Cochlear-Implant High Pulse Rate and Narrow Electrode Configuration Impair Transmission of Temporal Information to the Auditory Cortex

John C. Middlebrooks

Department of Otolaryngology Head and Neck Surgery, Kresge Hearing Research Institute, University of Michigan, Ann Arbor, Michigan

Submitted 8 October 2007; accepted in final form 25 April 2008

Middlebrooks JC. Cochlear-implant high pulse rate and narrow electrode configuration impair transmission of temporal information to the auditory cortex. J Neurophysiol 100: 92–107, 2008. First published April 30, 2008; doi:10.1152/jn.01114.2007. In the most commonly used cochlear prosthesis systems, temporal features of sound are signaled by amplitude modulation of constant-rate pulse trains. Several convincing arguments predict that speech reception should be optimized by use of pulse rates ≥2,000 pulses per second (pps) and by use of intracochlear electrode configurations that produce restricted current spread (e.g., bipolar rather than monopolar configurations). Neither of those predictions has been borne out in consistent improvements in speech reception. Neurons in the auditory cortex of anesthetized guinea pigs phase lock to the envelope of sine-modulated electric pulse trains presented through a cochlear implant. The present study used that animal model to quantify the effects of carrier pulse rate, electrode configuration, current level, and modulator waveform shape on transmission of temporal information from a cochlear implant to the auditory cortex. Modulation sensitivity was computed using a signal-detection analysis of cortical phase-locking vector strengths. Increasing carrier pulse rate in 1-octave steps from 254 to 4,069 pps resulted in systematic decreases in sensitivity. Comparison of sine-versus square-wave modulator waveforms demonstrated that some, but not all, of the loss of modulation sensitivity at high pulse rates was a result of the decreasing size of pulse-to-pulse current steps at the higher rates. Use of a narrow bipolar electrode configuration, compared with the monopolar configuration, produced a marked decrease in modulation sensitivity. Results from this animal model suggest explanations for the failure of high pulse rates and/or bipolar electrode configurations to produce hoped-for improvements in speech reception.

INTRODUCTION

Cochlear implants are highly successful neural prostheses that can restore hearing to severely and profoundly deaf people. The most commonly used speech processors stimulate implants with amplitude-modulated electrical pulse trains. The modulation waveform, derived from the envelope of band-passed acoustic input, carries essentially the only information about temporal characteristics of sound available to the implant user. Among human cochlear implant users, sensitivity to modulation tends to correlate with success in speech reception. That is, users with lower modulation-detection thresholds tend to show better speech-reception performance (vowel and consonant recognition: Caazals et al. 1994; Fu 2002; vowel and hearing-in-noise-test [HINT] sentence recognition: Colletti and Shannon 2005). It might be that poor speech reception is a necessary consequence of poor modulation sensitivity. Alternatively, both poor modulation sensitivity and poor speech reception might result from some third factor, such as poor neural survival or poor implant placement. Regardless of causality, it seems intuitively clear that stimulus conditions that impair modulation sensitivity would restrict the information about temporal aspects of speech available to implant users. It should be a goal of speech-processor design to optimize modulation sensitivity.

Two key parameters of the electrical stimulus that can be manipulated by designers of speech processors are the carrier pulse rate and the electrode configuration. The widely used SPEAK processing strategy uses a pulse rate of 250 pulses per second (pps), and many present-day processors use rates around 900 to 2,400 pps. It has been argued that higher pulse rates should provide more detailed sampling of speech waveforms and thus better transmission of temporal information about speech waveforms (Wilson 1997). Also, rates ≥2,000 pps are thought to elicit more natural, “stochastic” firing patterns from the auditory nerve (Rubinstein et al. 1999), which might enhance hearing. Despite these theoretical arguments, systematic studies of speech testing with pulse rates ≥1,400 pps have not yielded the consistent improvements that have been hoped for. The present study quantifies in an animal model the influence of carrier pulse rate on modulation sensitivity and suggests that the disappointing speech-reception results with high pulse rates might reflect impaired transmission of temporal information to the auditory cortex.

The most commonly used configurations for electrical stimulation through an intrascalar cochlear implant are monopolar and bipolar. In the monopolar configuration, the active electrode is a single intrascalar electrode and the return path is through an extracochlear electrode. In the bipolar configuration, the active electrode is an intrascalar electrode and the return is through another intrascalar electrode; “narrow” and “broad” bipolar refer to the spacing between the active and return electrodes. Monopolar stimulation is thought to produce relatively broad electrical fields and, in recordings from the central auditory system, has been shown to produce a broad spread of excitation across the tonotopic dimension (Bierer and Middlebrooks 2002; Rebscher et al. 2001; Snyder et al. 2004, 2008). Narrow bipolar stimulation produces more restricted electrical fields and more focal tonotopic excitation. Physiological and psychophysical studies have also shown that bipolar stimulation results in enhanced discrimination among channels stimulated sequentially (Busby et al. 1994; Middlebrooks and Bierer 2002) and decreased interference among channels stimulated simultaneously (Bierer and Middlebrooks 2004; Middlebrooks and Middlebrooks 2005). It might be that poor speech reception is a consequence of poor modulation sensitivity.

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Address for reprint requests and other correspondence: J. C. Middlebrooks, Department of Otolaryngology, University of California, Irvine, CA 92697-5310 (E-mail: jmidd@umich.edu).
Boëx et al. 2003a; Shannon 1983); results of forward masking in humans, however, do not consistently show differences between monopolar and bipolar configurations in terms of spread of excitation (e.g., Kwon and van den Honert 2006). Most physiological and psychophysical results suggest that bipolar stimulation would produce better transmission of spectral (i.e., cochlear place) information and thus better speech reception than would monopolar stimulation. Surprisingly, tests of speech reception generally show that speech-reception performance with bipolar stimulation is not consistently better—indeed, is sometimes worse—than with monopolar and that implant users generally prefer monopolar stimulation over bipolar (Kileny et al. 1998; Lehhardt et al. 1992; Pfingst et al. 1997, 2001; von Wallenberg et al. 1995; Zwolan et al. 1996). The present study addresses the possibility that narrow bipolar stimulation might have a negative impact on modulation sensitivity.

The companion paper (Middlebrooks 2008) demonstrated in an animal model that neurons in the auditory cortex can phase lock significantly to modulation waveforms across the range of modulation frequencies and depths relevant to speech reception. That paper also showed that a putative rate code for modulation frequency would be confounded by modulation depth, although that does not rule out a rate code for simple detection of modulation. The present study developed measures of modulation sensitivity in the animal model based on cortical phase locking and on spike count and it quantified effects of current level, pulse rate, modulation wave shape, and electrode configuration on modulation sensitivity. The report demonstrates substantial effects of each of those parameters and begins to explore mechanisms underlying those effects. Preliminary results were previously presented in abstracts (Middlebrooks 2005; Middlebrooks and Lee 2004).

METHODS
Stimulus presentation and data acquisition

Results presented in this study were drawn from experiments in the same 20 ketamine/xylazine-anesthetized guinea pigs presented in the companion paper (Middlebrooks 2008) plus one guinea pig in which only square-wave modulation waveforms were tested. All experiments with animals were conducted with approval of the University of Michigan Committee on Use and Care of Animals. All stimulus presentation and data acquisition procedures were identical to those in the companion paper, including electrical stimulation with intrascalar arrays of platinum-banded electrodes and cortical recording with a 16-site silicon-substrate recording probe. The present study analyzes the parametric dependence of modulation depth threshold using an expanded stimulus set. Stimuli consisted of sinusoidally amplitude modulated electrical pulse trains presented through an intrascalar electrode array. Individual pulses were symmetrically biphasic, initially cathodic, 40.96 μs per phase, with no interphase gap. Carrier pulse rates were 254, 509, 1,017, 2,034, and 4,069 pps. Those particular pulse rates were chosen because their pulse-to-pulse periods are integer multiples of the 40.96-μs phase duration. A square-wave modulation waveform was also tested, as described in RESULTS. All of the modulation frequencies tested with square-wave modulation were integer multiples of 21.19 Hz, which guaranteed an equal number of carrier pulses during high- and low-current phases of the square wave. For about half of the data collected using 254- and 4,069-pps carriers and sine-wave modulation, modulation frequencies were integer multiples of 20 Hz. No difference was noted in responses to 20- or 21.19-Hz-multiple modulation frequencies in sine-modulated conditions. In some of the summary plots, data collected with 20-, 40-, and 60-Hz sine-wave modulators are combined with those from 21.19-, 42.38-, and 63.57-Hz modulators, respectively, and are labeled “20–21,” “40–42,” and “60–64.” Modulator waveforms were presented in sine starting phase. Pilot experiments showed no difference in sensitivity to sine and cosine starting phases.

The monopolar stimulus configuration used the next-to-most-apical intrascalar electrode as active and a return path through a wire in a neck muscle. The present study adds also a narrow bipolar stimulation configuration condition in which the active electrode was the next-to-most-apical intrascalar electrode and the return electrode was the adjacent, most apical, electrode. Not all carrier rates or electrode configurations were tested in all animals; the numbers of animals and recording sites for each stimulus condition are indicated in each figure legend.

As in the companion paper, this report uses the term “unit activity” to indicate both single units and unresolved activity from two or more units. Many of the parametric comparisons involved recordings spanning several hours. For that reason, one cannot be sure that the sampled unit populations were identical throughout the recording period. Because of that uncertainty, quantitative comparisons of responses among stimulus conditions refer to particular recording sites rather than particular units. The order of testing of carrier rates, sine- and square-wave modulator waveforms, and electrode configurations was varied among animals to avoid bias that might have resulted from drift in the quality of physiological recording.

Electrical artifact could be detected at the cortical recording sites and, if not attenuated, potentially could have contaminated neural spike detection. For the 254- and 508-pps carriers, electrical artifact was eliminated using the sample-and-hold procedure described previously (Middlebrooks 2008). That procedure resulted in loss of data during 3.6 and 7.3%, respectively, of each period of the 254- and 509-pps pulse trains. For 1,017-, 2,034-, and 4,069-pps carriers, comb filters programmed into the recording path introduced spectral nulls at integer multiples of the pulse rates; a different filter was used for each carrier rate. That procedure eliminated the artifact entirely except for periods of <1 ms at pulse-train onset and offset as the filter charged and discharged. Those brief well-defined periods of artifact were eliminated during off-line spike sorting. The 1,017- and 2,034-Hz comb filters markedly distorted the spike waveforms, but that distortion did not interfere with the detection of spikes.

Data analysis

Thresholds for detection of stimulus modulation were determined using a procedure derived from Signal Detection Theory (Green and Swets 1966; Macmillan and Creelman 2005). The procedure relied on computations of the vector strength of cochlear neuronal phase locking to a modulator waveform (Goldberg and Brown 1969; Middlebrooks 2008). Analysis was based on the epoch 100 to 600 ms after onset of the pulse train; responses prior to 100 ms were excluded to avoid the erroneous impression of phase locking due to the temporally compact onset response. Response patterns on single trials often contained too few spikes to provide reliable estimates of vector strength. For that reason, the analysis was based on multiple “bootstrap” averages of four spike patterns (Efron and Tibshirani 1991): given spike patterns elicted by each of 20 presentations of a particular stimulus configuration, four patterns were drawn with replacement and an average spike pattern was formed. That procedure was repeated 20 times for each stimulus configuration. The vector strength was computed for each average of four spike patterns that contained four or more spikes; vector strengths computed from fewer than four spikes would have yielded erroneously high vector strengths that had no actual relationship to spike timing. Distributions were formed of the vector strengths computed on trials in which stimuli were or were not modulated. Those distributions were used to generate receiver-operating-characteristic (ROC) curves by systematically varying a vector-strength threshold.
criterion from 1 to 0 and plotting the fraction of the “modulated” and “nonmodulated” distributions that were greater than that criterion (Fig. 1A). The area under the ROC curve gave the fraction of the trials in which modulation was correctly discriminated from no modulation. That fraction was converted to a standard deviate (z score) and multiplied by \sqrt{2} to give the detection index (\(d'\)).

Neurometric functions were formed by plotting \(d'\) as a function of modulation depth (Fig. 1B); modulation depths of −40 to −5 dB, in steps of 5 dB, were tested routinely. In some cases the neurometric function increased monotonically with increasing modulation depth. In other cases, the neurometric function declined at the greatest modulation depths, as expected from the nonmonotonic dependence of vector strength on modulation depth that is seen in some recordings of cortical unit activity (Middlebrooks 2008). Regardless of whether a neurometric function was monotonic or nonmonotonic, the modulation threshold was taken as the lowest point (i.e., shallowest modulation depth) on the interpolated neurometric function at which \(d'\) was >1. In many cases, detection indices were <1 for all tested modulation depths. Those cases were indicated as NPL (for “not phase locked”) and their modulation thresholds were scored as 0 dB for the purpose of nonparametric statistics.

A parallel analysis was conducted in which detection of modulation was based on differences in spike count elicited on trials in which the stimulus was or was not modulated. The analysis was largely identical to that based on vector strength, except that trial-by-trial spike count was substituted for vector strength. The only other difference was that the spike-count analysis used data from all trials, not just the ones in which there were 4 or more spikes as in the vector-strength case. Other details were the same between both analyses, including exclusion of spikes from the first 100 ms after stimulus onset, use of bootstrapped averages, computation of \(d'\) from ROC curves, and so forth. Outcomes from the spike-count analysis are presented in the final section of RESULTS.

Base current levels (i.e., the mean current in modulated waveforms) were routinely set to 2, 4, and 6 dB above the minimum level of an unmodulated pulse train that elicited an onset response from units at the most sensitive recording site at each 16-site probe position. That minimum level varied with carrier pulse rate (Middlebrooks 2004), so stimulus levels were adjusted individually for each carrier rate. For each recording site and each tested stimulus condition, data are reported for the current level that produced the lowest modulation depth threshold, except when stated otherwise.

The presence of substantial numbers of nonphase-locked units precluded the use of parametric statistics for comparison of modulation thresholds among stimulus conditions. For that reason, nonparametric statistical procedures were used. The paired two-sided Wilcoxon sign-rank test was used to compare the effects of two stimulus conditions on units at a fixed set of recording sites. The Kruskal–Wallis test (a nonparametric form of one-way ANOVA) was used to test the effects of three or more stimulus values on a fixed set of sites.

**RESULTS**

**General characteristics of modulation sensitivity**

Cortical modulation detection thresholds (MDTs) varied widely among recording sites and stimulus conditions. Figure 2 shows distributions of MDTs as a function of stimulus current level (represented by each box and whiskers) and modulation frequency (each panel). The top row of panels represents modulation thresholds obtained with carrier pulse rates of 254 pps. For that carrier rate and each level and modulation frequency, MDTs ranged from not measurable (i.e., NPL: no phase locking at any tested modulation depth) down to −40 dB, which was the shallowest modulation depth tested. An MDT of −40 dB is comparable to the lowest thresholds measured in human psychophysical tests (Shannon 1992). The cortical excitation threshold was the lowest base current level that activated unit activity at the most sensitive site on the 16-site recording probe. At the 254-pps carrier rates, MDTs tended to be lowest when current levels were within a few decibels of the cortical excitation threshold. At each of the modulation frequencies shown, the distributions of MDTs increased significantly with increases in current level \(P < 0.0001\) at each modulation frequency). At the 4,069-pps carrier rate, the effect of current level on MDT was inconsistent. At that higher carrier rate, units at fewer than half of the recording sites showed significant phase locking at any current level. A slight but significant decrease in MDT with increasing level was observed for the 20- to 21-Hz modulation frequency \(P < 0.0005\), whereas a slight increase in MDT was observed for the 60- to 64-Hz frequency \(P < 0.05\). This figure shows that carrier rates of 254 pps tended to produce lower MDTs than did...
carrier rates of 4,069 pps. That issue is addressed in more detail in a later section.

Differences in the effects of stimulus level on MDTs tended to confound comparisons of other stimulus parameters. For that reason, we measured MDTs for each stimulus condition at levels 2, 4, and 6 dB above the cortical excitation threshold. Subsequent figures and analysis show for each stimulus condition the lowest MDT that was recorded across that range of levels.

The maximum modulation frequencies to which cortical neurons phase lock varies with the location of neurons relative to cortical lamina, with highest-frequency phase locking observed in thalamic afferent layers (Middlebrooks 2008). The dependence of phase locking on cortical laminar position likely is a major factor contributing to the observed wide variation in MDTs. Figure 3 shows the distributions of MDTs as a function of cortical depth relative to the thalamic afferent layers estimated from current-source-density analysis (Middlebrooks 2008). The shading indicates the interquartile ranges. At each modulation frequency, MDTs varied significantly with cortical depth (P < 0.01 to P < 0.0001), with the lowest MDTs recorded near the thalamic afferent layers. Median values of MDTs ranged between −20 and −30 dB in middle cortical layers for all modulation frequencies, whereas medians were above approximately −10 dB for more superficial or deeper layers. At any cortical depth, there were at least a few neurons that showed no significant phase locking.

![Figure 2](http://jn.physiology.org/)

![Figure 3](http://jn.physiology.org/)
Influence of carrier pulse rate on modulation sensitivity

Cortical phase locking was substantially more sensitive to sinusoidal modulation of 254-pps pulse trains than to sinusoidal modulation of 4,096-pps pulse trains. Figure 4, A, B, and C compares MDTs for those carrier rates at three sine-wave modulation frequencies ($P < 0.0001$ at each frequency); the bottom panels (Fig. 4, D, E, and F) show data for a square-wave modulation condition that are subsequently considered. For each of the three sine-wave modulation frequencies, data points in these scatterplots tend to lie above the positive diagonal, indicating that MDTs were higher (i.e., neurons were less sensitive) for the 4,069-pps rate than for the 254-pps rate ($P < 0.0001$ for each modulator frequency). The difference in threshold (indicated in the figures as $\Delta \text{thr}$) was reported as the median value of the difference in MDT between carrier-rate conditions for each recording site. Threshold differences computed in that way were 10.1, 8.0, and 6.3 dB for the 20- to 21-, 40- to 42-, and 60- to 64-Hz modulators, respectively.

Several factors might have contributed to the greater modulation sensitivity for the lower carrier rate. One such factor was the difference in the magnitudes of current steps between successive pulses in the modulated pulse trains. The stimulus characteristics underlying this hypothetical influence of step size factor are illustrated in the top panels of Fig. 5, which show 254-pps (Fig. 5A) and 4,096-pps (Fig. 5B) pulse trains modulated with 42-Hz sine waves at modulation depths of −10 dB. For the lower pulse rate, the sine-wave modulator was sampled at 3.94-ms intervals, and there could be substantial steps in current between successive pulses. At a modulation depth of −10 dB and modulation frequency of 42 Hz, for example, subsequent pulses differed in current by as much as 2.7 dB, which is as large as the entire dynamic range of many cortical neurons under conditions of cochlear-implant stimulation. For the higher pulse rate, in contrast, the modulator was sampled at 0.244-ms intervals and pulse-to-pulse changes in current were small, no more than 0.14 dB at the −10-dB modulation depth. The greater pulse-to-pulse step size for the lower pulse rate hypothetically might elicit stronger phase locking of the auditory pathway to the modulator waveform. The rationale for that hypothesis is elaborated in the DISCUSSION.

The contribution of step size to enhanced modulation sensitivity was estimated by testing cortical sensitivity to square-wave modulators, like those illustrated in Fig. 5, C and D. The square-wave modulator eliminated the dependence of step size on carrier rate; that is, current steps between subsequent pulses were either zero or some nonzero value that was independent of pulse rate. Given a −10-dB modulation depth, for example, the nonzero step size was 5.7 dB regardless of the carrier pulse rate. Figure 6 compares MDTs for square-versus-sine-wave modulators for carrier rates of 254 pps and 4,096 pps. For the high carrier rate (Fig. 6, B, E, and F), MDTs were significantly lower for the square-wave modulator ($P < 0.0001$ to 0.05). The median threshold differences were substantial for the 42- and 42-Hz modulation frequencies: −14.3 and −8.2 dB, re-

![Fig. 4](http://jn.physiology.org/)

**FIG. 4.** Comparison of MDT between 254- and 4,069-pps carrier pulse rates for sine-wave (top row of panels) and square-wave (bottom row) modulation waveforms. Within each panel, the area of each circle is proportional to the number of units at a particular combination of MDT for the 2 carrier rates. $n = 256$ single- and multiunit recordings in 20 guinea pigs for the sine-wave modulator, and $n = 146$ recordings in 10 guinea pigs for the square-wave modulator.
spectively; those values were larger than the 3-dB difference in the energy of square-versus-sinusoidal modulator waveforms. The median difference was zero for the 64-Hz modulator because more than half of the recording sites showed no phase locking for the high carrier rate regardless of the modulator wave shape. For the 254-pps carrier rate, the square-wave modulator significantly enhanced modulation sensitivity for the 21-Hz modulator, producing a 7.9-dB reduction in median MDT (Fig. 6A). No significant difference in MDTs was observed between sine- and square-wave modulators for the 42-

FIG. 5. Electrical stimulus waveforms. Modulated waveforms are shown for modulation frequencies of 41 Hz and depths of −10 dB. Left and right columns of panels show 254- and 4,069-pps carrier rates, respectively. Top and bottom rows of panels show sinusoidal and square-wave modulation waveforms, respectively.

FIG. 6. Comparison of MDT between sine- and square-wave modulation waveforms. Top and bottom rows of panels show 254- and 4,069-pps carrier rates. Other characteristics of the scatterplots are as in Fig. 4. $n = 136$ single- and multiunit recordings in 9 guinea pigs for the 254-pps carrier and $n = 146$ recordings in 10 guinea pigs for the 4,069-pps carrier.
and 64-Hz modulators at 254-pps carrier rates (Fig. 6, B and C). A likely explanation for that lack of significant effect is that at the higher modulation frequencies there is relatively little difference between sinusoidal and square waveforms sampled at the 254-pps pulse rate.

Use of square-wave modulators substantially reduced the effect of carrier rate on modulation sensitivity. As shown in Fig. 4, D, E, and F, MDTs still were significantly lower for the 254- than for the 4,096-pps carrier in the square-wave modulator condition \( P < 0.0001 \) to 0.01. Nevertheless, the magnitudes of the between-rate differences in MDTs were significantly less than those measured for the sine-wave modulator \( P < 0.0001, 0.0001, \) and 0.05 for 21-, 42-, and 64-Hz modulators, respectively; Wilcoxon rank-sum test of the between-rate differences computed for sine- and square-wave modulators). The differences between results obtained with sine- and square-wave modulators suggest that step size is a major factor accounting for the effect of carrier rate on modulation sensitivity. The presence of a significant difference in MDTs between carrier-rate conditions even for the square-wave modulator, however, indicates that factors additional to step size are involved.

Another factor that might contribute to the carrier-rate dependence of cortical modulation sensitivity is the degree to which the auditory nerve and auditory brain stem neurons entrain or phase lock to the pulse train carrier. For instance, auditory nerve fibers entrain (i.e., fire with every pulse) to cochlear electrical pulse trains around 254 pps (Javel et al. 1987) and many brain stem neurons likely phase lock (i.e., fire within a restricted range of stimulus phase) to that rate (Joris et al. 2004). In contrast, many auditory nerve fibers fire with non-Poisson statistics to about 4,000-pps unmodulated cochlear electrical pulse trains (Litvak et al. 2001) and, essentially, no brain stem neurons show phase locking to 4,000-pps rates. The range of carrier pulse rates between 254 and 4,069 pps was explored in 1-octave steps in an effort to gain insight into how auditory nerve and brain stem phase locking to the carrier pulse train might influence cortical phase locking to the modulator.

The distribution of MDTs measured for sine-wave modulation at various carrier rates is shown in Fig. 7, A, B, and C. For each of the three tested modulator frequencies, MDT varied significantly with increasing carrier rate; results of Kruskal–Wallis nonparametric ANOVA are shown in the bottom left of each panel. Pairwise tests between successively higher rates (indicated by asterisks and “ns” between adjacent pairs of boxes) indicate that the trend was toward consistently increasing MDT; that is, as carrier rates were increased all significant changes in MDT were positive. The trend was weaker for the highest modulation frequency, largely because around half of the neurons showed no phase locking at any modulation depth at any tested carrier rate. Comparison of carrier rates using sine-wave modulators was confounded by the step-size factor discussed earlier. For that reason, square-wave modulators were tested with the same range of carrier rates (Fig. 7, D, E, and F). Again, the ANOVA showed that MDT varied significantly with carrier rate, although in the square-wave case, MDTs were lower for the 508-pps carrier than for the 254-pps carrier. The MDTs increased monotonically with increases in carrier rate >508 pps.

Influence of electrode configuration on modulation sensitivity

The results presented earlier were all obtained with a monopolar stimulus condition consisting of the next-to-most-apical intracochlear electrode as the active electrode and a wire in a neck muscle as a distant ground. A narrow bipolar configuration was also tested in which the active electrode was the next-to-most-apical electrode and the return electrode was the adjacent, most apical, electrode; active and return electrodes were spaced 0.75 mm center to center. Consistent with our previous cortical studies (Bierer and Middlebrooks 2002), thresholds for cortical excitation were higher for the bipolar than for the monopolar configuration: for 40 \( \mu s / p h \) phase durations, thresholds for excitation of the most sensitive cortical units averaged 55.1 \( \pm \) 3.2 dB (re 1 \( \mu A \); mean \( \pm \) SD) for BP and 49.1 \( \pm \) 2.3 dB for MP \( n = 9 \) animals). Modulation sensitivity for each electrode configuration was measured at current levels 2, 4, and 6 dB above the excitation threshold for the corresponding configuration.

The MDTs for bipolar and monopolar configurations are compared in Fig. 8 for three sinusoidal modulation frequencies. Results for the 254-pps carrier are shown in the top row of panels (Fig. 8, A, B, and C). The MDTs for that carrier rate were significantly higher for the bipolar configuration at all modulation frequencies \( P < 0.0001 \), with threshold differences ranging from 5.5 to 8.8 dB. For the bipolar configuration, there was a considerable increase in the proportion of units that showed no phase locking at any modulation depth. A similar trend was observed for the 4,069-pps carrier (Fig. 8, D, E, and F). The MDTs were significantly higher for bipolar stimulation at all modulation frequencies \( P < 0.0001 \) to 0.005.

The elevated MDTs for the bipolar configuration might reflect, at least in part, an overall imprecision in transmission of temporal information from bipolar stimulation up the auditory pathway to the cortex. An independent measure of temporal precision was provided by the trial-by-trial SD in first-spike latency. That value averaged 1.40 \( \pm \) 0.78 ms for the monopolar configuration (mean \( \pm \) SD) and 1.87 \( \pm \) 1.48 ms for bipolar. The difference was significant \( t \)-test; \( P = 0.005, n = 127 \) recordings in 11 animals).

Modulation detection based on spike count

The companion paper (Middlebrooks 2008) argues that the magnitude of neural responses (i.e., spike count) could not accurately transmit information about modulation frequency because spike-count coding of modulation frequency would be confounded by sensitivity to modulation depth. For that reason, the present study of modulation sensitivity focused on the vector strength of phase locking as a measure of the presence or absence of modulation. Nevertheless, a change in response magnitude might be an adequate indicator of the presence of modulation in a task that did not require identification of modulation frequency. Modulation sensitivity based on spike count was compared with that based on vector strength, as described in METHODS. The MDTs computed by the two methods were remarkably similar. Across all stimulus conditions, the median difference between MDTs showed that modulation sensitivity based on spike counts was only 1.2 dB more sensitive than that based on vector strength \( P < 0.0001 \). With few exceptions, every conclusion regarding parametric sensi-
Activity presented in the present report based on vector strength was supported by the analysis of spike count. Statistical results from the two approaches are shown in Table 1. Regarding the increase in MDT with increasing stimulus level, the vector-strength measure for the 4,069-pps carrier gave inconsistent results across modulation frequencies, with a decrease in MDT with increasing level in the 20-Hz case. In contrast, the spike-count measure showed consistent significant increases in MDT. In addition, the spike-count measure showed a slightly weaker dependence of MDT on carrier rate than did the vector-strength measure, as indicated by a few instances in which the vector-strength measure gave a significant result that was not significant for the spike-count measure. Overall, modulation detection based on either the vector strength of phase locking or on the magnitude of responses showed similar sensitivity to base current level, modulation frequency and wave-shape, carrier pulse rate, and electrode configuration.

**Discussion**

**Relation to previous studies**

The companion paper (Middlebrooks 2008) reviewed studies of cortical and midbrain phase locking to modulated or unmodulated acoustic stimuli. Of those, only two studies of the inferior colliculus (ICC) using tonal stimuli applied statistical tests to estimate sensitivity to modulation. In the study by Krishna and Semple, significant phase locking to the best modulation frequencies was observed at 10% modulation depth (i.e., -20 dB) in 3 of 31 tested neurons, but most neurons...
required modulation depths of \( \geq 20\% \) (approximately \(-14\) dB). Nelson and Carney (2007) found that 28% of units in the ICC of unanesthetized rabbits showed significant phase locking to modulation depths less than or equal to \(-20\) dB. Those authors found that modulation sensitivity based on phase locking was substantially more sensitive than a measure based on spike counts. Moreover, phase locking, but not spike rate, was comparable in modulation sensitivity to values obtained in human psychophysical studies.

It is difficult (possibly invalid) to compare modulation depths between electrical and acoustical stimulation, but as a point of reference, 44.1% of cortical units in the present study showed MDTs of \(-20\) dB or lower under conditions of monopolar electrical stimulation with 20- to 21-Hz modulation of 254-pps carrier rates. In contrast to the ICC results by Nelson and Carney (2007), the present cortical results showed very similar modulation sensitivity based on either phase locking or spike count.

Animal physiological studies of central auditory phase locking to electrical cochlear stimuli are limited to two studies in the cortex (Schreiner and Raggio 1996; Wang et al. 1999) and two studies in the ICC (Snyder et al. 2000; Vollmer 2005), all reviewed in the companion paper. None of those studies quantified modulation depth sensitivity in detail.

The human psychophysical literature provides several demonstrations of modulation sensitivity under conditions of electrical cochlear stimulation. There are numerous caveats to consider in comparing cortical unit data in anesthetized guinea pigs with psychophysical data in humans. Among these concerns are the differences in species, differences in anesthetic state, differences in the physical fit of the stimulating electrode arrays in the cochleas, and any effects of chronic deafness and electrical stimulation in humans compared with the acute deafening and stimulation in the present experiments. Also, in most of the human studies (Busby et al. 1993; Fu 2002; Pfingst et al. 2007; Shannon 1992) the charge per phase of electrical pulses was modulated by varying pulse width rather than pulse current as in the present study. Two experimental differences deserve particular attention. The first is that the distributions of MDTs reported in the present study represent every unit that was encountered, including the most sensitive units as well as others that might make no contribution to modulation perception. In contrast, one might speculate that human psychophysical subjects make their perceptual decisions based on the activity of the most sensitive neurons. That speculation is supported by recordings in behaving monkeys showing that behavioral thresholds for sound detection correspond to the thresholds of the most sensitive single cortical neurons (Pfingst et al. 1976) and from studies in cat visual cortex showing that the most sensitive neural thresholds for changes in orientation correspond with behavioral thresholds (Bradley et al. 1987). Second, in human psychophysical studies stimulus levels are

**FIG. 8.** Comparison of MDT between monopolar and bipolar electrode configurations. Characteristics of the scatterplots are as in Fig. 4.  

- **A.** 20–21 Hz modulation 254 pps
- **B.** 40–42 Hz modulation 254 pps
- **C.** 60–63 Hz modulation 254 pps
- **D.** 20–21 Hz modulation 4069 pps
- **E.** 40–42 Hz modulation 4069 pps
- **F.** 60–63 Hz modulation 4069 pps

- Monopolar
- Bipolar

For each condition, the scatterplot shows the distribution of modulation depths and the statistical significance of the difference between monopolar and bipolar configurations.

**Table:**

<table>
<thead>
<tr>
<th>Frequency Range</th>
<th>Modulation Depth</th>
<th>Stimulus Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>20–21 Hz</td>
<td>(-8.8) dB</td>
<td>254 pps</td>
</tr>
<tr>
<td>40–42 Hz</td>
<td>(-7.1) dB</td>
<td>254 pps</td>
</tr>
<tr>
<td>60–63 Hz</td>
<td>(-5.5) dB</td>
<td>254 pps</td>
</tr>
<tr>
<td>20–21 Hz</td>
<td>(-2.3) dB</td>
<td>4069 pps</td>
</tr>
<tr>
<td>40–42 Hz</td>
<td>(0.0) dB</td>
<td>4069 pps</td>
</tr>
<tr>
<td>60–63 Hz</td>
<td>(0.0) dB</td>
<td>4069 pps</td>
</tr>
</tbody>
</table>

**Notes:**
- **A:** p<.0001 \( \Delta \) thr= 8.8 dB
- **B:** p<.0001 \( \Delta \) thr= 7.1 dB
- **C:** p<.0001 \( \Delta \) thr= 5.5 dB
- **D:** p<.0001 \( \Delta \) thr= 2.3 dB
- **E:** p<.005 \( \Delta \) thr= 0.0 dB
- **F:** p<.005 \( \Delta \) thr= 0.0 dB

- **NPL:** Normalized Pure Tone Level
- **pps:** Pulsed Per Second

**Legend:**
- Circles represent unit recordings
- Crosses represent multiunit recordings
- Triangles represent significant differences
- Diamonds represent trends that are not significant.
set relative to the fraction of the dynamic range from threshold to maximum comfortable loudness. In contrast, “maximum comfortable loudness” has no meaning in a study using anesthetized animals. Current levels in the present study were tested at three levels relative to cortical excitation thresholds and the modulation threshold for each condition was scored as the lowest obtained across the three levels.

Keeping these caveats in mind, there are several psychophysical studies with human cochlear-implant users that bear comparison with the present results. Shannon (1992) measured modulation sensitivity with 1,000-pps pulsatile or 1,000-Hz sinusoidal carriers. Modulation transfer functions (i.e., modulation threshold vs. modulation frequency) showed a low-pass characteristic with a cutoff around 140 Hz. In that study, MDTs obtained with the 1,000-pps pulsatile carrier ranged from about −40 to −25 dB across the range of modulator frequencies tested in the present study. Two similar studies (Busby et al. 1993; Cazals et al. 1994) reported broader ranges of performance across subjects, from about −10 to −40 dB; the study by Busby and colleagues used bipolar or common-ground electrode configurations, whereas the other studies used monopolar. In the study by Busby and colleagues, variations in carrier rate from 100 to 1,000 pps had no consistent effect on MDT. The most sensitive subjects showed MDTs of around −40 dB in the studies by Donaldson and Viemeister (2000), Fu (2002), and Pfingst and colleagues (2007). In the present study, the median values of MDTs of cortical units were around −15 dB for 20- to 64-Hz modulation of 254-pps carrier. The most sensitive 25% of units, however, had MDTs ranging from −40 to −25 dB, which was comparable to the ranges of MDTs obtained in previous human psychophysical studies.

Most human psychophysical studies have found that modulation sensitivity tends to improve with increasing stimulus level. In contrast, the present cortical unit results showed a decrease in sensitivity with increasing level for the 254-pps and no consistent trend for the 4,069-pps carrier. At least two factors might have contributed to the difference between the human psychophysical and animal physiological results. First, increases in stimulus level presumably increase the size of active cortical neuronal populations and thus in the perceptual studies, might enhance detection of modulation on the basis of coordinated activity within the population. In contrast, MDTs computed in the present study were based on the activity of single units or multiunit clusters and could not have exploited any benefits from growth in the size of active cortical populations. Second, the human data were collected at known points within perceptual dynamic ranges, whereas the upper end of the dynamic range was poorly defined in the animal studies.

| TABLE 1. Comparison of modulation detection based on vector strength (VS) and spike count |
|-----------------------------------------------|-------------------------------|-------------------------------|-------------------------------|
| MDT increases with increasing base level (Fig. 2) | 20–21 Hz | 40–42 Hz | 60–64 Hz |
| 254 pps | VS | P < 0.0001 | P < 0.0001 | P < 0.0001 |
| Count | P < 0.0001 | P < 0.0001 | P < 0.0001 |
| 4 kpps | VS | P < 0.005 (decreased) | P = 0.49 | P < 0.05 |
| Count | P < 0.005 | P < 0.0001 | P < 0.0001 |
| MDT varies with cortical depth, 254 pps (Fig. 3) | VS | P < 0.0001 | P < 0.005 | P < 0.01 |
| Count | P < 0.0005 | P < 0.05 | P < 0.0005 |
| MDT higher for 4 kpps than for 254 pps (Fig. 4) | Sine | VS | P < 0.0001, 10.1 dB | P < 0.0001, 8.0 dB | P < 0.0001, 6.3 dB |
| Count | P < 0.0001, 7.0 dB | P < 0.0001, 6.0 dB | P < 0.0001, 6.1 dB |
| Square | VS | P < 0.01, 1.6 dB | P < 0.0005, 0.9 dB | P < 0.0001, 3.0 dB |
| Count | P < 0.005, 1.0 dB | P = 0.32, 0.0 dB | P < 0.0001, 0.7 dB |
| MDT lower for square than for sine modulation (Fig. 6) | 254 pps | VS | P < 0.0001, −7.9 dB | P = 0.87, 2.4 dB | P = 0.34, −0.8 dB |
| Count | P < 0.0001, −6.7 dB | P < 0.05, −0.6 dB | P = 0.31, −1.5 dB |
| 4 kpps | VS | P < 0.0001, −14.3 dB | P < 0.0001, −8.2 dB | P < 0.05, 0.0 dB |
| Count | P < 0.0001, −10.1 dB | P < 0.0001, −3.4 dB | P = 0.89, 0.1 dB |
| MDT increases with carrier rate (Fig. 7) | Sine | VS | P < 0.0001 | P < 0.0001 | P < 0.05 |
| Count | P < 0.0005 | P = 0.13 | P < 0.05 |
| Square | VS | P < 0.0001 | P < 0.0001 | P < 0.0005; MDT lowest at 508 pps |
| Count | P < 0.005 | P = 0.14 | P < 0.01; MDT lowest at 508 pps |
| MDT higher for BP than for MP (Fig. 8) | 254 pps | VS | P < 0.0001, 8.8 dB | P < 0.0001, 7.1 dB | P < 0.0001, 5.5 dB |
| Count | P < 0.0001, 7.4 dB | P < 0.0001, 6.5 dB | P < 0.0001, 4.1 dB |
| 4 kpps | VS | P < 0.0001, 2.3 dB | P < 0.005, 0.0 dB | P < 0.005, 0.0 dB |
| Count | P = 0.12, 0.0 dB | P < 0.0005, 2.9 dB | P < 0.05, 0.7 dB |
might be that, in the present study, levels 6 dB or even 4 dB above cortical excitation thresholds drove neurons at multiple levels of the auditory pathway at saturated firing levels that were greater than would be tolerated by a human subject. One might expect saturated neural firing at brain stem levels to impair cortical sensitivity to stimulus modulation.

Two recent human psychophysical studies have tested the hypothesis that high carrier rates impair sensitivity to modulated electrical pulses. Galvin and Fu (2005) compared rates of 250 and 2,000 pps, using a modulation frequency of 20 Hz. In that study, the 250-pps rate produced MDTs that were 5 to 12 dB more sensitive than those at the 2,000 pps rate; the particular differences varied among stimulus levels. That compares with carrier-rate–dependent differences of 6 to 10 dB in the present physiological study. Pfingst and colleagues (2007) compared 250- and 4,000-pps pulse rates, using a modulation frequency of 40 Hz. Across 180 cases of subject, electrode site, and stimulus level, 70.8% of the cases showed lower MDTs at the lower carrier rate, averaging 5.5 dB lower across all the cases. The differences in MDTs associated with carrier rate were greatest for the most apical electrode site, averaging about 10 dB for all but the highest stimulus level.

The present cortical unit results showed a substantial effect of electrode configuration on MDT: for the 254-pps carrier, MDTs were 10.8, 7.9, or 5.6 dB lower for monopolar than for bipolar, depending on the modulation frequency. There is a precedent in the acoustical-stimulation literature for an effect of tonotopic spread of excitation on cortical phase locking. Eggermont (1994) demonstrated that best modulation frequencies for cortical phase locking to sinusoidal modulation waveforms were higher for broadband noise carriers than for tonal carriers, implying that modulation sensitivity was greater for broadband carriers. That is somewhat analogous to the present observation inasmuch as monopolar stimulation, like noise, produces more tonotopically widespread cochlear excitation than does bipolar stimulation or a tone.

Galvin and Fu (2005) tested the influence of electrode configuration (referred to as “mode” in that report) on modulation detection by human psychophysical subjects. Contrary to our cortical data, those investigators found no significant effect of electrode configuration on MDT. One likely explanation for the failure to find an effect of configuration is that the narrowest bipolar configuration that was tested by Galvin and Fu was a “BP +3” configuration in which active and return electrodes were 3.0 mm apart, center to center. In contrast, the bipolar configuration used in the present study used an active-to-return spacing of 0.75 mm. A direct comparison of stimulus configurations is difficult given the differing dimensions of human and guinea pig cochleas. Nevertheless, the broader bipolar configuration used in the Galvin and Fu study might not have achieved the focal electrical field that apparently impaired modulation sensitivity in the cortical unit study. In an ongoing study, Burkhode-Juhasz and colleagues (2007) demonstrated significantly greater modulation sensitivity for monopolar than for bipolar configurations. That group is using a 0.75-mm active-to-return bipolar spacing equal to that used in the present study.

We note that, in addition to modulation sensitivity, identification and discrimination of modulation frequency and wave shape presumably are of importance for speech discrimination and, likely, are influenced by carrier pulse rate and electrode configuration. Modulation-frequency discrimination has been addressed in a small number of psychophysical studies (e.g., Chatterjee and Chen 2008; Geurts and Wouters 2001; McKay and Carlyon 1999; McKay et al. 1994). The companion paper touched briefly on the issue of modulation-frequency coding (Middlebrooks 2008). Clearly, there is a need for further parametric physiological studies of modulation-frequency coding.

Mechanisms of carrier-rate and electrode-configuration effects on modulation sensitivity

From a viewpoint in the auditory cortex, one can attempt to infer, in most cases indirectly, the subcortical and cortical mechanisms underlying sensitivity to various electrical stimulation parameters. The influence of carrier rate on modulation sensitivity almost certainly involves multiple factors. One such factor that was explored in the present study was the effect of the magnitude of the current step between subsequent electrical pulses in a modulated waveform. The motivation for that analysis is as follows. In response to an unmodulated pulse train, or a train modulated at a subthreshold depth, a cortical or subcortical neuron likely is in a state of adaptation, in which the probability of a response to the nth pulse is masked by the preceding $n - 1$ pulses. An abrupt increase in the current on the nth pulse could counteract this effect and result in a time-locked spike. At a typical modulation threshold depth—for example, $-10$ dB for sinusoidal modulation at 40 Hz—current steps between the pulses of a 254-pps carrier could be as large as 2.7 dB. The largest steps would occur at a particular phase of the modulation waveform, leading to cortical phase locking to the modulator. In contrast, the corresponding maximum current step would be only 0.14 dB for a 4,069-pps carrier. It is likely that the gradual pulse-to-pulse changes at a high carrier rate would not result in responses that are time locked to specific pulses and, instead, could result in phase locking only to the across-pulse integrated activation. In the present study, the relationship between carrier pulse rate and maximum current step was dissociated by stimulating with square modulator waveforms. Consistent with the step-size hypothesis, use of the square-wave modulator lowered the MDTs for both 254- and 4,069-pps carriers and substantially reduced the difference in MDTs between those carrier rates. That observation is interpreted as an indication that differences in between-pulse step size contribute to the decrease in modulation sensitivity at high pulse rates. It has been argued that increased pulse rates in cochlear-implant speech processors would be beneficial because they would increase the temporal resolution of sampling of the speech waveform. The present results, however, suggest that high-resolution sampling of relatively low-frequency modulation waveforms might be detrimental to modulation sensitivity because the small increments of current (or charge) between subsequent high-rate pulses would not provide precise synchronization to particular carrier pulses and thus might reduce synchrony to the modulating waveform.

At least three properties of the cochlea and auditory nerve might contribute to the impairment of cortical modulation sensitivity at high carrier pulse rates. The first of these factors is temporal integration of spike initiation within the cochlea. Two or more electrical pulses presented to the cochlea in quick succession result in temporal integration of depolarization of...
auditory nerve cell bodies, with time constants around 260 to 352 μs (Cartee et al. 2006). Such integration would result in smearing of the pulses of a 4,069-pps pulse train having a 245.8-μs interpulse interval; the same would be true to a lesser degree with 2,034-pps pulse trains, with interpulse periods of 491.6 μs. Of course, 245.8 and 491.6 μs are shorter than the absolute refractory period of auditory neurons (~700 μs; Cartee et al. 2006), so one would not expect auditory nerve responses to each of two successive pulses. Even if a fiber had time to recover from refraction, as it might during the low-amplitude phase of a square-wave modulator, an interpulse interval shorter than the integration time constant would prevent auditory-nerve-fiber synchrony to any particular stimulus pulse. It remains to be tested whether failure of the auditory nerve to respond to the timing of individual pulses within a train would result in impaired sensitivity to modulation of that pulse train.

A second factor related to auditory nerve properties is auditory-nerve phase locking to the carrier pulses. Entrainment of auditory nerve spikes to each cycle of an electric sinusoid, pulse train, or square wave has been observed at rates as high as 800 pps (Javel et al. 1987; Kiang and Moxon 1972; Parkins 1989; van den Honert and Stypulkowski 1987). At higher rates of electrical stimulation, auditory nerve fibers fail to respond to every cycle, but can continue to fire within a restricted fraction of the stimulus repetition period (i.e., to phase lock) at rates of ≥2 kpps (Dynes and Delgutte 1992; Litvak et al. 2001; van den Honert and Stypulkowski 1987). The range of carrier pulse rates between 254 and 4,069 pps was explored in the present study in 1-octave steps. Use of a square-wave modulator to eliminate the confound between pulse rate and step size (Fig. 7, D, E, and F) revealed minimum MDTs at a carrier rate of 508 pps. One might speculate that at carrier pulse rates of ≥1,016 pps, auditory nerve fibers become less tightly synchronized to the carrier pulses and that decrease of synchrony in some way degrades phase locking to the modulating waveform. The relationship between phase locking to the carrier and phase locking to the modulator is as yet unknown. One possible connection, however, is that a fiber that is phase locked to the carrier pulses might have an increased probability of firing to the largest pulse within a modulator period, which would enhance phase locking to the modulator. At a higher carrier rate, the fiber could not respond to particular pulses.

A final factor related to auditory-nerve responses is the tendency of nerve fibers to adapt at high stimulus rates. Javel and colleagues (1987) showed essentially no adaptation of cat auditory nerve fibers for 100-ms bursts of 800-pps pulses but reported that fibers would “cease to respond” after about 2 to 5 s to pulse rates >600 pps; that cessation of response was attributed to depolarization block. Similarly, Litvak and colleagues (2003) demonstrated prominent auditory-nerve adaptation to 5,000-pps pulse trains after about 100 s. Litvak and colleagues (2001) showed that spike rates dropped by about 47% after 145 ms of stimulation at 4.8 pps compared with 50% at a rate of 1.2 pps. Zhang and colleagues (2007) evaluated spike rates at onset (0–12 ms poststimulus onset) and at asymptotic rate (200–300 ms poststimulus onset). At a stimulus pulse rate of 5,000 pps, about 33% of fibers adapted their spike rates by >90%, whereas only about 1% of fibers (1 of 80 fibers) showed such strong adaptation to 250-pps pulse rate; the adaptation to 1,000-pps stimuli was similar to that at 5,000 pps. Miller and colleagues (2008) also observed that mean interspike intervals and Fano factor (a measure of trial-by-trial variance in spike counts) both increased dramatically over the first 100 ms of a 5,000-pps pulse train, whereas relatively little change in those measures was observed for 250-pps trains. A study of electrically evoked compound action potential (ECAP) in human subjects also suggests potent auditory-nerve adaptation at pulse rates ≥1,000 pps (Wilson et al. 1997). In that study, ECAP magnitudes showed a striking pulse-by-pulse alternation across the range of pulse rates from about 400 to 1,000 pps. The period of that alternation was short relative to the modulation periods on the present study, but the overall adaptation, if present in the present animal preparation, might have influenced phase locking to the modulator. More complex patterns in the human ECAPs were observed up to rates of about 1,500 pps, with deep adaptation evident at higher rates. In the present study, the decline in modulation sensitivity at pulse rates might have reflected an inability of highly adapted auditory nerve fibers to transmit envelope waveforms.

Phase locking of neurons in the central auditory pathway to carrier pulses might also contribute to the carrier-rate sensitivity of modulation detection. In response to acoustic tones, brain stem neurons phase lock at frequencies corresponding to the lower half of the range of carrier rates tested in the present study. For instance, neurons in the guinea pig cochlear nucleus phase lock to tone frequencies up to about 1.5 kHz (Winter and Palmer 1990), and neurons in the guinea pig ICC phase lock to tones at frequencies up to about 1 kHz (Liu et al. 2006). The brain stem neurons showing phase locking up to about 1 kHz, however, generally have characteristic frequencies ≤1 kHz and thus contribute little or no input to the generally high-characteristic-frequency neurons recorded in the present study and in other studies using cochlear-implant stimulation. In responses to unmodulated cochlear-implant pulse trains, neurons in the cat ICC phase lock to cochlear-implant pulse trains at rates up to ~100 pps (Snyder et al. 1995, 2000); our ongoing studies (Middlebrooks and Snyder, unpublished) are demonstrating significant phase locking in the ICC at electrical pulse rates ≤300 pps. The maximum modulation sensitivity observed at a 508-pps carrier rate in the present study might reflect robust phase locking to the carrier at the level of the cochlear nucleus or pons. As is the case for phase locking in the auditory nerve, however, it is not clear whether or how phase locking of neurons to carrier pulses would influence phase locking to the modulating waveform.

The factors hypothesized to account for the increase in cortical MDTs obtained with high carrier pulse rates may be summarized as follows. The effect of the decreasing current step associated with decreasing interpulse interval is evident across the entire range of pulse rates that were tested; that effect could be eliminated experimentally by the use of square-wave modulation waveforms. At pulse rates of 254 and 508 pps, at which modulation sensitivity was maximal, auditory-nerve fibers are entrained or at least phase locked to the carrier pulse train and many neurons in the lower brain stem are phase locked to the carrier. As the carrier rate is increased to 1,017 pps, the increase in MDT could reflect the loss of brain stem phase locking to the carrier and decrease in auditory nerve phase locking. At pulse rates ≥2,034 pps, adaptation of the auditory nerve to prolonged stimulation might reduce the strength of representation of the modulation waveform. At
the highest tested pulse rate, temporal integration would have smeared any contribution of auditory nerve responses to the timing of individual pulses.

The monopolar electrode configuration tended to enhance modulation sensitivity compared with that obtained with bipolar stimulation. A likely explanation for that effect is that monopolar stimulation elicited a highly synchronous pattern of input across a broad range of the tonotopic range, whereas bipolar stimulation produced rather steep latency gradients. Previous studies have shown that single pulses presented in a monopolar configuration can activate neurons over a >2-octave range of the tonotopic axis in the ICC (Snyder et al. 2004, 2008) and auditory cortex (Bierer and Middlebrooks 2002), whereas the spread of activation typically is <1 octave in response to a narrow bipolar stimulus. Moreover, responses to bipolar stimuli tend to show rather steep gradients in first-spike latency along the tonotopic axis of the ICC (Middlebrooks and Snyder, unpublished) and auditory cortex (Bierer and Middlebrooks 2002). The classical frequency bandwidths of area-A1 cortical neurons measured with acoustic tones 20 dB above threshold are <0.5-octave wide at most levels of the tonotopically organized auditory pathway in the high-frequency range of interest in the present study (Calford et al. 1983). Nevertheless, there is substantial subthreshold input converging onto cortical neurons from across several octaves (Metherate et al. 2005). One can infer that a monopolar stimulus would elicit highly synchronous, tonotopically wide-spread patterns of ascending input that would converge simultaneously onto individual cortical neurons. Conversely, one might expect narrow bipolar stimuli to elicit temporally dispersed, tonotopically more restricted patterns of input that would be less effective in transmitting signals to cortical neurons with high temporal fidelity. The loss of fidelity in the bipolar condition was evident in an increased trial-by-trial variation in first-cortical spike latencies in response to bipolar stimulation. One might expect that loss of fidelity to be exaggerated in the tonic, adapted phase of cortical responses during which modulation sensitivity was measured.

Implications for cochlear-implant speech-processor design

In the pulsatile processing schemes used by the majority of cochlear-implant speech processors, temporal information in speech waveforms is transmitted exclusively by modulation of fixed-rate carriers. Users’ success in recognizing speech generally correlates with their modulation sensitivity (Cazals et al. 1994; Colletti and Shannon 2005; Fu 2002). In normal-hearing listeners, “smearing” of temporal envelopes by low-pass filtering the envelopes impairs speech reception, particularly of consonants (Druilman et al. 1994). Inasmuch as the present results demonstrate influences of carrier rate and electrode configuration on modulation sensitivity in the auditory cortex, it seems likely that those parameters would influence speech reception in human cochlear-implant users.

There are, of course, many differences between speech perception in a conscious, long-term deaf human cochlear-implant user and cortical responses in the anesthetized, acutely deafened guinea pigs used in the present study. Those differences preclude detailed quantitative comparisons of stimulus parameters. Nevertheless, one can identify several qualitative factors that might have some impact on speech perception.

Specifically, the present data predict that carrier pulse rates ≥1,000 to 2,000 pps would impair modulation sensitivity and thus might degrade speech reception. Similarly, the data predict that the impairment of modulation sensitivity resulting from use of bipolar stimulation might counteract any other benefits of focal stimulation patterns.

CARRIER PULSE RATES. Several arguments have been put forward in favor of use of relatively high carrier pulse rates. One such argument is that stimulus envelopes would be represented more precisely as a result of a higher sampling rate and that aliasing of envelopes would be avoided by pulse rates several times that of the highest envelope frequency (Wilson 1997). Indeed, physiological studies of responses to sinusoidal modulation have shown that sample rates less than four times that of the modulation frequency produce marked distortions of the modulation waveforms (Snyder et al. 2000). Moreover, reductions in carrier rates <150 pps produce significantly poorer speech reception (Fu and Shannon 2000). Another anticipated benefit of higher pulse rates is that rates ≥2,000 pps might overcome the hypersynchrony of auditory-nerve firing that is observed at lower rates, resulting in more stochastic auditory-nerve firing patterns (Rubinstein et al. 1999; Wilson et al. 1997). Stochastic firing is predicted to result in wider dynamic ranges and more faithful transmission of envelope waveforms. A third possible benefit of high pulse rates is that dynamic ranges between thresholds and maximum comfortable loudness tend to be broader at high rates (Pfingst et al. 2007), and broadened dynamic ranges presumably would reduce distortion resulting from compression of the broad dynamic range of acoustic hearing to the more limited dynamic of electric hearing.

Despite these predicted benefits of higher pulse rates, tests of speech performance using high pulse rates have been somewhat disappointing. Brill and colleagues (1997) evaluated speech reception as a function of carrier pulse rate; in that study, the number of stimulating channels decreased as the pulse rate per channel increased. Speech-reception performance was essentially constant up to a rate of 6,060 pps with four channels. Using a four-channel speech processor, Fu and Shannon (2000) observed an improvement in speech reception with increases in pulse rates from 50 to 150 pps, but saw no further improvement as rates were increased to 500 pps. Vandali and colleagues (2000) observed no significant difference in speech-reception performance at rates between 250 and 807 pps and observed a significant decrease in performance at 1,615 pps, largely because of the performance of one of the five subjects. Holden and colleagues (2002) showed no consistent trend in speech performance at rates between 720 and 1,800 pps: some subjects performed significantly better in some conditions with the higher rate, some did better with the lower rate, and some subjects showed no significant difference. Loizou and colleagues (2000) reported that word and consonant reception showed a small but significant improvement at rates between 800 and 2,100 pps but that performance at 2,100 was not significantly better than that at 1,400. Overall, tests of effects of elevated pulse rates on speech reception show no effect, small decreases in performance, or small increases in performance under limited conditions.

The failure of elevated pulse rate to show the hoped-for improvements in speech reception might reflect to some degree
the loss of modulation sensitivity at higher pulse rates shown in
the present study. Another factor that could have prevented
improvements in speech reception is an increase in between-
channel interference that was demonstrated in a previous
guinea pig auditory cortex study using interleaved pairs of
pulse trains (Middlebrooks 2004). That study showed that, for
a given interpulse interval, there was more interference among
thresholds between two high-rate (4,069-pps) pulse trains than
that between two low-rate (254-pps) trains. Also, the maximum
available interpulse delay was greater for low-rate trains.
Between-channel interference was essentially abolished with
2-ms interpulse intervals, but such long interpulse intervals are
not possible at high rates.

Although speech reception fails to show the predicted im-
provements with increasing pulse rates (Rubinstein et al. 1999;
Wilson 1997), there similarly is a failure to show the marked
decay in speech performance that one might expect from the
present results showing that high pulse rates impair modulation
sensitivity. One possibility is that a decrease in modulation
sensitivity was counterbalanced by an improvement in the
precision of temporal sampling. An alternative explanation is
that speech tests in humans have not yet fully explored very
high sampling rates. With the exception of the study by Brill
and colleagues (1997), there have been no systematic studies of
speech reception using the high pulse rates (i.e., >2,000 pps)
that hypothetically would result in stochastic firing properties
of the auditory nerve (Rubinstein et al. 1999; Wilson 1997)
and, according to the present results, would impair modulation
sensitivity most dramatically.

The present study demonstrates that MDT measured in
decibels tends to increase with increasing carrier pulse rate,
although dynamic ranges of cochlear-implant users also tend to
increase with increasing carrier rate (e.g., Pfingst et al. 2007).
For that reason, speech processors using higher carrier rates
tend to apply less compression of the speech waveform, with
the result that the signal to the cochlear-implant electrodes
shows somewhat deeper modulation than that used with a
lower carrier rate. The failure of high-carrier-rate speech pro-
cessors to show consistent improvements in speech recognition
might be interpreted as evidence that the increase in available
modulation depth is not sufficient to counteract the loss in
modulation sensitivity at high rates. It is clear that further
studies of the effect of carrier rate on speech recognition are
warranted.

ELECTRODE CONFIGURATION. Several lines of evidence suggest
that bipolar electrical configurations would permit more pre-
cise transmission of spectral (i.e., cochlear place) information
than would monopolar configurations. Geometrical models and
measurements in saline tanks and cadavers demonstrate that
bipolar electrode configurations produce more restricted elec-
tric field distributions (Black and Clark 1980; Kral et al. 1998).
In animal studies, physiological recordings from auditory
nerve fibers (Kral et al. 1998; van den Honert and Styp-
pulkowski 1987) and neurons in the ICC (Rebscher et al. 2001;
Snyder et al. 2004, 2008) and auditory cortex (Bierer and
Middlebrooks 2002) show that bipolar stimulation produces
more restricted spread of excitation. Also, interference between
pulses presented simultaneously through two channels is re-
duced in a narrow bipolar configuration compared with mo-
нопolar (Bierer and Middlebrooks 2004). Results from human
psychophysical studies are somewhat less consistent. Busby
and colleagues (1980) showed that bipolar stimulation resulted
in a broader range of pitch ranks than did monopolar, but von
Wallenberg and colleagues (1995) reported essentially identi-
cal ranges of pitch ranks for monopolar and bipolar stimulation.
In forward-masking studies, Boëx and colleagues (2003b)
found more restricted spread of activation with bipolar than
with monopolar stimulation, whereas Kwon and van den Hon-
ert (2006) reported that masking with bipolar stimulation was
not consistently more restricted. Measures of loudness (Shan-
non 1983) or threshold (Boëx et al. 2003a) summation between
pairs of channels stimulated simultaneously showed reduced
between-channel interference with bipolar configurations.

Despite encouraging results from animal physiological stud-
ies, speech studies in humans have demonstrated no benefit from
bipolar stimulation. Indeed, reported speech results with
bipolar stimulation are never better than those with monopolar
and, in some cases, are significantly worse (Kileny et al. 1998;
animal results demonstrating elevated MDTs using bipolar
stimulation might account for the disappointing human speech
results. If human perceptual performance follows the present
animal cortical results, use of narrow bipolar stimulation would
reduce sensitivity to the envelopes of speech waveforms and
would thus impair speech reception.

Concluding remarks

Several well-reasoned arguments predict that signaling of
temporal information by a cochlear implant should be en-
hanced by use of high (e.g., >2,000 pps) carrier rates and that
signaling of spectral information should be enhanced by use of
relatively focal electrode configurations (e.g., bipolar rather
than monopolar). Neither of those predictions has been borne
out in the form of giant improvements in speech reception. The
present results in an animal model demonstrate that high pulse
rates and bipolar electrode configurations both tend to impair
modulation detection at the cortical level. Impairment in mod-
ulation sensitivity, if present in humans, almost certainly would
influence speech reception and might counterbalance any other
benefits of high pulse rates or bipolar stimulation. Although
there have been some human psychophysical confirmations of
effects of pulse rate and electrode configuration on modulation
sensitivity (Burkholder-Juhasz et al. 2007; Galvin and Fu
2005; Pfingst et al. 2007), tests of broader parameter ranges are
needed. Further animal physiological studies are needed to
explore the mechanisms by which adaptation in the auditory
pathway and phase locking to carrier pulses might influence
transmission of modulation waveforms. A dialogue between
animal research and human psychophysical and speech testing
is needed for optimal design of speech processing strategies for
auditory prosthesis.

A C K N O W L E D G M E N T S

I thank C.-C. Lee, K. Otto, and A. Kirby for participation in some of the
experiments. A. Kirby, B. Pfingst, E. Macpherson, and R. Snyder made helpful
comments on versions of the manuscript. J. Wiler provided skillful technical
assistance and Z. Onsan helped with the illustrations.

G R A N T S

This work was supported by National Institute on Deafness and Other
Communication Disorders Grants R01 DC-04312 and P30 DC-05188.
REFERENCES


