The effect of parametric variations of cochlear implant processors on speech understanding

Philipos C. Loizou[a)] and Oguz Poroy
Department of Electrical Engineering, University of Texas at Dallas, Richardson, Texas 75083-0688

Michael Dorman
Department of Speech and Hearing Science, Arizona State University, Tempe, Arizona 85287

(Received 13 January 2000; accepted for publication 1 May 2000)

This study investigated the effect of five speech processing parameters, currently employed in cochlear implant processors, on speech understanding. Experiment 1 examined speech recognition as a function of stimulation rate in six Med-El/CIS-Link cochlear implant listeners. Results showed that higher stimulation rates (2100 pulses/s) produced a significantly higher performance on word and consonant recognition than lower stimulation rates (<800 pulses/s). The effect of stimulation rate on consonant recognition was highly dependent on the vowel context. The largest benefit was noted for consonants in the /uCu/ and /iCi/ contexts, while the smallest benefit was noted for consonants in the /aCa/ context. This finding suggests that the /aCa/ consonant test, which is widely used today, is not sensitive enough to parametric variations of implant processors. Experiment 2 examined vowel and consonant recognition as a function of pulse width for low-rate (400 and 800 pps) implementations of the CIS strategy. For the 400-pps condition, wider pulse widths (208 μs/phase) produced significantly higher performance on consonant recognition than shorter pulse widths (40 μs/phase). Experiments 3–5 examined vowel and consonant recognition as a function of the filter overlap in the analysis filters, shape of the amplitude mapping function, and signal bandwidth. Results showed that the amount of filter overlap (ranging from -20 to -60 dB/oct) and the signal bandwidth (ranging from 6.7 to 9.9 kHz) had no effect on phoneme recognition. The shape of the amplitude mapping functions (ranging from strongly compressive to weakly compressive) had only a minor effect on performance, with the lowest performance obtained for nearly linear mapping functions. Of the five speech processing parameters examined in this study, the pulse rate and the pulse width had the largest (positive) effect on speech recognition. For a fixed pulse width, higher rates (2100 pps) of stimulation provided a significantly better performance on word recognition than lower rates (<800 pps) of stimulation. High performance was also achieved by jointly varying the pulse rate and pulse width. The above results indicate that audiologists can optimize the implant listener's performance either by increasing the pulse rate or by jointly varying the pulse rate and pulse width. © *2000 Acoustical Society of America.* [S0001-4966(00)02308-0]

PACS numbers: 43.71.Ky, 43.66.Sr [CWT]

INTRODUCTION

To account for the variability in performance among cochlear implant subjects, cochlear implant manufacturers started incorporating several speech-processing strategies in their implant processors. The Advanced Bionics Corporation, for example, supports among other strategies the SAS and the continuous interleaved sampling (CIS) strategies in their Clarion device, while the Cochlear Corporation supports the ACE, SPEAK, and CIS strategies in their Nucleus 24 device. Increasing the number of speech strategies available in implant processors not only increases the chances that one of those strategies might be more beneficial than others (e.g., Osberger and Fisher, 1999), but also increases the complexity in choosing the right set of parameters associated with each strategy.

Most speech processing strategies today can be configured using a multitude of parameters, which can be easily modified by the audiologist using existing fitting software. In the SPEAK strategy, for example, one can change the pulse width, the number of maxima selected, the pulse rate, the filter allocation table, etc. In the CIS strategy, currently supported by all implant devices (Nucleus 24, Med-El and Clarion) in the United States, one can change the pulse width, the pulse rate, the electrode stimulation order, and the compression function that maps the acoustical input to electrical output. A series of studies by Wilson and colleagues (e.g., Wilson *et al.*, 1991, 1993, 1995, 1999) has shown that the CIS parameters (e.g., pulse rate, pulse width, etc.) can be varied to optimize individual subjects' performance. They reported, for example, significant improvements in speech recognition for one subject over the course of three successive visits to their laboratory. The subject's mean score on consonant recognition improved from 56% correct (at the initial visit) using a 167-μs/phase, 500-pps processor with a staggered order of stimulation, to 79% correct (at the second visit) using 33-μs/phase, 833-pps processor with a staggered order of stimulation, to 85% correct (at the third visit) using 33-μs/phase, 833-pps processor with an apex-to-base order of stimulation. Dorman and Loizou (1997) also showed simi-

[a)]Electronic mail: loizou@utdallas.edu

lar improvements for another Ineraid subject fitted with a CIS processor. The subject's sentence scores improved from 73% correct using a 100-μs/phase, 823-pps processor with a staggered stimulation order to 100% correct using a 40-μs/phase, 2020-pps processor with an apex-to-base stimulation order. These studies demonstrated how the electrode stimulation order and the stimulation rate could greatly affect the subjects' performance. Other parameters that were found to affect performance include the number of channels (Wilson *et al.*, 1991, 1995; Lawson *et al.*, 1996), pulse duration (Wilson *et al.*, 1993), envelope cutoff frequency (Lawson *et al.*, 1993), and signal bandwidth (Zerbi *et al.*, 1998). As it turns out, some of these parameters interact with each other. Wilson *et al.* (1993) reported that subjects obtained the highest performance at different combinations of pulse rate and pulse width. Other studies (Brill *et al.*, 1997; Kiefer *et al.*, 1997) showed a tradeoff between number of channels and pulse rate. Brill *et al.* (1997) found that trading channels for higher stimulation rates improved performance. Higher stimulation rates produced significantly higher performance than lower stimulation rates. For implant devices with a large number of electrodes, the choice of electrodes to be stimulated is yet another parameter. The principal advantage of such electrode arrays having many electrodes is not that they can support a large number of channels, but rather that they allow the selection of subsets of electrodes to optimize individual subject's performance. Lawson *et al.* (1996) and Zwolan *et al.* (1997) have demonstrated that the selection of electrodes can significantly affect performance.

Given the large number of speech processing parameters available and the effect of some of these parameters on speech understanding, it is becoming increasingly more important to identify the set of parameters which is most likely to affect speech recognition. This is an extremely important issue particularly for fitting implant patients, since in a clinical setting the audiologists cannot devote too much time to select these parameters.

The aim of this study is to identify the speech processing parameter(s) that affects speech recognition the most. The results of our study will greatly facilitate the fitting of new implant patients, as it will provide to audiologists a good starting point for fitting. We use the CIS speech processing strategy in this article, however, the parameters examined here are employed in all speech processing strategies, and most of these parameters can be changed manually by audiologists using existing fitting software. Five parameters will be varied systematically: pulse rate, pulse width, filter overlap, compression function, and signal bandwidth. The effect of these parameters on speech recognition will be examined.

I. EXPERIMENT 1: EFFECT OF PULSE RATE

Cochlear implant devices present information to the electrodes in analog or pulsatile form. In pulsatile stimulation, the information is delivered to the electrodes using a set of narrow pulses. In some devices, the amplitudes of these pulses are extracted from the envelopes of the filtered waveforms. The advantage of this type of stimulation over analog stimulation is that the pulses can be delivered in a nonoverlapping fashion, thereby minimizing channel interactions. One of the issues associated with pulsatile stimulation is how fast, in terms of number of pulses per second, we need to stimulate the electrodes. Whether higher rates of stimulation provide more benefit to speech understanding than lower rates of stimulation has been the subject of debate. Presently, the stimulation rate employed by commercial implant processors varies from a low of 200 pulses/s to a high of 2400 pulses/s per electrode (see review by Loizou, 1998). Does it make a difference in speech understanding, however, whether we stimulate the electrodes at a low rate of 200 pulses/s or at a high rate of 2400 pulses/s? Wilson *et al.* (1998) demonstrated, using intracochlear-evoked potential recordings, that higher rates of stimulation have the potential of disrupting the high synchrony observed with low-rate stimulation, allowing more normal ''stochastic'' firing patterns. Higher rates of stimulation also provide a better temporal representation of the speech envelope, but how important is that for speech understanding? These are the questions we are trying to address in this experiment. To answer these questions, we varied systematically the pulse rate from 400 to 2100 pulses/s and examined vowel, consonant, and monosyllabic word recognition.

A. Method

1. Subjects

The subjects were six postlingually deafened adults who had used a six-channel CIS processor for periods ranging from 3 to 4 years. All the patients had used a four-channel, compressed-analog signal processor (Ineraid) for at least 4 years before being switched to a CIS processor. The patients ranged in age from 40 to 68 years and they were all native speakers of American English. Biographical data for each patient are presented in Table I.

2. Speech material

The test material included monosyllabic words, consonants in /vCv/ environment, and vowels in /hVd/ environment. The word test consisted of four different lists with 50 monosyllabic (CNC) words in each list. The consonant test consisted of 20 /vCv/ consonants in three vowel environments, /i a u/, produced by a single female speaker, and was taken from the consonant database recorded at the House Ear Institute (Shannon *et al.*, 1999). The 20 consonants were /b p d t g k f v s z ʃ ð t ʃ d ʒ m n r l j w/ in i/C/i, u/C/u, and a/C/a format. We chose to use consonants produced by a female speaker to avoid possible ceiling effects, since the vowels produced by female speakers were found by implant patients to be relatively harder to identify than the vowels produced by male speakers (Loizou *et al.*, 1998). The vowel material consisted of the vowels in the words: ''heed, hid, hayed, head, had, hod, hud, hood, hoed, who'd, heard.'' Each word was produced by three men, three women, and three girls. The stimuli were drawn from a set used by Hillenbrand *et al.* (1995).

3. Experimental setup

All the experiments were performed on our laboratory cochlear implant processor, which was based on the design

TABLE I. Biographical data of the six cochlear-implant users who participated in this study.

Subject	Gender	Age (years) at detection of hearing loss	Age at which hearing aid gave no benefit	Age fit with Ineraid	Age at testing	Etiology of hearing loss	Score on H.I.N.T sentences in quiet	Score on NU-6 words in quiet
S1	F	10	46	47	55	unknown	44	20
S2	M	5	43	48	58	unknown	92	43
S3	F	7	31	33	40	unknown/ hereditary	100	80
S4	F	23	48	51	57	unknown	100	71
S5	M	20	46	63	68	unknown	88	46
S6	M	19	19	29	41	Cogan's syndrome	100	93

of the Geneva/RTI/MEEI wearable processor (Francois *et al.*, 1994). Several modifications were made to the Geneva design, the most important of which was the addition of five current sources. (The wearable Geneva/RTI/MEEI processor was originally designed with one current source and could therefore provide only nonsimultaneous stimulation.) These modifications enabled us to investigate nonsimultaneous as well as simultaneous stimulation. The block diagram of the laboratory processor is shown in Fig. 1. The input analog circuit consists of an audio multiplexer that selects the source of the input signal to the processor, several fixed-gain amplifiers, one variable-gain amplifier (adjusted externally by a sensitivity knob), an antialiasing filter, and a 16-bit A/D converter. The sampling rate of the A/D converter is controlled by the DSP chip, and for this study it was fixed at 22 kHz. The cutoff frequency of the antialiasing filter was set at 6.7 kHz. Once the signal is digitized, it is transmitted to the Motorola DSP56002 chip, where it is processed through the CIS strategy (see description in the following section). The CIS outputs are finally fed through a SSI port to the current sources built around a digital-to-analog converter. Biphasic pulses are generated, with amplitudes equal to the CIS envelope outputs, and sent to the electrodes for stimulation. The pulse width as well as the stimulation rate was controlled through software. More information about the hardware of the laboratory processor can be found in Poroy and Loizou (2000).

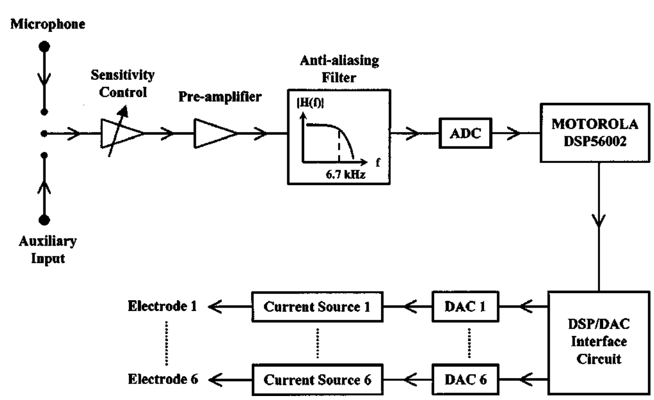

FIG. 1. Block diagram of the laboratory cochlear implant processor used in this study.

4. CIS implementation

Signals were first processed through a preemphasis filter (2000-Hz cutoff), with a 3-dB/oct roll-off, and then bandpassed into six frequency bands using sixth-order Butterworth filters. The center frequencies of the six bandpass filters were 461, 756, 1237, 2025, 3316, and 5428 Hz. The envelopes of the filtered signals were extracted by full-wave rectification and low-pass filtering (second-order Butterworth) with a 100-, 200-, or 400-Hz cutoff frequency, depending on the pulse rate. The six envelope amplitudes A_i ($i=1,2,...,6$) were mapped to electrical amplitudes E_i using a power-law transformation:

$$E_i = cA_i^p + d, \tag{1}$$

where c and d are constants chosen so that the electrical amplitudes fall within the range of threshold and most-comfortable level, and p is the power exponent. The power exponent p was set equal to -0.0001 to obtain a compression function similar to the logarithmic function found in the Med-El/CIS link device.[1] The power-law mapping was implemented using a table-lookup procedure using a table with 512 entries (for each electrode). The mapped envelope amplitudes were finally used to modulate biphasic pulses of duration 40 μs/phase at stimulation rates ranging from 400 to 2100 pulses/s. The electrodes were stimulated in the same order as in the subject's daily processors. For most subjects, the electrodes were stimulated in ''staggered'' order.

The stimulation rates investigated were 400, 800, 1400, and 2100 pulses/s. This range was chosen because it corresponds to the range of pulse rates currently supported by the three commercial implant devices, Clarion, Nucleus 24, and Med-El. The envelope (low-pass) cutoff frequencies were set to 100 Hz for the 400-pps processor, 200 Hz for the 800-pps processor, and 400 Hz for the 1400- and 2100-pps processors. These frequencies were chosen to avoid aliasing effects. The pulse duration for all four rate conditions was fixed at 40 μs/phase. Note that the 400 or 800-pps rate conditions could alternatively be implemented by widening the pulse width (see experiment 2). Changing the pulse width, however, would not only change the pulse rate but would also affect the threshold and most-comfortable level (MCL) values, and consequently the electrical dynamic range. In this experiment, we only wanted to vary a single parameter, the pulse rate.

5. Procedure

The test was divided into four sessions, one for each rate condition. The four conditions were counterbalanced among subjects to avoid any order effects. Each session consisted of a consonant, a vowel, and a monosyllabic-word test. In the vowel test, there were 12 repetitions of each vowel, and in the consonant test there were 9 repetitions of each consonant. The stimuli were presented in blocks of three repetitions each. The monosyllabic words were presented only once. A different word list was used for each condition. The vowels and the consonants were completely randomized. All test sessions were preceded by one practice session in which the identity of the vowels/consonants was indicated to the listeners.

The stimuli were presented directly to the subjects through our laboratory processor at a comfortable listening level. To collect responses, a graphical interface was used that allowed the subjects to identify the words they heard by clicking on the corresponding button in the graphical interface. For the monosyllabic-word test, the subjects wrote down the word they heard.

B. Results and discussion

The results on monosyllabic word, vowel, and consonant recognition are shown in Fig. 2.

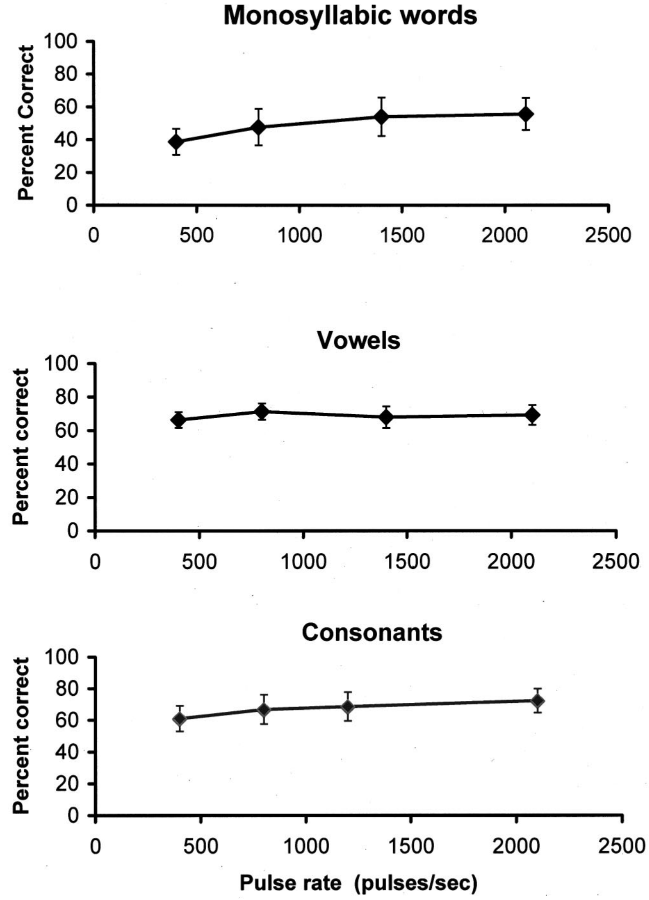

FIG. 2. Speech recognition as a function of stimulation rate. The word and consonant recognition scores obtained at 2100 pps were significantly higher than the corresponding scores obtained at 800 or 400 pps. Higher stimulation rates did not provide any benefits for vowel recognition. Error bars indicate standard errors of the mean.

1. Monosyllabic words

The monosyllabic words were scored in percent words correct [Fig. 2(a)]. Repeated measures analysis of variance indicated a significant main effect of rate [$F(3,15)=7.197$, $p<0.005$]. *Post hoc* analysis (according to Scheffe) showed that the score obtained at 2100 pulses/s was significantly ($p<0.05$) higher than the score obtained at 800 pulses/s. The scores at 2100 and 1400 pulses/s did not differ.

The individual subject's performance is given in Fig. 3. All subjects benefited from higher rates, some more than others. The rate at which the open-set performance reached an asymptote varied across subjects. Some subjects (S5, S6) showed a significant improvement starting at a rate of 1500 pulses/s, while other subjects (S3, S4) showed a significant improvement starting at a rate of 800 pulses/s. Other subjects (S1, S2) did not show an improvement until using the rate of 2100 pulses/s.

The above results clearly show that cochlear implant users can receive a significant benefit on open-set speech recognition from higher rates of stimulation. There does not seem to be a single ''critical'' rate, however, above which all subjects receive significant benefit. The ''critical'' rate seems to be dependent on the subject. In our study, all patients achieved maximum performance with a rate of 2100 pulses/s.

2. Multi-talker vowels

The results for vowel recognition, scored in terms of percent correct, are given in Fig. 2(b). The mean vowel scores for the rates 400, 800, 1400, and 2100 pulses/s were 66.3%, 71.3%, 68%, and 69.3%, respectively. Repeated measures analysis of variance indicated no significant main effect [$F(3,15)=1.623$, $p=0.226$] of rate on vowel recognition. The finding that rate does not seem to affect vowel recognition is not surprising, since higher rates do not improve spectral resolution, which is needed for the perception of vowels. Vowels are characterized by slowly changing formant transitions which can be adequately captured even with slow

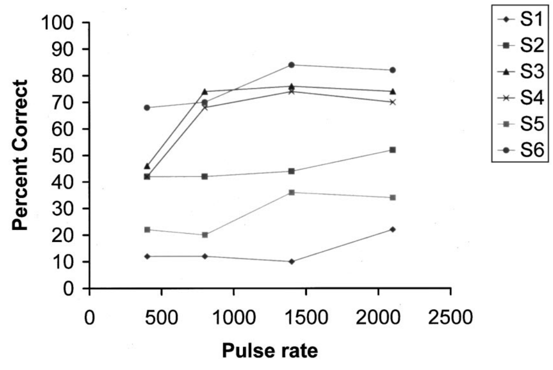

FIG. 3. Individual subject's performance on word recognition as a function of stimulation rate.

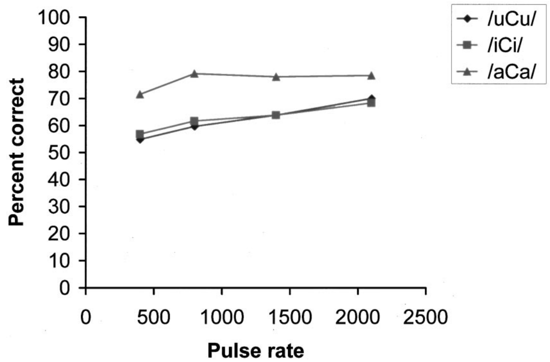

FIG. 4. Consonant recognition as a function of stimulation rate for three different vowel contexts. The standard errors of the mean for /uCu/ recognition were 9.7%, 10.7%, 12.3%, and 10.3% for the rates of 400, 800, 1400, and 2100 pps, respectively. The standard errors of the mean for /iCi/ recognition were 8.3%, 9.3%, 8.2%, and 6.0%, and for /aCa/ recognition were 7.3%, 8.5%, 7.1%, and 7.2%.

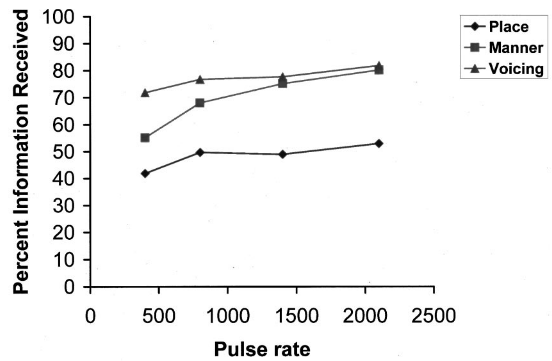

FIG. 5. Mean percent information received on the features of voicing, manner, and place of articulation as a function of stimulation rate. The standard errors of the mean for ''place'' were 8.6%, 10.1%, 5.5%, and 10.2% for the stimulation rates of 400, 800, 1400, and 2100 pps, respectively. The standard errors of the mean for ''manner'' were 9.9%, 11.7%, 5.5%, and 8.6%, and for ''voicing'' were 7.8%, 8.2%, 4.5%, and 6.9%.

rates of stimulation. Higher rates affect the temporal representation of the speech envelopes which is more important for the perception of consonants.

3. Consonants

The results on consonant recognition, scored in terms of percent correct, are given in Fig. 2(c). Repeated measures analysis of variance indicated a significant main effect [$F(3,15)=12.273$, $p<0.0001$] of rate on consonant recognition. *Post hoc* analysis (Scheffe's test) showed a significant ($p=0.05$) difference between the mean scores obtained at 2100 and 800 pulses/s, and a significant ($p=0.022$) difference between the scores at 800 and 400 pulses/s. There was no significant difference between the scores obtained at 1400 and 2100 pulses/s.

To examine the benefits of high-rate stimulation on multi-vowel consonant recognition, we analyzed the results for each vowel context separately. Figure 4 shows the mean scores on consonant recognition for each of the three vowel contexts. As can be seen, the effect of rate on consonant recognition is highly dependent on the vowel context. Higher rates of stimulation seem to benefit mostly consonants in the /iCi/ and /uCu/ contexts. The effect of higher rates of stimulation on the recognition of consonants in the /aCa/ context was small. Interestingly enough, the /aCa/ consonant test is probably the most widely used test today in the speech community. Yet, this test is not sensitive to parametric variations, such as pulse rate, of CIS processors.

We also analyzed the consonant confusion matrices using information transmission analysis (Miller and Nicely, 1955). The consonant features ''manner of articulation,'' ''place of articulation,'' and ''voicing'' were computed and scored in terms of percent information transferred. The results from the feature analysis are shown in Fig. 5. Large (monotonic) improvements in ''manner'' were obtained as the pulse rate increased, while moderate improvements in ''place'' were obtained. There was a small improvement in ''voicing.'' Repeated measures analysis of variance showed a significant main effect of rate on manner [$F(3,15)=22.2$, $p<0.0005$], a significant main effect of rate on place [$F(3,15)=5.38$, $p=0.01$], and a nonsignificant effect of rate on voicing [$F(3,15)=2.26$, $p=0.123$]. *Post hoc* tests showed that the manner scores reached an asymptote at 1400 pulses/s. Place scores improved and reached a plateau at 800 pulses/s, however, the improvement from 400 to 800 pps was nonsignificant ($p=0.06$).

Given the large improvements obtained in ''manner'' with higher rates of stimulation, we decided to analyze the confusion matrices further in terms of consonant class identification. We wanted to know whether higher rates of stimulation improve, say, stop, fricative or semivowel identification. The 20 /vCv/ consonants were divided into five consonant classes: stops /b p d t g k/, fricatives /f v s z ʃ ð/, affricates /tʃ dʒ/, nasals /m n/, and semivowels /r l j w/, and scored in terms of percentage of consonants identified correctly within each class. The mean scores for each vowel context are given in Fig. 6. As shown, consonant class (manner) identification improves with higher rates of stimulation and that improvement seems to be dependent on the vowel context. In the /aCa/ context, for instance, stop and affricate identification does not seem to be affected by the rate of stimulation. In contrast, in the /iCi/ context, there were large improvements in stop and affricate identification with increasing rate of stimulation. There was improvement in fricative identification in all vowel contexts with increasing rate of stimulation. For the /aCa/ and /iCi/ contexts, fricative identification reached an asymptote at 800 pulses/s, whereas for the /uCu/ context fricative identification kept increasing even up to 2100 pulses/s. There was also a large improvement in nasal discrimination in the /uCu/ context. In summary, the identification of stops, fricatives, affricates, and nasals improved with higher rates of stimulation. The improvement was more evident in the /uCu/ and /iCi/ contexts.

The above results demonstrated that higher rates of stimulation benefit consonant identification. This benefit stemmed largely from improved manner identification. We believe that this is because higher rates improve the temporal representation of the speech envelopes, thereby accentuating temporal cues that are important for consonant perception (Fig. 7). Envelope information is known to cue manner distinction. Stops, for instance, are characterized by a period of

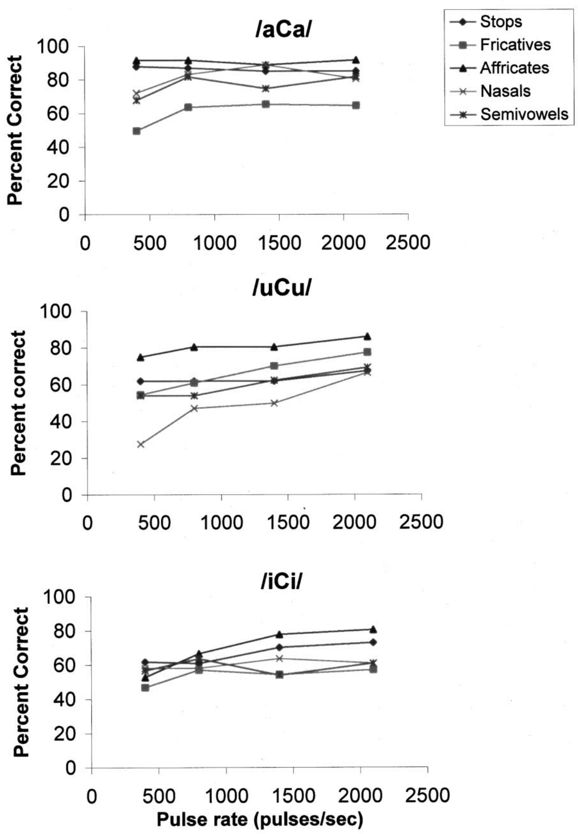

FIG. 6. Consonant class identification as a function of stimulation rate for three vowel contexts.

silence (closure) prior to signal onset. In contrast, nasals, semivowels, and fricatives lack that distinctive interval of silence before signal onset. In addition, nasals, fricatives, and semivowels are marked with relatively larger envelope amplitudes (containing more energy) during the release compared to the stop consonants. Therefore, assuming that implant devices transmit enough temporal information, the stop consonants should not be confused with fricatives, nasals, or semivowels. Yet, the low scores in manner identification obtained with low rates of stimulation suggest that some patients did confuse stops with fricatives, or stops with nasals, etc. Higher rates improved the temporal representation of the speech envelopes (see the example in Fig. 7), which in turn improved manner identification.

It is also important to note that since the speech envelopes are not sampled very often at the lower rates, it is possible that certain short-duration segments (e.g., burst) of the speech waveform may be missed altogether. This is illustrated in Fig. 7, which shows the pulsatile waveforms of the syllable /t i/ obtained at different rates. As shown in Fig. 7, the unvoiced stop consonant /t/ is marked by a period of silence (closure), followed by a burst and aspiration. As the pulse rate increases, the burst becomes more distinctive, and perhaps more salient perceptually. There seems to be no evidence of the burst at low rates, 200 or 400 pulses/s. This example clearly demonstrates that lower rates do not provide a good, if any at all, temporal representation of the burst in stop consonants. We believe that this is the reason we observed larger improvements in stop-consonant identification in the /iCi/ context than in the /aCa/ context with higher rates of stimulation. It is known from the speech perception literature that the burst is the dominant cue for the perception of stops in front-vowel environments (Smits *et al.*, 1996; Dorman *et al.*, 1977). In contrast, in the /a/ context the formant transitions are perceptually the most important for stop identification.

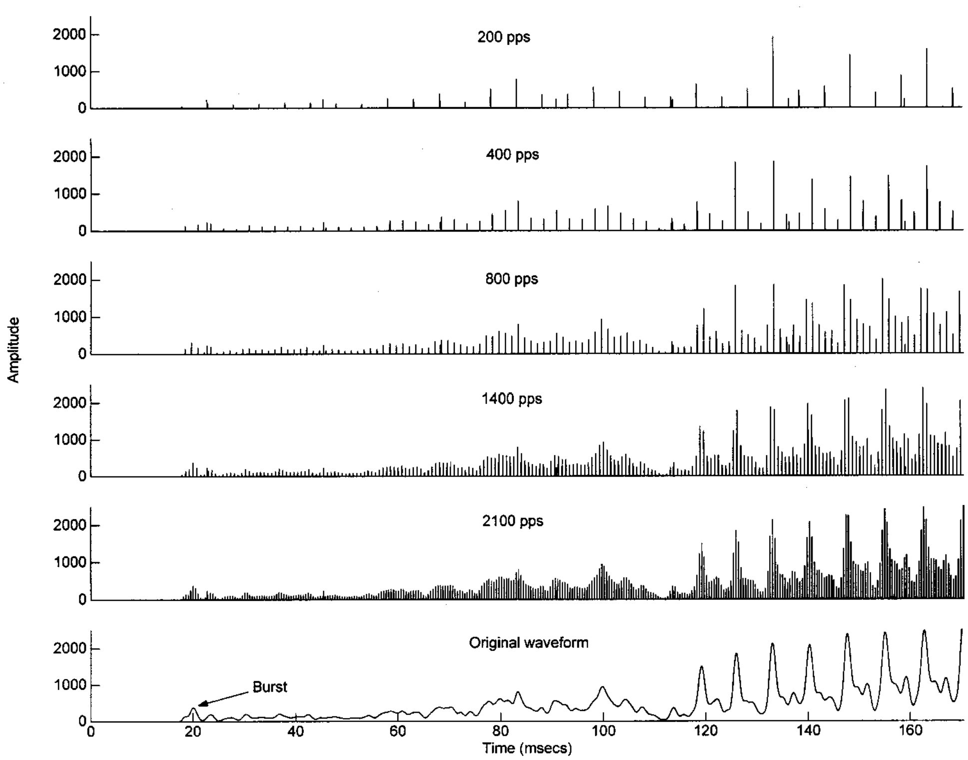

FIG. 7. The pulsatile waveforms for channel 5 of the syllable /t i/ obtained at five different stimulation rates. These waveforms were obtained by bandpass filtering the syllable /t i/ into six channels, performing envelope detection, and sampling the rectified envelopes at the rates indicated. Only the waveforms for channel 5 (with a center frequency of 3316 Hz) are shown. The bottom panel shows the original speech envelopes of channel 5. This figure shows the effect of stimulation rate in detecting short-duration segments (e.g., burst) of speech. As the pulse rate increases, the burst becomes more distinctive, and perhaps more salient perceptually.

4. Comparison with previous studies

Our findings on the effect of pulse rate on speech understanding are consistent with those by Wilson and colleagues (Wilson *et al.*, 1994; Lawson *et al.*, 1996). Lawson *et al.* (1996) reported that five Nucleus-22 patients with percutaneous access, who were fitted with a six-channel CIS processor, obtained a maximum performance on consonant recognition with a pulse rate of 833 pulses/s. There was a significant increase in performance when the rate increased from 250 to 833 pps. There were small, but not significant, differences in performance when the rate increased from 833 to 2525 pps. We suspect that the reason that Lawson *et al.* (1996) did not observe further improvements in performance for rates above 833 pps is because they used medial consonants in the /a/ context for testing. As discussed earlier, the /aCa/ consonant test is not sensitive enough to parametric variations of CIS processors.

Lawson *et al.* (1996) also compared the performance of the SPEAK strategy with a high-rate *n*-of-*m* strategy. Consonant recognition performance obtained with the high-rate (833 pps) 6-of-18 strategy was found to be significantly higher than the performance obtained with the 250-pps SPEAK strategy. The benefits in using high rates of stimulation with the CIS strategy compared to the relatively lower rates with the SPEAK strategy has been demonstrated by others in between-subjects comparisons (e.g., Kiefer *et al.*, 1996; Loizou *et al.*, 1997). Loizou *et al.* (1997) compared the performance of 11 Nucleus-22 patients using the SPEAK strategy with the performance of 7 Med-El/CIS-Link patients using the CIS strategy on consonant recognition in quiet and in +5-, +10-, and +15-dB S/N noise. Performance on consonant recognition was significantly higher with the high-rate CIS strategy than the SPEAK strategy in both quiet and noisy conditions. Feature analysis revealed that the information transmitted in manner and voicing was significantly higher with the CIS strategy. This is consistent with our findings in this study of improved manner identification with higher rates of stimulation.

Kiefer *et al.* (1997) also noted similar improvements in speech understanding with higher rates of stimulation for 13 patients using Med-El's COMBI-40 and COMBI-40+ devices. With eight channels active, Kiefer *et al.* (1997) examined vowel, consonant, and monosyllabic word recognition as a function of pulse rate (600, 1200, and 1500 pps). They found a statistically significant difference between the mean scores for monosyllabic word and consonant recognition at 600 and 1500 pps. There was no statistically significant difference between the mean vowel recognition scores at 1500 and 600 pps; an outcome consistent with our findings on vowel recognition. Brill *et al.* (1997) also examined the effect of stimulation rate for four subjects using Med-El's COMBI-40+ device, with only four channels active. Only four (out of 12) channels were activated in their study in order to examine extremely high rates of stimulation. Significant differences in word recognition were found between stimulation rates of 4545 and 800 pps.

Fu and Shannon (2000) recently investigated the effect of stimulation rate in six Nucleus-22 implant listeners fitted with a four-channel CIS processor. They varied the stimulation rate from 50 pulses/s to a maximum of 500 pulses/s and examined vowel and consonant recognition. Their results showed no effect on phoneme recognition for rates between 150 to 500 pulses/s, and a significant decrement in performance for rates lower than 150 pulses/s. Our results cannot be compared with theirs, since the maximum rate of stimulation in their study was constrained, due to hardware limitations, to rates lower than 500 pulses/s. In our study, the stimulation rate of 400 pulses/s was the lowest rate examined.

II. EXPERIMENT 2: EFFECT OF PULSE DURATION

In nonsimultaneous pulsatile stimulation, the pulse rate is directly related to the pulse duration. In general, short-duration pulses allow high rates of stimulation, and the smaller the pulse duration, the higher the rate. Therefore, in order to implement high-rate processors short-duration pulses need to be used. In contrast, low-rate processors can be implemented either by using short duration pulses and inserting an appropriate temporal gap between consecutive cycles (as in experiment 1), or by widening the pulse width. For instance, the 400-pps processor can be implemented either by using 40-μs/phase duration pulses and inserting a 2020-μs temporal gap between consecutive cycles, or by using 208-μs/phase pulses with no temporal gap between consecutive cycles. In both implementations the rate is the same; however, the pulse widths are different. This raises the question then: ''Do wider pulses offer any advantages over short pulses for a fixed rate?'' It is known from psychophysics studies (Shannon, 1993) that as the pulse duration increases, the thresholds decrease and the dynamic range increases. Given the increase in dynamic range associated with wider pulses, it is therefore of question as to whether wider pulses might be more beneficial for speech recognition for low-rate implementations of the CIS strategy. This question is addressed in this experiment by implementing 400- and 800-pps processors using wide and short duration pulses.

A. Method

1. Subjects

The subjects were the same as in experiment 1. Five of the six subjects were available for this experiment.

2. Speech material

The same /vCv/ consonants and multi-talker vowels of experiment 1 were used.

3. Procedure

The thresholds and most-comfortable levels were measured for the 104 and 208 μs/phase pulses using the ascending and descending methods of limits. We measured most-comfortable levels with an ascending method of limits, asking the subjects to tell us when the stimulus levels became most comfortable to them. Trains of biphasic pulses were presented in 50-ms bursts with a 500-ms interval between bursts.

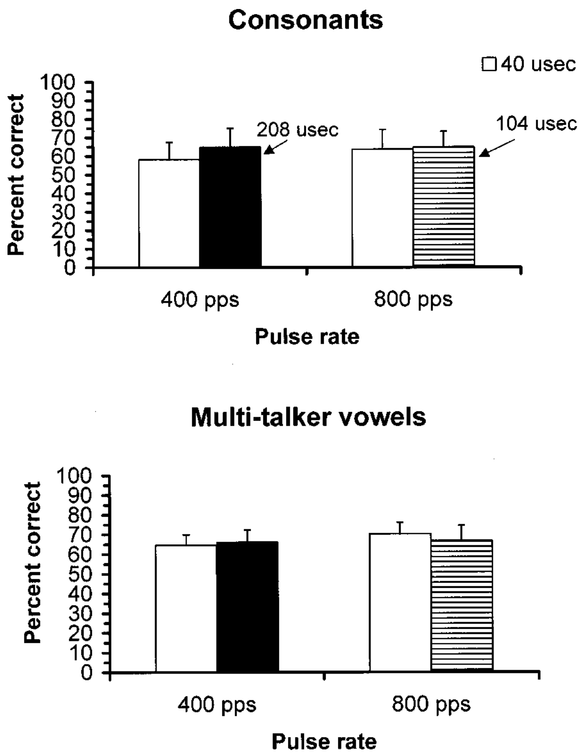

FIG. 8. Vowel and consonant recognition for 400- and 800-pps processors using short and wide pulse widths. Error bars indicate standard errors of the mean.

B. Results and discussion

The mean vowel and consonant scores for the two pulse-width conditions (208 and 104 μs/phase) are shown in Fig. 8, and are compared with the corresponding 400 and 800 rate conditions obtained using 40-μs/phase pulses in experiment 1. A two-factor (rate and pulse width) ANOVA with repeated measures showed a significant main effect of pulse width $[F(1,4)=10.4,\ p<0.05]$, a nonsignificant effect of pulse rate $[F(1,4)=2.7,\ p=0.174]$, and no interaction between pulse rate and pulse width $[F(1,4)=2.8,\ p=0.169]$ on consonant recognition. *Post hoc* tests performed on the 400-pulses/s rate condition showed that wider pulses produced a significantly ($p=0.02$) higher score on consonant recognition. There was no significant difference between the scores obtained for the 800-pulses/s condition. A two-factor ANOVA with repeated measures showed no main effect of pulse width $[F(1,4)=10.4,\ p<0.05]$, no effect of pulse rate $[F(1,4)=2.7,\ p=0.174]$, and no interaction between pulse rate and pulse width $[F(1,4)=2.8,\ p=0.169]$ on vowel recognition.

The above results demonstrated an advantage in using wider pulses at low rates for consonant recognition. For the same rate of stimulation (400 pps) conveying the same amount of temporal-envelope information within each channel, wider pulses (208 μs/phase) yielded a significantly higher performance on consonant recognition than short-duration (40 μs/phase) pulses. We suspect that this benefit is partly due to the increased dynamic range associated with wider pulses (Shannon, 1993). This outcome is in agreement with the findings by Loizou *et al.* (2000) about the significant effect of reduced dynamic range on speech recognition. Feature analysis of the consonant confusion matrices showed that the dynamic range affected mostly the reception of place information.

In this study, we only explored the effect of two different pulse widths per given rate. Other combinations of pulse rate and pulse duration are possible. In fact, one may construct a two-dimensional parameter space, in which the first dimension represents pulse duration and the second dimension represents pulse rate, and find the pulse rate–width combination that will yield the highest performance. Wilson and colleagues (Wilson *et al.*, 1993; Lawson *et al.*, 1993) performed such a two-dimensional parametric study of pulse rate and pulse duration. They reported an improvement in consonant recognition for 400-pps processors when wider pulses were used (Lawson *et al.*, 1993). For a four-channel processor, consonant recognition improved from 76% correct using 33-μs/phase pulses to 89% correct using 100-μs/phase pulses. For a six-channel processor consonant recognition improved from 85% correct with 33-μs/phase pulses to 90% correct with 200-μs/phase pulses. These results are consistent with our findings, in that wider pulses benefit consonant recognition for low-rate processors. The optimal combination of rate and pulse width, however, seemed to be dependent on the speaker (male versus female voice) as well as on the subject (Lawson *et al.*, 1993; Wilson *et al.*, 1993).

III. EXPERIMENT 3: EFFECT OF FILTER OVERLAP

The envelope signals in implant processors are typically extracted by processing the acoustic signal through a bank of bandpass filters spanning the signal bandwidth. The filter order (i.e., the number of filter coefficients) of the bandpass filter affects the filter roll-off, and consequently the overlap between contiguous filters. Higher-order filters have a steep roll-off, hence smaller overlap between adjacent channels. Conversely, lower-order filters are characterized by a shallow roll-off and therefore larger overlap between adjacent channels. A large filter overlap affects the spectral representation of the signal since it smears spectral information. In fact, broad (overlapping) filters have been used in the past to simulate broadened auditory filters in impaired hearing (e.g., Moore *et al.*, 1992) or channel interaction in cochlear implants (e.g., Shannon *et al.*, 1998). In this experiment we investigated the effect of spectral smearing on consonant identification by systematically varying the filter order to produce filters with varying degrees of overlap.

A. Subjects and test materials

Four implant subjects participated in this experiment. The same /vCv/ consonants in experiment 1 were used.

B. Signal processing and procedure

The CIS implementation was the same as in experiment 1. The only parameter that changed in this experiment was the filter order. Three different filter orders, fourth, eighth, and tenth, were investigated [sixth-order filters were used in experiment 1]. All the filters were Butterworth, and the roll-off of the fourth, sixth, eighth, and tenth filters was -20, -30, -45, and -60 dB/oct, respectively.

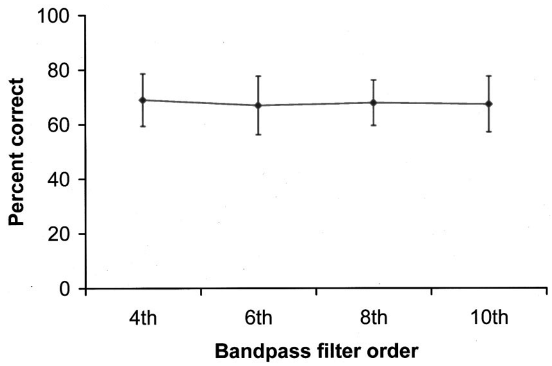

FIG. 9. Consonant recognition as a function of bandpass filter order. Error bars indicate standard errors of the mean.

The test was divided into three sessions, one for each filter condition. The three conditions were counterbalanced across subjects. All test sessions were preceded by one practice session in which the identity of the consonants was indicated to the listeners. The stimuli were presented directly to the subjects, through our laboratory processor, at a comfortable listening level. The pulse width was fixed at 40 μs/phase in all filter conditions yielding a stimulation rate of 2100 pulses/s.

C. Results and discussion

The mean consonant scores as a function of the filter order are shown in Fig. 9 and are compared with the scores obtained in Experiment 1 using sixth-order filters. As can be seen, the filter overlap seems to have no effect on consonant recognition. Repeated measures analysis of variance indicated no significant main effect of filter order on consonant recognition [$F(3,9)=0.276$, $p=0.841$].

Shannon *et al.* (1998), using noise-band simulations, did a similar experiment investigating the effect of filter overlap. Speech material was processed through four bandpass filters with varying degrees of overlap, and the envelope of each speech band was extracted through rectification and low-pass filtering. The four envelopes were modulated with band-limited noise, recombined, and presented to normal-hearing listeners for identification. Increasing the filter overlap produced a significant effect on consonant and sentence recognition; however, the absolute level of performance remained quite high even with −18-dB/oct filters. This is in agreement with our results obtained using −20-dB/oct filters (fourth order). The mean consonant recognition scores remained high even with −20-dB/oct filters.

The findings of this experiment have implications for the implementation of the CIS strategy. The filter order seemed to have no significant effect on consonant recognition, and the subjects reported no degradation in speech quality or timbre with lower-order filters. This suggests that one could implement the CIS strategy using fourth-order filters without sacrificing performance. This could reduce considerably the amount of computation required per cycle, and consequently reduce battery power.

This experiment has only investigated the effect of filter overlap on consonant recognition. It may well be that the amount of filter overlap is more important for vowel recognition. One would expect that steeper filters would be better in capturing formant information, particularly when the formants are very close to each other. Further studies are therefore needed to investigate the effect of filter overlap on vowel recognition.

IV. EXPERIMENT 4: EFFECT OF AMPLITUDE MAPPING FUNCTION

A major concern in designing cochlear implant processors is in the proper transformation of acoustic amplitude to electrical amplitude. Speech sounds in a normal conversation can range from 40 to 60 dB; however, implant listeners have only a dynamic range of 6–20 dB in electrical current. To accommodate for the smaller electrical dynamic range, the acoustic amplitudes are typically compressed using a logarithmic or power-law transformation. Logarithmic compression has been used because it has been shown to match loudness between electrical and acoustical amplitudes (Eddington *et al.*, 1978; Dorman *et al.*, 1993). Although the choice of compression function is important, only a few studies (e.g., Fu and Shannon, 1998; Wilson *et al.*, 1999) were performed to investigate the effect of the shape of the compression function on speech recognition.

The present study investigated the importance of the shape of the amplitude mapping function on consonant recognition. A range of amplitude mapping functions was created, from a strongly compressive function to a weakly—almost linear—compressive function, by varying the exponents of the power-law transformation [Eq. (1)]. Consonant recognition was examined as a function of the power exponent.

A. Subjects and test materials

Four implant subjects participated in this experiment. The same /vCv/ consonants in experiment 1 were used.

B. Signal processing and procedure

The CIS implementation was the same as in experiment 1. The pulse width was fixed at 40 μs/phase for all mapping conditions yielding a stimulation rate of 2100 pulses/s. The only parameter that changed in this experiment was the power exponent, p, used in Eq. (1). The exponent of the power function was systematically changed to −0.1, 0.2, and 0.6. Figure 10 shows the amplitude mapping functions obtained using different power exponents. The most compressive function was obtained using $p=-0.1$ and the least compressive (nearly linear) function was obtained using $p=0.6$. The function obtained using $p=-0.0001$ was logarithmic, and was the default exponent value used in our CIS implementation (experiment 1).

The test was divided into three sessions, one for each mapping condition. The three conditions were counterbalanced across subjects. All test sessions were preceded by one practice session in which the identity of the consonants was indicated to the listeners. The stimuli were presented directly to the subjects, through our laboratory processor, at a comfortable listening level.

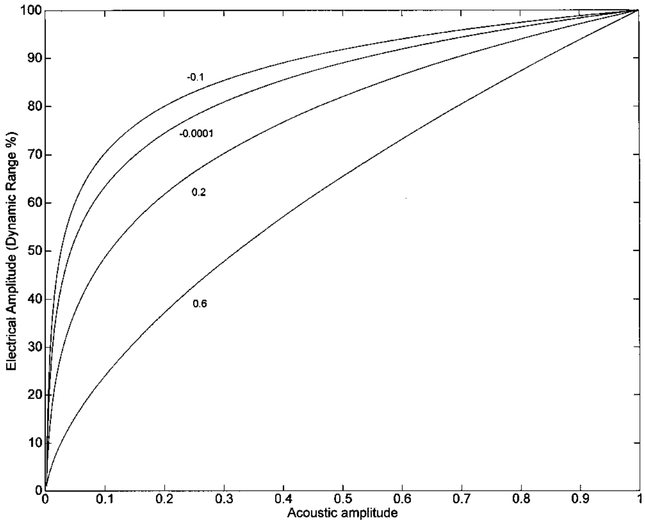

FIG. 10. The set of amplitude mapping functions used in this study.

C. Results and discussion

The results are shown in Fig. 11. Performance remained constant at about 70% correct when the value of the power exponent ranged from -0.1 to 0.2, and dropped to 40% correct when $p=0.6$. Repeated measures ANOVA showed a significant main effect $[F(3,9)=27.335,\ p<0.0005]$ of power exponent on consonant recognition. *Post hoc* tests showed no significant differences $(p>0.5)$ between the scores obtained with $p=-0.0001$ (from experiment 1), $p=-0.1$, and $p=0.2$. There was, however, a significant difference $(p<0.05)$ between the scores obtained with $p=0.2$ and $p=0.6$.

The present results indicate only a mild dependence of performance on the value of the power exponent, and consequently the shape of the amplitude mapping function, over a broad range of exponents. These findings are consistent with those by Boëx *et al.* (1995), Fu and Shannon (1998), Wilson *et al.* (1999), and Zeng and Galvin (1999). Boëx *et al.* (1995, 1997) investigated the use of mapping functions that preserved normal loudness growth in cochlear implants. They fitted four subjects with wearable Geneva/RTI/MEEI processors, and they reported that two subjects preferred and performed better with normal-loudness growth functions than the standard logarithmic functions. Fu and Shannon (1998) varied the exponent of the power-law mapping function from a strongly compressive $(p=0.05)$ to a weakly compressive value $(p=0.75)$ and investigated the effect of compression on phoneme recognition using 3 Nucleus-22 implant users fitted with a four-channel CIS processor. Their study showed that the function relating power exponent and performance was relatively flat, suggesting that phoneme recognition was only mildly affected by the shape of the compression function, at least for power exponents $p\leqslant 0.5$. There was a significant drop in performance when the mapping function was nearly linear $(p=0.75)$, consistent with our findings. Best performance was obtained for power exponents that restored normal loudness growth. Wilson *et al.* (1999) recently replicated the study by Fu and Shannon (1998) and also found a relatively uniform performance on consonant recognition over a wide range of exponent values for two subjects wearing the Med-El/CIS-link device. Vowel recognition was hardly affected by the different manipulations in mapping functions. Zeng and Galvin (1999) reported similar findings when they varied the amount of compression by manipulating the Q-value in the Nucleus-22 processor with four implant listeners. The Q-values used (20, 30, 40, and 50) corresponded to power exponents 0.24, 0.4, 0.51, and 0.63, respectively. Vowel and consonant recognition were not significantly affected by the amount of compression.

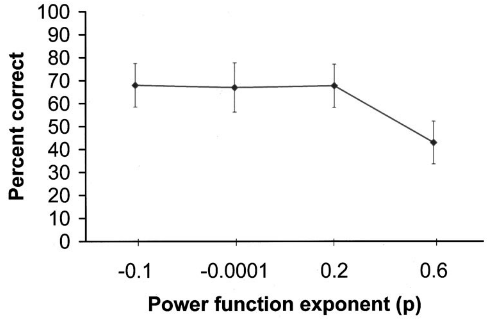

FIG. 11. Consonant recognition as a function of the power exponent of the mapping function. Error bars indicate standard errors of the mean.

The above experiments were performed in quiet. A different pattern of performance was observed for experiments conducted in noise. Results from a single subject in the Wilson *et al.* (1999) study indicated that less compressive mapping functions might be more helpful for listening to speech in noise. For consonants in +15-dB S/N speech-shaped noise, the power exponent of 0.4 produced the highest score. Fu and Shannon (1999) observed a different outcome in a similar study performed in noise. Maximum consonant performance in both quiet and noisy conditions was obtained with the exponent of 0.2 corresponding to logarithmic mapping. The overall effect of the amplitude mapping function, however, was different in noise. Performance with weakly compressive mapping functions declined mildly in noise, whereas performance with strongly compressive amplitude mapping declined dramatically in noise. In contrast to the above two studies, Zeng and Galvin (1999) found no significant effect of compression in noise, for consonant and vowel recognition.

In summary, our results and those by others confirm that in quiet logarithmic mapping functions are needed for accurate consonant and vowel recognition. More studies are needed, however, to investigate how the acoustic signal mapping should be done in noise. It may well be that different mapping functions are needed for noisy conditions.

V. EXPERIMENT 5: EFFECT OF SIGNAL BANDWIDTH

The signal bandwidth is constrained, according to Nyquist's theorem, to half the sampling frequency. Accordingly, the higher the sampling frequency, the higher the signal bandwidth. Although a high bandwidth (>10 kHz) is needed for music perception (and enjoyment), a considerably smaller bandwidth is needed for speech perception. For vowels, a bandwidth of 3 kHz is needed for the perception of the first three formants (Peterson and Barney, 1952; Hillenbrand *et al.*, 1995). A relatively higher bandwidth is needed for the perception of consonants, such as fricatives and stops. The spectral peaks of /f/ and /s/, for instance, could be as high as 8.5 kHz (Manrique and Massone, 1981). Limiting the spectrum to a small bandwidth, such as the 3.2-kHz telephone bandwidth, for instance, can potentially produce consonant confusions between f/s, t/k, etc. It seems reasonable, then, to expect that a wide bandwidth might be beneficial for consonant recognition.

Indeed, recent experiments by Zerbi *et al.* (1998) showed that a wide bandwidth was beneficial for the perception of consonants, and, in particular, consonants produced by female talkers. Significant improvements on consonant recognition were obtained when the bandwidth increased from 5.5 to 9.5 kHz. Zerbi *et al.* (1998) found that the addition of a high-frequency band spanning from 5.6 to 9.5 kHz helped reduce principal confusions (e.g., t/k, f/s, p/t) among medial consonants uttered by a female talker. No significant improvements were found, however, for consonants produced by a male talker.

The two bandwidths (5.5 and 9.5 kHz) examined in the Zerbi *et al.* study were one octave apart, hence it was not clear whether any bandwidth in between would improve consonant recognition. We therefore investigated two additional bandwidths, in the present experiment, in the range of 5.5 to 9.9 kHz. The results on the 6.7-kHz bandwidth were reported in experiment 1, and are included here for comparative purposes. The two additional bandwidths examined were 8.4 and 9.9 kHz. At question in this experiment was whether a high (> 5.5 kHz) signal bandwidth would benefit consonant and/or vowel recognition.

A. Subjects and test materials

Five implant subjects participated in this experiment. The same /vCv/ consonants and multi-talker vowels of experiment 1 were used.

B. Signal processing and procedure

The CIS implementation was the same as in experiment 1. Two different antialiasing filters were used to limit the signal bandwidth to 8.4 and 9.9 kHz, respectively. The sampling frequency was fixed at 22 kHz for both bandwidth conditions. The center frequencies of the six analysis filters were obtained by dividing the bandwidth into six logarithmically equal bands. The center frequencies for the 8.4-kHz bandwidth were 472, 803, 1365, 2320, 3943, and 6703 Hz. The center frequencies for the 9.9-kHz bandwidth were 481, 839, 1464, 2555, 4460, and 7786 Hz. Sixth-order Butterworth filters were used.

The test was divided into two sessions, one session for each bandwidth condition. The two conditions were counterbalanced across subjects. All test sessions were preceded by one practice session in which the identity of the consonants/vowels was indicated to the listeners. The stimuli were presented directly to the subjects, through our laboratory processor, at a comfortable listening level. The pulse width was fixed at 40 μs/phase in all bandwidth conditions yielding a stimulation rate of 2100 pulses/s.

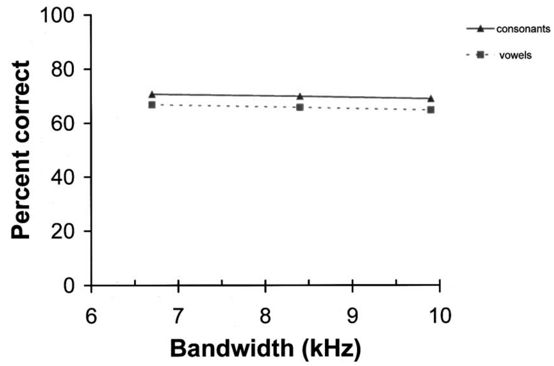

FIG. 12. Mean vowel and consonant recognition performance as a function of signal bandwidth. The standard errors of the mean for the consonants were 9.1%, 8.9%, and 7.3% for the bandwidths of 6.7, 8.4, and 9.9 kHz, respectively. The standard errors of the mean for the vowels were 6.5%, 9.1%, and 7.3% for the bandwidths of 6.7, 8.4, and 9.9 kHz, respectively.

C. Results and discussion

The results are shown in Fig. 12. Performance on consonant recognition remained constant at roughly 70% correct for all three bandwidth conditions, while performance on vowel recognition remained constant at roughly 65% correct. Repeated measures ANOVA showed no significant main effect [$F(2,8)=0.53$, $p=0.608$] of signal bandwidth on consonant recognition, and no significant main effect [$F(2,8)=0.54$, $p=0.5$] on vowel recognition.

Our results indicate that the recognition of consonants and vowels is not affected by the signal bandwidth, at least in quiet and for bandwidths in the range of 6.7–9.9 kHz. Zerbi *et al.* (1998) found a significant difference between the performance on consonant recognition obtained at 5.5 and 9.5 kHz. In the present study we showed that a 6.7-kHz bandwidth is sufficient for accurate consonant recognition. Using a higher (>6.7 kHz) bandwidth did not seem to benefit consonant recognition. We suspect that this is because the number of filters allotted for high frequencies was not increased when the bandwidth increased. In all three conditions, two filters were allotted for frequencies above 3 kHz. One could conceivably choose a different filter allocation allowing more filters in the high frequencies. Such an arrangement, however, would most likely impair vowel recognition where more filters are needed to capture low-frequency information in the F1–F2 region.

VI. CONCLUSIONS

(i) Of the five speech processing parameters examined in this study, the pulse rate and the pulse width had the largest (positive) effect on speech recognition. For a fixed pulse width, higher rates (2100 pps) of stimulation provided significantly more benefit on speech recognition compared to lower rates (400 pps) of stimulation, consistent with the findings by Lawson *et al.* (1996), Brill *et al.* (1997), and Kiefer *et al.* (1997). When we jointly varied the pulse rate and the pulse width, we found that the pulse width had a significant effect on consonant recognition, at least for low-rate (400 pps) processors. This interaction between pulse rate and pulse width suggests that it is possible to achieve high levels of consonant recognition by a slower pulse rate (400 ps) and an appropriate choice of pulse width. Audiologists or clinicians can therefore optimize a subject's performance either by (1) using a short-duration pulse (say, 40 μs/phase) and a high pulse rate (as per experiment 1), or (2) by jointly varying the pulse width and pulse rate (as per experiment 2). Of the two methods that can be used to optimize implant listener's performance, the former is less time consuming.

(ii) The finding that the 2100-pps processors produced a significantly higher performance on open-set recognition than the 800-pps processors suggests that implant manufacturers ought to support high stimulation rates (>800 pps) to improve the benefits of cochlear implants.

(iii) There was a significant effect of stimulation rate on consonant recognition. Higher rates of stimulation produced a significant benefit to consonant recognition. This benefit stemmed primarily from improved manner identification.

(iv) The effect of stimulation rate on consonant recognition was highly dependent on the vowel context. Higher rates benefited mostly consonants in the /uCu/ and /iCi/ contexts. A small benefit was observed for consonants in the /aCa/ context. This finding strongly recommends the use of multi-vowel consonants for testing parametric variations of implant processors. The /aCa/ consonant test is not sensitive enough to parametric variations of implant processors.

(v) For low-rate (<800 pps) processors, wider pulse widths provided a significantly higher benefit on consonant recognition than narrower pulse widths. We suspect that this is due to the increased dynamic range associated with wider pulses. There was no effect of pulse width on vowel recognition. As demonstrated in this study, as well by others (Wilson *et al.*, 1993; Lawson *et al.*, 1993), it is possible to optimize a subject's performance by varying both pulse rate and pulse width. The best combination of pulse rate and pulse width, however, might be subject dependent.

(vi) The amount of filter overlap seemed to have no effect on consonant recognition. Filters with a −20-dB/oct roll-off produced the same performance as filters with a −60-dB/oct roll-off.

(vii) The shape of the amplitude mapping function had only a minor effect on performance, with the lowest performance obtained when a nearly linear mapping function ($p=0.6$) was used. Performance did not seem to be affected by the value of the power exponents over a broad range of exponents, −0.1 to 0.2, including the exponent −0.0001 corresponding to logarithmic mapping. The present logarithmic mapping function used in current implant devices appears to be a good choice for many subjects, at least in quiet.

(viii) Consonant and vowel recognition was not affected by the signal bandwidth. There was no statistically significant difference in performance between the 6.7-, 8.4-, and 9.9-kHz signal bandwidths.

ACKNOWLEDGMENTS

This research was supported by Grant No. R01 DC03421 from the National Institute of Deafness and other Communication Disorders, NIH.

[1]The power exponent p was set equal to -0.0001 to match the logarithmic compression function, of the form $y=A\log(1+cx)+B$, used in the Med-El device, where A and B are constants used for mapping the acoustic signal x between threshold and most-comfortable level, and c is a constant that controls the shape of the compression function ($c=512$, in our case). It should be noted that the value of the power exponent p corresponding to logarithmic mapping depends on the input range, and, in particular, on the minimum value of x (Xmin). If Xmin=0, then the value of $p=0.2$ will yield a logarithmic mapping similar to the one used in the Med-El device, whereas if Xmin=1, then the value of $p=-0.0001$ will yield a logarithmic mapping. In our implementation, Xmin=1.

Boëx, C., Pelizzone, M., Piloux, V., and Montandon, P. (**1995**). ''Use of loudness scaling measurements to determine compressive mapping in speech processing for cochlear implants,'' Abstracts of 1995 Conference on Implantable Auditory Prostheses, p. 57.

Boëx, C., Eddington, D., Noel, V., Rabinowitz, W., Tierney, J., and Whearty, W. (**1997**). ''Restoration of normal loudness growth for CIS sound coding strategies,'' Abstracts of 1997 Conference on Implantable Auditory Prostheses, p. 26.

Brill, S., Gstottner, W., Helms, J., Ilberg, C. v., Baumgartner, W., Muller, J., and Kiefer, J. (**1997**). ''Optimization of channel number and stimulation rate for the fast CIS strategy in the COMBI 40+,'' Am. J. Otol. **18**, S104–S106.

Dorman, M., and Loizou, P. (**1997**). ''Changes in speech intelligibility as a function of time and signal processing strategy for an Ineraid patient fitted with CIS processors,'' Ear Hear. **18**(2), 147–155.

Dorman, M., Smith, L., and Parkin, J. (**1993**). ''Loudness balance between acoustic and electric stimulation by a patient with a multichannel cochlear implant,'' Ear Hear. **14**, 290–292.

Dorman, M., Studdert-Kennedy, M., and Raphael, L. (**1977**). ''Stop-consonant recognition: Release burst and formant transitions as functionally equivalent, context-dependent cues,'' Percept. Psychophys. **22**, 109–122.

Eddington, D., Dobelle, W., Brachman, D., Mladevosky, M., and Parkin, J. (**1978**). ''Auditory prosthesis research using multiple intracochlear stimulation in man,'' Ann. Otol. Rhinol. Laryngol. **87**(S53), 1–39.

Francois, J., Tinembart, J., Bessat, C., Leone, P., Rossman, F., and Pelizzone, M. (**1994**). ''Implants cochle aires: Un processeur portable pour le developpement de l'algorithme CIS,'' *Actes de le conference DSP 94*, Paris.

Fu, Q-J., and Shannon, R. (**1998**). ''Effect of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners,'' J. Acoust. Soc. Am. **104**, 2570–2577.

Fu, Q-J., and Shannon, R. (**1999**). ''Phoneme recognition by cochlear implant users as a function of signal-to-noise ratio and nonlinear amplitude mapping,'' J. Acoust. Soc. Am. **106**, L18–L23.

Fu, Q-J., and Shannon, R. (**2000**). ''Effect of stimulation rate on phoneme recognition by Nucleus-22 cochlear implant listeners,'' J. Acoust. Soc. Am. **107**, 589–597.

Hillenbrand, J., Getty, L., Clark, M., and Wheeler, K. (**1995**). ''Acoustic characteristics of American English vowels,'' J. Acoust. Soc. Am. **97**, 3099–3111.

Kiefer, J., Muller, J., Pfennigdorff, T., Schon, F., Helms, J., von Ilberg, C., Baungartner, W., Gstottner, W., Ehrenberger, K., Arnold, W., Stephan, K., Thumfart, W., and Baur, S. (**1996**). ''Speech understanding in quiet and in noise with the CIS speech coding strategy (Med-El Combi 40) compared to the Multipeak and Spectral Peak strategies (Nucleus),'' ORL, **58**, 127–135.

Kiefer, J., Ilberg, C. v., Rupprecht, V., Hubnet-Egener, J., Baumgartner, W., Gstottner, W., Forgasi, K., and Stephan, K. (**1997**). ''Optimized speech understanding with the CIS speech coding strategy in cochlear implants:

the effect of variations in stimulus rate and number of channels,'' Abstracts of Vth International Cochlear Implant Conference, New York, NY.

Lawson, D., Wilson, B., and Zerbi, M. (**1993**). ''Speech Processors for Auditory Prostheses,'' NIH Project N01-DC-2-2401, Second Quarterly Progress Report.

Lawson, D., Wilson, B., Zerbi, M., and Finley, C. (**1996**). ''Speech Processors for Auditory Prostheses,'' NIH Project N01-DC-5-2103, Third Quarterly Progress Report.

Loizou, P. (**1998**). ''Mimicking the human ear: An overview of signal processing techniques for converting sound to electrical signals in cochlear implants,'' IEEE Signal Process. Mag. **15**(5), 101–130.

Loizou, P., Dorman, M., and Fitzke, J. (**2000**). ''The effect of reduced dynamic range on speech understanding: Implications for patients with cochlear implants,'' Ear Hear. **21**, 25–31.

Loizou, P., Dorman, M., and Powell, V. (**1998**). ''The recognition of vowels produced by men, women, boys and girls by cochlear implant patients using a six-channel CIS processor,'' J. Acoust. Soc. Am. **103**, 1141–1149.

Loizou, P., Graham, S., Dickins, J., Dorman, M., and Poroy, O. (**1997**). ''Comparing the performance of the SPEAK strategy (Spectra 22) and the CIS strategy (Med-El) in quiet and in noise,'' Abstracts of 1997 Conference on Implantable Auditory Prostheses.

Manrique, A., and Massone, M. (**1981**). ''Acoustic analysis and perception of Spanish fricative consonants,'' J. Acoust. Soc. Am. **69**, 1145–1153.

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

Moore, B., Glasberg, B., and Simpson, A. (**1992**). ''Evaluation of a method of simulating reduced frequency selectivity,'' J. Acoust. Soc. Am. **91**, 3402–3423.

Osberger, M., and Fisher, L. (**1999**). ''SAS-CIS preference study in postlingually deafened adults implanted with the Clarion cochlear implant,'' Ann. Otol. Rhinol. Laryngol. **108**, 74–79.

Peterson, G., and Barney, H. (**1952**). ''Control methods used in a study of vowels,'' J. Acoust. Soc. Am. **24**, 175–184.

Poroy, O., and Loizou, P. (**2000**). ''Development of a speech processor for laboratory experiments with cochlear implant patients,'' *IEEE International Conference on Acoustics Speech and Signal Processing*, Istanbul, Turkey.

Shannon, R. (**1993**). ''Psychophysics,'' in *Cochlear Implants: Audiological Foundatations*, edited by R. Tyler (Singular Publishing Group).

Shannon, R., Zeng, F-G., and Wygonski, J. (**1998**). ''Speech recognition with altered spectral distribution of envelope cues,'' J. Acoust. Soc. Am. **104**, 2467–2476.

Shannon, R., Jensvold, A., Padilla, M., Robert, M., and Wang, X. (**1999**). ''Consonant recordings for speech testing,'' J. Acoust. Soc. Am. **106**, L71–L74.

Smits, R., Bosch, L., and Collier, R. (**1996**). ''Evaluation of various sets of acoustic cues for the perception of prevocalic stop consonants. I. Perception experiment,'' J. Acoust. Soc. Am. **100**, 3852–3864.

Wilson, B., Lawson, D., and Zerbi, M. (**1993**). ''Speech Processors for Auditory Prostheses,'' NIH Project N01-DC-2-2401, Fifth Quarterly Progress Report.

Wilson, B., Lawson, D., and Zerbi, M. (**1995**). ''Advances in coding strategies for cochlear implants,'' Advances in Otolaryngology–Head and Neck Surgery **9**, 105–129.

Wilson, B., Finley, C., Lawson, D., and Zerbi, M. (**1998**). ''Temporal representations with cochlear implants,'' Am. J. Otology **18** (Suppl.), S30–S34.

Wilson, B., Lawson, D., Finley, C., and Zerbi, M. (**1991**). ''Speech Processors for Auditory Prostheses,'' NIH Project N01-DC-9-2401, Tenth Quarterly Progress Report.

Wilson, B., Lawson, D., Zerbi, M., and Wolford, R. (**1999**). ''Speech Processors for Auditory Prostheses,'' NIH Project N01-DC-8-2105, Third Quarterly Progress Report.

Zeng, F-G., and Galvin, J. (**1999**). ''Amplitude mapping and phoneme recognition in cochlear implant listeners,'' Ear Hear. **20**, 60–74.

Zerbi, M., Lawson, D., and Wilson, B. (**1998**). ''Speech Processors for Auditory Prostheses,'' NIH Project N01-DC-5-2103, Tenth Quarterly Progress Report.

Zwolan, T., Collins, L., and Wakefield, G. (**1997**). ''Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects,'' J. Acoust. Soc. Am. **102**, 3673–3685.