Appendix 1: Summary of reporting activity for this quarter

---

The main objective of this project is to design, develop, and evaluate speech processors for implantable auditory prostheses. Ideally, such processors will represent the information content of speech in a way that can be perceived and utilized by implant patients. An additional objective is to record responses of the auditory nerve to a variety of electrical stimuli in studies with patients. Results from such recordings can provide important information on the physiological function of the nerve, on an electrode-by-electrode basis, and also can be used to evaluate the ability of speech processing strategies to produce desired spatial or temporal patterns of neural activity.

- Studies with Clarion subject MI-5 (April 13). The main purpose of these studies was to characterize the current sources in the Clarion implant with recordings of scalp potentials produced with different commanded levels of pulse amplitude. Speech reception with alternative choices of parameter values for CIS processors also was measured.
- Ongoing studies with Ineraid subject SR2, who now is working with us for one or two days

, 1997). The studies have included experiments to identify the most important regions of the mapping function for speech reception. In broad terms, the results have shown or suggested that (a) the upper part of the mapping function, corresponding to relatively high envelope signals and relatively high loudnesses, is more important than the lower part of the mapping function, (b) clipping of relatively high envelope signals should be strictly avoided, (c) relatively weak consonant sounds such as the bursts of plosives need to be mapped well above auditory threshold in order to be recognized in a speech context, and (d) manipulations in the amount of compression produce results that often depend on the subject and the tests used to evaluate speech reception performance. In some cases, manipulations over a wide range exert almost no effect on speech reception scores, whereas in other cases, with other subjects or other tests, the same manipulations can produce large differences in

Recently, Fu and Shannon have published results from a study in which effects of changes in the amount of compression were evaluated in tests with three subjects (Fu and Shannon, 1998). The subjects were users of the Nucleus-22 implant. They listened to simulations of 4-channel CIS processors whose outputs were presented to the transmitting coil of the Nucleus implant with a custom interface system. The exponent in the power function used for mapping was varied in nine steps from a highly compressive mapping (exponent of 0.05) to an almost linear mapping (exponent of 0.75). They found only a mild dependence of performance on the value of the exponent over a broad range of exponents (exponents from 0.1 to 0.5 for consonant identification, and exponents from 0.1 to

This result was somewhat surprising to us, so we undertook a similar study using subjects with percutaneous access to their electrode arrays. The tests with one of our two subjects also included presentations of speech tokens in conjunction with noise, allowing us to evaluate possible interactions between manipulations in the mapping function and the speech-to-noise ratio of the test items.

Ineraid subjects SR2 and SR9 participated in the present studies. SR2 enjoys extremely good results with his implant and a CIS processor, whereas SR9 has much lower speech reception scores with her implant and a CIS processor. Both are highly experienced subjects and both have used their CIS

The mapping function used in standard implementations of CIS processors is of the form

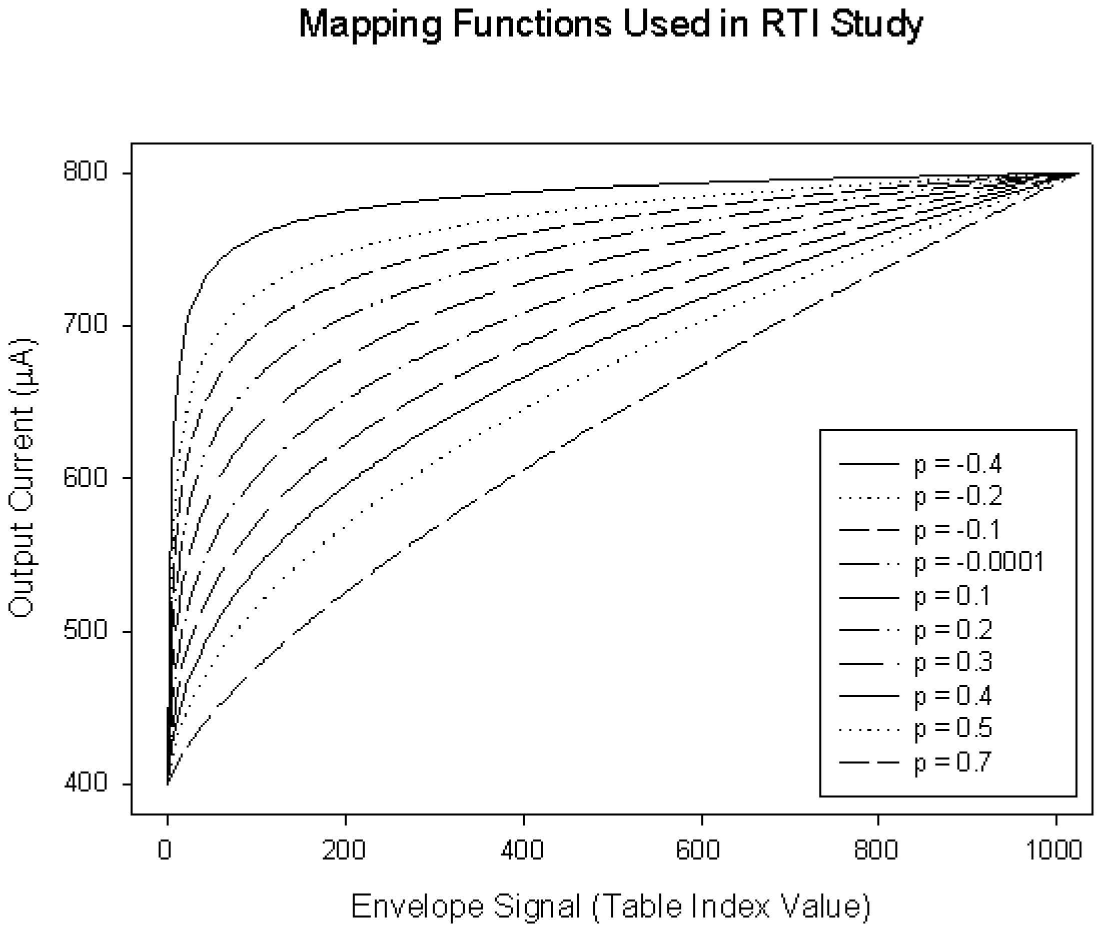


*Fig. 1. Mapping functions used in the RTI study.*

All processors for subject SR2 used four channels, whose outputs were delivered to the four most-apical electrodes in SR2's implant. Additional parameters that were held constant across processors th order bandpass filters, a 350 to 9500 Hz range of frequencies spanned by the bandpass filters, half wave rectifiers in the envelope detectors, 8 order lowpass filters with a corner frequency at 160 Hz in those detectors, a pulse rate of 500 pulses/s/electrode, a pulse duration of 18 µs/phase, and a "staggered" sequence of electrode stimulation. Most of these parameters were selected to approximate or match those used in the Fu and Shannon study. One exception was the pulse duration, which was much shorter in our study. The choice of a reduced number of channels for subject SR2 also was made to reduce his overall level of performance to a sensitive range for speech reception

All processors for subject SR9 used six channels, whose outputs were delivered to the six intracochlear electrodes in her Ineraid implant. The order of the bandpass filters, the range of frequencies spanned by the filters, the characteristics of the lowpass filters in the envelope detectors, and the update order were the same as those used in the processors for SR2. The pulse rate used for SR9 was 417 pulses/s/electrode, and the pulse duration used for SR9 was 33 µs/phase. In addition, a full wave rectifier was used in the envelope detectors. The choices of channel number, pulse rate, and pulse duration for SR9 were made to obtain an overall level of performance in a sensitive range for

The exponent in the mapping function was varied between -0.4 and 0.7 in ten processors tested with

The processors for both subjects were evaluated with our standard test of consonant identification, with each of the consonants presented in an /a/-consonant-/a/ context in randomized orders and with multiple exemplars from one male and one female talker. Twenty four different consonants were used in the tests with SR2, and 16 were used in the tests with SR9. The higher number was used for SR2 to bring his scores down into a sensitive range. All tests were conducted with hearing alone, and no feedback was given as to correct or incorrect responses. At least 10 replications of each consonant were used in the test for each condition and for each of the talkers.

The consonants were presented in quiet for both subjects and also in noise for subject SR2. CCITT noise was used, which has a spectrum that matches the long-term spectrum of speech. The additional conditions for SR2 included the speech-to-noise ratios of +15 and +10 dB.

The processors for SR2 also were evaluated with tests of vowel identification in a /h/-vowel-/d/ context. Multiple exemplars of each of eight vowels recorded from a male talker were included in these tests. As with the consonant tests, the vowel tests used randomized orders and were conducted with hearing alone and no feedback as to correct or incorrect responses. The vowel tests included presentation of the vowels in quiet and at the speech-to-noise ratios of 15 and 10 dB. Fifteen replications of each vowel were used in the test for each condition.

## Results

***Percent-correct scores for consonants.*** Percent-correct scores from the tests of consonant identification are presented in Figs. 2 and 3. Figure 2 shows the scores on a scale of 0 to 100, and Fig. 3 shows the scores on a scale of 40 to 90. Scores for subject SR2 are presented in the left column of each figure, and scores for SR9 are presented in the right column.

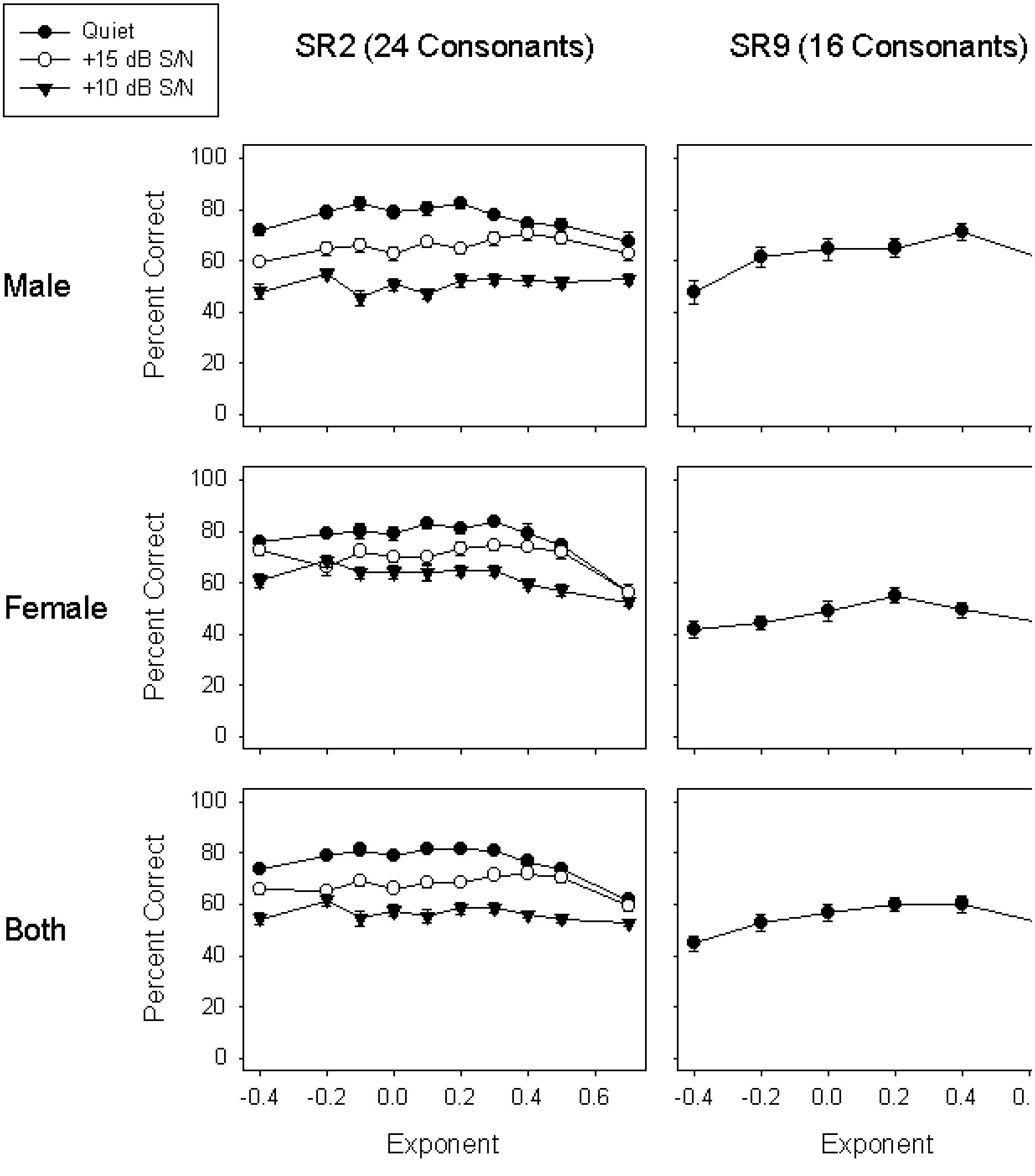


*Fig 2. Consonant identification for various exponents in the mapping function used for CIS processors, subjects SR2 and SR9. The exponent of -0.4 produces a highly compressive mapping function, and the exponent of 0.7 produces an almost-linear function (see Fig. 1). The exponent of -*

*0.0001 produces a function that closely approximates the logarithmic function used in standard CIS*

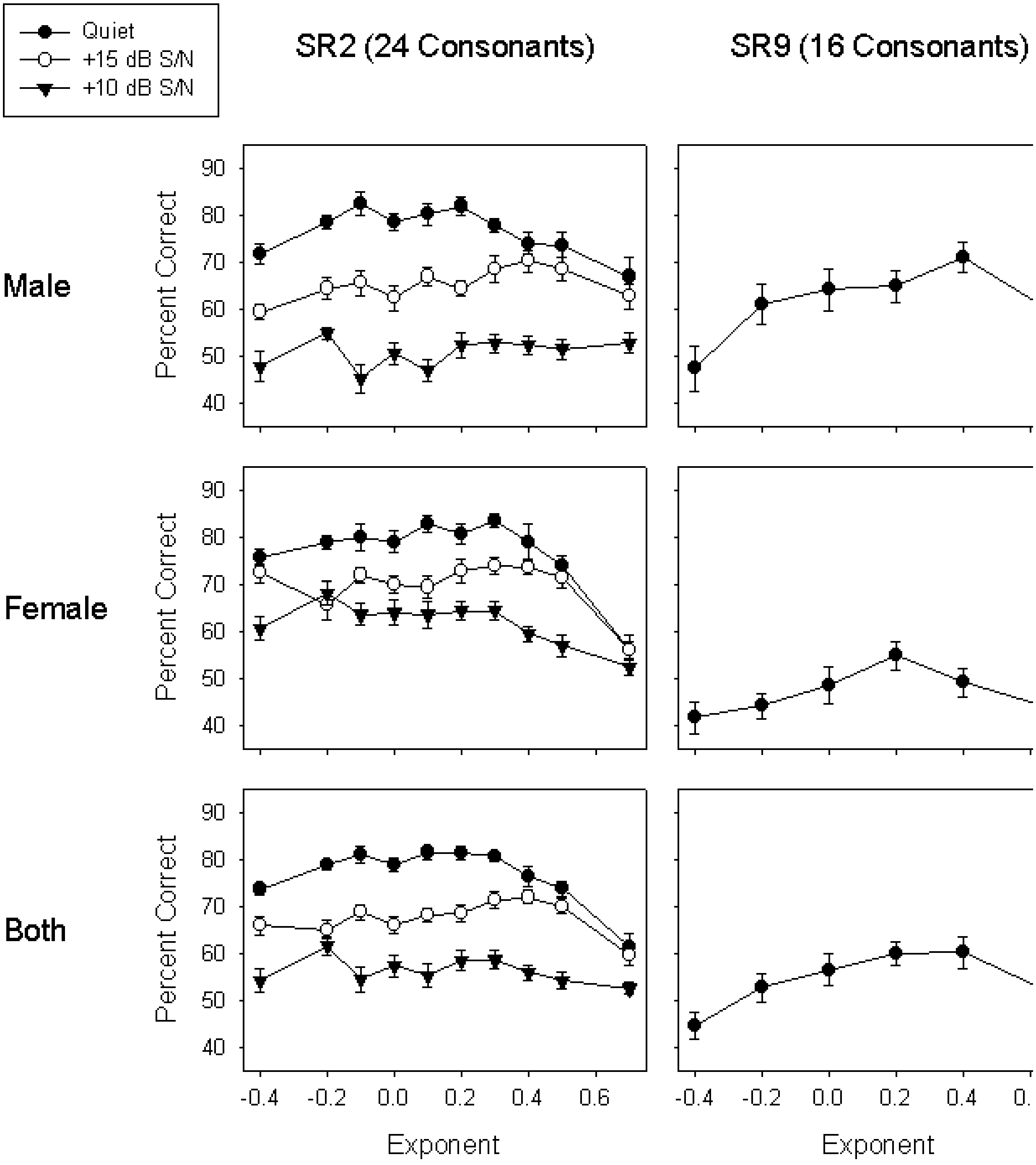


*Fig. 3. Same as Fig. 2, but with a different scale for percent correct.*

comparisons among percent correct scores for subject SR2.

| | | | *Post hoc* test | Significant differences (numbers refer to |
|---|---|---|---|---|
| | | | | -0.1 and 0.2 0.7, -0.4 0.1 0.7 |
| | | | | |
| | | | | |
| | | | | -0.4 through 0.5 0.7 |
| | | | | -0.4 and -0.1 through 0.5 0.7 |
| | | | | -0.2 through 0.3 0.7 -0.2 0.5 |
| | | | | -0.4 through 0.5 0.7 -0.1 and 0.1 through |
| | | | | -0.1 and 0.1 through 0.5 0.7 |
| | | | | |
| | | | | -0.2 through 0.3 0.7 -0.2 through 0.2 -0.4 -0.1 and 0.2 0.4, 0.5 |
| | | | | 0.3 through 0.5 -0.4 0.4 -0.0001, 0.7 |
| | | | | |
| | | | | |
| | | | | -0.4 through 0.5 0.7 -0.4 and 0.2 through |
| | | | | -0.4 through 0.4 0.7 -0.2 through 0.3 0.5 |

| | | | | |
|---|---|---|---|---|
| Both | inf | $p < 0.001$ | | -0.4 through 0.5 0.7 -0.2 through 0.3 -0.4, 0.5 0.1 and 0.2 0.4 |
| | | | | -0.4 through 0.5 0.7 0.3 through 0.5 -0.2 0.3 and 0.4 -0.4, -0.0001 |
| | | | | |

Table 2. ANOVA and comparisons among percent correct scores for subject SR9.

| | | | | Significant differences (numbers refer to |
|---|---|---|---|---|
| | | | | |
| | | | | |
| | | | | |
| | | | | -0.2 through 0.4 -0.4 0.4 0.7 |
| | | | | |
| | | | | 0.4 and 0.2 -0.4, 0.7 -0.0001 -0.4 |

In general, the percent-correct scores are relatively uniform over wide ranges of exponent values for both subjects. For presentation of the consonants in quiet, the ANOVAs and
show that exponents in the range of -0.2 to 0.4 usually produce scores that are not significantly different from each other, for all talker conditions and each subject. In most cases, the best score (obtained with an exponent in the range of -0.1 to 0.4) is significantly greater than the score or scores

For SR2, exponents of -0.1 and 0.2 produced significantly higher scores for the male talker than the exponents at the extremities, -0.4 and 0.7. Also, the exponent of 0.1 produced a higher score than the exponent of 0.7. For the female talker, all exponents from -0.4 through 0.5 produced higher scores than the exponent of 0.7. This also was the case for the combined talkers. In addition, the exponent of -0.1, and the exponents from 0.1 through 0.3, produced higher scores than the exponent of -0.4. The exponent of 0.1 produced a higher score than the exponent of 0.5 for the combined talkers as well.

For SR9, exponents of 0.2 and 0.4 produced significantly higher scores for the male talker than the exponent of -0.4. Differences among scores for the female talker were not significant. Scores for the combined talkers showed the same pattern as that found for the male talker,
with the exponents of 0.2 and 0.4 were significantly higher than the score obtained with the exponent

The region of highest scores for SR2 included the exponent value approximating the standard mapping function (exponents in the region of -0.0001). The highest scores for SR9, on the other hand, were obtained for less-compressive mapping functions, with exponents in the range of 0.2 to 0.4.

Results from the tests with SR2 involving presentation of the consonants in noise were somewhat different from the results obtained for presentation of the consonants in quiet. For the male talker and the +15 dB S/N, the exponent of 0.4 produced a higher score than the exponent of -0.4. For the female talker, all exponents from -0.4 through 0.5, except for the exponent of -0.2, produced higher scores than the exponent of 0.7. For combined speakers, the exponent of -0.1, and the exponents from 0.1 through 0.5, produced higher scores than the exponent of 0.7.

For the +10 dB S/N conditions, no significant differences were found among the scores for the different exponents either for the male talker or for the combined talkers. For the female talker, the exponents from -0.2 through 0.3 produced higher scores that the exponent of 0.7. In addition, the exponent of -0.2 produced a higher score than the exponent of 0.5.

An interesting aspect of the results for the quiet and +15 dB S/N conditions is that the curves for both talkers do not differ significantly over the range of exponent values from 0.4 to 0.7. The scores are relatively high with the 0.4 exponent. These observations suggest that a choice of a 0.4 exponent might be advantageous for listening to speech at speech-to-noise ratios of +15 dB or higher. This also would not be a bad choice for more adverse speech-to-noise ratios, inasmuch as the results for the +10 dB S/N conditions either show no dependence of scores on the choice of exponent (male and combined speakers) or only a weak dependence (female speaker). In the case of the weak dependence, the exponent of 0.4 does not produce a significant decrement in performance compared with the highest performance among exponent choices (exponent of -0.2), at least according to the Tukey tests. (The Fisher LSD tests do indicate a significant difference in scores for the -0.2 and 0.4 exponents.)

***Feature-transmission scores for consonants.*** Figures 4 and 5 show feature-transmission scores for the consonant tests conducted with subjects SR2 and SR9, respectively. For subject SR2, the scores for overall information transmission, voicing and manner of articulation are quite similar over the tested range of exponent values, for all talker and speech-to-noise conditions. Transmission of manner information appears to be somewhat more susceptible to noise interference than voicing or overall information, but the curves for the three features are still similar at the speech-to-noise ratio of +10

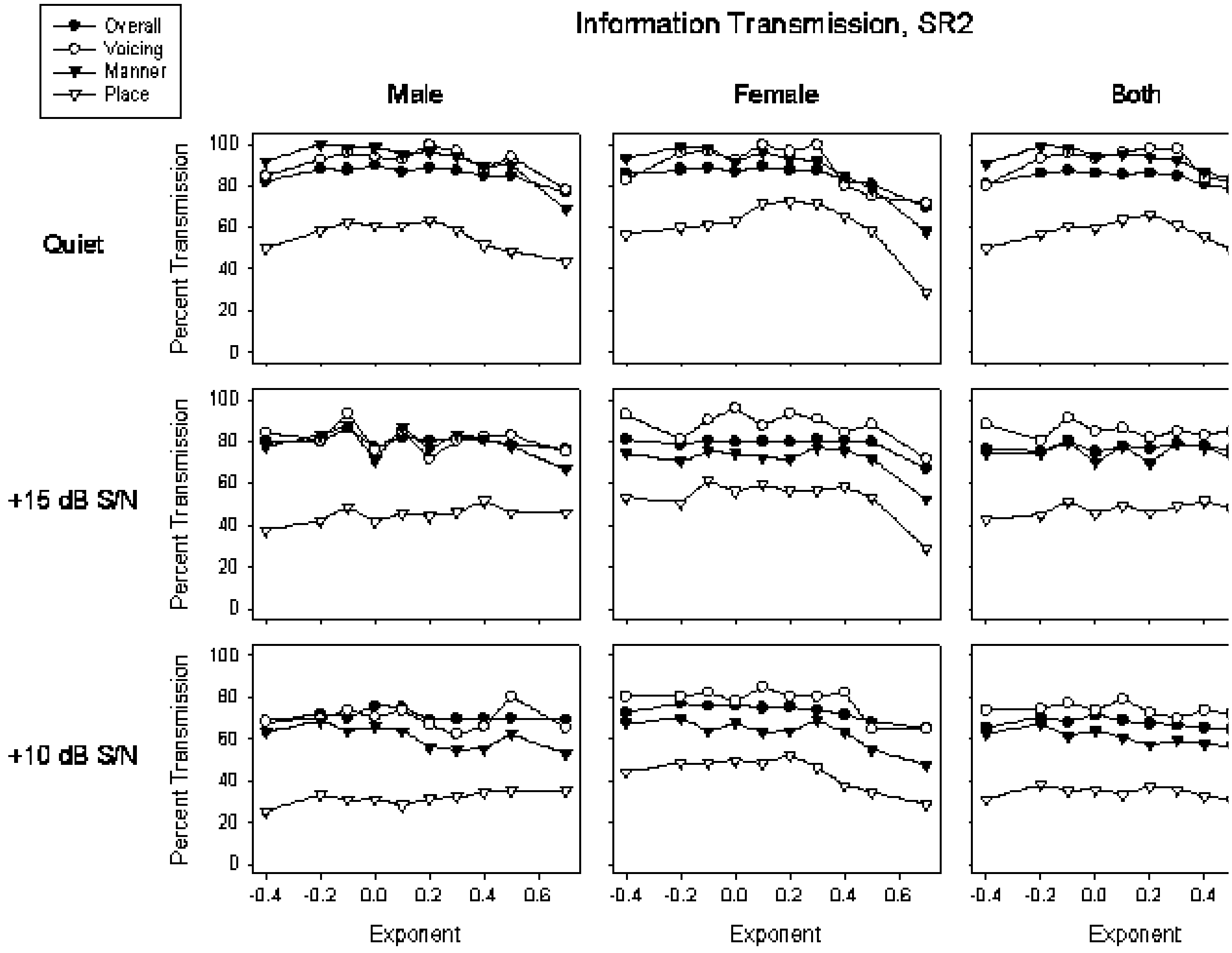
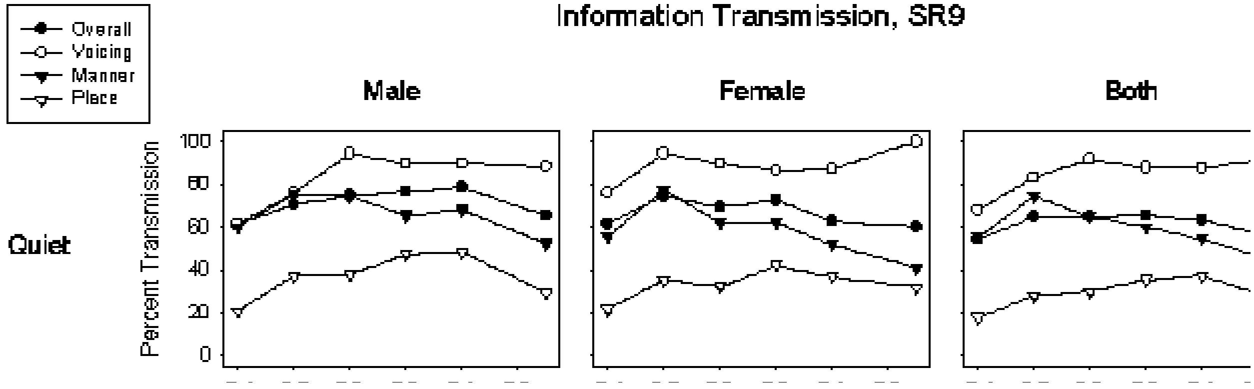


*Fig. 4. Feature transmission scores from the tests of consonant identification with subject SR2.*



Results from the tests of vowel identification with SR2 are presented in Fig. 6. Note that the number of exponent values included in these tests is smaller than the number included in the consonant tests for this subject. Three values were included in the vowel tests for the quiet and +15 dB conditions, and five values were included in the tests at the +10 dB speech-

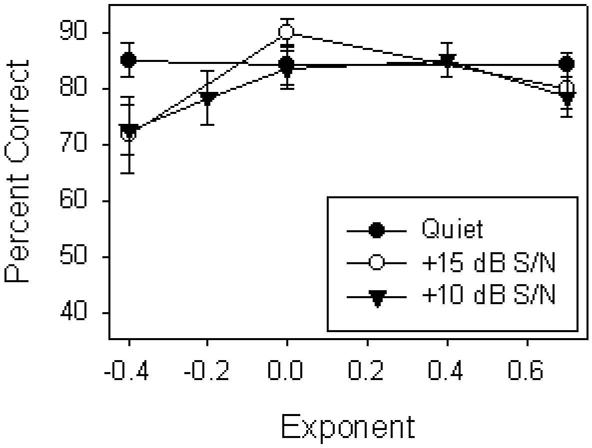


*Fig. 6. Vowel identification for various exponents in the mapping function used for CIS processors,*

The scores across exponent values are relatively uniform for presentation of the vowels in quiet and for each of the speech-to-noise ratios. Indeed, neither the ANOVA for quiet, nor the ANOVAs for vowels with noise, indicated a significant difference among scores for the different exponent values.

## Comparisons with findings from other laboratories

As noted before, Fu and Shannon recently have published results from a study to evaluate effects of manipulations in mapping functions for cochlear implant speech processors (Fu and Shannon, 1998). A graph of the mapping functions used in that study is presented in Fig. 7. (These curves also were derived with a threshold value of 400 µA and a MCL value of 800 µA.) The range of exponent values used by Fu and Shannon produces a set of functions that are similar to the set used in our study (compare Figs. 7 and 1). The exponents themselves are quite different from ours, because their functions map envelope values from 0 to 1000, whereas ours map envelope values from 1 to 1024. The starting value for envelope strongly affects the values of A and k in the mapping function equation (see page 6), and this in turn produces quite different curves for the same exponent value.

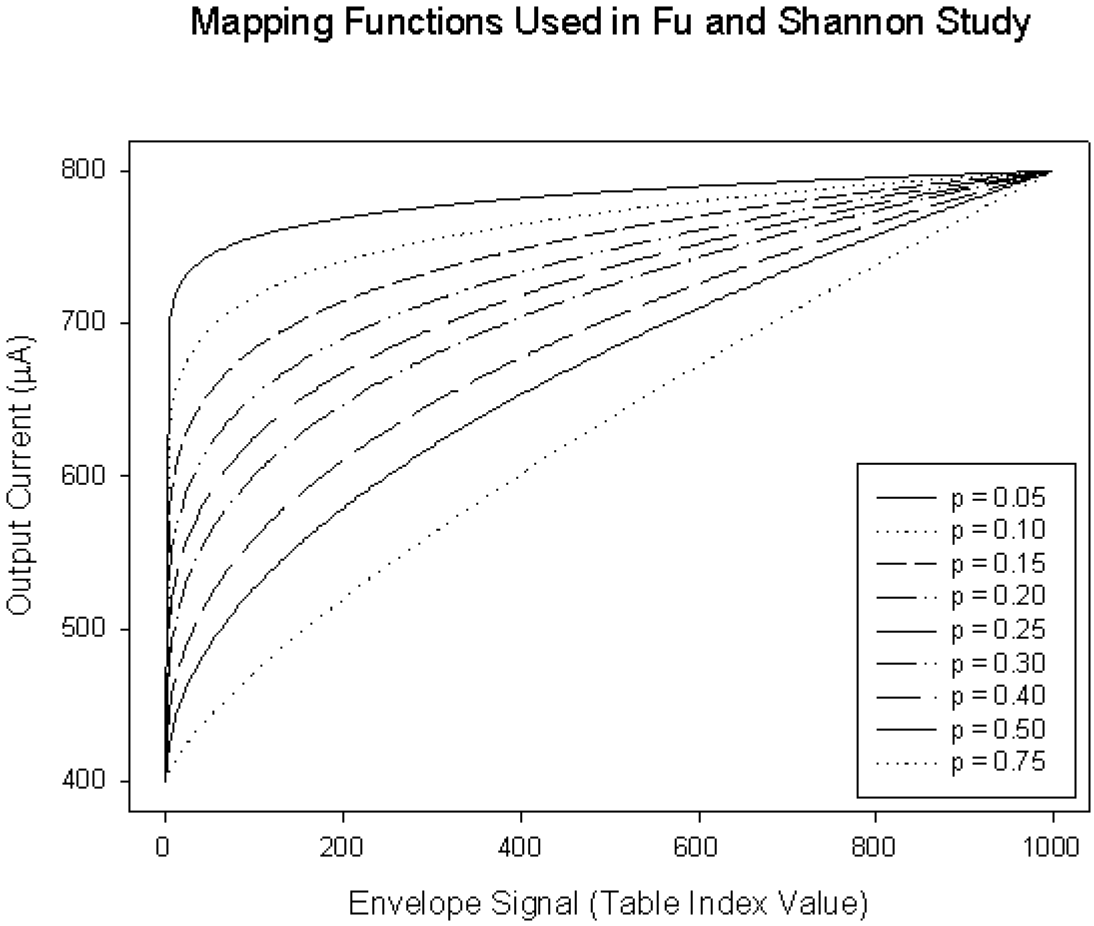


*Fig. 7. Mapping functions used in the Fu and Shannon study.*

Mapping functions corresponding to the highest and lowest exponents in each of the studies are shown in Fig. 8. As is evident from the figure, the overall ranges of functions for the two studies approximate each other. The most compressive function used by Fu and Shannon is somewhat more compressive than the most compressive function used in our study. The least compressive functions for the two

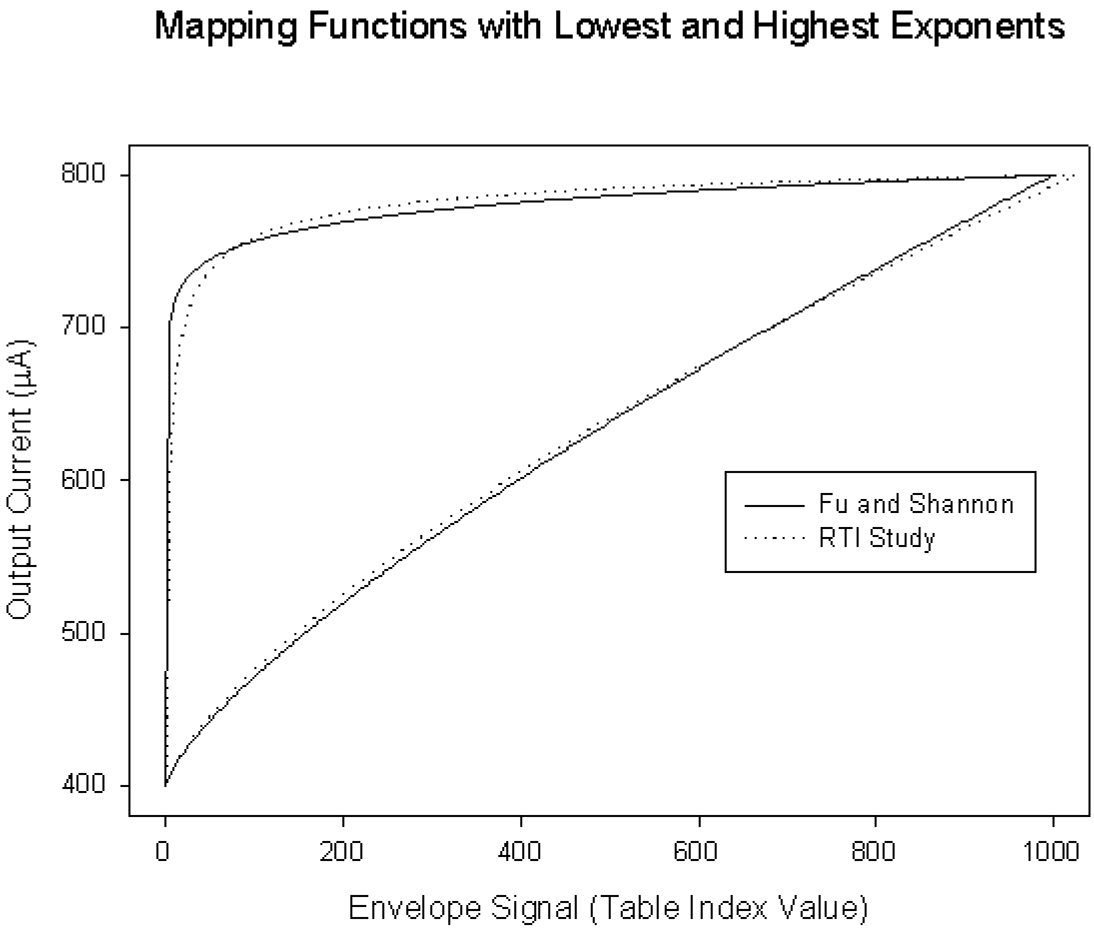


*Fig. 8. Mapping functions with lowest and highest exponents, Fu and Shannon study (solid lines) and*

Fu and Shannon used 4 channel CIS processors, as implemented in custom software and with use of their research interface for laboratory control of the Nucleus CI22 implant. Three subjects with this implant participated in the studies. The overall range spanned by the bandpass filters was 100 to 6000 Hz, compared with the range of 350 to 9500 Hz in our study. Stimuli were directed to monopolar electrodes in the Ineraid implant in our study, and to bipolar pairs of electrodes in the Nucleus implant in the Fu and Shannon study (a "BP+1" configuration was used, involving electrodes separated by 1.5 mm from center to center along the electrode array). With the exception of pulse duration, other parameters were identical between the processors used for SR2 in our study and the processors used for the three subjects in the Fu and Shannon study. The processors used for SR9 in our study had six channels, full wave rectifiers, and a pulse rate and phase duration that were somewhat different from those used in the processors for the Fu and Shannon study. In general, the processors were similar but

Figure 9 shows the percent-correct scores from the tests of consonant identification conducted by Fu and Shannon. The exponents are plotted along both linear (left panel) and logarithmic (right panel) scales. Fu and Shannon presented the logarithmic plot in their paper, but plots with a linear scale facilitate comparisons between the Fu and Shannon results and our results, which are plotted along a

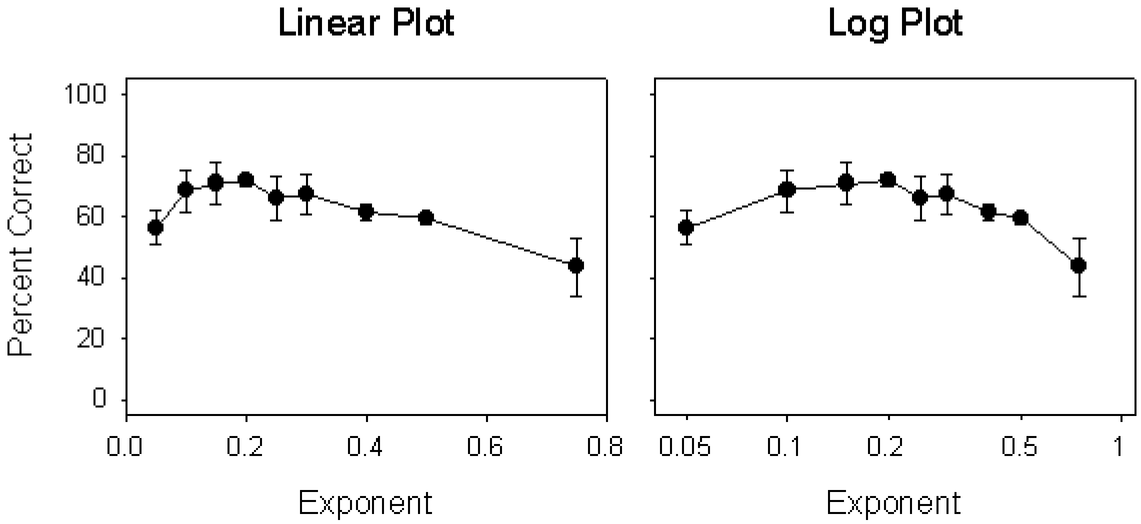


*Fig. 9. Percent-correct scores from tests of consonant identification in the study conducted by Fu and Shannon (1998). The exponents used in the study are shown along both linear (left panel) and logarithmic (right panel) scales. The tests included 16 consonants presented in an /a/-consonant-/a/*

compressive than the best function in their data for quiet) might be helpful for listening to speech in noise. The speculation was based on results from acoustic simulation studies, using subjects with

Figure 10 shows feature transmission scores for the Fu and Shannon study. They plotted scores for voicing, manner and place. As in our results, the scores for transmission of place information are lower than the scores for transmission of manner or voicing information (compare Fig. 10 with Figs. 4 and 5). Unlike our results, the Fu and Shannon results indicate lower scores for the transmission of voicing information than for the transmission of manner information at low exponent values (exponent values

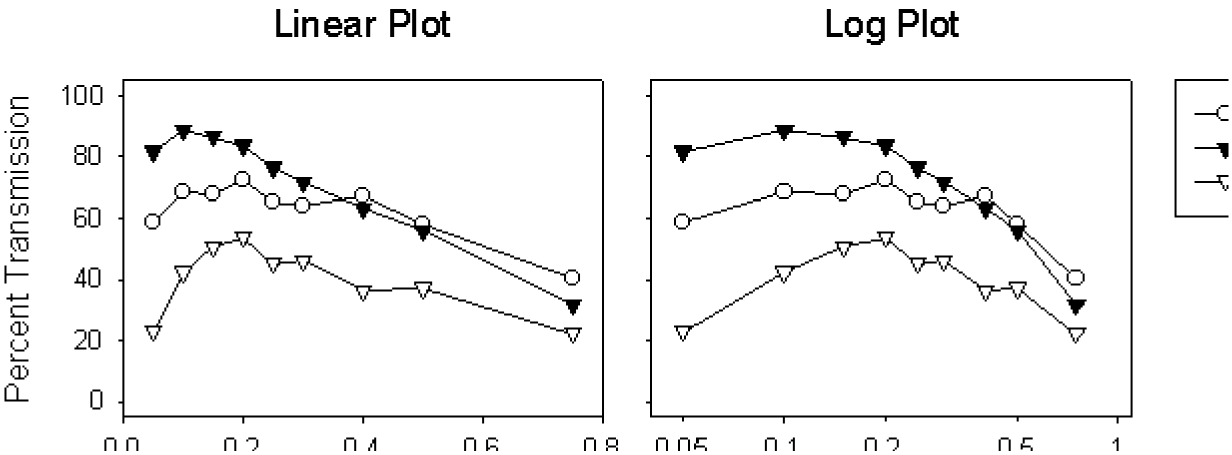
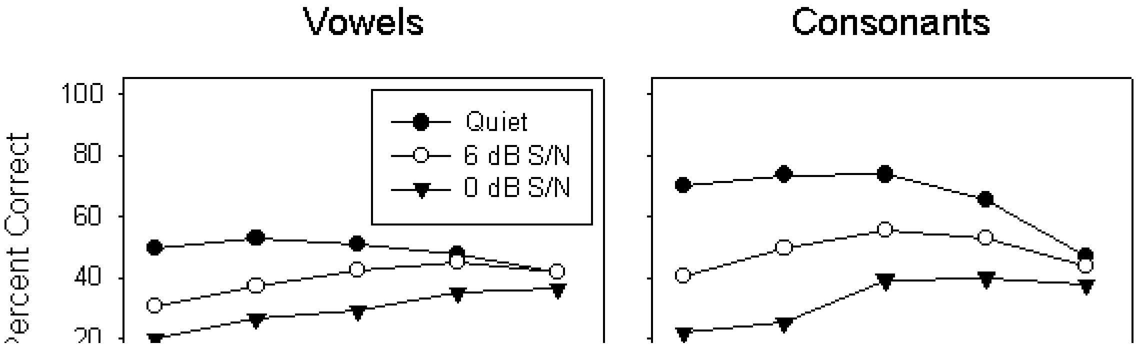


*Fig. 10. Feature-transmission scores from the study conducted by Fu and Shannon (1998). The exponents used in the study are shown along both linear (left panel) and logarithmic (right panel)*

The Fu and Shannon results also show a correspondence between the shape of the curve for place transmission and the shape of the curve for percent-correct scores (compare Figs. 10 and 9). The maximum in the curve for place is somewhat "sharper" (more peaked) than in the curve for percent correct in the Fu and Shannon results. The shapes of the curves are more similar in our data.

Most recently, Fu and Shannon have extended their initial studies to include presentation of consonants and vowels in noise (Fu and Shannon, 1999). The three subjects who participated in the initial studies also participated in these subsequent studies. The processors were the same as those of the initial studies except that a BP+5 configuration was used for the electrodes (with the two electrodes of each bipolar pair separated by 4.5 mm). The noise used was an approximation to speech-spectrum noise, derived by filtering wide-band (spectrally flat) noise with a first-order lowpass filter

The principal results from the subsequent studies are presented in Fig. 11. (Only the averages of the percent-correct scores across the three studied subjects are shown, in that error bars were not reported by Fu and Shannon.) As in our results for subject SR2, manipulation of the exponent over a broad range has almost no effect on the identification of vowels in quiet. Also, the addition of noise, even at the high levels used by Fu and Shannon (+6 and 0 dB speech-to-noise ratios), produces only relatively small decrements in vowel identification, especially for high values of the exponent.



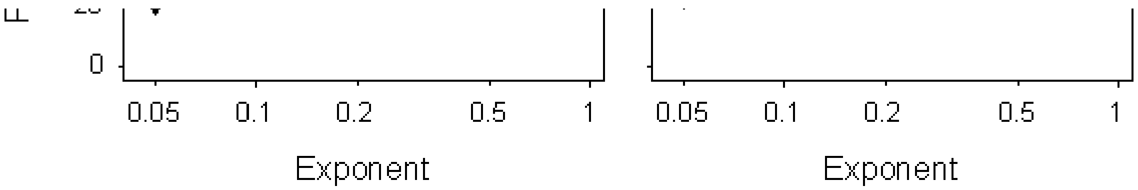
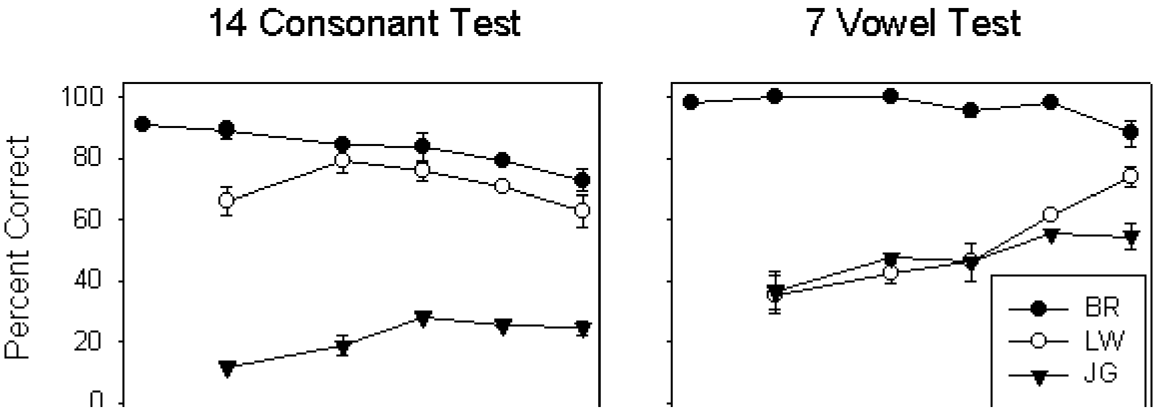


*Fig. 11. Percent-correct scores from an additional study conducted by Fu and Shannon (1999), that included presentation of vowels and consonants in noise. Exponents are shown along a logarithmic scale only. The vowel tests included 12 vowels presented in a /h/-vowel-/d/ context, and the consonant tests included 16 consonants presented in an /a/-consonant-/a/ context.*

Results from the tests of consonant identification also are broadly similar to our results with subject SR2 (compare the right panel of Fig. 11 with the left column of Fig. 2). In quiet, performance declines in both sets of results with use of the least-compressive mapping function. In noise, scores are more similar between the least-compressive mapping function and somewhat more-compressive functions. For the relatively adverse speech-to-noise ratios used in the Fu and Shannon study, consonant identification is maximized across the quiet and speech-in-noise conditions with the exponent of 0.2. That exponent produces a mapping function that approximates the mapping function used in standard CIS processors (the curve for the 0.2 exponent in Fig. 7 is similar to the curve for the -0.0001

Boëx and coworkers also have conducted a study to evaluate effects of changes in the mapping function on the performance of CIS processors. The principal results are presented in Fig. 12 (data from Boëx, 1995). The three subjects were long-term users of the Ineraid device. CIS processors were used, with five channels and a relatively high pulse rate, 2000 pulses/s/electrode. As in our study, monopolar coupling was used. The mapping functions, and the dependence of the mapping functions on the value of the exponent, were similar to ours.



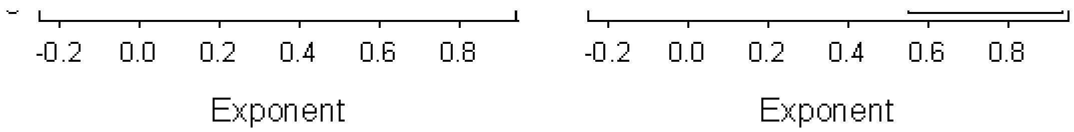


*Fig. 12. Percent-correct scores from the study conducted by Boëx (1995). Scores for each of the three subjects are indicated by the different symbols. The consonant tests included presentation of 14 French consonants in an /a/-consonant-/a/ context, and the vowel tests included presentation of 7 French vowels in isolation, without bracketing consonants.*

Results from the consonant tests (left panel of Fig. 12) show gradual decrements in performance with increases in exponent value for subject BR, a shallow peak in performance at the exponent of 0.3 for subject LW, and a plateau in performance with exponents of 0.5 and higher for subject JG. The patterns of scores across exponent values for subjects LW and JG are somewhat similar to the overall pattern observed with our subject SR9. For these subjects, a peak or plateau in performance is found at exponent values in the range of 0.2 to 0.5. In contrast, the pattern for subject BR is somewhat similar to the overall pattern observed with our subject SR2. This pattern is one of relatively uniform performance for exponents in the range of -0.2 to about 0.5, and of lower performance at higher

The results for vowel identification (right panel of Fig. 12) differ from the results for consonant identification for two of the subjects in the Boëx study. Consonant scores for subject LW indicate a peak in performance at the exponent value of 0.3, but the vowel scores indicate monotonic increases in performance with increases in exponent values up to the tested limit of 0.9. Consonant scores for subject JG show a plateau in performance at and above the exponent value of 0.5, but the vowel scores indicate improvements in performance with increases in exponent values beyond 0.5. Scores for subject BR are relatively uniform across exponent values, as in our results for subject SR2 (Fig. 6) and as in the results for the three subjects studied by Fu and Shannon (see,

Results from the recent study of Loizou and Poroy (1999) are presented in Fig. 13. Six subjects participated in this study. All were users of CIS-Link processors in conjunction with their Ineraid implants. A separate laboratory system was used to implement CIS processors for the tests of Fig. 13. The processors for each subject used six channels, a pulse rate of 800

μs/phase, and a staggered update order. As in our study and in the study of Boëx and coworkers, monopolar coupling was used. The dependence of the mapping functions on the value of the exponent was identical to that in our study (see Fig. 1). Exponents used in the study of Loizou and Poroy included -0.1, -0.0001, 0.2 and 0.6.

Results from the present study, along with results from other studies, show that the performance of CIS processors can be relatively insensitive to manipulations in mapping functions over a rather broad range. That range usually includes the default mapping function for CIS processors, a logarithmic mapping function or a power function with an exponent of -0.0001. Typically, some range of higher and lower values of the exponent can support equivalent scores in tests of consonant identification. Beyond that range, which varies from subject to subject, consonant identification scores decline.

Some subjects show a peak or asymptote in performance at an exponent that is somewhat higher than the default value. In the studies to date, these subjects appear to be in the low- or mid-performance categories with their implants. Subject SR9 achieved her best scores with an exponent in the range of

- A visit by Chris van den Honert, of Cochlear Corporation, to discuss details of upcoming cooperative studies involving new electrode designs (July 12).
- Continued development of the Access database mentioned in the Introduction, to bring together in a database format a complete record of speech processor designs and results for our

- Initial entries into the database.

We expect that Stefan's main work with us will involve studies with recipients of bilateral implants. He will play a major role in upcoming studies with recipients of COMBI 40+ implants on both sides, in cooperation with the University Hospital in Würzburg, Germany, and with recipients of CI24M implants on both sides, in cooperation with

, Würzburg, Germany, June 30 through July 4, 1999.