND B, COTTER N, VAN COMPERNOLLE D, WHITE RL. Pitch perception by cochlear implant subjects. *J. Acoust. Soc. Am.* 82:106–115, 1987.
- WHITFORD LA, SELIGMAN PM, EVERINGHAM CE, ANTOGENELLI T, SKOK RD, HOLLOW RD, PLANT KL, GERIN ES, STALLER SJ, McDERMOTT WR, 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.
- ZENG FG, GALVIN JJ. Amplitude mapping and phoneme recognition in cochlear implant listeners. *Ear Hear.* 20:60–74, 1999.
- 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.