Research Article

Modulation Frequency Discrimination with Modulated and Unmodulated Interference in Normal Hearing and in Cochlear-Implant Users

HEATHER A. KREFT,[1] DAVID A. NELSON,[1] AND ANDREW J. OXENHAM[1,2]

[1]*Clinical Psychoacoustics Laboratory, Department of Otolaryngology, University of Minnesota, MMC 396, Rm 8-323 PWB, 420 Delaware St. SE, Minneapolis, MN 55455, USA*
[2]*Department of Psychology, University of Minnesota, Elliott Hall, 75 East River Parkway, Minneapolis, MN 55455, USA*

ABSTRACT

Differences in fundamental frequency (F0) provide an important cue for segregating simultaneous sounds. Cochlear implants (CIs) transmit F0 information primarily through the periodicity of the temporal envelope of the electrical pulse trains. Successful segregation of sounds with different F0s requires the ability to process multiple F0s simultaneously, but it is unknown whether CI users have this ability. This study measured modulation frequency discrimination thresholds for half-wave rectified sinusoidal envelopes modulated at 115 Hz in CI users and normal-hearing (NH) listeners. The target modulation was presented in isolation or in the presence of an interferer. Discrimination thresholds were strongly affected by the presence of an interferer, even when it was unmodulated and spectrally remote. Interferer modulation increased interference and often led to very high discrimination thresholds, especially when the interfering modulation frequency was lower than that of the target. Introducing a temporal offset between the interferer and the target led to at best modest improvements in performance in CI users and NH listeners. The results suggest no fundamental difference between acoustic and electric hearing in processing single or multiple envelope-based F0s, but confirm that differences in F0 are unlikely to provide a robust cue for perceptual segregation in CI users.

Correspondence to: Andrew J. Oxenham · Department of Psychology · University of Minnesota · Elliott Hall, 75 East River Parkway, Minneapolis, MN 55455, USA. Telephone: +1-612-6242241; fax: +1-612-6252101; email: oxenham@umn.edu

Keywords: cochlear implants, pitch, F0, modulation frequency discrimination interference, pitch discrimination interference

INTRODUCTION

Cochlear implant (CI) users often experience difficulty understanding speech when it is presented in a background of other sounds, such as competing speech. Pitch differences are known to assist in segregating competing voices (e.g., Brokx and Nooteboom 1982; Darwin et al. 2003). Therefore, a potentially important factor in explaining this deficit is the relatively poor pitch perception of most CI users. The loss of pitch information may be in part due to poorer spectral resolution (caused by a limited number of electrodes, non-uniform survival of spiral ganglion cells, and spread of current) and to the loss of temporal fine structure information within individual frequency channels, which in turn may impede speech perception in complex backgrounds (e.g., Qin and Oxenham 2003; Stickney et al. 2007).

Some pitch information is conveyed to CI users via periodicity in the temporal envelope of the pulse trains or, in the case of low pulse rates (less than about 300 Hz), by the pulse rate itself (e.g., Busby and Clark 1997; Kong et al. 2009). Carlyon and colleagues have shown that CI users are often sensitive to changes in pulse rates in ways that resemble the sensitivity shown by normal-hearing (NH) listeners when presented with acoustic pulse trains that are high-pass filtered to remove potentially resolved spectral components (e.g., Carlyon et al. 2002, 2008).

Also, NH listeners have been shown to make use of temporal-envelope pitch cues when listening to speech against a background of a competing talker (Oxenham and Simonson 2009), suggesting that such cues might, in principle, be available to CI users. On the other hand, a study using pulse-train-excited vocoder simulations of CI processing in NH listeners found that performance with different pulse rates for the target and interfering talker (presented in separate frequency channels) was no better than performance with the same pulse rate for both (Deeks and Carlyon 2004). Similarly, a direct study of CI users' ability to make use of differences in pulse rate to perceptually segregate the pulse train presented on one electrode from the pulse trains presented to neighboring electrodes found that CI users did not benefit from the use of different pulse rates between the target and other pulse trains (Carlyon et al. 2007).

The idea that fundamental frequency (F0) and pitch differences can aid in perceptual segregation relies to some extent on the assumption that listeners are able to extract one F0 in the presence of another. There has been some formal study of this ability in NH listeners, using the phenomenon known as pitch discrimination interference (PDI) (Gockel et al. 2004, 2005; Micheyl and Oxenham 2007, Gockel et al. 2009a, b, c). Even when the target and interferer are both comprised solely of unresolved harmonics (similar to the temporal pitch experienced by CI users), pitch discrimination remains relatively good, so long as the target and interferer are presented in separate spectral regions (Gockel et al. 2005). On the other hand, if the target and interferer are comprised of unresolved harmonics and are presented in the same spectral region, pitch discrimination becomes essentially impossible and listeners report the percept of an unmusical "crackle" rather than two pitches (Carlyon 1996a, b; Micheyl et al. 2006, 2010).

To our knowledge, there are no published studies with CI users on pitch discrimination in the presence of interfering pitch information. Some studies have investigated the aggregate pitch percept produced by two peripherally overlapping temporal patterns (McKay and McDermott 1996), but none has studied the ability to discriminate one pitch stimulus in the presence of another. The extent to which temporal PDI is observed in CI users should help determine whether, at least in principle, CI users might benefit from differences in F0 (or pulse rate) between simultaneously presented sources.

GENERAL METHODS AND RATIONALE

We tested both CI users and NH listeners on their ability to discriminate the modulation frequency of temporal envelopes that were half-wave-rectified sinu-

soids. For the CI users, the temporal envelopes were imposed on a pulse train presented via a single electrode in the center of the electrode array at a rate of 2,000 pulses per second (Kreft et al. 2010). For the NH listeners, the temporal envelope was imposed on a carrier frequency of 6.3 kHz to create a "transposed stimulus" (van de Par and Kohlrausch 1997; Bernstein and Trahiotis 2002; Oxenham et al. 2004). The transposed stimulus is intended to elicit a temporal response in the auditory nerve that is similar to that produced by a low-frequency sinusoid (corresponding to the modulator frequency), but at a tonotopic location corresponding to the carrier frequency. These half-wave rectified sinusoids have been used previously for modulation frequency discrimination in both NH listeners (Oxenham et al. 2004) and CI users (Kreft et al. 2010). Transposed stimuli have been shown to improve sensitivity to interaural time differences in NH listeners, relative to that found for sinusoidal amplitude modulation (Bernstein and Trahiotis 2002); in CI users, transposed stimuli have been found to produce modulation frequency or rate discrimination thresholds that are very similar to those found for many other waveforms, including sinusoidal or square-wave modulation (Landsberger 2008; Kreft et al. 2010).

Modulation frequency discrimination thresholds were measured in isolation and in the presence of an interferer that was located basally or apically to the target cochlear location, or was centered on the same location as the target. Based on earlier results in NH listeners (e.g., Carlyon 1996a; Gockel et al. 2005; Micheyl et al. 2006), we might expect very strong interference effects when the target and interferer are at the same spectral location, and weaker effects when they are spectrally well separated. In addition, we explored the use of onset and offset asynchrony, by gating the target on 200 ms after the interferer, and gating it off 200 ms before the offset of the interferer. On one hand, asynchrony should provide a strong segregation cue that can reduce PDI when the target and interferer are presented in separate spectral regions (Gockel et al. 2004). On the other hand, asynchrony does not help when the stimuli are in the same overlapping spectral region (Carlyon 1996b), and it may be that a loss of spectral resolution can explain why asynchrony has not been shown to provide a strong segregation cue to CI users (Carlyon et al. 2007). We tested this prediction by using very remote spatial locations for the interferer (i.e., apical or basal electrodes, with the target presented to a middle electrode), in the expectation that direct interference would be relatively small. All subjects provided written informed consent, and the protocols were approved by the Institutional Review Board of the University of Minnesota.

METHODS: NORMAL-HEARING LISTENERS

Subjects

The six subjects, all female, ranged in age from 21 to 37 years (mean age=30.2 years) with audiometric thresholds of less than 20 dB HL at octave frequencies from 250 to 8,000 Hz. In addition, they all met our additional inclusion criterion of having an absolute threshold (measured using an adaptive two alternative forced-choice procedure) for a pure tone at 10,080 Hz (the highest carrier frequency used) of less than 30 dB sound pressure level (SPL).

Stimuli

The target stimulus was created by multiplying a 6,350-Hz pure tone carrier with a modulator that was a half-wave rectified sinusoid, low-pass filtered (fourth-order Butterworth) with a cutoff frequency of 20 % of the carrier frequency (i.e., 1,270 Hz) to limit the spectral extent of the stimulus. The starting phase of the modulator was randomized on each presentation, and the modulation depth was always 100 %. The target duration was 300 ms (including onset and offset ramps). The nominal frequency of the target modulation was 115 Hz. The target level was centered on 50 dB SPL (root mean square, rms), but was roved on each presentation within a range of ±3 dB (with uniform distribution) to reduce potential loudness cues.

The interfering stimulus, when present, was also created from a pure-tone carrier with a frequency of 4,000, 6,350, or 10,080 Hz, termed "apical," "middle," and "basal," respectively, for comparison with the CI conditions described below. The carriers were limited to relatively high frequencies to reduce the potential for spectral sidebands from the modulation to be spectrally resolved. It is generally believed that harmonics above about the 12th are unresolved (Houtsma and Smurzynski 1990; Shackleton and Carlyon 1994; Bernstein and Oxenham 2003), so a modulator of 115 Hz produces sidebands around 4,000 Hz (corresponding to a harmonic number around 35) that are clearly unresolved. The starting phase of the interfering carrier led that of the target carrier by 90 °, so that the intensities, rather than the amplitudes, of the two carriers added, even when they were at the same frequency. The interferer was either unmodulated (UNMD condition), or was multiplied with a low-pass-filtered half-wave rectified sinusoid, as with the target. Four different modulation frequency conditions were tested: (1) 9 semitones below the nominal target modulation frequency, or about 68 Hz (LOW condition); (2) the same frequency as the nominal target frequency, i.e., 115 Hz (SAME condition); (3) 9 semitones above the nominal target modulation frequency, or about 193 Hz (HIGH condition); and (4) a modulation frequency selected at

random in each interval with uniform distribution between 12 and 6 semitones below the nominal target modulation frequency, i.e., between 57.5 and 83.1 Hz (RAND condition). The interferer was always presented at an rms level of 50 dB SPL. The interferer was either gated synchronously with the 300-ms target, or was gated on 200 ms earlier and gated off 200 ms later, for a total duration of 700 ms.

The stimuli were presented in a background of threshold-equalizing noise (TEN), as defined by Moore et al. (2000). The TEN was band pass filtered between 20 and 2,500 Hz and was presented at a level per equivalent rectangular auditory-filter bandwidth (ERB$_N$) at 1 kHz of 40 dB SPL. The noise was designed to limit the audibility of distortion products, particularly those corresponding to the modulation frequency and its lower (spectrally resolved) harmonics. The noise was gated on 50 ms before the onset of the first interval and was gated off 50 ms after the end of the second interval in each trial. All stimuli, including the background noise, were gated on and off with 10-ms raised-cosine ramps.

Procedure

Experiments were controlled by a personal computer running custom MATLAB programs, including the AFC routines developed by Stefan Ewert (University of Oldenburg). Stimuli were generated digitally and were output by a 24-bit soundcard (Lynx22, Lynx Studio Technology, Costa Mesa, CA) at a sampling rate of 48 kHz, via a headphone buffer (HB6, Tucker-Davis Technologies, Alachua, FL) to headphones (HD580, Sennheiser USA, Old Lyme, CT) to listeners who were seated in a double-walled sound-attenuating chamber. All stimuli were presented monaurally to the left ear.

Thresholds were obtained using a two-interval two-alternative forced-choice task with a three-down one-up adaptive procedure that tracks the 79.4 % correct point on the psychometric function (Levitt 1971). In each trial, one of the two intervals had the higher target modulation frequency (selected at random with equal probability), and the listener was asked to decide which of the two target stimuli had the higher pitch. The two intervals were separated by a 200-ms inter-stimulus interval. Correct-answer feedback was provided after each trial.

The target modulation frequencies in the two intervals were geometrically centered on the nominal frequency, f_{mod}, of 115 Hz, so that the lower and higher frequencies were $f_{mod}(1+\Delta f_{mod}/100)^{-1/2}$ and $f_{mod}(1+\Delta f_{mod}/100)^{+1/2}$, respectively, where Δf_{mod} is the frequency difference, expressed as a percentage of the lower frequency. The initial frequency difference was 100 %. Initially Δf_{mod} was increased or decreased by a factor of 2. After the first reversal in the direction of the change in the tracking variable from "up" to "down," Δf_{mod} was changed in

factors of $\sqrt{2}$. After a further two reversals, the factor was decreased to the fourth root of 2, which was the final step size. The tracking procedure did not allow the value of Δf_{mod} to exceed 400 %. The geometric mean of Δf_{mod} from the final four turn points was taken as the threshold for each run. A minimum of three such threshold estimates were averaged to obtain a single modulation frequency discrimination threshold. Conditions were tested in a blocked randomized order, with all conditions tested once before any were repeated. The order of presentation was selected randomly between subjects and between repetitions. No training was provided; however, four of the six subjects had previous extensive psycho-acoustic testing experience. In a few instances, performance continued to improve with each new block of testing for the most difficult conditions. Testing continued until performance stabilized and a minimum of three additional thresholds were collected.

METHODS: COCHLEAR-IMPLANT USERS

Subjects and implants

The subjects were three post-lingually deafened adults with either a Clarion C-II or Hi-Res90K cochlear implant. Two subjects were bilaterally implanted, and each ear was tested separately, resulting in a total of five test ears. Full insertion of the electrode array (25 mm) was achieved in all cases. Table 1 provides additional subject information. For the present study, stimulation was monopolar, with the active intracochlear electrode referenced to an electrode on the case of the internal receiver–stimulator.

Stimuli and procedure

Experiments were controlled by a personal computer running custom programs written for the Bionic Ear Data Collection System (BEDCS; Advanced Bionics, Valencia, CA). The target stimuli were 300-ms trains of 32 μs/phase, cathodic-first biphasic pulses, presented in monopolar mode at a rate of 2,000 pulses per second (pps). The amplitude of the pulses was modulated by a

half-wave rectified sinusoid with random starting phase on each trial, and a modulation depth of 100 %.

First, the absolute threshold (THS) and the maximum acceptable loudness (MAL) were measured for each CI user, and for each electrode location (electrodes 2, 8, and 14), using the methods described in Kreft et al. (2010). In these measurements, the modulation frequency was set to 115 Hz, which was the nominal modulation frequency of the target. Earlier work has shown only small and unsystematic variations in THS or MAL as a function of modulation frequency (e.g., Kreft et al. 2010). The dynamic range (DR) was then determined using the difference in current (microamperes, μA) between MAL and THS for each CI user and each electrode location.

In the discrimination experiments, the same interference conditions were tested as described for the NH listeners, with the LOW, SAME, HIGH, and RAND modulation conditions, as well as the UNMD condition, and with the interferer either synchronously gated with the 300-ms target, or gated on for 700 ms, with the target temporally centered within the interferer. The target was presented on the middle electrode (electrode 8) at a nominal level corresponding to 40 % DR. In order to reduce any potential loudness cues, the current level for the target was roved across intervals by ±10 % of the nominal level with uniform distribution. The interferers were also trains of 32 μs/phase, cathodic-first biphasic pulses, presented in monopolar mode at a rate of 2,000 pps, also at a level corresponding to 40 % DR. The only exception was the unmodulated interferer, for which the pulse amplitude was set to the same level as the maximum pulse amplitude of the modulated stimuli, leading to a higher rms level overall. This was done so that the maximum peripheral interference produced by the unmodulated interferer would match that produced by the modulated interferers; a similar technique was employed by Chatterjee (2003) when investigating modulation detection interference. The target and interferer were interleaved such that the pulses were offset by half a period from the other. In other words, the pulse rate was effectively doubled to 4,000 pps when both were present on the same electrode. Interfering electrodes were selected at three points across the array,

TABLE 1						
Subject information						
Subject code	M/F	Age (years)	CI use (years)	Order	Etiology	Duration HL prior to implant (years)
D11	M	79.5	7.1	1st	Unknown	16
D19	F	50.9	6.2	1st	Unknown	11
D20	M	79.3	6.1	2nd	Unknown	16
D24	M	59.7	2.2	NA	Unknown progressive	27
D26	F	50.7	2.0	2nd	Unknown	11

Subject code, gender, age when tested for the present study, duration of implant use prior to the study, order of implantation for the bilateral users, etiology of deafness, and duration of bilateral severe-to-profound hearing loss prior to implantation. Implants 1 and 3 are two implants on the same subject, as are implants 2 and 5

corresponding to apical (electrode 2), middle (electrode 8), and basal (electrode 14) locations. These locations corresponded to frequency ranges in the clinical map of 338–445, 1,160–1,278, and 3,490–4,114 Hz, respectively. A wide spacing between apical, middle, and basal electrodes, together with a relatively low current level, was selected to reduce as far as possible the peripheral interactions between the stimulated electrodes (e.g., Nelson et al. 2011). Due to inherent delays in the BEDCS interface as implemented in our lab, a relatively long inter-stimulus interval of 700 ms was required. No onset or offset ramps were used.

The same adaptive tracking procedure was used as described for the NH listeners. No training was provided; however, all of the subjects had previous extensive psychophysical testing experience.

RESULTS

The results from both groups of listeners are reported below. All analysis was carried out on the log-transformed difference limens (DLs). Statistical significance includes a Huynh-Feldt correction for lack of sphericity, as appropriate, but with the original degrees of freedom reported.

Modulation frequency discrimination without interference

Figure 1 shows the DLs for modulation frequency for both NH (left panel) and CI groups (right panel). Geometric mean data are shown as larger colored stars (blue for NH and red for CI groups), and

individual data are shown as different smaller symbols, as shown in the panel legends. The different modulation conditions are shown along the abscissa, and within each modulation condition, the different spectral locations for the interferer are shown with the mean data connected. Considering first the DLs without interference (left-most condition in each panel), the mean DL for the NH group of 7.14 % is lower than that found by Oxenham et al. (2004) using similar stimuli. The mean DL of 12.1 % for the CI users seems somewhat higher (and more similar to the NH results reported by Oxenham et al. 2004), although the difference between the NH and CI groups was not statistically significant $[t(9)=1.9, p=.087]$, and appears to be driven by two poorer performers. Indeed, the median DL for the groups is quite similar (7.4 % and 9.3 % for the NH and CI users, respectively).

Modulation frequency discrimination with synchronous interference

Average DLs in the presence of unmodulated interferers (UNMD condition) were generally between 10 and 20 %, and did not appear to vary systematically with interferer location or between the two subject groups. A one-way repeated-measures analysis of variance (RMANOVA) was performed on the NH data, with the unmodulated interferer (absent, apical, middle, or basal) as the within-subjects factor. There was a significant effect of interferer $[F(3,15)=.32, p=.022]$. Contrast analysis revealed that the DL with no interferer (NO) was significantly lower (better) than the pooled DL estimate from the three unmodulated interferer (UNMD) conditions ($p<0.012$). Considering
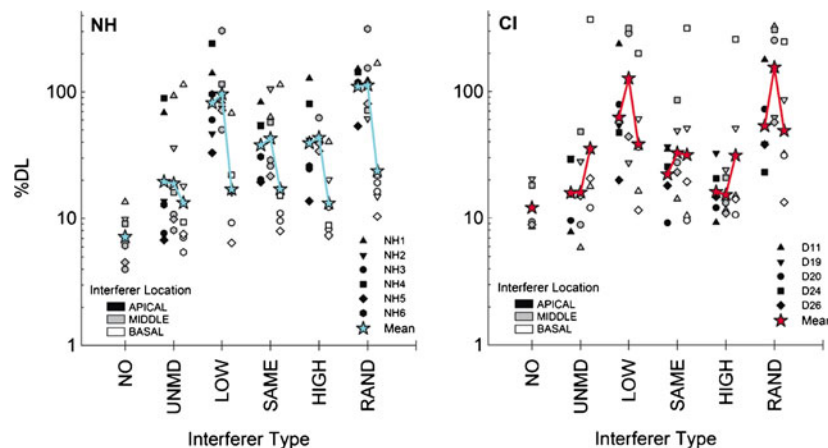

FIG. 1. Modulation frequency DLs, expressed as a percentage of the baseline modulation frequency (115 Hz), with a synchronous interferer. Both mean (*large stars*) and individual (*smaller symbols*) are shown. The left and right panels show results from NH listeners and CI users, respectively. The apical, middle, and basal interferer locations are depicted with *different shading* of the symbols. The different interferer conditions are shown separately. NO denotes DLs for the condition with no interferer; UNMD denotes DLs with the unmodulated interferer; LOW denotes DLs with the interferer modulated at 68 Hz; SAME denotes DLs with the interferer modulated at 115 Hz—the nominal modulation frequency of the target; HIGH denotes DLs with the interferer modulated at 193 Hz; and RAND denotes DLs with the interferer modulated with a different frequency in each interval, selected with uniform distribution on a semitone (log) scale from between 57.5 and 81.3 Hz.

just the three unmodulated interferer (UNMD) conditions, no significant effect of interferer location [$F(2,10)=0.799$, $p>0.4$] was found.

The same analysis was undertaken with the CI data. In contrast to the results from NH listeners, an RMANOVA with the unmodulated interferer as the within-subjects factor (levels: absent, apical, middle, and basal) failed to show a significant effect of interferer [$F(3,12)=3.94$, $p=0.087$], suggesting that there was no robust effect produced by an unmodulated interferer, regardless of its location.

Comparing across the two groups using a mixed-model RMANOVA, there was no significant main effect of group [$F(1,9)=0.34$, $p>0.5$], suggesting no overall difference in sensitivity between the NH and CI groups, but the interaction between interferer location and group did reach significance [$F(3,27)=3.11$, $p=0.043$], supporting the finding from the within-group ANOVAs of an effect of interference in the NH group but not in the CI group. In the NH group, the location of the interferer did not strongly affect the amount of interference (as indicated by the lack of a significant difference in DLs between apical, middle, and basal locations, described above), which is surprising, given that the apical and basal interferers were relatively remote from the target location (2/3 octave). Note, however, that group level analyses need to be treated with some caution, given the large inter-subject variability observed in the data, with some subjects in both groups performing consistently more poorly than others.

Consider next the effects of the modulated interferers (LOW, SAME, HIGH, and RAND). For the NH group, performance was generally somewhat poorer than with the unmodulated interferers (UNMD), particularly for the apical and middle interferer locations. A two-way RMANOVA on the NH data with factors of interferer location (apical, middle, basal) and interferer modulation type (UNMD, LOW, SAME, HIGH, RAND) revealed a main effect of location [$F(2,10)=14.5$, $p=0.001$] and modulation type [$F(4,20)=13.2$, $p<0.001$], as well as a significant interaction between location and modulation type [$F(8,40)=3.86$, $p=0.002$]. These effects presumably reflect the generally lower DLs with the basal (higher spectral location) interferer, the generally lower DLs with the unmodulated interferer (UNMD), as well as the generally higher (poorer) DLs with the lower frequency (LOW and RAND) modulation, and the fact that the effects of modulation and modulation type seem more pronounced in the apical and middle regions than in the basal region, where DLs are often similar to those for the unmodulated interferer (UNMD). A contrast analysis revealed that DLs with the unmodulated interferers (UNMD) were significantly lower than the pooled DL estimates from the conditions with modulated interferers

[$F(1,5)=15.0$, $p=0.012$], suggesting that in general modulation produced interference. To address the effect of using a random frequency modulator, we compared the DLs from the 68-Hz and random modulators across the three spectral locations in a separate RMANOVA. The results showed no significant effect of modulation type [$F(1,5)=3.08$, $p=0.14$], and no interaction between modulation type and interferer location [$F(2,10)=0.36$, $p=0.71$]. Thus, the introduction of random variations in modulation frequency from trial to trial did not result in greater impairment than fixed low-frequency modulation interference.

The same two-way RMANOVA using the data from the CI group, with factors of interferer modulation type and interferer location, showed a main effect of modulation type [$F(4,16)=7.59$, $p=0.001$] but no main effect of interferer location [$F(2,8)=.53$, $p=.61$], although the interaction was significant [$F(8,32)=6.05$, $p<0.001$], presumably reflecting the fact that there appeared to be little overall effect of modulation type for the basal interferer. A contrast analysis of the DLs in the unmodulated condition with DLs in all other conditions failed to reach significance [$F(1,4)=6.13$, $p=0.069$], again presumably because of the lack of effect with the basal interferer. However, the overall level of performance and pattern of results were quite similar to those observed in data from the NH group. Comparing just the low-frequency (68-Hz) and random frequency modulators, there was no significant effect of modulation type [$F(1,4)=.84$, $p=.41$], and no interaction with interferer location [$F(2,8)=1.52$, $p=0.28$], suggesting that (as with the NH group) random variations in interferer modulation frequency did not further impair performance.

Comparing the two groups using a mixed-model RMANOVA, there was no main effect of subject group [$F(1,9)=0.008$, $p=0.93$], although the interaction between subject group and location did reach significance [$F(2,18)=4.52$, $p=0.027$], presumably reflecting the impression that the effect of interferer location was more pronounced and more systematic for the NH group than for the CI group. Neither the interaction between modulation type and subject group [$F(4,36)=0.914$, $p=0.41$] nor the three-way interaction [$F(8,72)=1.75$, $p=0.14$] was significant.

Overall, performance in the presence of a modulated interferer was often very poor, with many subjects (both NH and CI) obtaining DLs greater than 100 %, suggesting little or no rate discrimination ability in the presence of an interferer, particularly when the interferer modulation frequency was lower than that of the target. Spectral (or spatial) separation between the target and the interferer did not lead to robust improvements in performance, particularly when compar-

ing performance between the apical and middle interferer locations.

Effects of temporal asynchrony between the target and interferer

Figure 2 shows average modulation frequency DLs expressed as a percentage for the five modulation conditions in the cases where the target was gated on 200 ms after the interferer (and gated off 200 ms before the interferer), as well as the data for the no interferer (NO) from Figure 1 replotted to facilitate comparisons. As with Figure 1, data from the NH and CI groups are shown in the left and right panels, respectively.

For the NH data, a RMANOVA showed that both main effects of modulation type [$F(4,20)=10.79$, $p<0.001$] and interferer location [$F(2,10)=49.9$, $p<0.001$] were significant, as was their interaction [$F(8.40)=3.83$, $p=0.02$]. In contrast to the results with the synchronous interferer, the middle asynchronous interferer seemed to produce consistently higher (poorer) thresholds than the interferer at the other two locations. Contrast analysis revealed a significant difference between DLs in the unmodulated conditions compared with pooled estimates from the modulated conditions [$F(1,5)=31.8$, $p=0.002$], suggesting that overall modulation produced interference.

In many respects, the pattern of the CI data was similar to that of the NH data. A RMANOVA again revealed significant main effects of modulation type [$F(4,16)=6.69$, $p=0.002$] and interferer location [$F(2,8)=12.44$, $p=0.004$], along with a significant interaction [$F(8,32)=6.64$, $p<0.001$]. The interaction seems to be due to the lack of effect of modulation or modulation type for either the basal or apical interferers. In contrast to the data from the NH group (but in line with the synchronous interferer conditions of the CI data), contrast analysis showed no significant difference between DLs in the unmodulated conditions compared with pooled estimates from the modulated conditions [$F(1,4)=2.33$,

$p=0.2$], again presumably because neither the apical or basal interferers seem to show much effect of modulation type. Despite these differences, the between-subjects main effect of group (NH versus CI) was not significant [$F(1,9)=1.284$, $p=0.286$], suggesting again that the overall level of performance across the two groups was quite similar. However, the two-way and three-way interactions with subject group were significant ($p<0.01$ in all cases), reflecting the somewhat different pattern of results across locations and modulation types in the two groups seen in Figure 2.

Figure 3 compares the synchronous with the asynchronous conditions directly using the individual ratios of the DLs for all five modulation conditions and three locations. A ratio of 1 indicates that the DLs for that subject in that condition are the same for both synchronous and asynchronous conditions. Ratios greater than 1 indicate a larger DL for the synchronous condition, and thus a benefit of onset and offset asynchrony.

Considering first the data from the NH group, a three-way RMANOVA (with factors modulation type, location, and asynchrony) showed that asynchrony did not have a significant main effect on DLs [$F(1,5)=2.65$, $p=0.17$]. However, there was a significant interaction between asynchrony and modulation condition [$F(4,20)=5.47$, $p=0.004$], and between asynchrony and interferer location [$F(2,10)=38.4$, $p<0.001$]. These interactions seem to reflect the apparent benefit of asynchrony in the apical interferer location, particularly for the LOW and RAND conditions, as well as the detrimental effect of asynchrony in the case of the same interferer spectral location. No clear effect of asynchrony was observed in the basal (higher spectral) location, where little effect of interferer modulation was observed even in the synchronous conditions. These observations were supported by separate RMANOVAs in the three interferer locations: for the apical (low-frequency) interferer, the effect of asynchrony was highly signif-
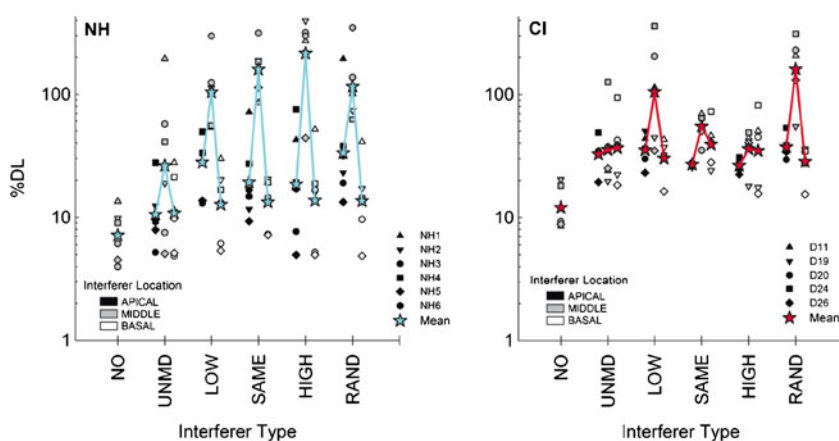

FIG. 2. Modulation frequency DLs, expressed as a percentage of the baseline modulation frequency (115 Hz), with an asynchronous interferer that began 200 ms before and ended 200 ms after the target in each interval. Data are presented as described in Figure 1.
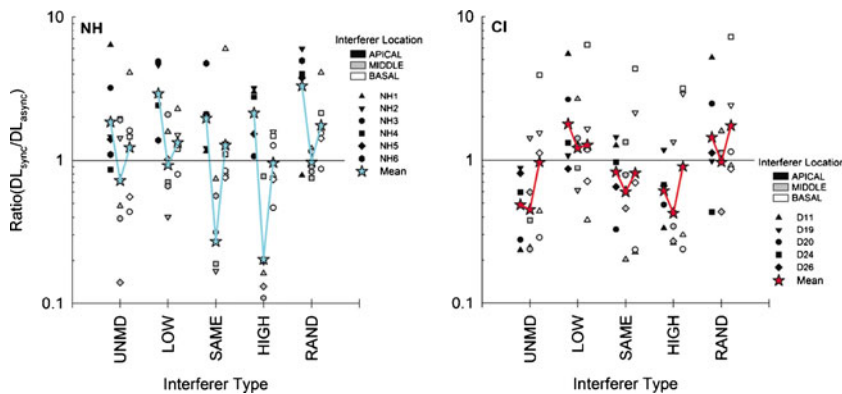
icant [$F(1,5)=410$, $p<0.001$], with no interaction between asynchrony and modulation type ($p>0.1$), suggesting *improved* overall performance in the presence of target interferer asynchrony. For the middle (same frequency) interferer, the effect of asynchrony was also significant [$F(1,5)=23.8$, $p=0.005$], confirming the *deterioration* in performance due to asynchrony, with an interaction between asynchrony and modulation type, reflecting the fact that DLs in some conditions (such as the high-frequency modulation) were affected more than others. For the basal (higher-frequency) interferer, neither the main effect of asynchrony, nor its interaction with modulation type, reached significance ($p>0.2$ in both cases).

For the CI group, a three-way RMANOVA again showed no significant main effect of asynchrony [$F(1,4)=0.60$, $p=0.483$], and in this case, the interaction between asynchrony and location was also not significant [$F(2,8)=0.82$, $p=0.476$], suggesting a different pattern from that observed in the NH group, although the interaction between asynchrony and modulation type did reach significance [$F(4,16)=3.56$, $p=0.029$].

A mixed-model RMANOVA including data from both groups revealed no significant main effect of subject group [$F(1,9)=0.331$, $p=0.579$], and no interaction between subject group and asynchrony [$F(1,9)=2.28$, $p=0.165$]; however, the three-way interaction between subject group, asynchrony, and interferer location did reach significance [$F(2,18)=4.84$, $p=0.036$], suggesting that the interferer location influenced the effect of asynchrony more for the NH group than for the CI group.

DISCUSSION

Summary of results

Modulation frequency discrimination interference was measured in both CI users and NH listeners. Substantial interference, relative to no interferer, was

observed. In many cases, modulation of the interferer resulted in greater interference than was found with the unmodulated interferer, particularly when the interferer was apical to the target or at the same location as the target, and when the interferer modulation frequency was lower than the target. Asynchronous gating of the target, relative to the interferer, seemed to provide some benefit to NH listeners when the interferer was apical to the target, but resulted in poorer performance when the target and interferer were at the same spectral location, and provided little or no benefit when the interferer was basal relative to the target. Little evidence for any systematic effect of asynchronous gating was found for the CI group.

Comparison with previous studies

To our knowledge, there are no previous studies of modulation frequency discrimination interference in NH listeners using transposed tones. Similarly, there seem to be no previous published results on modulation frequency discrimination with modulated interference in CI users. Chatterjee (2003) reported interference in modulation *detection* produced by the presence of interfering modulation at different electrode locations, and has shown that this interference exceeds the amount produced by an unmodulated interfering pulse train. Chatterjee and Oberzut (presented at the 2009 Conference on Implantable Auditory Prostheses, Lake Tahoe, CA) have reported that interfering modulation can enhance modulation frequency discrimination on a remote electrode. However, this enhancement may be because their baseline condition involved target and interferer modulation that was at the same frequency and in-phase in one interval of the forced-choice task, and was at different frequencies in the other interval, meaning that discrimination did not require the extraction of either modulation frequency, and could have been achieved by

detecting any form of incoherence in the stimulation on the two electrodes.

Comparing pitch discrimination interference in NH and CI listeners

The stimuli and experimental design that were used for the NH and CI groups were selected to be as comparable as possible. Nevertheless, some important differences should be born in mind when directly comparing the results between the two groups. The first factor is that the spacing between the interferer and the target carriers was different: for the CI group, the spacing was selected to be relatively wide, to reduce peripheral interactions as much as possible; for the NH group, the spacing was constrained by having to ensure that the carriers were above currently accepted limits of phase-locking to the carrier frequency, and to ensure that modulation frequencies were sufficiently high to avoid spectral contributions of the sidebands generated by the modulation (e.g., Santurette and Dau 2011; Santurette et al. 2012). Therefore, the spacing between carriers of 2/3 octave, although probably wide enough to limit peripheral interactions at the relatively low stimulus levels used in this study, was not as wide, in terms of cochlear locations, as that used in the CI group. On the other hand, spatial spread of excitation is generally thought to be wider in CI users than in NH users, due to factors such as current spread, particularly in monopolar stimulation mode, and this may counteract the effects of wider carrier spacing in the CI group. A second factor is that no attempt was made to equate the loudness of the stimuli between the two groups. In particular, the unmodulated interferers were likely to have been perceived as louder than the modulated interferers by the CI group; the decrease in loudness introduced by the modulation may have helped to counteract the interference produced by the modulation. Note, however, that interference for modulation detection was observed under similar conditions by Chatterjee (2003). A third factor is the carrier frequency. For the NH group, the carrier frequency and place of stimulation covary; for the CI group, the place of stimulation is determined by the electrode location, and the carrier frequency (pulse rate) was held constant at 2,000 Hz. We do not believe that this difference is likely to be material in interpreting the outcomes. In particular, the high pulse rate was selected to be well above the rates at which CI users are generally sensitive to changes in pulse rate, and the rate at which changes in rate induce changes in reported pitch. Because of this, it is unlikely that an even higher pulse rate would have had any material effect on the results. A fourth factor is that the average age of the CI users was substantially higher than that of the NH listeners, which may have led to poorer performance by the CI group based on age.

Keeping these potential differences in mind, one of the most striking outcomes of this study was the similarity of the results from the CI users and the NH listeners. This outcome is generally consistent with findings from Carlyon et al. (2002), who have reported similar results for rate discrimination of electric and acoustic pulse trains with CI users and NH listeners, and from Carlyon et al. (2007), who showed a similar inability to use rate difference cues to segregate the pulses from one spectral region or electrode from those on other neighboring regions or electrodes. The present findings extend those of previous studies by showing that modulation frequency discrimination is impaired by the presence of an interferer, even in conditions where there is very little possibility of strong peripheral overlap of the target and interferer. The separation of six electrodes (in CI users) or 2/3 octave (in NH listeners) should have been sufficient to limit the influence of spread of excitation in most listeners at the stimulation levels used in this study. For instance, stimulation at a distance of six electrodes and at a level of 40 % DR falls outside the forward-masked tuning curves of most of the 15 CI users measured by Nelson et al. (2011), suggesting that the degree of peripheral interaction would have been relatively small.

The poor performance of the NH listeners seems at odds with the results of Gockel et al. (2004), who showed relatively good performance with unresolved harmonics ($d' \approx 1$ with $\Delta F0 = 3.5$ %), and relatively little effect of interference from unresolved harmonics in a lower spectral region. One important difference may be the bandwidth of the stimuli; in Gockel et al. (2004), the target was a band-pass-filtered harmonic tone complex (in sine phase) filtered between 3,900 and 5,400 Hz (about one half octave) with an F0 of 88 Hz, providing a total of about 17 components in the passband. In the present study, the 10-dB bandwidth of the target at its nominal modulation frequency was only about 500 Hz, or one tenth octave, and included only five components. The narrower bandwidth (and smaller number of components) may have resulted in lower pitch salience, and hence more susceptibility to interference.

Overall, it appears that the temporal cues provided by modulated electrical pulse trains (or a modulated high-frequency acoustic carrier) are relatively poor and are not robust to spectrally remote interference, often leading to DLs that are so high as to be probably unusable for the purposes of perceptual segregation. This outcome is common to both CI users and NH listeners, and so does not appear to be due to deficits specific to CI users, but instead may reflect a fundamental limitation of the periodicity information that is conveyed temporally within a band-limited spectral region.

Effects of interferer modulation frequency on discrimination

Although the presence of the interferer elevated thresholds overall, the effect of interferer *modulation* was not uniform, and varied as a function of both electrode/spectral location, modulation frequency, and subject group. When modulation interference (or PDI) was observed, it was found typically for apical (low) and middle (same) interferers, but not for the basal (high) interferer. In general, interference was greatest for the lower modulation frequencies, in both the fixed (68-Hz) and random frequency conditions.

The difference between the low and same modulation frequency may reflect the fact that overall (composite) rate cues could be used more readily when the target and interferer had similar rates (e.g., Carlyon 1996a). Note that the target and interferer never had exactly the same rate, as the target modulation frequencies were always centered on the interferer modulation frequency of 115 Hz, and that the starting modulation phases were selected at random on each presentation. The reduced interference often observed at the highest modulation frequency (193 Hz) may reflect a reduction in pitch strength (and discrimination) often observed at higher frequencies. For instance, also using half-wave rectified sinusoidal modulation presented at 40 % DR, Kreft et al. (2010) found that frequency DLs in CI users increased from around 10 % to over 50 % as the modulation frequency increased from 115 to 230 Hz. Similar deteriorations in performance with increasing rate were observed by Oxenham et al. (2004) in NH listeners. It seems likely that PDI decreases as the pitch salience of the interferer decreases. This outcome is in line with the results of Bernstein and Trahiotis (2002), who found that sensitivity to interaural time differences imposed on the envelopes of transposed stimuli decreased within increasing modulation rate. All these results are consistent with decreasing detection sensitivity to monaural amplitude modulation at modulation frequencies above about 150 Hz (Kohlrausch et al. 2000).

Asynchrony as a segregation cue

Overall, gating the target asynchronously with the interferer did not improve performance. In fact, when the interferer was in the same spectral region as the target, asynchronous gating led to poorer performance and higher DLs, particularly in the NH group. This detrimental effect of asynchronous gating is consistent with the findings from NH listeners by Carlyon (1996a, b), who showed a similar effect with acoustic pulse trains

filtered into the same spectral region (see also Micheyl et al. 2006). The newer aspect of our data is the fact that asynchronous gating did not produce a pronounced benefit to performance, even in cases where the target and interferer were well spectrally separated: although some improvement was observed in NH listeners with the apical interferer, there was not with the basal interferer. It may be that when pitch cues are as weak as they are in most CI users (and in NH listeners when presented with "transposed" stimuli) that any form of interference is sufficient to impair performance in ways that are not mitigated by asynchronous gating.

Implications for CI processing schemes

Discrimination thresholds without interference were generally poor (between about 4 % and 20 %, or one and three musical semitones), and were further impaired in the presence of simultaneous interference. In cases where the interfering modulation frequency was low (LOW or RAND), DLs often exceeded 100 %, suggesting little or no modulation frequency discrimination ability. The fragility of pitch discrimination in the presence of any interference leads to the conclusion that differences in modulation (or pulse) rate are unlikely to serve as an effective segregation cue in current cochlear implants. The fact that similar results were observed in NH listeners suggests that the limitations are of a fundamental nature pertaining to the utility (or lack thereof) of temporal envelope cues, when presented within a limited spectral region. The conclusion that temporal rate differences are unlikely to serve as a robust segregation cue in CIs is consistent with the conclusions of Carlyon et al. (2007), and extend them to conditions where the target and interferer are relatively remote in tonotopic location. The lack of robust benefit of asynchronous gating provides a further indication that rate differences are unlikely to serve as a usable segregation cue in current CI processors.

ACKNOWLEDGMENTS

REFERENCES

BERNSTEIN JG, OXENHAM AJ (2003) Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number? J Acoust Soc Am 113:3323–3334

BERNSTEIN LR, TRAHIOTIS C (2002) Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli". J Acoust Soc Am 112:1026–1036

BROKX JP, NOOTEBOOM SG (1982) Intonation and the perceptual separation of simultaneous voices. J Phonetics 10:23–36

BUSBY PA, CLARK GM (1997) Pitch and loudness estimation for single and multiple pulse per period electric pulse rates by cochlear implant patients. J Acoust Soc Am 101:1687–1695

CARLYON RP (1996A) Encoding the fundamental frequency of a complex tone in the presence of a spectrally overlapping masker. J Acoust Soc Am 99:517–524

CARLYON RP (1996B) Masker asynchrony impairs the fundamental-frequency discrimination of unresolved harmonics. J Acoust Soc Am 99:525–533

CARLYON RP, LONG CJ, DEEKS JM, MCKAY CM (2007) Concurrent sound segregation in electric and acoustic hearing. J Assoc Res Otolaryngol 8:119–133

CARLYON RP, MAHENDRAN S, DEEKS JM, LONG CJ, AXON P, BAGULEY D, BLEECK S, WINTER IM (2008) Behavioral and physiological correlates of temporal pitch perception in electric and acoustic hearing. J Acoust Soc Am 123:973–985

CARLYON RP, VAN WIERINGEN A, LONG CJ, DEEKS JM, WOUTERS J (2002) Temporal pitch mechanisms in acoustic and electric hearing. J Acoust Soc Am 112:621–633

CHATTERJEE M (2003) Modulation masking in cochlear implant listeners: envelope versus tonotopic components. J Acoust Soc Am 113:2042–2053

DARWIN CJ, BRUNGART DS, SIMPSON BD (2003) Effects of fundamental frequency and vocal-tract length changes on attention to one of two simultaneous talkers. J Acoust Soc Am 114:2913–2922

DEEKS JM, CARLYON RP (2004) Simulations of cochlear implant hearing using filtered harmonic complexes: implications for concurrent sound segregation. J Acoust Soc Am 115:1736–1746

GOCKEL H, CARLYON RP, MOORE BCJ (2005) Pitch discrimination interference: the role of pitch pulse asynchrony. J Acoust Soc Am 117:3860–3866

GOCKEL H, CARLYON RP, PLACK CJ (2004) Across-frequency interference effects in fundamental frequency discrimination: questioning evidence for two pitch mechanisms. J Acoust Soc Am 116:1092–1104

GOCKEL HE, CARLYON RP, PLACK CJ (2009A) Further examination of pitch discrimination interference between complex tones containing resolved harmonics. J Acoust Soc Am 125:1059–1066

GOCKEL HE, CARLYON RP, PLACK CJ (2009B) Pitch discrimination interference between binaural and monaural or diotic pitches. J Acoust Soc Am 126:281–290

GOCKEL HE, HAFTER ER, MOORE BC (2009C) Pitch discrimination interference: the role of ear of entry and of octave similarity. J Acoust Soc Am 125:324–327

HOUTSMA AJM, SMURZYNSKI J (1990) Pitch identification and discrimination for complex tones with many harmonics. J Acoust Soc Am 87:304–310

KOHLRAUSCH A, FASSEL R, DAU T (2000) The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers. J Acoust Soc Am 108:723–734

KONG YY, DEEKS JM, AXON PR, CARLYON RP (2009) Limits of temporal pitch in cochlear implants. J Acoust Soc Am 125:1649–1657

KREFT HA, OXENHAM AJ, NELSON DA (2010) Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users. J Acoust Soc Am 127:656–659

LANDSBERGER DM (2008) Effects of modulation wave shape on modulation frequency discrimination with electrical hearing. J Acoust Soc Am 124:EL21–27

LEVITT H (1971) Transformed up-down methods in psychoacoustics. J Acoust Soc Am 49:467–477

MCKAY CM, MCDERMOTT HJ (1996) The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. J Acoust Soc Am 100:1081–1092

MICHEYL C, BERNSTEIN JG, OXENHAM AJ (2006) Detection and F0 discrimination of harmonic complex tones in the presence of competing tones or noise. J Acoust Soc Am 120:1493–1505

MICHEYL C, KEEBLER MV, OXENHAM AJ (2010) Pitch perception for mixtures of spectrally overlapping harmonic complex tones. J Acoust Soc Am 128:257–269

MICHEYL C, OXENHAM AJ (2007) Across-frequency pitch discrimination interference between complex tones containing resolved harmonics. J Acoust Soc Am 121:1621–1631

MOORE BCJ, HUSS M, VICKERS DA, GLASBERG BR, ALCANTARA JI (2000) A test for the diagnosis of dead regions in the cochlea. Br J Audiol 34:205–224

NELSON DA, KREFT HA, ANDERSON ES, DONALDSON GS (2011) Spatial tuning curves from apical, middle, and basal electrodes in cochlear implant users. J Acoust Soc Am 129:3916–3933

OXENHAM AJ, BERNSTEIN JGW, PENAGOS H (2004) Correct tonotopic representation is necessary for complex pitch perception. Proc Natl Acad Sci U S A 101:1421–1425

OXENHAM AJ, SIMONSON AM (2009) Masking release for low- and high-pass filtered speech in the presence of noise and single-talker interference. J Acoust Soc Am 125:457–468

QIN MK, OXENHAM AJ (2003) Effects of simulated cochlear-implant processing on speech reception in fluctuating maskers. J Acoust Soc Am 114:446–454

SANTURETTE S, DAU T (2011) The role of temporal fine structure information for the low pitch of high-frequency complex tones. J Acoust Soc Am 129:282–292

SANTURETTE S, DAU T, OXENHAM AJ (2012) On the possibility of a place code for the low pitch of high-frequency complex tones. J Acoust Soc Am 132:3883–3895

SHACKLETON TM, CARLYON RP (1994) The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. J Acoust Soc Am 95:3529–3540

STICKNEY GS, ASSMANN PF, CHANG J, ZENG FG (2007) Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences. J Acoust Soc Am 122:1069–1078

VAN DE PAR S, KOHLRAUSCH A (1997) A new approach to comparing binaural masking level differences at low and high frequencies. J Acoust Soc Am 101:1671–1680