# Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies

Gail S. Donaldson and David A. Nelson
*Department of Otolaryngology, University of Minnesota, Minneapolis, Minnesota 55455*

(Received 13 November 1998; revised 28 July 1999; accepted 19 November 1999)

Two related studies investigated the relationship between place-pitch sensitivity and consonant recognition in cochlear implant listeners using the Nucleus MPEAK and SPEAK speech processing strategies. Average place-pitch sensitivity across the electrode array was evaluated as a function of electrode separation, using a psychophysical electrode pitch-ranking task. Consonant recognition was assessed by analyzing error matrices obtained with a standard consonant confusion procedure to obtain relative transmitted information (RTI) measures for three features: stimulus (RTI stim), envelope ($\mathrm{RTI\,env}_{[\mathrm{plc}]}$), and place-of-articulation ($\mathrm{RTI\,plc}_{[\mathrm{env}]}$). The first experiment evaluated consonant recognition performance with MPEAK and SPEAK in the same subjects. Subjects were experienced users of the MPEAK strategy who used the SPEAK strategy on a daily basis for one month and were tested with both processors. It was hypothesized that subjects with good place-pitch sensitivity would demonstrate better consonant place-cue perception with SPEAK than with MPEAK, by virtue of their ability to make use of SPEAK's enhanced representation of spectral speech cues. Surprisingly, all but one subject demonstrated poor consonant place-cue performance with both MPEAK and SPEAK even though most subjects demonstrated good or excellent place-pitch sensitivity. Consistent with this, no systematic relationship between place-pitch sensitivity and consonant place-cue performance was observed. Subjects' poor place-cue perception with SPEAK was subsequently attributed to the relatively short period of experience that they were given with the SPEAK strategy. The second study reexamined the relationship between place-pitch sensitivity and consonant recognition in a group of experienced SPEAK users. For these subjects, a positive relationship was observed between place-pitch sensitivity and consonant place-cue performance, supporting the hypothesis that good place-pitch sensitivity facilitates subjects' use of spectral cues to consonant identity. A strong, linear relationship was also observed between measures of envelope- and place-cue extraction, with place-cue performance increasing as a constant proportion ($\sim$0.8) of envelope-cue performance. To the extent that the envelope-cue measure reflects subjects' abilities to resolve amplitude fluctuations in the speech envelope, this finding suggests that both envelope- and place-cue perception depend strongly on subjects' envelope-processing abilities. Related to this, the data suggest that good place-cue perception depends both on envelope-processing abilities and place-pitch sensitivity, and that either factor may limit place-cue perception in a given cochlear implant listener. Data from both experiments indicate that subjects with small electric dynamic ranges ($<$8 dB for 125-Hz, 205-$\mu$s/ph pulse trains) are more likely to demonstrate poor electrode pitch-ranking skills and poor consonant recognition performance than subjects with larger electric dynamic ranges. © *2000 Acoustical Society of America.* [S0001-4966(00)01403-X]

PACS numbers: 43.66.Ts, 43.66.Hg, 43.66.Jh, 43.71.Ky [JWH]

## INTRODUCTION

Recent developments in cochlear implant speech processing strategies have resulted in significant improvements in implant listeners' speech recognition performance. As a result, users of the newest generation of speech-processing strategies commonly achieve high levels of speech recognition under favorable listening conditions (e.g., Dorman, 1993; Skinner *et al.*, 1994; Tyler *et al.*, 1996). At present, we have only a limited understanding of how such high levels of speech recognition are achieved by the better implant performers or, conversely, why some implant listeners still perform poorly with the same prostheses and speech-processing strategies. Related to this, there exists relatively little information concerning the psychophysical abilities underlying speech recognition in electric hearing, or how such abilities may vary as a function of stimulation parameters that differ among speech-processing strategies.

One psychophysical ability thought to be important for speech recognition in electric hearing is place-pitch sensitivity, i.e., the ability to distinguish among electrodes on the basis of tonotopically mediated pitch, or timbre. Theoretically, cochlear implant listeners with good place-pitch sensitivity can make use of the spectral information in speech if it is well-represented in the electrical stimulus. Such spectral information contributes primarily to discrimination of the vowel height and frontness features (related to $f1$ and $f2$ frequency, respectively) and the consonant place-of-articulation feature, since these phonemic features are coded predominantly by spectral cues. Townshend *et al.* (1987) were the first investigators to examine place-pitch sensitivity in cochlear implant listeners, using psychophysical electrode pitch-ranking and electrode discrimination tasks. They showed that some implantees exhibit relatively strong place-pitch sensitivity consistent with the normal tonotopic organization of the cochlea, whereas others perceive little change in pitch as a function of electrode location. This finding, confirmed in several later studies (see below), suggests that place-pitch sensitivity may be an important factor limiting speech recognition ability in some cochlear implant users.

Only a few studies have examined the relationship between place-pitch sensitivity and speech recognition, and these studies are limited to subjects using the Nucleus $f0/f1/f2$ or MPEAK speech processing strategies. Busby *et al.* (1993) showed that discrimination of electrode trajectories by four late-deafened cochlear implant subjects predicted the degree to which speech recognition performance improved with a multichannel device ($f0/f1/f2$ strategy) relative to a single-channel device. Nelson *et al.* (1995) evaluated the relationship between place-pitch sensitivity and consonant recognition in eight postlingually deafened subjects using the Nucleus MPEAK or $f0/f1/f2$ strategy. Several subjects exhibited good or excellent place-pitch sensitivity; however, none of the subjects demonstrated particularly good place-of-articulation performance. This led the authors to conclude that spectral cues to consonant identity are not coded very effectively by the MPEAK and $f0/f1/f2$ strategies. Even so, a weak correlation between place-pitch sensitivity and consonant place-cue performance was reported. Nelson *et al.*'s impression that the $f0/f1/f2$ and MPEAK strategies provide only minimal place-cue information was subsequently supported by place-cue data reported by Parkinson *et al.* (1996) in a comparison study of the $f0/f1/f2$ and MPEAK strategies. Zwolan *et al.* (1997) examined the relationship between electrode discrimination and speech recognition in 11 users of the Nucleus MPEAK strategy. A considerable range of electrode discrimination performance was observed across subjects, with several subjects demonstrating excellent performance; however, no significant relationship between electrode discrimination and speech recognition was observed. In general, the above studies suggest that many cochlear implant patients possess good place-pitch sensitivity, but that place-pitch sensitivity is not strongly related to speech recognition performance with the $f0/f1/f2$ or MPEAK speech-processing strategies.

The present experiments were motivated by the MPEAK-SPEAK comparison study of Skinner *et al.* (1994), which showed that some but not all users of the MPEAK speech-processing strategy can achieve improved speech recognition performance with SPEAK. We were especially interested in the individual differences apparent in Skinner *et al.*'s data, in particular, the demonstration that subjects who achieved equivalent levels of performance with MPEAK sometimes exhibited considerably different levels of performance with SPEAK. Because SPEAK provides a more detailed spectral characterization of the speech waveform than MPEAK, we hypothesized that subjects showing the greatest gains with SPEAK were those subjects who were best able to use this spectral information, i.e., those with good place-pitch sensitivity. Skinner *et al.* did not analyze speech recognition data in terms of spectral versus temporal speech cues; thus it is not possible to determine whether the improvements demonstrated by their subjects were primarily related to spectral cues as our hypothesis would predict.

The studies described here specifically evaluated the relationship between place-pitch sensitivity and consonant recognition in cochlear implant subjects using the MPEAK and SPEAK strategies. The first experiment compared consonant recognition performance with MPEAK and with SPEAK in a single group of subjects. Subjects were experienced users of the MPEAK strategy, and their performance with SPEAK was evaluated after they used the SPEAK strategy on a daily basis for one month. The findings from this first experiment were somewhat surprising: Although subjects demonstrated a wide range of place-pitch sensitivity as measured by our electrode pitch-ranking task, their place-cue performance was equally poor with the SPEAK strategy as with the MPEAK strategy, and there was no systematic relationship between place-pitch sensitivity and consonant place-cue performance with either strategy.

Subsequent examination of several cochlear implant listeners' performance over time with the SPEAK strategy showed that performance on the consonant place-cue measure sometimes continued to improve well beyond the first month of use. This suggested that the relatively short SPEAK trial used in our first experiment may have provided an inaccurate picture of the relationship between place-pitch sensitivity and consonant recognition. As a result, we reevaluated this relationship in a second study using subjects with substantially more SPEAK experience. Data from the second study showed the expected positive relationship between place-pitch sensitivity and consonant place-cue performance, consistent with our hypothesis. In addition, this study indicated that consonant place-cue performance depends not only on place-pitch sensitivity but also on subjects' ability to resolve amplitude fluctuations in the speech signal.

## EXPT 1: PLACE-PITCH SENSITIVITY AND CONSONANT RECOGNITION WITH MPEAK VERSUS SPEAK

### Methods

#### *Subjects*

Fourteen postlingually deafened adult subjects participated in the first experiment. All were experienced users of the Nucleus 22-electrode implant, having used their implants continuously for at least one year. Twelve were clinical users of the MPEAK speech-processing strategy implemented on the Mini Speech Processor and two were clinical users of the $f0/f1/f2$ strategy implemented on the Wearable Speech Processor. All subjects were native speakers of American English. Subjects provided informed consent to participate in the study and were paid on an hourly basis for their participation. Demographic data and other information describing the subjects are provided in Table I.

TABLE I. Description of 14 cochlear implant subjects who participated in expt. 1: Subject identifying code, gender, age, etiology of deafness (implanted ear), duration of bilateral severe-to-profound hearing loss prior to implantation, depth of electrode array insertion (mm from the round window, with 25 mm representing complete insertion), duration of implant use prior to the study, and duration of MPEAK use prior to the study. To provide readers with an indication of subjects' clinical performance levels with MPEAK (or $f0/f1/f2$), subjects' scores for the NU-6 monosyllabic word test (% correct words and % correct phonemes) are also shown.

| Subj | m/f | Age (yrs) | Etiology of deafness | Duration (yrs) | Depth (mm) | CI use (yrs) | MPEAK use (yrs) | Nu-6 % words | Nu-6 % phons |
|---|---|---|---|---|---|---|---|---|---|
| AJA | *m* | 48 | skull fracture | 10 | 20 | 6.4 | 4.7 | 22 | 43 |
| AMB | *m* | 51 | progressive SNHL | 1 | 25 | 1.6 | 1.6 | 40 | 63 |
| BRL | *f* | 49 | progressive SNHL | 25 | 20 | 3.2 | 3.2 | 8 | 27 |
| EJQ | *f* | 52 | mumps, progressive SNHL | 9 | 22 | 8.0 | 8.0 | 0 | 12 |
| FXC | *m* | 68 | progressive SNHL | 4 | 25 | 4.1 | 4.1 | 16 | 44 |
| JMS | *f* | 48 | progressive SNHL | 36 | 25 | 4.9 | 4.9 | 8 | 34 |
| JWB | *m* | 60 | cochlear otosclerosis | 4 | 20 | 8.5 | ⋯ | 6 | 22 |
| KRK | *m* | 66 | familial SNHL noise exp. | 5 | 24 | 5.1 | 5.1 | 6 | 27 |
| LMF | *f* | 24 | meningitis | 12 | 21 | 6.8 | 6.8 | 0 | 4 |
| MAS | *f* | 65 | genetic/progressive SNHL | 10 | 25 | 2.4 | 2.4 | 28 | 52 |
| PLF | *f* | 67 | otosclerosis | $<1$ | 25 | 1.4 | 1.4 | 14 | 40 |
| REC | *m* | 71 | traumatic noise exposure | 15 | 25 | 5.5 | 5.5 | 6 | 27 |
| RFM | *m* | 62 | Meniere's disease | 1 | 22 | 6.4 | ⋯ | 8 | 28 |
| TVB | *m* | 46 | progressive SNHL | 8 | 22 | 4.7 | 4.7 | 12 | 37 |

### Study design

Each subject completed a one-month trial with the SPEAK strategy implemented on a loaner Spectra processor, and underwent consonant confusion testing four times. Ten subjects who were clinical users of the MPEAK strategy completed the following standard (ABBA) protocol: (1) testing with MPEAK 2–4 weeks prior to the beginning of the SPEAK trial (**MPK-1** condition); (2) testing with SPEAK on the first day of the one-month trial (**SPK-1** condition); (3) repeat testing with SPEAK at the conclusion of the one-month trial (**SPK-2** condition); and (4) repeat testing with MPEAK approximately 4 weeks later (**MPK-2** condition). Comparisons of consonant recognition with the MPEAK and SPEAK strategies, described below, are based on data from these ten subjects. Four additional subjects (two MPEAK users and two $f0/f1/f2$ users) provided data for the SPK-1 and SPK-2 conditions but were unable to complete the standard test protocol for the MPK-1 and/or MPK-2 conditions. Data for these subjects are included in analyses that compare electrode pitch-ranking performance and consonant recognition with the SPEAK strategy. A psychophysical electrode pitch-ranking task was used to measure place-pitch sensitivity across the electrode array in each subject. In most cases, this testing was performed within the three-month time period spanned by consonant recognition testing.[1]

### Speech processor maps

Subjects used their usual, clinical maps for testing with MPEAK. The SPEAK map created for each subject initially used the same thresholds (*T*-levels) and most comfortable loudness levels (*C*-levels) as the MPEAK map. In a few cases, *C*-levels were subsequently reduced by a constant small percentage across electrodes to compensate for increased overall loudness with the SPEAK strategy. Other SPEAK parameters were set to their default values, including the frequency allocation table which varied according to the number of electrodes activated in a subject's map. All subjects used the Cochlear Corp. ''Stimulus Level'' intensity coding scheme for both the MPEAK and SPEAK strategies. Over the range of levels used here, this coding scheme holds current amplitude constant at approximately 1 mA, and varies pulse duration in logarithmic steps between 19 and 400 $\mu$s/ph (Cochlear Corp., 1996).

The MPEAK and SPEAK speech-processing strategies typically employ bipolar stimulation in which the active and return members of a given electrode pair are closely spaced electrodes along the cochlear array. In this report, electrodes are numbered in increasing order from the most apical (1) to the most basal (22) electrode along the array, and a given bipolar electrode pair is referred to by its more basal member. Two expt. 1 subjects were programmed with a bipolar separation of 0.75 mm (BP mode); all others were programmed with a bipolar separation of 1.5 mm (BP+1).

### Consonant recognition procedures

**Stimuli**. Stimuli used in the consonant confusion procedure were 19 /aCa/ disyllables spoken by each of three female talkers and three male talkers, where $C$ =/p,t,k,b,d,g,f,θ,s,ʃ,v,ð,z,ʒ,m,n,r,l,j/. These were identical to the stimuli used by Van Tasell *et al.* (1992) in their ''unprocessed'' condition. The 114 tokens (19 stimuli×6 talkers) were digitized by Van Tasell *et al.* at a sampling rate of 10 kHz with 12-bit resolution.

**Test procedure**. The subject was seated approximately 1 meter in front of a pair of high-quality loudspeakers and a video screen in a sound-isolated room. Speech tokens were played from computer memory at 10 kHz, amplified, and presented through the speakers. The presentation level of individual tokens varied over a 5-dB range between 60–65 dB SPL (slow response, A-weighting scale) in the sound field at the location of the subject's head. This level is consistent with a conversational or slightly raised vocal effort by a speaker 1 m from the listener (Pearsons *et al.*, 1976; Skinner *et al.*, 1997). The speech processor was set to the ''normal''

TABLE II. Envelope and place categories of Van Tasell *et al.* (1992) that were used for information transmission analyses in the present study.

| Category # | Envelope feature | Place feature |
|---|---|---|
| 1 | /b,d,g,v,ð,ʒ,z/<br>(voiced fricatives and plosives) | /b,p,f,v,m/<br>(front) |
| 2 | /p,t,k/<br>(voiceless plosives) | /d,s,j,t,n,z,ð,1,θ/<br>(mid) |
| 3 | /f,s,θ,ʃ/<br>(voiceless fricatives and affricates) | /k,ʃ,ʒ,r,g/<br>(back) |
| 4 | /m,n,r,l,j/<br>(nasals and glides) | |

(rather than the noise-reduction) setting, and subjects were instructed to adjust the processor's sensitivity control such that stimuli were comfortably loud. A standard test block consisted of one presentation each of all 114 tokens, in randomized order. The stimulus was presented once on each trial, and the subject used a computer mouse to select his or her response from a list of 19 alternatives displayed on the video screen. Correct-answer feedback was provided immediately after each stimulus presentation, as recommended by Van Tasell *et al.* (1992). A practice block of 38 randomly selected tokens was presented initially, followed by 5 standard blocks. Testing was usually completed in a single 2–3 hour session; however, it was occasionally necessary to carry over testing to a second session. In the latter case, a practice block was obtained at the beginning of each test session. A pooled confusion matrix was created from the five standard blocks of data for a particular subject and test condition. Each pooled matrix represented 30 observations (5 blocks×6 tokens) per stimulus.

**Analysis**. Information transmission analysis was performed on the pooled consonant confusion matrices to obtain measures of *relative transmitted information for stimulus* (RTI stim), *relative transmitted information for envelope conditional on place* ($\mathrm{RTI\,env}_{[\mathrm{plc}]}$), and *relative transmitted information for place conditional on envelope* ($\mathrm{RTI\,plc}_{[\mathrm{env}]}$). RTI stim represents the proportion of all available stimulus information that is successfully transmitted to the listener, with possible values ranging from 0 to 1. It was computed in the manner described by Miller and Nicely (1955). $\mathrm{RTI\,env}_{[\mathrm{plc}]}$ represents the subject's ability to extract and use low-frequency temporal information in the speech waveform (envelope and periodicity), whereas $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ represents the subject's ability to utilize available spectral cues to consonant place-of-articulation. $\mathrm{RTI\,env}_{[\mathrm{plc}]}$ and $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ were computed via sequential information analysis (SINFA), using the modified envelope categories described by Van Tasell *et al.* (1992) and the place categories *front, middle*, and *back*. Category membership for the envelope and place features used in SINFA analyses are shown in Table II. The conditional values $\mathrm{RTI\,env}_{[\mathrm{plc}]}$ and $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ were computed because the envelope and place feature sets shown in Table II are not completely orthogonal. These conditional measures can be viewed as independent indicators of subjects' abilities to use temporal (envelope) and spectral cues in the acoustic speech waveform (Van Tasell *et al.*, 1992).

### *Electrode pitch-ranking procedures*

The stimuli and procedures used for the electrode pitch-ranking task were similar to those described by Nelson *et al.* (1995). Stimuli were 500-ms trains of 125 Hz, 205 $\mu$s/ph biphasic pulses presented at current amplitudes yielding constant, comfortable loudness across electrodes as determined by the loudness balancing procedure described below. Stimulation was bipolar, with bipolar mode matched to that used in the subject's speech processor (BP or BP+1). Electrodes were stimulated directly (bypassing the speech processor) using a specialized interface (Shannon *et al.*, 1990) controlled through the parallel port of an 80–486 computer running custom software.

Current amplitudes used in the electrode pitch-ranking task were determined by balancing loudness to a common level of ''medium loud'' across electrodes. To accomplish this, estimates of threshold and maximum acceptable loudness were first obtained for each usable electrode in the subject's array, using an ascending method of limits procedure. In this procedure, the current amplitude of the pulse train stimulus (500-ms pulse trains separated by 500-ms silent intervals) was raised slowly from below threshold to a level where the subject first reliably heard the tone (i.e., could tap his or her finger in synchrony with the tone). After this level was recorded, the current amplitude of the stimulus was gradually increased further. The subject indicated loudness changes by sliding his or her finger along a printed loudness scale, and stimulation was terminated when maximum acceptable loudness was reached. After a short pause, a second ascent was completed, and average threshold and maximum acceptable loudness values were computed from the two values of each. The range of current amplitudes yielding a loudness percept of ''medium loud'' was then determined for one electrode near the middle of the array by slowly raising and lowering current levels over the range of amplitudes yielding medium to loud percepts. The current amplitude on this reference electrode was then set to the approximate midpoint (in logarithmic amplitude units) of the medium loud range. Following this, current amplitudes on each of the remaining electrodes were adjusted to produce similar loudness by presenting the reference stimulus alternately with an adjustable stimulus on each new electrode. The current amplitude on the nonreference electrode was adjusted by the experimenter until the stimulus on that electrode was judged to be equally loud as the stimulus on the reference electrode. Once loudness matches were obtained on all usable electrodes in the subject's array, the loudness-balanced stimuli were played back to the subject in random order to ensure that no stimulus was noticeably louder or softer than the others. If any irregularities in loudness were noted by the subject, the loudness balance procedure was repeated on the electrodes in question and the stimuli were again checked for equal loudness using the playback procedure.

Loudness balancing was intended to ensure that loudness differences between stimuli would not distract subjects from judging pitch differences between stimuli presented on two different electrodes. However, it could not guarantee that all discriminable loudness differences between electrode pairs had been eliminated. For this reason, additional steps were taken to ensure that subjects could not use loudness cues to improve their pitch-ranking performance. First, subjects were not given feedback on the electrode pitch-ranking task. Thus even if small loudness differences could be discriminated on a given stimulus pair, the subject could not use the loudness cue to improve performance. Second, stimuli were presented in randomized blocks consisting of 10 trials each of 6 to 10 electrode pairs, and the ordering of individual pairs was also randomized.

Pitch-ranking data were obtained using a 2IFC procedure in which two electrodes were stimulated in sequence, separated by a 500-ms silent interval. On each two-interval trial, the subject's task was to select the stimulus that was ''higher in pitch'' or ''sharper.'' The term ''sharpness'' was included in our instructions because subjects in our previous study (Nelson *et al.*, 1995) reported that some stimulus pairs differed in sharpness rather than pitch. A correct response was scored when the subject chose the stimulus presented to the more basal electrode. Feedback was withheld in order to eliminate the possibility that subjects could correctly order stimuli in the case of pitch reversals, i.e., when stimuli were perfectly discriminable but the more apical electrode produced a higher pitch. As indicated above, this also removed the possibility that any remaining loudness differences between stimuli could be used to improve performance. All possible pairs of electrodes separated by distances of 0.75, 1.5, 3.0, and 4.5 mm (1st, 2nd, and 4th diagonals of the comparison matrix) were tested in each subject. Additional electrode separations of 6.0 and 7.5 mm (8th and 10th diagonals) were tested in most subjects whose data did not approach perfect performance at 4.5 mm separation. Note that the term ''electrode separation,'' as used here, refers to the distance between the basal (or apical) members of each of the two electrode pairs stimulated on a given trial, and not to the distance between electrodes in a single electrode pair, which we refer to as ''bipolar separation.''

Subjects were initially trained to perform the pitch-ranking task using widely spaced pairs of electrodes. Following training, data were obtained for 5 to 8 electrode pairs at a time, with 10 trials per comparison pair presented in random order within the block of 50 to 80 trials. Comparisons within a given block of 50 to 80 trials involved a limited region of the electrode array (basal, middle, or apical) and a fixed electrode separation. Blocks of trials with larger and smaller electrode separations were alternated, so that subjects were not required to endure long stretches of trials that involved very difficult comparisons. After one complete data set (10 trials/comparison) was obtained for all electrode separations, two additional data sets were obtained. This resulted in a total of 30 trials per comparison. Three or four two-hour sessions were typically required to complete the entire electrode pitch-ranking procedure.

A merged comparison matrix was constructed for each subject's pitch-ranking data, with comparison scores expressed as percent correct responses. The average percent correct scores obtained in the 1st, 2nd, 4th, and 6th diagonals of the matrix were computed to arrive at mean percent correct pitch-ranking scores for electrode separations of 0.75, 1.5, 3.0 and 4.5 mm. These mean percent correct scores were then translated to units of perceptual sensitivity ($d'$) using the conversion tables of Hacker and Ratcliffe (1979). A $d'$ value of 3.29 was assigned to perfect performance.

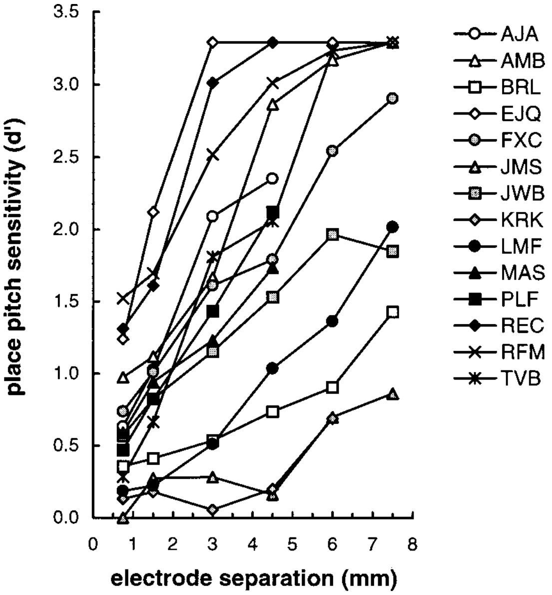

FIG. 1. Average place-pitch sensitivity ($d'$) across the electrode array as a function of electrode separation, for 14 subjects in expt. 1. Perfect discrimination corresponds to $d' = 3.29$.

## Results and discussion

### *Electrode pitch ranking*

Figure 1 shows average place-pitch sensitivity as a function of electrode separation for each of the 14 subjects who participated in expt. 1. There was considerable variability in subjects' performance at all electrode separations. At the narrowest separation (0.75 mm, corresponding to the distance between adjacent electrodes in the Nucleus array), place-pitch sensitivity ($d'$) ranged from 0.13 to 1.52; however, only 3 of 14 subjects achieved performance better than $d' = 1$. Performance improved systematically as electrode separation increased from 0.75 to 4.5 mm for most subjects, with the result that 11 of 14 subjects demonstrated place-pitch sensitivity of $d' = 1$ or better at an electrode separation of 4.5 mm. Two subjects (JMS and KRK) demonstrated unusually poor pitch-ranking performance. For these subjects, performance was near chance for electrode separations of 0.75 to 4.5 mm and improved only slightly for wider electrode separations.

It should be noted that pitch reversals were common at narrow electrode separations, but occurred less often at wider electrode separations. Specifically, 13 of 14 subjects demonstrated one or more pitch reversals at an electrode separation of 0.75 mm, whereas only 2 of 14 subjects demonstrated any pitch reversals at the 4.5-mm separation. It is also noteworthy that place-pitch sensitivity did not vary systematically with distance along the electrode array in most cases. Only two subjects demonstrated clear differences in place-pitch sensitivity in the apical versus basal halves of the implanted array and, even for these subjects, differences were not dramatic.

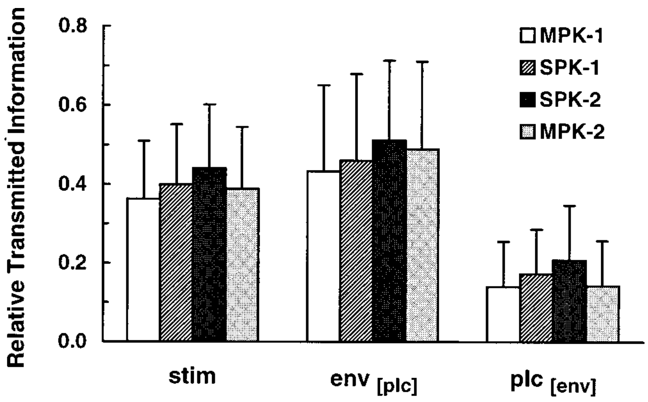

FIG. 2. Mean consonant recognition data for ten subjects who followed the standard expt. 1 protocol. Relative transmitted information (RTI) measures for the stimulus, envelope, and place features are shown for the four consonant recognition tests obtained with MPEAK and SPEAK, in the order obtained (see text). Error bars represent 1 standard deviation (s.d.).

### *Consonant recognition*

**Group data.** Figure 2 shows mean consonant recognition data as a function of test condition for the ten subjects who followed the standard testing protocol. Mean scores for RTI stim ranged from 36.3% to 44.1%, depending on test condition (MPK-1, SPK-1, SPK-2 or MPK-2). Mean scores for RTI $\mathrm{env}_{[\mathrm{plc}]}$ (50.1%–60.6%) were considerably higher than those for RTI $\mathrm{plc}_{[\mathrm{env}]}$ (12.9%–20.7%) for all test conditions, reflecting subjects' strong reliance on envelope cues to consonant identification.

Comparison of mean scores for the first and second tests with MPEAK showed a small but significant improvement in envelope-cue recognition, with RTI $\mathrm{env}_{[\mathrm{plc}]}$ increasing from 43.4% for MPK-1 to 49.0% for MPK-2 (paired *t*-test, $p < 0.05$). This improvement may be attributable to the experience that subjects accrued with the test materials between the MPK-1 and MPK-2 time points or, possibly, to a more general improvement in envelope-cue perception that occurred as the result of the subjects' intervening experience with SPEAK (Skinner *et al.*, 1994, p. 24). Mean scores for RTI stim and RTI $\mathrm{plc}_{[\mathrm{env}]}$ were comparable for the MPK-1 and MPK-2 conditions. As expected, mean performance with SPEAK improved between the first and last days of the one-month trial period (SPK-1 versus SPK-2 conditions). Improvements were statistically significant ($p < .01$) for all three features (RTI stim, RTI $\mathrm{env}_{[\mathrm{plc}]}$ and RTI $\mathrm{plc}_{[\mathrm{env}]}$) both for the subset of ten subjects shown in Fig. 1 and for the entire group of 14 subjects who were tested with SPEAK.

Comparisons between MPEAK and SPEAK performance were made using data from the second test with each processor (MPK-2 and SPK-2 conditions, respectively). Overall transmission of stimulus information (RTI stim) and transmission of spectral information (RTI $\mathrm{plc}_{[\mathrm{env}]}$) were slightly higher with SPEAK than with MPEAK and these differences were statistically significant (paired *t* tests, $p < .05$). Transmission of envelope information (RTI $\mathrm{env}_{[\mathrm{plc}]}$) was similar for the two strategies. In general, these findings indicate that consonant place-of-articulation cues were transmitted slightly better with the SPEAK strategy than with the MPEAK strategy, and that improvements in place-cue transmission with SPEAK resulted in slightly better overall consonant recognition with that strategy.

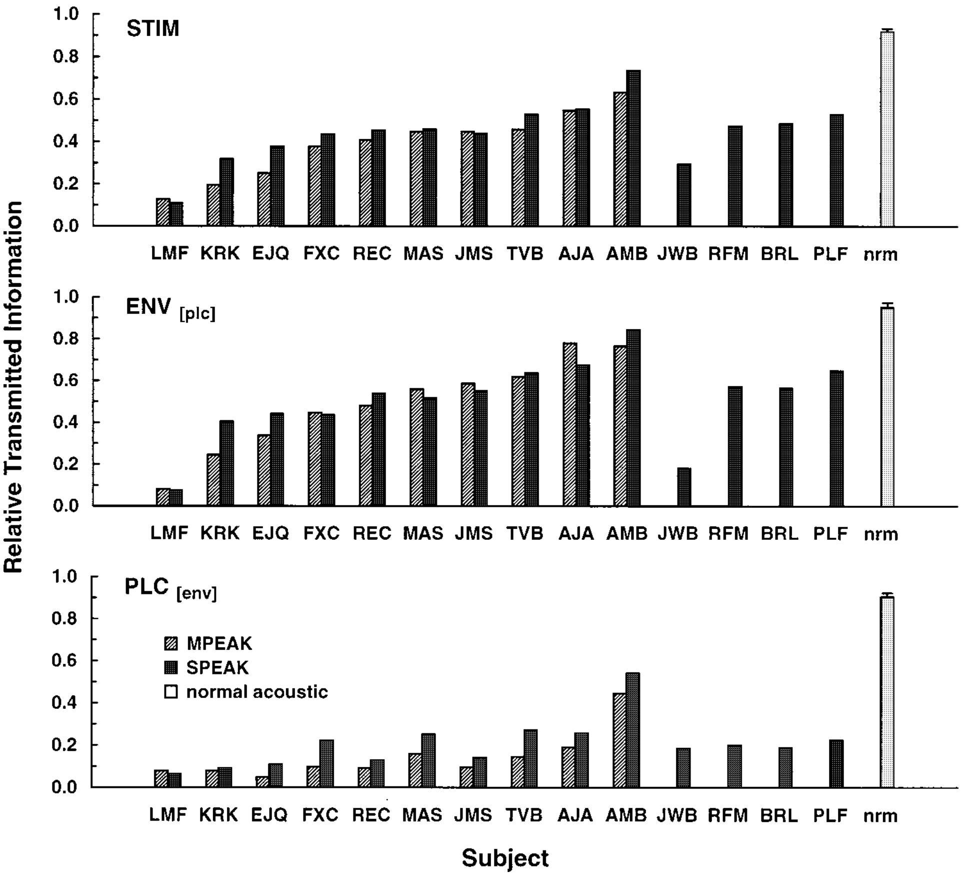

FIG. 3. Consonant recognition data for ten subjects who followed the standard expt. 1 protocol (MPEAK and SPEAK data), and for four subjects who provided SPEAK data only. Relative transmitted information (RTI) values for stimulus, envelope (conditional on place), and place (conditional on envelope) are shown in the top, middle, and bottom panels, respectively, as a function of speech-processing strategy. Mean values for four normal-hearing acoustic listeners (nrm) are also shown. Error bars indicate 1 s.d.

**Individual data.** Figure 3 shows individual subjects' consonant recognition data for the MPEAK and SPEAK processors. Data for the stimulus, envelope, and place features are shown in the top, middle, and bottom panels, respectively. Within each panel, data are shown for the ten subjects who followed the standard testing protocol (in order of increasing performance for the stimulus feature) and for the four subjects tested with SPEAK only (in similar order). Mean data for four normal-hearing, acoustic subjects are shown to the far right of each panel for reference purposes.

Performance on the stimulus feature (RTI stim), which represents overall transmitted information for the consonant stimuli, varied considerably across subjects with scores ranging from 11.8% (subject LMF with MPEAK) to 73.4% (subject AMB with SPEAK). A similar pattern of scores across subjects was obtained for the envelope feature. In contrast, scores for the place feature were almost uniformly low, with only one subject (AMB) scoring above 30% with either the MPEAK or the SPEAK processing strategy. Differences in individual subjects' performance with the MPEAK versus SPEAK strategies were small, with only three subjects (KRK, EJQ, and AMB) demonstrating more than 5% improvement in overall performance (RTI stim) with SPEAK. Two of these subjects, KRK and EJQ, were relatively poor performers. For these subjects, the overall improvement with SPEAK stemmed primarily from improved transmission of envelope cues ($\text{RTI env}_{[\text{plc}]}$). The third subject, AMB, was a considerably better performer. In his case, improved performance with SPEAK appeared to stem from small improvements in both envelope- and place-cue transmission ($\text{RTI env}_{[\text{plc}]}$ and $\text{RTI plc}_{[\text{env}]}$).

### Place-pitch sensitivity and consonant place-cue performance

Although six of the ten subjects who were tested with MPEAK demonstrated good or excellent place-pitch sensitivity as estimated by the electrode pitch-ranking task, only one of these six (AMB) extracted more than 20% of the available place-of-articulation information from the consonant stimuli with the MPEAK strategy. Consistent with this, correlations between place-pitch sensitivity ($d'$) and place-cue perception ($\text{RTI plc}_{[\text{env}]}$ scores) with MPEAK all failed to reach significance. It is noteworthy that the correlations between place-pitch sensitivity and $\text{RTI plc}_{[\text{env}]}$ became systematically stronger as electrode separation was increased from 0.75 to 4.5 mm for the pitch-ranking measure, since this suggests that place-pitch sensitivity within relatively narrow frequency regions may be less important to consonant recognition than place-pitch sensitivity across wider frequency distances. The left panel of Fig. 4 demonstrates the relationship between $\text{RTI plc}_{[\text{env}]}$ and place-pitch sensitivity at the 4.5-mm electrode separation where the strongest relationship was observed.

Correlations between place-pitch sensitivity and consonant place-of-articulation ($\text{RTI plc}_{[\text{env}]}$) were somewhat stronger for SPEAK than for MPEAK; however, they still did not approach statistical significance. Once again, correlations between place-pitch sensitivity and $\text{RTI plc}_{[\text{env}]}$ increased as electrode separation increased from 0.75 to 4.5 mm for the place-pitch measure, suggesting that consonant place-of-articulation cues predominantly involve spectral contrasts across relatively broad frequency regions. The strongest relationship between $\text{RTI plc}_{[\text{env}]}$ and place-pitch sensitivity was again observed for an electrode separation of 4.5 mm ($r=0.350$, $p=0.220$). This relationship is illustrated in the right panel of Fig. 4.

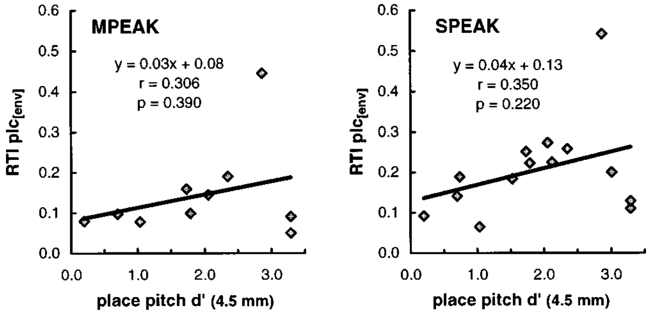

FIG. 4. Scatterplots showing average place-pitch sensitivity for electrode pairs separated by 4.5 mm versus performance on the consonant place-of-articulation feature ($\text{RTI plc}_{[\text{env}]}$). Data for the MPEAK and SPEAK speech-processing strategies are shown in the left and right panels, respectively.

There was no systematic relationship between place-pitch sensitivity at any electrode separation and improvement on the place-of-articulation feature ($\text{RTI plc}_{[\text{env}]}$) with SPEAK relative to MPEAK. Thus our hypothesis that subjects with the best place-pitch sensitivity would achieve the greatest improvements in consonant place-of-articulation performance with SPEAK was not supported by the expt. 1 data.

### Summary and explanation of findings

The 14 subjects who participated in expt. 1 demonstrated a wide range of performance on the electrode pitch-ranking task, with several subjects exhibiting excellent place-pitch sensitivity even at small electrode separations. However, only one subject (AMB) demonstrated any substantial ability to extract consonant place-of-articulation cues with either the MPEAK or SPEAK speech-processing strategy. With respect to MPEAK, these results echo the findings of Nelson *et al.* (1995) and support their conclusion that consonant place-cue information is not well represented in the MPEAK-encoded stimuli. The finding that good place-pitch sensitivity did not translate into good place-cue perception with SPEAK was more surprising, since the SPEAK strategy provides considerably more detail concerning the spectral characteristics of speech than MPEAK.

A possible explanation for subjects' poorer-than-expected place-cue performance with SPEAK was that a relatively short duration (one month) of SPEAK use was provided in our protocol. It is possible that subjects require a longer period of daily experience with SPEAK to make full use of the additional spectral information that it provides. To examine this possibility, we evaluated the effect of additional experience on consonant recognition performance in three subjects (AMB, FXC, and TVB) who upgraded permanently to SPEAK following their participation in expt. 1. Consonant recognition had been obtained several times from each of these subjects over a 12–18-month time period following

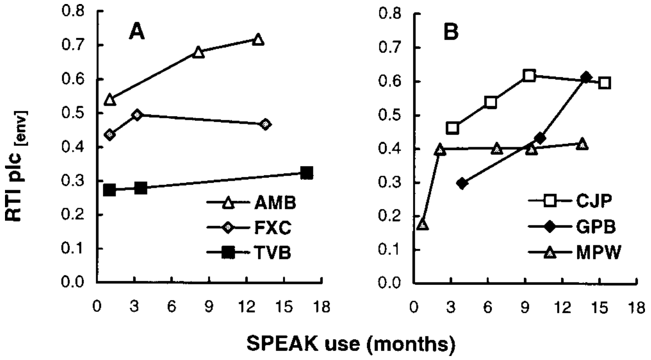

FIG. 5. Relative transmitted information for consonant place-of-articulation (RTI plc$_{[\mathrm{env}]}$) as a function of duration of SPEAK use. (a) Data for three subjects who participated in expt. 1 and subsequently upgraded to SPEAK. Data points at 1 month are the SPK-2 data from expt. 1; subsequent data points represent the number of months of continuous use following upgrade to SPEAK. (b) Data for three subjects who received SPEAK at initial hookup. Additional information concerning these subjects is provided in Table III.

their processor upgrade, in conjunction with another experiment. Scores for the consonant place-of-articulation feature (RTI plc$_{[\mathrm{env}]}$) derived from these data are shown in Fig. 5(a). It is apparent that subject AMB improved substantially in his ability to extract place-of-articulation cues as he gained additional experience with SPEAK over a time course of 12 months. In contrast, subjects FXC and TVB demonstrated relatively stable performance for RTI plc$_{[\mathrm{env}]}$ as experience with SPEAK increased from 1 month to 14 months and 17 months, respectively. It should be noted that RTI plc$_{[\mathrm{env}]}$ scores for each of these subjects were nearly identical for the MPK-1 and MPK-2 test conditions in expt. 1. This suggests that subjects had achieved stable performance with the test materials prior to time that the data in Fig. 5 were obtained, and argues against the possibility that the improvements shown for subject AMB in Fig. 5 were related to increased familiarity with the test materials rather than improved speech perception *per se*.

Corresponding data for three other subjects who received SPEAK as their first speech processor are shown in Fig. 5(b). One subject (MPW) achieved stable performance on the RTI plc$_{[\mathrm{env}]}$ feature within three months of initial hookup; however, the remaining two subjects (CJP and GPB) showed clear improvements in the place measure over the first 9 months and 14 months of use, respectively. In general, then, the data in Fig. 5 suggest that a subject's ability to extract spectral information from SPEAK-encoded speech stimuli may continue to improve with experience over the course of several months to a year, or even longer. This suggests that the relatively short period of SPEAK experience provided in expt. 1 may have been a factor in our failure to identify a relationship between place-pitch sensitivity and consonant place-cue performance. To determine whether this was the case, we reevaluated the relationship between place-pitch sensitivity and consonant recognition in expt. 2, using subjects with greater SPEAK experience.

## EXPT 2. PLACE-PITCH SENSITIVITY AND CONSONANT RECOGNITION IN EXPERIENCED SPEAK USERS

### Subjects

Expt. 2 participants were 12 adult users of the Nucleus 22-electrode implant who had used the SPEAK processing strategy on a daily basis for at least one year. As in expt. 1, all subjects were postlingually deafened and were native speakers of American English. Seven subjects in this group had upgraded to SPEAK after using the $f0/f1/f2$ or MPEAK processing strategy; the remaining five subjects had used the SPEAK strategy continuously since implant hookup. Again, all subjects provided informed consent to participate in the study and were paid on an hourly basis for their participation. Additional information concerning expt. 2 subjects is provided in Table III.

### Design and procedures

Each subject underwent consonant recognition testing and electrode pitch-ranking testing, using procedures identical to those described in expt. 1. Subjects used their usual

TABLE III. Description of 12 cochlear implant subjects who participated in expt. 2: Subject identifying code, gender, etiology of deafness (implanted ear), duration of bilateral severe-to-profound hearing loss prior to implantation, age at implantation, depth of electrode array insertion (mm from the round window, with 25 mm representing complete insertion), duration of implant use prior to the study, and duration of SPEAK use prior to consonant testing. To provide readers with an indication of subjects' clinical performance levels with SPEAK, scores for the NU-6 monosyllabic word test (% correct words and % correct phonemes) are also shown.

| Subj | m/f | Age (yrs) | Etiology of deafness | Duration (yrs) | Depth (mm) | CI use (yrs) | SPEAK use (yrs) | Nu-6 % words | Nu-6 % phons |
|---|---|---|---|---|---|---|---|---|---|
| AGF | *m* | 70 | noise exposure | 25 | 20 | 8.2 | 1.9 | 0 | 23 |
| AMB | *m* | 49 | progressive SNHL | 1 | 25 | 4.9 | 1.1 | 68 | 81 |
| CJP | *m* | 29 | maternal rubella | <1 | 23 | 1.3 | 1.3 | 70 | 87 |
| DAW | *f* | 57 | otosclerosis | 10 | 25 | 1.0 | 1.0 | 36 | 57 |
| EES | *f* | 54 | Cogan's syndrome | 4 | 17 | 8.7 | 1.8 | 12 | 40 |
| FXC | *m* | 64 | progressive SNHL | 4 | 25 | 5.7 | 1.1 | 8 | 30 |
| GPB | *m* | 57 | meningitis | <1 | 25 | 1.2 | 1.2 | 56 | 75 |
| JPB | *m* | 52 | progressive SNHL | 4 | 24 | 6.4 | 1.8 | 52 | 76 |
| MPW | *m* | 31 | genetic/progressive SNHL | <1 | 20 | 1.1 | 1.1 | 66 | 83 |
| RFM | *m* | 56 | Meniere's disease | 1 | 22 | 8.9 | 1.2 | 4 | 25 |
| TVB | *m* | 41 | progressive SNHL | 8 | 22 | 6.3 | 1.4 | 24 | 58 |
| WPS | *m* | 66 | noise exposure | <1 | 25 | 2.9 | 2.9 | 32 | 52 |

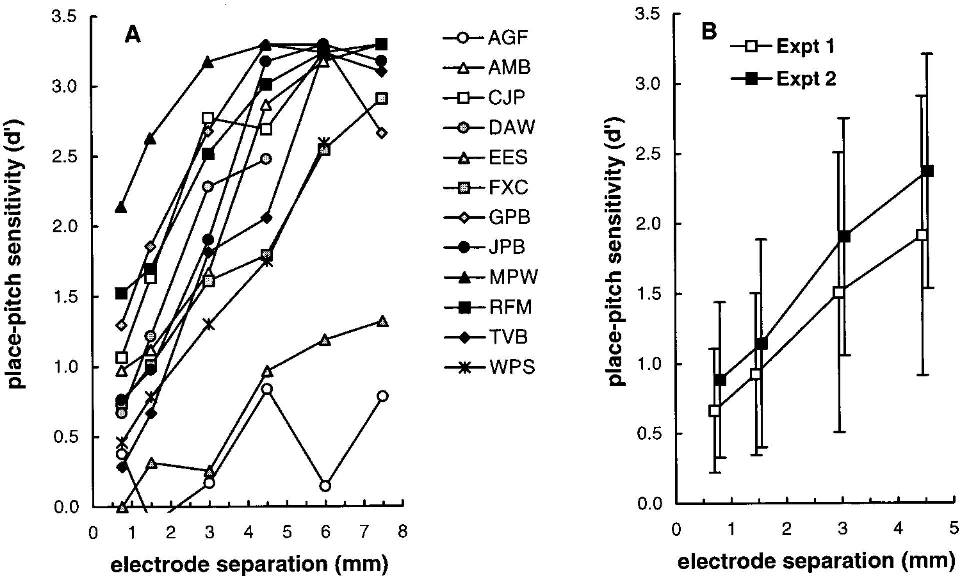

FIG. 6. (a) Average place-pitch sensitivity ($d'$) across the electrode array as a function of electrode separation, for 12 subjects in expt. 2. Perfect discrimination corresponds to $d'=3.29$. (b) Comparison of mean place-pitch sensitivity obtained by 14 expt. 1 subjects and 12 expt. 2 subjects. Error bars represent 1 s.d.

(clinical) SPEAK maps implemented on their own Spectra 22 speech processors for consonant recognition testing. Speech processor parameters were set to their default values, as in expt. 1. One subject was programmed with a bipolar separation of 0.75 mm (BP mode); all others were programmed with a bipolar separation of 1.5 mm (BP+1). As in expt. 1, bipolar separations used in the electrode pitch-ranking task matched those used in subjects' SPEAK maps.

## Results and discussion

### *Electrode pitch ranking*

Figure 6(a) shows average place-pitch sensitivity as a function of electrode separation for each of the 12 subjects who participated in expt. 2. Data for subjects AMB, FXC, RFM, and TVB, who participated in expt. 1, are replotted from Fig. 1. Subjects demonstrated a wide range of place-pitch sensitivity, similar to that observed in expt. 1. One subject (MPW) demonstrated exceptionally good pitch-ranking performance, particularly for narrow electrode separations; two others (EES and AGF) exhibited unusually poor pitch-ranking performance at all electrode separations. As noted later, both of these subjects also demonstrated small electrical dynamic ranges. Figure 6(b) compares mean electrode pitch-ranking performance for subjects in expts. 1 and 2. Average place-pitch sensitivity was somewhat higher for the expt. 2 participants at all electrode separations; however, intersubject variability was similar for the two groups.

### *Consonant recognition*

Figure 7 summarizes consonant recognition performance for individual subjects in expt. 2. Subjects are ordered along the $x$-axis according to their performance on the stimulus feature. As in expt. 1, subjects demonstrated a wide range of overall performance and achieved considerably higher transmitted information scores for the envelope feature ($\mathrm{RTI\,env}_{[\mathrm{plc}]}$) than for the place feature ($\mathrm{RTI\,plc}_{[\mathrm{env}]}$). Average performance for the stimulus feature increased 27% relative to the SPK-2 performance obtained in expt. 1 (44.6% for expt. 1 versus 56.8% for expt. 2). Corresponding increases for the envelope and place features were 20.2% and 76.4%, respectively, suggesting that additional experience with SPEAK had a considerably greater impact on place-cue performance than on envelope-cue performance. Related to this, 8 of 12 subjects in this experiment exhibited $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ scores greater than 30%, as compared to a single subject (AMB) in expt. 1.

### *Place-pitch sensitivity and place-cue perception*

In contrast to the findings of expt. 1, a positive relationship was observed between place-pitch sensitivity and consonant place-cue perception ($\mathrm{RTI\,plc}_{[\mathrm{env}]}$). As in the first experiment, correlations between place-pitch sensitivity and $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ became stronger as electrode separation increased from 0.75 to 4.5 mm. They approached statistical significance at electrode separations of 1.5 and 3.0 mm ($r=0.535$, $p=0.07$ and $r=0.563$, $p=0.06$, respectively) and reached statistical significance at an electrode separation of 4.5 mm ($r=0.711$, $p<0.05$). A scatterplot of $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ versus average place-pitch sensitivity at 4.5-mm electrode separation is shown in Fig. 8. Note that there was no systematic relationship between place-pitch sensitivity for adjacent electrodes (0.75-mm electrode separation) and place-cue perception, even though subjects demonstrated a wide range of place-pitch sensitivity for this condition. This suggests that fine spectral resolution is not necessary for consonant place-cue discrimination with the SPEAK strategy.

### *Envelope- versus place-cue performance*

Another result that emerged from the expt. 2 data was a strong positive relationship between the envelope- and place-cue measures. This relationship is implied in Fig. 3, where it can be seen that subjects who achieved the highest scores on the envelope feature ($\mathrm{RTI\,env}_{[\mathrm{plc}]}$) also tended to achieve the highest scores on the place feature ($\mathrm{RTI\,plc}_{[\mathrm{env}]}$.) Figure 9 demonstrates the relationship between $\mathrm{RTI\,env}_{[\mathrm{plc}]}$ and $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ more directly. Here $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ is plotted as a function of $\mathrm{RTI\,env}_{[\mathrm{plc}]}$ and a linear function fitted to the plotted data indicates that $\mathrm{RTI\,plc}_{[\mathrm{env}]}$ increases as a constant

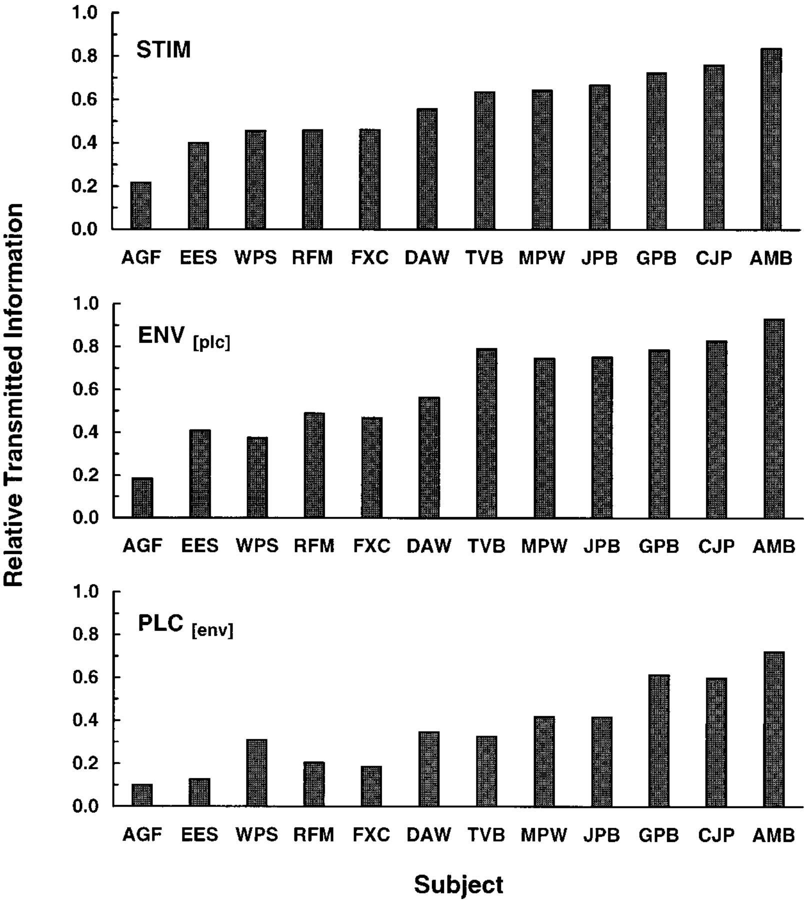

FIG. 7. Consonant recognition data for 12 experienced users of the SPEAK strategy. Relative transmitted information (RTI) values for stimulus, envelope (conditional on place), and place (conditional on envelope) are shown in the top, middle, and bottom panels, respectively.

proportion ($\sim$0.8) of RTI $\text{env}_{[\text{plc}]}$. This finding suggests that both envelope and place-cue measures may depend on a common underlying ability, possibly the ability to detect amplitude fluctuations in the envelope of the speech waveform. A mutual dependence of envelope-cue and place-cue perception on envelope following seems reasonable given that envelope-cue perception involves the detection of amplitude fluctuations in a single-channel representation of the stimulus and place-cue perception involves the detection of patterns of amplitude fluctuations across several frequency channels or electrodes. In this sense, RTI $\text{env}_{[\text{plc}]}$ can be viewed as a measure of overall envelope-processing ability, independent of place-pitch sensitivity (channel separation), and RTI $\text{plc}_{[\text{env}]}$ can be viewed as a measure of subjects' combined ability to resolve envelope fluctuations and to utilize place-pitch cues.

To get a better feel for the relationship between place-pitch sensitivity, envelope following, and place-cue perception, it is helpful to consider the data for individual subjects shown in Figs. 6(a) and 7. In general, these individual data support the contention that subjects' place-cue perception reflects both their envelope-processing abilities (as indicated by scores on the consonant envelope-cue measure) and their place-pitch sensitivity. First, consider the five subjects who achieved the highest scores on the consonant place-cue measure RTI $\text{plc}_{[\text{env}]}$: MPW, JPB, GPB, CJP, and AMB (Fig. 7, bottom panel). Each of these subjects demonstrated high envelope-cue scores (Fig. 7, middle panel), suggesting that they possess good envelope-processing abilities. In addition, each exhibited good or excellent place-pitch sensitivity at all electrode separations. For these subjects, then, it appears that the combination of good envelope-processing skills and good place-pitch sensitivity permitted relatively high performance on the consonant place-cue feature. There is no obvious explanation for differences in place-cue performance within this group, in particular to explain the fact that GPB, CJP, and AMB achieved higher place-cue scores (60%–72%) than MPW and JPB (42%). However, it is noteworthy that subject AMB, who achieved the highest score on the place-cue measure (72%), was distinguished from the others in this group by his very high score on the envelope-cue measure (93%) but not by his place-pitch sensitivity, which was second poorest among the subjects in this group at narrow electrode separations (0.75–3.0 mm). This supports our impression that envelope-processing is at least as important as place-pitch sensitivity in determining consonant place-cue perception.

The next group of five subjects in the bottom panel of Fig. 7 (WPS, RFM, FXC, DAW, and TVB) exhibited a reduced ability to extract consonant place cues (RTI $\text{plc}_{[\text{env}]}$ scores between 19% and 35%). One of these subjects, TVB, demonstrated excellent envelope-cue performance but only moderate place-pitch sensitivity, suggesting that place-cue extraction may have been limited by place-pitch sensitivity rather than envelope-processing ability. Two others in this group, RFM and DAW, demonstrated excellent place-pitch sensitivity but only moderate envelope-cue performance. For these subjects, envelope-processing skills rather than place-

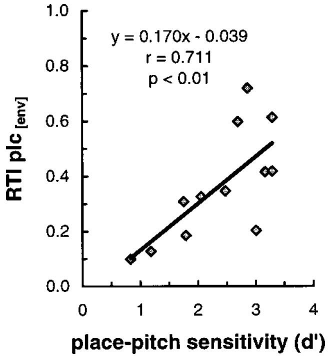

FIG. 8. Relative transmitted information for consonant place-of-articulation (RTI $\mathrm{plc}_{[\mathrm{env}]}$) as a function of average place-pitch sensitivity ($d'$) for electrode pairs separated by 4.5 mm. Data are for 12 experienced users of the SPEAK strategy.

pitch sensitivity may have been the factor limiting place-cue extraction. The remaining subjects in this group, WPS and FXC, demonstrated low-to-moderate envelope cue scores and only moderate place-pitch sensitivity, suggesting that both envelope-processing skills and place-pitch sensitivity may have limited place-cue performance.

The final two subjects shown in Fig. 7, AGF and EES, were unable to extract meaningful amounts of place-cue information from the consonant stimuli (RTI $\mathrm{plc}_{[\mathrm{env}]}$ scores <13%). These subjects demonstrated unusually poor place-pitch sensitivity, in addition to poor (AGF) or moderately poor (EES) performance on the envelope-cue measure. In effect, it appears that these subjects had little chance of extracting spectral information from the consonant stimuli, given the dual limitations of poor envelope processing and poor place-pitch sensitivity.

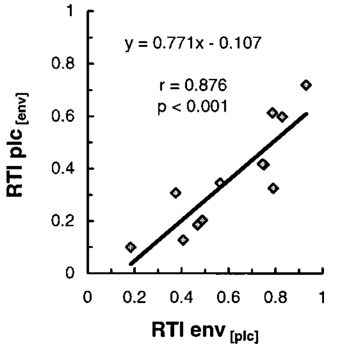

FIG. 9. Spectral versus envelope measures of consonant recognition (RTI $\mathrm{plc}_{[\mathrm{env}]}$) versus (RTI $\mathrm{env}_{[\mathrm{plc}]}$) for 12 experienced users of the SPEAK strategy. A strong linear relationship between the two measures is observed, as indicated by the regression fit (heavy line).

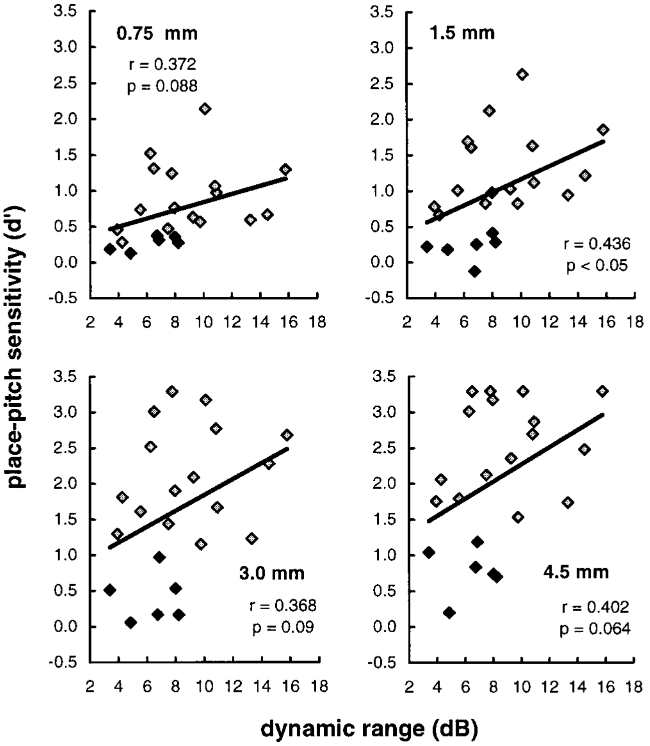

FIG. 10. Place-pitch sensitivity ($d'$) at four electrode separations (0.75, 1.5, 3.0, and 4.5 mm) as a function of dynamic range (dB $\mu$A). Dynamic range data were taken from measurements obtained for the electrode pitch-ranking procedure, and represent average values across electrodes for 205-$\mu$s/ph, 125-Hz, 500-ms pulse trains. Data are shown for all 22 subjects who participated in expts. 1 and 2. Filled symbols indicate data for six subjects who demonstrated the poorest place-pitch sensitivity (AGF, BRL, EES, JMS, KRK, and LMF).

In general, then, the individual data support the notion that both envelope-processing abilities and place-pitch sensitivity are prerequisites for place-cue extraction with the SPEAK strategy. It appears that both factors are necessary and that neither alone is sufficient for achieving good consonant place-cue perception. An important implication of these findings is that spectral information cannot provide an alternative source of information about consonant identity in those cochlear implant subjects who obtain limited temporal information. This suggests that improved strategies for encoding spectral speech features will be most effective among listeners who demonstrate good temporal-cue recognition. It further suggests that efforts to improve consonant perception among poorer-performing subjects (who have limited recognition for both temporal and spectral cues) should focus on improving subjects' perception of temporal cues.

### *Dynamic range versus place-pitch sensitivity and consonant recognition*

Figure 10 shows place-pitch sensitivity at each of four electrode separations as a function of dynamic range for 22 subjects who participated in expts. 1 and 2. There is a clear trend for place-pitch sensitivity to increase with dynamic range at each electrode separation, although the correlation between place-pitch sensitivity and dynamic range is statistically significant only at an electrode separation of 1.5 mm (upper right panel). Note that the six subjects with the poorest place-pitch sensitivity across electrode separations (filled symbols) all possess average dynamic ranges less than 8 dB. This indicates that subjects with small dynamic ranges may be most ''at risk'' for poor place-pitch sensitivity and, therefore, poor consonant place-cue perception.

Relationships between dynamic range and the consonant recognition measures RTI $\mathrm{env}_{[\mathrm{plc}]}$ and RTI $\mathrm{plc}_{[\mathrm{env}]}$ are illustrated in Fig. 11 for expt. 2 participants. Subjects' scores for the envelope feature (RTI $\mathrm{env}_{[\mathrm{plc}]}$) tended to increase with dynamic range (panel A). However, linear regression failed

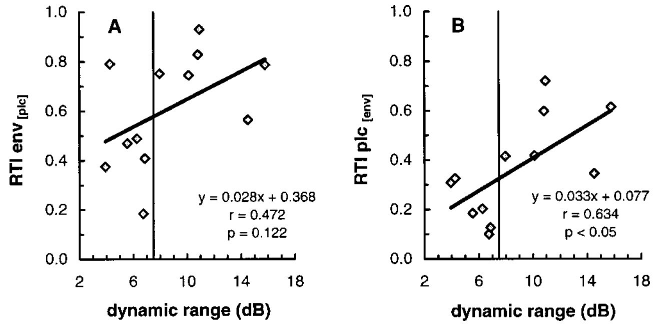

FIG. 11. (a) RTI $\mathrm{env}_{[\mathrm{plc}]}$ as a function of dynamic range (dB $\mu$A) for 12 experienced users of the SPEAK strategy. Dynamic range data represent the average values across electrodes for 125-Hz, 205-$\mu$s/ph, 500-ms pulse trains. The vertical line indicates a dynamic range of 7.5 dB (see text). (b) As in a, but showing RTI $\mathrm{plc}_{[\mathrm{env}]}$ versus dynamic range.

to yield a statistically significant regression coefficient, primarily due to the performance of one subject (TVB) who exhibited a small average dynamic range (3.9 dB $\mu$A) but good envelope-cue performance (79.0%). Note that five of six subjects with average dynamic ranges less than 7 dB exhibited envelope scores less than 50% (data points to the left of the vertical line), whereas all six subjects with dynamic ranges greater than 7 dB achieved scores greater than 50% (points to the right of the vertical line). Place-cue performance also increased with dynamic range for these 12 subjects (panel B) and linear regression of RTI $\mathrm{plc}_{[\mathrm{env}]}$ on dynamic range yielded a significant correlation coefficient ($r=0.634$, $p<0.05$). Again, there was a clear performance difference between subjects whose dynamic ranges were 7 dB and smaller versus those whose dynamic ranges were 8 dB or greater. This finding, that subjects with small dynamic ranges perform poorly on place-cue recognition, is not altogether surprising, since these subjects (with one exception) demonstrated low scores for the envelope-cue measure [Fig. 11(a)] and since a tight coupling was previously demonstrated between envelope- and place-cue performance (Fig. 9).

### *Summary*

The 12 subjects who participated in expt. 2 had considerably greater experience with the SPEAK speech-processing strategy (1.0–2.9 years) than the subjects in expt. 1 (1 month), and also demonstrated considerably better consonant place-cue performance. Mean place-pitch sensitivity for the expt. 2 subjects was slightly higher than that observed for expt. 1 subjects, but the range of performance was similar. In contrast to the expt. 1 results, a significant relationship was observed between place-pitch sensitivity and place-cue performance. This supports our hypothesis that the detailed spectral information provided by the SPEAK encoding strategy is best utilized by subjects with good place-pitch sensitivity. A strong linear relationship was observed between individual subjects' scores on the consonant envelope- and place-cue measures, RTI $\mathrm{env}_{[\mathrm{plc}]}$ and RTI $\mathrm{plc}_{[\mathrm{env}]}$. This suggests that place-cue perception depends both on envelope-processing abilities and place-pitch sensitivity. Finally, the combined electrode pitch-ranking data from expts. 1 and 2 show that poor place-pitch sensitivity is most likely to occur in subjects with small average dynamic ranges ($<$ 8 dB for the test stimulus used here), indicating that subjects with small dynamic ranges are most ''at risk'' for poor consonant recognition.

## GENERAL DISCUSSION

The present results indicate that the SPEAK processing strategy can provide meaningful levels of consonant spectral-cue transmission in cochlear implant listeners who possess adequate place-pitch sensitivity and envelope-processing skills. They also confirm earlier findings of Nelson *et al.* (1995) and Parkinson *et al.* (1996) which showed that the MPEAK strategy provides very limited consonant spectral-cue information. Prior to summarizing our conclusions from the present research, we would like to comment briefly on several issues related to the effects of experience on speech recognition performance and place-pitch sensitivity, and concerning the generalizability of the present findings to other speech stimuli and listening conditions.

### Effects of experience

Results from our first experiment highlight the importance of experience with a new speech-processing strategy as a factor in speech recognition performance. Several studies have shown that performance increases over time following a change in speech-processor configuration (Tyler *et al.*, 1986; Parkinson *et al.*, 1996; Rosen *et al.*, 1998) and suggest that improvements may depend on the nature of the processing change as well as the speech materials being evaluated. However, many important issues concerning the effects of experience on implant speech recognition have not been addressed. For example, it is not known whether the time course of improvement to asymptotic performance for consonant temporal cues is more rapid than that for spectral cues, as suggested by the present data, or whether rates of improvement vary as a function of speech materials or speech-processing strategy. Such issues have considerable importance for the evaluation and design of speech-processing strategies and also for the evaluation of post-implant speech recognition performance in individual patients. Thus it will be important to address them in future studies. It should be noted that, although experience was a limiting factor in the present experiments, it may or may not have similar importance in other experiments involving cochlear implant speech recognition.

A possible weakness of the present study is that it did not evaluate the influence of SPEAK experience on place-pitch sensitivity as measured with our electrode pitch-ranking task. Since the SPEAK strategy provides an enhanced representation of spectral speech cues, it is possible that use of this strategy could sharpen the spatial resolution of neural responses in the electrically stimulated auditory system. Such improvements could parallel or underlie the improvements in consonant place-cue perception observed in Fig. 5 as a function of increasing experience with the SPEAK strategy. In the present experiments, electrode pitch-ranking testing was most often performed prior to consonant identification testing with SPEAK; thus if SPEAK experience served to improve place-pitch sensitivity, electrode pitch-ranking measures would have underestimated place-pitch sensitivity at the time of consonant testing. We suspect that this would have been a small effect, since longitudinal measures of place-pitch sensitivity that we have obtained in several SPEAK subjects indicate minimal or no improvement over time. Nonetheless, such effects could exist in some subjects. It is important to note that underestimation of place-pitch sensitivity would not have altered the study's major finding related to place-pitch sensitivity, i.e., that most subjects have moderate or good place-pitch sensitivity but that only a subset of these achieve meaningful amounts of consonant place-cue perception.

### Generalizability of the present findings

The present research evaluated subjects' use of spectral cues under optimal listening conditions (moderately loud stimuli in quiet) and for a particular subset of speech stimuli (consonants). Thus we can only speculate as to whether findings would generalize to other speech stimuli or would hold true under less optimal listening conditions. The dependence of spectral cue transmission on envelope processing in addition to place-pitch sensitivity may apply uniquely to consonant stimuli, owing to their brief durations and their low intensities relative to the vowel segments of speech. Vowel stimuli may depend more strongly on place-pitch sensitivity alone, such that subjects with poor performance on consonant envelope cues (indicating poor envelope-processing abilities) may be able to achieve high levels of vowel recognition on the basis of spectral information only. A recent study by Fishman *et al.* (1997) suggests that fine spectral resolution may be more important for vowel recognition than for consonant recognition. This is generally consistent with the present finding that consonant place-cue performance is more strongly related to place-pitch sensitivity at wider electrode separations than at narrower ones. It also suggests that relationships between place-pitch sensitivity and spectral-cue performance might exhibit a different pattern for vowel stimuli than for consonant stimuli. With respect to the issue of listening conditions, recent research has suggested that fine spectral resolution is more important to speech recognition in noise than in quiet (Delhorne *et al.*, 1997; Dorman *et al.*, 1997). Consistent with this, subjects in the Skinner *et al.* (1994) study showed the largest speech recognition improvements with SPEAK relative to MPEAK for sentence materials presented in a background of speech babble. In general, we would expect place-pitch sensitivity at narrow electrode separations to be important for perception of spectral cues to vowel identity and for the perception of consonant spectral cues under unfavorable signal-to-noise conditions. Additional research is needed to evaluate these predictions.

### Clinical implications

The present findings indicate that moderate- and poor-performing SPEAK users rely almost exclusively on temporal cues to consonant identity, most often because they have limited envelope-processing abilities but occasionally because they possess inadequate place-pitch sensitivity. Consonant discrimination does not appear to rely strongly on fine spectral resolution, thus place-pitch sensitivity is probably not the factor limiting consonant place-cue perception in most of these listeners. This suggests that attempts to improve spatial resolution, for example by excluding indiscriminable electrodes from subjects' maps (Zwolan *et al.*, 1997), are unlikely to result in improved consonant recognition among poorer performers. Instead, it may be more beneficial to focus on enhancing temporally based cues. There are currently no well-defined strategies for improving consonant temporal-cue recognition in cochlear implant listeners. However, it is possible that the use of specific stimulus parameters and speech-processing strategies could enhance the transmission of such cues for some individuals. Future research should evaluate this possibility.

## CONCLUSIONS

(1) Spectral cues to consonant identity are poorly represented by the MPEAK speech-encoding strategy. Even experienced cochlear implant subjects with excellent place-pitch sensitivity and good envelope-cue performance exhibit very limited consonant place-cue perception with MPEAK.

(2) Experience is an important factor influencing consonant recognition performance with SPEAK. A subjects' ability to make use of the spectral information provided by the SPEAK strategy may improve substantially over the first few months of daily use.

(3) Cochlear implant listeners' ability to extract spectral cues from consonant stimuli with the SPEAK strategy depends on both place-pitch sensitivity and envelope-processing ability. Place-cue performance may be limited by either one of these factors, or both, in a given individual.

(4) Cochlear implant subjects with small electrical dynamic ranges are considerably more likely than other subjects to exhibit poor place-pitch sensitivity and poor consonant recognition with the SPEAK strategy.

(5) Attempts to increase consonant recognition among poorer-performing cochlear implant listeners should focus primarily on improved transmission of temporally based (envelope) cues. Better-performing subjects are more likely to benefit from improved transmission of spectral (place-of-articulation) cues.

## ACKNOWLEDGMENTS

This research was supported by NIDCD Grant No. DC00110 and the Lions 5M International Hearing Foundation. Preliminary findings were presented at the 19th Midwinter Research Meeting of the Association for Research in Otolaryngology (February 1996). The authors wish to thank Dianne Van Tasell for providing the digitized stimuli used in consonant recognition experiments; Cochlear Corporation for providing two loaner Spectra processors for use in expt. 1; John Van Essen for writing the computer programs used for the electrode pitch-ranking and consonant recognition procedures; Sharon Smith for creating SPEAK maps for the expt. 1 subjects; and Tanya Grann and Tara Khetrapal for assisting in data collection and data analysis. We also thank Dr. Rich Tyler and an anonymous reviewer for their helpful comments on an earlier version of this manuscript. Finally, we would like to express our sincere appreciation to the 22 cochlear implant subjects who participated in this research project.

[1] Ten of 14 subjects completed electrode pitch-ranking (EPR) testing within the three months spanned by their consonant recognition test sessions. One additional subject underwent EPR testing six months following his completion of consonant testing, due to scheduling constraints. The remaining three subjects had completed electrode pitch-ranking testing as part of an earlier study (Nelson *et al.*, 1995). These subjects were not retested on the pitch-ranking task as part of this experiment, but were tested on the electrode pitch-ranking task again about a year following the conclusion of this study. In each case, pitch-ranking performance was similar for the earlier and later test points, and the earlier data were used in the present analyses.

Busby, P. A., Tong, Y. C., and Clark, G. M. (**1993**). ‘‘Electrode position, repetition rate, and speech perception by early- and late-deafened cochlear implant patients,’’ J. Acoust. Soc. Am. **93**, 1058–1067.

Cochlear Corporation (**1996**). Technical reference manual: Nucleus 22 channel cochlear implant system.

Delhorne, L. A., Eddington, D. K., Noel, V. A., Rabinowitz, W. M., Tierney, J., and Whearty, M. E. (**1997**). ‘‘Speech reception with CIS processing: Longitudinal evaluations and implications from acoustic simulations,’’ presented at the 1997 Conference on Implantable Auditory Prostheses, Pacific Grove, California (unpublished).

Dorman, M. F. (**1993**). ‘‘Speech recognition by adult users of cochlear implants,’’ in *Cochlear Implants: Audiological Foundations*, edited by R. Tyler (Singular, San Diego), pp. 145–190.

Dorman, J. F., Loizou, P. C., Shannon, R. V., and Brill, S. M. (**1997**). ‘‘Converging evidence from simulations and patients on the number of channels necessary for high levels of speech understanding,’’ presented at the 1997 Conference on Auditory Prostheses, Pacific Grove, California (unpublished).

Fishman, K. E., Shannon, R. V., and Slattery, W. H. (**1997**). ‘‘Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor,’’ J. Speech Hear. Res. **40**, 1201–1215.

Hacker, M. J., and Ratcliffe, R. (**1979**). ‘‘A revised table of $d'$ for *M*-alternative forced choice.’’ Percept. Psychophys. **26**, 168–170.

Miller, G. A., and Nicely, P. E. (**1955**). ‘‘An analysis of perceptual confusions among some English consonants,’’ J. Acoust. Soc. Am. **27**, 338–352.

Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., and Levine, S. (**1995**). ‘‘Electrode ranking of ‘place-pitch’ and speech recognition in electrical hearing,’’ J. Acoust. Soc. Am. **98**, 1987–1999.

Parkinson, A. J., Tyler, R. S., Woodworth, G. G., Lowder, M. W., and Gantz, B. J. (**1996**). ‘‘A within-subject comparison of adult patients using the Nucleus F0F1F2 and F0F1F2B3B4B5 speech processing strategies,’’ J. Speech Hear. Res. **39**, 261–277.

Pearsons, K. S., Bennett, R. L., and Fidell, S. (**1976**). ‘‘Speech levels in various environments,’’ Bolt Beranek and Newman Report 321, Canoga Park, CA.

Rosen, S., Faulkner, A., and Wilkinson, L. (**1998**). ‘‘Perceptual adaptation by normal listeners to upward shifts of spectral information in speech and its relevance for users of cochlear implants,’’ Abstracts of the 21st Mid-winter Research Meeting, Association for Research in Otolaryngology, p. 69.

Shannon, R. V., Adams, D. D., Ferrel, R. L., Polumbo, R. L., and Grandgenett, M. (**1990**). ‘‘A computer interface for psychopysical and speech research with the Nucleus cochlear implant,’’ J. Acoust. Soc. Am. **87**, 905–907.

Skinner, M. W., Clark, G. M., Whitford, L. A., Seligman, P. M., Staller, S. J., Shipp, D. B., Shallop, J. K., Everingham, C., Menapace, C. M., Adrndt, P. L., Antogenelli, T., Brimacombe, J. A., Pijl, S., Daniels, P., George, C. R., McDermott, H. J., and Beiter, A. L. (**1994**). ‘‘Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system,’’ Am. J. Otol. **15**, 15–27.

Skinner, M. W., Holden, L. K., Holden, T. A., Demorest, M. E., and Fourakis, M. S. (**1997**). ‘‘Speech recognition at simulated soft, conversational and raised-to-loud vocal efforts by adults with cochlear implants,’’ J. Acoust. Soc. Am. **101**, 3766–3782.

Townshend, B., Cotter, N., Van Compernolle, D., and White, R. L. (**1987**). ‘‘Pitch perception by cochlear implant subjects,’’ J. Acoust. Soc. Am. **82**, 106–115.

Tyler, R. S., Gantz, B. J., Woodworth, G. G., Parkinson, A. J., Lowder, M. W., and Schum, L. K. (**1996**). ‘‘Initial independent results with the Clarion cochlear implant,’’ Ear Hear. **17**, 528–536.

Tyler, R. S., Preece, J. P., Lansing, C. R., Otto, S. R., and Gantz, B. J. (**1986**). ‘‘Previous experience as a confounding factor in comparing cochlear-implant processing schemes,’’ J. Speech Hear. Res. **29**, 282–287.

Van Tasell, D. J., Greenfield, D. G., Logemann, J. L., and Nelson, D. A. (**1992**). ‘‘Temporal cues for consonant recognition: Training, talker, generalization and the use in evaluation of cochlear implants,’’ J. Acoust. Soc. Am. **92**, 1247–1257.

Zwolan, T. A., Collins, L. M., and Wakefield, G. H. (**1997**). ‘‘Electrode discrimination and speech recognition in postingually deafened cochlear implant subjects,’’ J. Acoust. Soc. Am. **102**, 3673–3685.