

The upper limit of temporal pitch for cochlear-implant listeners: Stimulus duration, conditioner pulses, and the number of electrodes stimulated

Robert P. Carlyon and John M. Deeks

MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, United Kingdom

Colette M. McKay

School of Psychological Sciences, University of Manchester, Manchester M13 9PL, United Kingdom

(Received 8 May 2009; revised 15 October 2009; accepted 18 December 2009)

Three experiments studied discrimination of changes in the rate of electrical pulse trains by cochlear-implant (CI) users and investigated the effect of manipulations that would be expected to substantially affect the pattern of auditory nerve (AN) activity. Experiment 1 used single-electrode stimulation and tested discrimination at baseline rates between 100 and 500 pps. Performance was generally similar for stimulus durations of 200 and 800 ms, and, for the longer duration, for stimuli that were gated on abruptly or with 300-ms ramps. Experiment 2 used a similar procedure and found that no substantial benefit was obtained by the addition of background 5000-pps “conditioning” pulses. Experiment 3 used a pitch-ranking procedure and found that the range of rates over which pitch increased with increasing rate was not greater for multiple-electrode than for single-electrode stimulation. The results indicate that the limitation on pulse-rate discrimination by CI users, at high baseline rates, is not specific to a particular temporal pattern of the AN response.

© 2010 Acoustical Society of America. [DOI: 10.1121/1.3291981]

PACS number(s): 43.66.Ts, 43.66.Hg [RYL]

Pages: 1469–1478

I. INTRODUCTION

An important finding in the cochlear-implant (CI) literature concerns the ability of CI users to process differences in the repetition rate of an electrical stimulus, such as a pulse train. For pulse rates lower than about 300 pps, listeners can detect rate differences of a few percent, and some listeners can use pulse-rate differences to recognize melodies and to estimate or even produce musical intervals (Pijl and Schwarz, 1995; McDermott and McKay, 1997; Moore and Carlyon, 2005). However, at higher pulse rates, these abilities usually decline dramatically; although marked individual differences exist, the vast majority of CI users do not reliably associate increases in pulse rate above about 400 pps with increases in pitch, and in many listeners this “upper limit” is considerably lower (Shannon, 1983; Townshend *et al.*, 1987; McKay *et al.*, 2000; Zeng, 2002; but see Kong *et al.*, 2009).

The upper limit of temporal pitch in CI users is of both clinical and scientific importance. Clinically, attempts to encode “temporal fine structure” in CIs (Nie *et al.*, 2005; Stickney *et al.*, 2005; Riss *et al.*, 2008) are likely to be limited by the range of temporal repetition rates that can be processed by CI users. Scientifically, several authors have pointed out the apparent paradox between the inability of CI users to encode high pulse rates and both the accurate encoding of such stimuli in the auditory nerve of deafened animals, and evidence that normal-hearing (NH) listeners use phase locking cues to estimate the frequency of pure tones up to at least 2000 Hz (Moore, 1973; Hartmann *et al.*, 1990; Shepherd and Javel, 1997; Micheyl *et al.*, 1998). This in turn has led to suggestions that accurate temporal processing may be subjected to additional requirements that are not met by CI

stimulation; these have included (i) stochastic activity in the auditory nerve (AN) (Rubinstein *et al.* 1999), (ii) comparison of timing differences across different AN fibers produced by the basilar-membrane (BM) traveling wave (Shamma, 1985; Moore and Carlyon, 2005), (iii) a “match” between the temporal pattern of activity in a given nerve fiber and the place on the BM that it innervates (Moore, 1982; Oxenham *et al.*, 2004; Moore and Carlyon, 2005), and (iv) stimulation of auditory nerve fibers that innervate very apical regions of the cochlea (Middlebrooks and Snyder, 2009).

Despite its importance, the locus of the high-rate limitation remains largely unknown. Circumstantial evidence for a peripheral limitation, at the level of the AN, comes from measures of the electrically evoked compound action potential (ECAP). Wilson *et al.* (1997a) reported that although the ECAP to each pulse of a low-rate (e.g., 100 pps) pulse train was of roughly equal size, at higher pulse rates (e.g., 1000 pps) the ECAP to odd-numbered pulses was larger than that to even-numbered pulses. This “alternating-amplitude” pattern was attributed to a large number of neurons being activated by the first pulse, refractory when the second pulse was presented, recovered by the third, and so on. It could affect pitch perception at high rates by conveying two intervals to the brain, corresponding to the time between every pulse and to that between every second pulse [Fig. 1(a)]. For example, if the compound activity of the AN were processed by an array of “more central” neurons having different thresholds (Carlyon *et al.*, 2008b), those with higher thresholds would respond only to the odd-numbered pulses and therefore convey a pulse rate an octave lower than that actually presented. If the depth of the alternating-amplitude pattern increased with pulse rate, then more and more of these neurons would

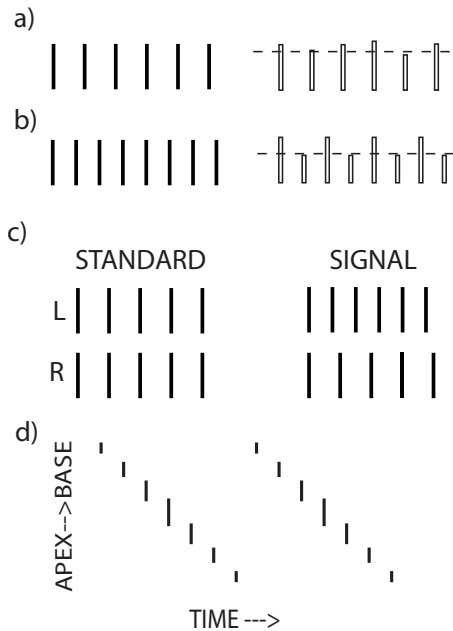


FIG. 1. (a) and (b) show schematic pulse trains (left-hand side, filled bars) together with hypothetical “alternating-amplitude” ECAPs (right-hand side, open bars). More central neurons having thresholds above the dashed lines would respond to every other pulse. If the ECAP becomes more modulated with increasing rate (b) then the increase in the number of more central neurons responding only to every other pulse might counteract the increase in pitch resulting from the shorter interpulse interval. (c) shows pulse trains presented in the standard and signal intervals of the “dichotic” conditions studied by van Hoesel and by Carlyon (van Hoesel and Clark, 1997; van Hoesel, 2007; Carlyon *et al.*, 2008a). The pulse rate presented to the left ear in the signal interval differs not only from that presented to the same ear in the standard interval but also from that presented to the right ear in the signal interval. The resulting binaural cue aids performance at low, but not at high, pulse rates (Carlyon *et al.*, 2008a). (d) represents the pulse trains presented on seven electrodes in the “peaked” condition of experiment 3.

convey rates corresponding to every second pulse, and this could counteract the increase in pitch expected from shorter interpulse intervals [Fig. 1(b)].

Evidence also exists for a more central limitation. Oxenham *et al.* (2004) required NH listeners to discriminate differences in the modulation rate of so-called “transposed tones” (Bernstein and Trahiotis, 2002)—high-frequency tones multiplied by a half-wave-rectified low-frequency sinusoid. They argued that these stimuli should produce similar patterns of AN activity to a pure tone, but in fibers having characteristic frequencies (CFs) corresponding to the carrier (but see Dreyer and Delgutte, 2006). Frequency difference limens (DLs) were considerably higher than for pure tones, and, importantly, listeners could not extract a “missing fundamental” from three harmonically related modulators each applied to a separate high-frequency carrier. They argued that the processing of phase locking at the level of the brainstem and above required either a match between the place and rate of stimulation or a comparison between the relative timing of activity in different AN fibers. More recently, Kong *et al.* (2009) showed that although the variation in rate discrimination as a function of baseline rate differed across listeners, for a given listener this variation was similar for rate discrimination of simple pulse trains and for the discrimination of the rate of sinusoidal amplitude modulation applied to a

5000-pps carrier. Presumably, the two types of stimulus produced different patterns of AN activity, tentatively suggesting that the across-listener differences may have arisen from more central processes.

The present study evaluates stimulus manipulations that would be expected to markedly affect the pattern of AN activity and which could feasibly extend the upper limit of temporal pitch. Experiment 1 examines the effect of increasing stimulus duration from 200 to 800 ms, and of turning the pulse trains on and off with long (300-ms) ramps. Physiological data indicate that the alternation in the ECAP amplitude to successive pulses gets smaller later in the pulse train (Wilson *et al.*, 1997b), so using a longer stimulus should provide a portion of the stimulus during which the alternation is reduced. Furthermore, we would expect the use of gradual onsets to also reduce this alternation because the effective “first pulse” will differ across neurons according to their threshold. Experiment 2 investigated whether the addition of a background 5000-pps pulse train would improve rate discrimination of moderate-rate (100–500 pps) pulses. The use of these so-called “conditioning pulses” has been shown to abolish the alternating-amplitude ECAP pattern (Rubinstein *et al.*, 1999). Experiment 3 studied whether pitch increased monotonically to higher overall pulse rates when multiple electrodes, rather than a single electrode, were stimulated (McKay *et al.*, 2005). Overall, none of the manipulations performed in the three experiments produced a substantial improvement in performance at the upper end of the rates studied. Section V discusses the generality of the present findings and considers additional steps that may be necessary to extend the range of rates that can be accurately encoded by CI users.

II. EXPERIMENT 1: EFFECTS OF SIGNAL DURATION AND ONSET/OFFSET ENVELOPE

A. Method

Six users of the Nucleus CI24M implant (Table I) were presented with biphasic pulse trains to electrode 11, in monopolar (MP1+2) mode, with a pulse duration of 45 μ s per phase and an interphase gap of 8 μ s. In different conditions, the pulse trains were either turned on abruptly and had a duration of 200 or 800 ms or, for the 800-ms duration, were turned on and off with 300-ms ramps. The ramps were linear in terms of Cochlear Corporation’s (roughly logarithmic) clinical “current units” (CUs), with end points corresponding to threshold and the steady-state amplitude of the central portion of the pulse train (see below). Stimuli were controlled using the APEX software package (Laneau *et al.*, 2005) and checked using a test implant and digital storage oscilloscope.

Rate discrimination was measured for 100-, 200-, 300-, 400-, and 500-pps standards using the “mixed-block” procedure described by Kong *et al.* (2009). On each trial, the listener was presented with the standard and a signal, having a 35% higher rate, in random order, and was required to indicate which interval contained the higher pitch. Responses were scored as correct when the signal interval was chosen, and correct-answer feedback was provided. The order in

TABLE I. Details of the cochlear-implant users who took part in the experiments.

Subject	Age (years)	Etiology	Duration of profound		BKBq (%)	BKBn (%)
			deafness (years)	Duration of implant use		
S1	61	CSOM	10	8 years	90	85
S2	77	Otosclerosis/noise induced	22	5 years	90	69
S3	41	Congenital progressive	35	2 years	97	46
S4	76	Progressive unknown	5	5 years	84	78
S5	66	Sudden, unknown	21	6 years	89	56
S6	71	Otosclerosis	>30	7 years	49	22
S7	60	Sudden, viral	>30	1 year	58	14
M1	44	Progressive unknown	2 years	18 months	98.7%	88.0%
M2	72	Unknown	Unknown	2 years	Unknown	Unknown

which the different standard rates were presented from trial to trial was randomized.¹ Each block consisted of 10 trials per standard rate, and the results from 10 blocks were averaged; each data point presented here therefore corresponds to the mean of 100 trials.

Prior to the main part of the experiment, the stimuli were loudness balanced using the method described by Landsberger and McKay (2005). A pair of stimuli was presented to each subject. At the initial presentation, the first stimulus in the pair was fixed at “comfort (“C”) level” and the second was presented at a much quieter level. The subject was instructed to raise the level of the variable sound (second stimulus) until it was slightly louder than the fixed sound and then reduce the level until the two sounds were equally loud. All baseline rates were loudness matched to the baseline rate of 100 pps, and then each signal was loudness matched to the corresponding baseline.

B. Results

Data for each listener are shown separately in the six panels of Fig. 2(a). Listeners S1–S5 show the “classic” pattern in most conditions, with very good performance at the lowest standard rates, that deteriorates to near-chance levels at the highest rates tested. Exceptions are listener S1, who performs close to ceiling at all rates tested for the 200-ms duration, and listener S3, who for the 200-ms duration shows the nonmonotonic pattern recently described by Kong *et al.* (2009). Listener S6 shows an unusual pattern, in all three conditions, with below-chance performance at intermediate rates. Note that this occurred despite the fact that correct-answer feedback was provided after every trial. This result is consistent with a decrease in pitch with increasing pulse rate over some range of rates, and we will return to this finding in Sec. IV.

Data averaged across listeners are plotted in Fig. 2(b), which shows that mean performance declines monotonically with increasing standard rate. Although performance appears to be slightly better at high rates for the 200-ms duration, a two-way (rate \times condition) repeated-measures analysis of variance (ANOVA) found no interaction between rate and condition [$F(8,5)=1.6$, $p=0.2$], and the main effect of condition just failed to reach significance [$F(2,10)=3.7$, $p=0.06$].

The rationale for experiment 1 was that the alternating-ECAP pattern should be reduced by turning the pulse trains on gradually and that the reduction in the alternation observed later in the pulse train in animal experiments might reduce the influence of the alternation on perception, by allowing the listener to focus on the later portions of the stimulus. However, neither of these two manipulations helped. Overall, performance at high rates was, if anything, slightly better for the shorter (200-ms) stimulus, and the addition of 300-ms ramps produced only a small and nonsignificant improvement (compare upright and inverted triangles in Fig. 2).

III. EFFECT OF HIGH-RATE “CONDITIONING” PULSES

A. Rationale

As noted in the Introduction, it has been suggested that the deterministic nature of the AN response to electrical stimulation may impair the encoding of temporal information. Two approaches—the addition of noise to the signal and the addition of high-rate background pulses—have been proposed (Morse and Evans, 1996; Rubinstein *et al.*, 1999; Zeng *et al.*, 2000; Chatterjee and Robert, 2001). In both cases, it has been suggested that there is an optimal level of the added stimulus that increases the stochasticity of the response to the signal without “swamping” it. Experiment 2 was motivated by suggestion of Rubinstein *et al.* (1999) that a more stochastic pattern of firing could be introduced by the addition of high-rate background pulses, and by their demonstration that the alternating-amplitude ECAP pattern in response to a 1016-pps pulse train could be eliminated by the addition of a 5081-pps conditioner. Note that, although the high-rate pulse trains used to study the effects of conditioners on threshold and on neural responses are typically presented continuously (Runge-Samuelson *et al.*, 2004; Hong and Rubinstein, 2006), this is not crucial for an effect on the ECAP amplitude alternation to be observed; the conditioner used by Rubinstein *et al.* (1999) was turned on only 29 ms before the 1016-pps pulse train. Experiment 2 used the same “mixed-block” procedure as in experiment 1, with the following differences.

Four users of the Nucleus CI24M implant and two users of the MedEl Pulsar implant (M1 and M2, see Table I) took part. In both cases the stimulus duration was 300 ms. For the CI24M users, rate discrimination was measured for baseline

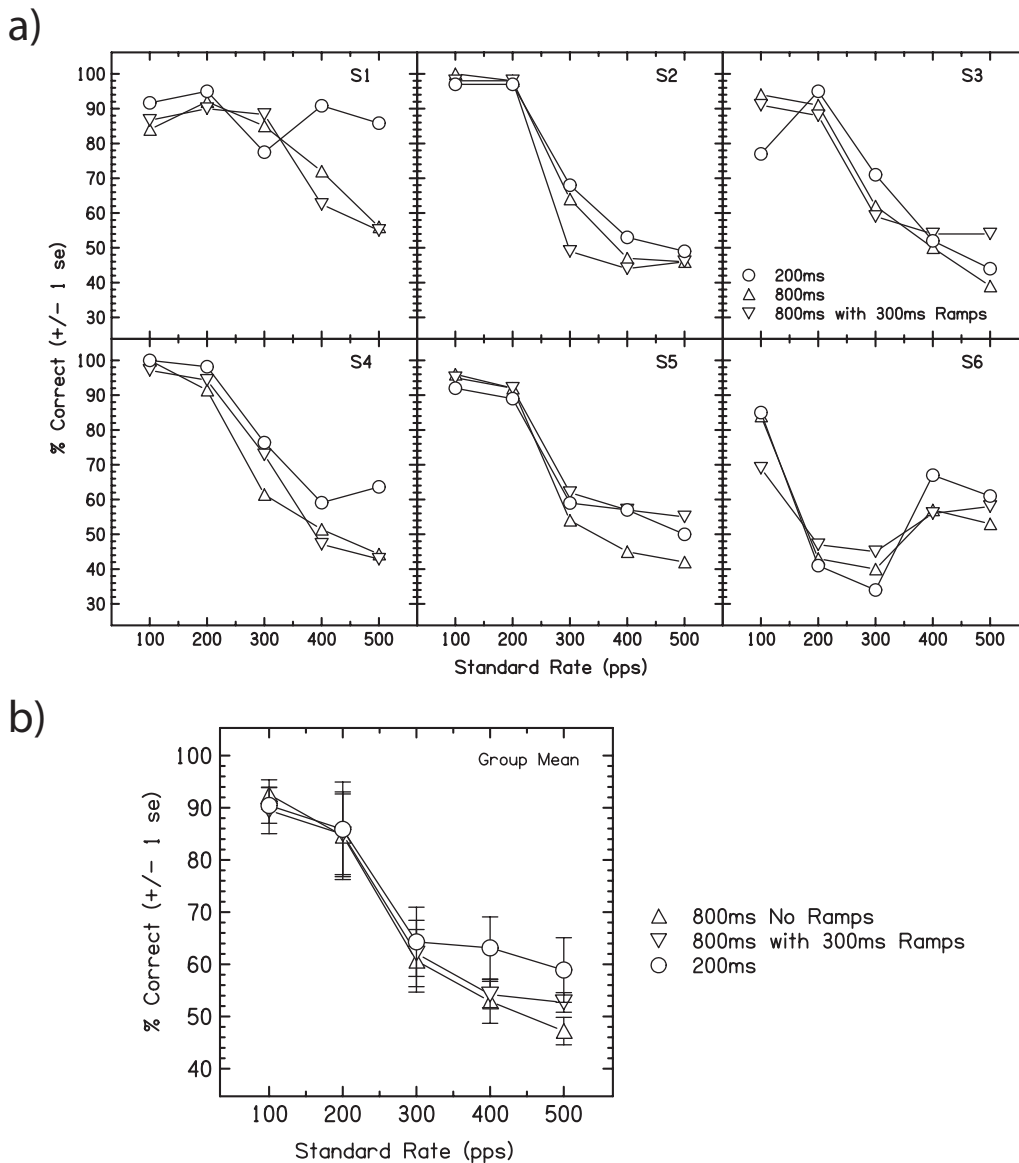


FIG. 2. (a) shows rate discrimination, obtained using the “mixed rate” procedure, for the six subjects of experiment 2. Data are shown for signal durations of 200 and 800 ms, and, in the latter case, with and without 300-ms onset and offset ramps. (b) shows the data averaged across listeners, with standard errors indicated by error bars.

rates of 100–500 pps in the absence of a conditioner and in the presence of a 5000-pps conditioner that, in different conditions, could have one of a range of different levels. These levels were based on a preliminary experiment that measured the threshold and comfort (“C”) level for a 300-ms 5000-pps pulse train. The conditioner levels were then chosen to be at -20% , 0% , 20% , 40% , 60% , and 80% of the dynamic range, where 0% corresponds to threshold and 100% to C level. These percentages were calculated in terms of Cochlear Corporation’s CUs, which are a roughly logarithmic function of current. The conditioners were turned on 300 ms before and off 200 ms after, each standard and signal pulse train. Inter-pulse intervals were filled with an integer number of conditioner pulses at the conditioner rate. The different conditioner levels were run in separate blocks of 10 trials per point, in a counterbalanced order, until 100 trials per point had been collected for each subject and condition. Pulse duration was $45 \mu\text{s}/\text{phase}$ and the interphase duration was $8 \mu\text{s}$. Stimuli were checked as in experiment 1.

Stimulation of the MedEl implant was implemented using the “RIB2” software and hardware produced by the University of Innsbruck, and checked by connecting the output of a “detector box” whose internal electronics are the same as for a Pulsar implant, to a digital storage oscilloscope. The experimental method was similar to that for the CI24M users, except that dynamic range was defined in terms of MedEl’s (approximately linear) current units, and that the conditioner pulse trains were turned on 4 min before each block of trials.² For these listeners, within each session two blocks of ten trials/point were run for each conditioner level before moving to the next-highest level, to avoid any residual effects from the conditioner in one block of trials on the neural responses during the next block. In addition, one listener (M1) was subsequently retested about 1 month later using conditioners turned on and off before each stimulus, as for the CI24M users. Pulse duration was $43 \mu\text{s}/\text{phase}$ with no interphase gap. For both groups of listener, the stimuli were loudness balanced in the absence of a conditioner using

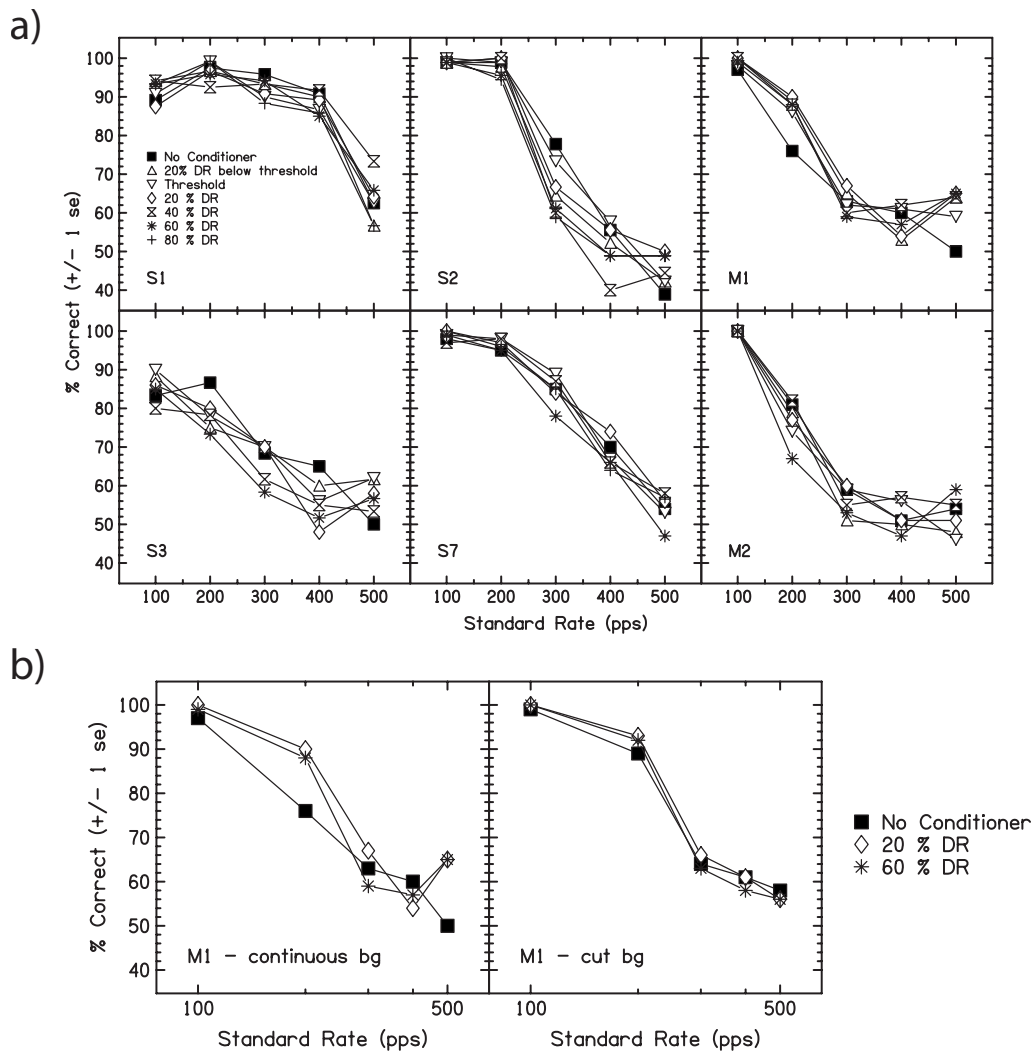


FIG. 3. (a) shows rate discrimination data from experiment 2, for conditions without a conditioner (solid squares) and for conditioners of various levels (open symbols and asterisks; color online). (b) shows the original data for listener M2 obtained with a continuous conditioner in the left-hand panel, together with, in the right-hand panel, a retest in which the conditioner was turned on and off before each stimulus.

the same method as for experiment 1. These loudness balanced levels were used regardless of the presence or level of the conditioner.

B. Results

Results from the six listeners who took part in experiment 2 are shown in Fig. 3(a). Performance in the absence of a conditioner, shown by the filled squares, was universally good at the lowest rate tested (100 pps) and declined monotonically at higher standard rates. Generally speaking, the conditioner pulses had no substantial effect on performance. Minor exceptions were listeners S2 and S3, for whom most conditioners impaired performance, and listener M1, for whom all levels of conditioner produced slightly better performance at 200 and 500 pps than when no conditioner was present. To test whether this small improvement depended on the conditioner pulses being presented continuously, this listener was retested several months later with conditioners that were turned on 300 ms before and off 200 ms after, each standard and signal pulse train. The retest was performed with no conditioner and with conditioner levels correspond-

ing to 20% and 60% of the dynamic range. The results are shown in the right-hand panel of Fig. 3(b), with the original data for the same conditions replotted in the left-hand panel. The main difference between the two plots is that performance at 200 and 500 pps, in the absence of the conditioner, was better in the second set of measures. Combined with the fact that performance for all levels of the continuous conditioner was similar, the most parsimonious explanation for the apparent effect of the continuous conditioner is that the initial estimate of no-conditioner performance was, for some reason, too low at 200 and 500 pps, and that there was no reliable effect of adding a conditioner.³

IV. EXPERIMENT 3: CONCURRENT STIMULATION OF MULTIPLE ELECTRODES

A. Rationale and method

McKay *et al.* (2005) investigated an informal clinical observation that, when a patient's CI is reprogrammed to use a speech-processing strategy with a faster pulse rate, the patient sometimes reports hearing a change in pitch even though both the original and new pulse rates are much higher

than 300 pps. They considered the possibility that the reported pitch change arose from the change in current level necessitated by the new rate, which may have resulted in “place-of-excitation” differences. To test this idea, they generated nine seven-electrode stimuli, comprising three “spectral profiles” (with current maxima at apical, middle, and basal locations) and three pulse rates of 450, 900, and 1800 pps per electrode. A multidimensional scaling study revealed separate perceptual dimensions for spectral profile and pulse rate. The authors concluded that the increases in pulse rate above 450 pps, when applied to multiple electrodes, produced detectable changes in a perceptual dimension that were not conveyed by place-of-excitation cues. This dimension could, of course, be pitch, but other possibilities exist, including the greater adaptation observed for higher rates of stimulation (Zhang *et al.*, 2007) and changes in other temporal properties such as perceived roughness.

Experiment 3 tested whether increases in pulse rate produce increases in pitch over a wider range with multiple electrodes, compared to single-electrode, stimulation. In the single-electrode condition, 500-ms 25 μ s/phase pulse trains applied to electrode 11 of a CI24M implant in monopolar (“MP1+2”) mode. The interphase gap was 25 μ s. Nine pulse rates, ranging from 112.5 to 1800 pps in half-octave steps, were used. In the “flat” and “peak” spectral profile conditions, the same pulse rates were applied concurrently to electrodes 8–14 inclusive. Electrodes were stimulated in a basal to apical order and the interpulse interval between electrodes corresponded to 1/8 of the stimulus period. Before testing, comfort (C) levels were obtained for each subject, spectral profile, and rate. Initially, subjects indicated the comfort level of each electrode presented in isolation and at a rate of 450 pps. For multiple-electrode configurations, C level was determined by presenting all seven electrodes at their respective relative levels and at a rate of 450 pps, initially at a level well below expected comfort level. For “flat” spectral profiles, each electrode was presented at the C level determined for single-electrode presentation minus a global adjustment (in CUs) compensating for the combined loudness effects of several electrodes. “Peaked” profiles were based on these same individual electrode levels but with attenuations of 25, 18, 5, 0, 5, 18, and 25 CUs applied to electrodes 8–14, respectively. The levels of all electrodes were then varied together until the subject reported that the C level had been reached. Using the C level obtained at 450 pps as a reference, the C level for 112.5, 225, 900, and 1900 pps trains were obtained by the method of adjustment described above.

For each condition, the pitches of the different pulse-rate stimuli were compared in experiment 3a using the “Midpoint comparison” (MPC) procedure described by Long *et al.* (2005). The MPC procedure ranks stimuli along a single perceptual dimension (e.g., pitch) in an optimally efficient manner, by means of paired comparisons. Initially, two stimuli are selected at random and the subject indicates which has the higher pitch. Next, a new stimulus is compared to the higher ranked of these two and, if judged higher, is put in first place; otherwise it is then compared to the lower-ranked stimulus and positioned either second or third. As an ex-

ample, later on in the procedure, a provisional ranking of seven stimuli might be [C I E A F B D]. The new stimulus, G, would first be compared to the middle rank (A), and if judged higher the list is bisected so that it would be compared to B; if judged lower than B, it would then be compared to F. The procedure is continued until a rank is obtained for all electrodes and was repeated 20 times for each condition, allowing us to estimate a mean and standard error for each rank. For each repetition, the various rates were introduced to the procedure in a fresh random order. The procedure has several desirable properties, that it shares with the mixed-block procedure used in experiments 1 and 2: (i) it is not subjected to many of the biases that plague magnitude-estimation (“pitch scaling”) tasks (Poulton, 1979); (ii) there is no single stimulus that is presented repeatedly throughout each block, as is the case with adaptive procedures or the method of constant stimuli; this reduces the likelihood of the subject focusing on some idiosyncratic feature of one stimulus; (iii) pitch varies substantially throughout each run, helping the subject focus on that dimension; and (iv) each run includes some easily discriminable pitches, preventing the subject from becoming discouraged. We chose to use it instead of the mixed-block procedure for three reasons: (i) the absence of feedback makes the procedure more appropriate for revealing pitch reversals, (ii) it provides an overview of how pitch varies over a wide range of closely spaced rates, and (iii) because we were studying a wide range of rates, many of which we expected to produce very similar pitches, the mixed-block procedure, in which only adjacent rates are compared, would result in the subject spending much of the time guessing. This problem was partially alleviated by the midpoint comparison procedure because nonadjacent rates were often compared.

To determine whether the different pulse rates were discriminable, experiment 3b used the mixed-block procedure described in Sec. II, with the important difference that a three-interval two-alternative forced-choice (3I2AFC) trial structure was used. This “odd-man-out” procedure allowed the subject to use any difference between the stimuli to identify the stimulus with the higher rate, which could occur in either the second or third intervals. Both single-electrode and peaked stimuli were used, with standard rates of 112.5, 225, 450, 900, and 1272 pps. Signal rates were always half an octave higher than the corresponding standard, and correct-answer feedback was provided after every trial.

B. Results

The results of the MPC procedure of experiment 3a are shown in Fig. 4(a). For the single-electrode condition (filled triangles), the results are generally consistent with those obtained using the mixed-block procedure of experiments 1 and 2. Listeners S1, S2, S4, and S7 all show an increase in pitch up to some value, above which the functions asymptote. Furthermore, listener S2, whose functions asymptote at a lower rate than the others, also showed discrimination performance that declined at a low rate in experiments 1 and 2. Listeners S1 and S7, whose pitch rankings increased up to rates of about 636 and 318 pps, respectively, also showed good per-

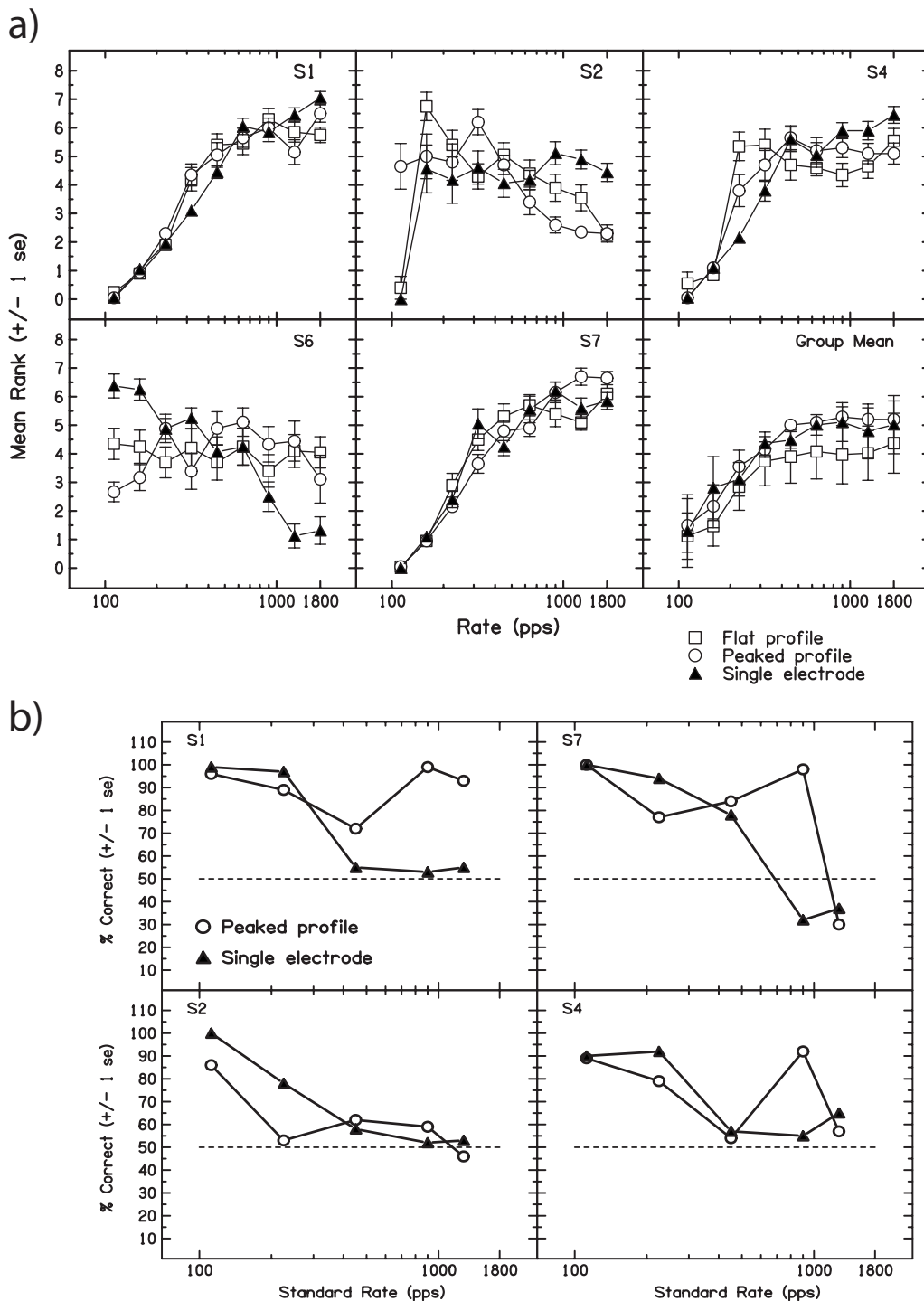


FIG. 4. (a) shows the results of the midpoint-comparison procedure of experiment 3a for single-electrode and multielectrode stimulation. (b) shows the results of the odd-man-out-discrimination task used in experiment 3b.

formance up to relatively high rates in experiment 2 [Fig. 3(a)]. Listener S6 shows an unusual pattern, in that pitch appears to decrease with increasing pulse rate between 159 and 225 pps, and also above 636 pps. This listener also showed an unusual pattern in experiment 1 [Fig. 2(a)], with below-chance performance for standard rates of 200 and 300 pps.

The topic of primary interest concerns the comparison between the functions obtained with single-electrode and multielectrode stimuli. Inspection of Fig. 4(a) shows that in

no case does pitch continue to increase up to a higher rate for the multielectrode (open symbols) stimuli than for the single-electrode (filled triangles) stimulus. The only discernable effects are a lower asymptote for listener S4, flat functions for S6, and a decrease in pitch at higher pulse rates for S2. A two-way repeated-measures ANOVA on the mean ranks for each subject, excluding the anomalous data of S6, revealed the expected main effect of pulse rate [$F(8,24)=11.78$; $p < 0.02$], but no main effect of spectral profile [$F(2,6)=1.0$, n.s.] and no interaction [$F(16,48)=1.22$, n.s.].⁴

The discrimination results from the odd-man-out task of experiment 3b are shown in Fig. 4(b). For three listeners (S1, S7, and S4), there are one or more high rates where discrimination is better with multielectrode (peaked) than with single-electrode stimulation, even though pitch ranking over these range of rates had reached an asymptote in experiment 3a. This may help reconcile the negative findings of experiment 3a with observation of McKay *et al.* (2005) that changes to pulse rate applied to multiple electrodes gave rise to a discriminable perceptual dimension, even though rate changes to a single electrode were not detectable in a forced-choice task. That is, adding electrodes may allow a rate change to become discriminable, without necessarily providing a consistent change in pitch. One way in which this could happen comes from finding of Zhang *et al.* (2007) that adaptation in AN fibers depends on pulse rate: with interleaved stimulation on different electrodes the “effective” pulse rate for any given neuron is increased, and this could result in more adaptation. Alternatively, multielectrode stimulation could influence the temporal pattern of firing, e.g., by abolishing the alternating-amplitude ECAP pattern, and this difference could be detectable without improving pitch perception. Another possible explanation is that subjects used small loudness differences in the forced-choice procedure, which did not lead to differences in pitch, and that the salience of these differences depended on the number of electrodes being stimulated.

V. DISCUSSION

A. Neural basis of temporal pitch perception by CI users

The results presented here show that a number of stimulus manipulations, that one would reasonably expect to have a substantial impact on the AN response to a pulse train, do not produce a marked improvement in rate discrimination at high rates. A related finding was obtained by Kong *et al.* (2009), who compared rate discrimination of regular pulse trains with the ability to discriminate changes in modulation rate imposed on a 5000-pps carrier. They found that, in both conditions, the variation in performance as a function of the standard (modulation) rate differed markedly across listeners, but that, for a given listener and electrode, the pattern of results was similar for the two stimuli. Similarly, in experiment 2 of the present study, the upper limit differed across listeners in a way that was immune to the addition of conditioning pulses of a range of levels. Overall, the results lead to the conclusion that the upper limit of temporal pitch is not specific to any particular temporal pattern of AN activity. However, this does not, of course, mean that features of the auditory nerve response can never affect pitch perception. Indeed, we have argued elsewhere that, at low rates, neural refractoriness does affect temporal pitch (Carlyon *et al.*, 2002; van Wieringen *et al.*, 2003; Carlyon *et al.*, 2008b), and the presence of some pitch reversals in experiment 3 could result from an alternating-amplitude ECAP without this particular pattern being the reason why temporal pitch generally asymptotes at high rates.

Although the present data argue against an explanation

of the upper limit in terms of peripheral phenomena such as reflected by the alternating-amplitude ECAP pattern, data from other studies indicate that the limitations to high-rate processing are not specific to pitch judgments. For bilateral CI users, discrimination of interaural time differences between pulse trains presented to each ear also deteriorates with increases in pulse rate (van Hoesel and Clark, 1997; van Hoesel and Tyler, 2003; van Hoesel, 2007), and, even when a rate difference between matched electrodes in the two ears provides a usable binaural cue at low rates, this is not the case when the pulse rate is high enough for monaural rate discrimination to break down [Carlyon *et al.*, 2008a: see Fig. 1(c) of the present article]. Hence we can tentatively conclude that the upper limitation arises at a site of processing that is central to the AN but common to binaural and pitch processes. One possibility is suggested by tantalizing observation of Snyder *et al.* (1995) that cells in the inferior colliculus (IC) differ in the highest rate to which they will phase lock, but that this “upper limit” is similar, for a given cell, regardless of whether the phase locking is to the pulse rate of a simple pulse train or to the modulation rate imposed on a high-rate carrier. Hence a stimulus-independent limitation may occur at or before the IC, although we should of course stress that this suggestion remains speculative.

The data presented here and elsewhere are also consistent with the idea, mentioned in the Introduction, that the electrical stimulus would need to be modified in order for more central processes to effectively process the temporal activity present in the AN. As long ago as 1982, Moore (1982) suggested that the range of interpulse intervals that is effectively processed by the brain is linked to the characteristic frequencies of the AN fibers that convey that information, and this idea has been more recently implemented in a quantitative model (Bernstein and Oxenham, 2005). Interestingly, although this idea suggests that high-rate information should be more effectively conveyed by stimulating the base of the cochlea, recent recordings from the IC suggest that, conversely, stimulation of AN fibers that innervate the apex of the cochlea may better convey temporal information (Middlebrooks and Snyder, 2009). Another suggestion is that the brain may exploit the phase differences in the response of AN fibers having different CFs, which arise from the slowing of the traveling wave on the BM (Shamma, 1985; Moore and Carlyon, 2005; Cedolin and Delgutte, 2007). Clearly, the stimuli used here and in other studies have neither reproduced these timing differences nor produced either an exact “place-rate” match or stimulation of apical AN fibers. Hence the present results should not be interpreted as meaning that *no* stimulus manipulation can improve temporal processing of high-rate electrical stimuli. It is also true that none of our listeners fell into the “star” category who can reliably discriminate rates higher than 500 pps (Hochmair-Desoyer *et al.*, 1983; Wilson *et al.*, 1997b; Kong *et al.*, 2009). It is therefore possible that either our method of stimulation or the neural status of our particular listeners imposed some additional limitation on performance that masked any potential improvement in the peripheral neural code. This issue is currently under active investigation in our laboratory (Macherey and Carlyon, 2009). What is clear, however, is that for

typical CI users, and in the absence of special types of stimulation that are yet to be implemented, manipulations that we would expect to substantially change the temporal pattern of auditory nerve activity do not have a marked or consistent effect on temporal pitch perception at high rates.

ACKNOWLEDGMENT

We thank Antje Ihlefeld for programming the RIB2 system used in experiment 3.

¹As discussed by Kong *et al.* (2009), this procedure has several strengths: (i) unlike an adaptive procedure, it is possible to measure performance in conditions where it is likely to be close to chance, (ii) the variation in the standard from trial to trial helps keep the listener's attention focused on the pitch dimension and encourages him/her to use that perceptual dimension for all judgments, (iii) the presence of some "easy" (low-rate) standards in each block prevents discouragement, and (iv) the use of a two-interval, rather than odd-man-out procedure, makes it less likely that the listener will use some extraneous cue such as loudness or roughness differences. A weakness is that performance at some rates may be obscured by ceiling effects. However, this was considered not to be critical for the present study, whose purpose was to study performance at rates that, in some conditions, were expected to be close to chance.

²For listener M1, all conditioners having a level above threshold (0% of dynamic range) were still audible at the end of this 4-min period, and no marked reduction in loudness was observed. Listener M2 described conditioner levels of 20% and 40% as just audible at the start and inaudible at the end of the 4-min adaptation period, with some reduction in loudness for the 60% level.

³The conclusion that conditioner pulses did not improve rate discrimination was supported by the results of a two-way repeated-measures ANOVA. The use of this statistic is complicated by the facts that the "optimal" conditioner listener may vary across listeners (Hong and Rubinstein, 2006) and that different conditioner levels were used for the Cochlear and MedEl listeners. We therefore used only two levels of the "conditioner" variable: no conditioner at all, and the level that, for each listener, gave the highest percent-correct score averaged across all rates. Such a method runs the risk of obtaining a significant effect of conditioner by chance because we are selecting the conditioner level producing the best performance from a set of six. Nevertheless, the ANOVA, while producing a main effect of rate [$F(4,20)=15.575$; $p<0.001$], produced no effect of conditioner [$F(1,5)=2.96$, n.s.] and no significant interaction [$F(4,20)=1.86$, n.s.].

⁴When all subjects were included, there was still no effect of profile and no interaction, but the main effect of rate just failed to reach significance.

- Bernstein, J. G. W., and Oxenham, A. J. (2005). "An autocorrelation model with place dependence to account for the effect of harmonic number on fundamental frequency discrimination," *J. Acoust. Soc. Am.* **117**, 3816–3831.
- Bernstein, L. R., and Trahiotis, C. (2002). "Enhancing sensitivity to interaural delays at high frequencies by using "transposed stimuli"," *J. Acoust. Soc. Am.* **112**, 1026–1036.
- Carlyon, R. P., Long, C. J., and Deeks, J. M. (2008a). "Pulse-rate discrimination by cochlear-implant and normal-hearing listeners with and without binaural cues," *J. Acoust. Soc. Am.* **123**, 2276–2286.
- Carlyon, R. P., Mahendran, S., Deeks, J. M., Long, C. J., Axon, P., Baguley, D., Bleeck, S., and Winter, I. M. (2008b). "Behavioral and physiological correlates of temporal pitch perception in electric and acoustic hearing," *J. Acoust. Soc. Am.* **123**, 973–985.
- Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M., and Wouters, J. (2002). "Temporal pitch mechanisms in acoustic and electric hearing," *J. Acoust. Soc. Am.* **112**, 621–633.
- Cedolin, L., and Delgutte, B. (2007). "Spatio-temporal representation of the pitch of complex tones in the auditory nerve," in *Hearing—From Sensory Processing to Perception*, edited by B. Kollmeier, G. Klump, V. Hohmann, U. Langemann, M. Mauermann, S. Uppenkamp, and J. Verhey (Springer, Berlin), pp. 61–70.
- Chatterjee, M., and Robert, M. E. (2001). "Noise enhances modulation sensitivity in cochlear implant listeners: Stochastic resonance in a prosthetic sensory system?" *J. Assoc. Res. Otolaryngol.* **2**, 159–171.
- Dreyer, A., and Delgutte, B. (2006). "Phase locking of auditory-nerve fibers

- to the envelopes of high-frequency sounds: Implications for sound localization," *J. Neurophysiol.* **96**, 2327–2341.
- Hartmann, W. M., McAdams, S., and Smith, B. K. (1990). "Hearing a mistuned harmonic in an otherwise periodic complex tone," *J. Acoust. Soc. Am.* **88**, 1712–1724.
- Hochmair-Desoyer, I. J., Hochmair, E. S., Burian, K., and Stiglbanner, H. K. (1983). "Percepts from the Vienna cochlear prosthesis," *Ann. N.Y. Acad. Sci.* **405**, 295–306.
- Hong, R. S., and Rubinstein, J. T. (2006). "Conditioning pulse trains in cochlear implants: Effects on loudness growth," *Otol. Neurotol.* **27**, 50–56.
- Kong, Y.-Y., Deeks, J. M., Axon, P. R., and Carlyon, R. P. (2009). "Limits of temporal pitch in cochlear implants," *J. Acoust. Soc. Am.* **125**, 1649–1657.
- Landsberger, D. M., and McKay, C. M. (2005). "Perceptual difference between low and high rates of stimulation on single electrodes for cochlear implantees," *J. Acoust. Soc. Am.* **117**, 319–327.
- Laneau, J., Boets, B., Moonen, M., van Wieringen, A., and Wouters, J. (2005). "A flexible auditory research platform using acoustic or electric stimuli for adults and young children," *J. Neurosci. Methods* **142**, 131–136.
- Long, C. J., Nimmo-Smith, I., Baguley, D. M., O'Driscoll, M., Ramsden, R., Otto, S. R., Axon, P. R., and Carlyon, R. P. (2005). "Optimizing the clinical fit of auditory brain stem implants," *Ear Hear.* **26**, 251–262.
- Macherey, O., and Carlyon, R. P. (2009). "Effect of intracochlear stimulation site on the upper limit of temporal pitch," in Conference on Implantable Auditory Prostheses, Lake Tahoe, CA.
- McDermott, H. J., and McKay, C. M. (1997). "Musical pitch perception with electrical stimulation of the cochlea," *J. Acoust. Soc. Am.* **101**, 1622–1631.
- McKay, C. M., Henshall, K. R., and Hull, A. E. (2005). "The effect of rate of stimulation on perception of spectral shape by cochlear implantees," *J. Acoust. Soc. Am.* **118**, 386–392.
- McKay, C. M., McDermott, H. J., and Carlyon, R. P. (2000). "Place and temporal cues in pitch perception: Are they truly independent?," *ARLO* **1**, 25–30.
- Micheyl, C., Moore, B. C. J., and Carlyon, R. P. (1998). "The role of excitation-pattern cues and temporal cues in the frequency and modulation-rate discrimination of amplitude-modulated tones," *J. Acoust. Soc. Am.* **104**, 1039–1050.
- Middlebrooks, J., and Snyder, R. (2009). "Enhanced transmission of temporal fine structure using penetrating auditory nerve electrodes," in Association for Research in Otolaryngology, 32nd Midwinter Research Meeting, Baltimore, MD.
- Moore, B. C. J. (1973). "Frequency difference limens for short-duration tones," *J. Acoust. Soc. Am.* **54**, 610–619.
- Moore, B. C. J. (1982). *An Introduction to the Psychology of Hearing*, 2nd ed. (Academic, London).
- Moore, B. C. J., and Carlyon, R. P. (2005). "Perception of pitch by people with cochlear hearing loss and by cochlear implant users," in *Springer Handbook of Auditory Research: Pitch Perception*, edited by C. J. Plack and A. J. Oxenham (Springer, New York), pp. 234–277.
- Morse, R. P., and Evans, E. F. (1996). "Enhancement of vowel coding for cochlear implants by addition of noise," *Nat. Med.* **2**, 928–932.
- Nie, K. B., Stickney, G., and Zeng, F. G. (2005). "Encoding frequency modulation to improve cochlear implant performance in noise," *IEEE Trans. Biomed. Eng.* **52**, 64–73.
- Oxenham, A. J., Bernstein, J. G. W., and Penagos, H. (2004). "Correct tonotopic representation is necessary for complex pitch perception," *Proc. Natl. Acad. Sci. U.S.A.* **101**, 1421–1425.
- Pijl, S., and Schwarz, D. W. F. (1995). "Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes," *J. Acoust. Soc. Am.* **98**, 886–895.
- Poulton, E. C. (1979). "Models for the biases in judging sensory magnitude," *Psychol. Bull.* **86**, 777–803.
- Riss, D., Arnoldner, C., Baumgartner, W. D., Kaider, A., and Hamzavi, J. S. (2008). "A new fine structure speech coding strategy: Speech perception at a reduced number of channels," *Otol. Neurotol.* **29**, 784–788.
- Rubinstein, J. T., Wilson, B. S., Finley, C. C., and Abbas, P. J. (1999). "Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation," *Hear. Res.* **127**, 108–118.
- Runge-Samuels, C. L., Abbas, P. J., Rubinstein, J. T., Miller, C. A., and Robinson, B. K. (2004). "Response of the auditory nerve to sinusoidal electrical stimulation: Effects of high-rate pulse trains," *Hear. Res.* **194**,

- Shamma, S. (1985). "Speech processing in the auditory system: II. Lateral inhibition and the central processing of speech evoked activity in the auditory nerve," *J. Acoust. Soc. Am.* **78**, 1622–1632.
- Shannon, R. V. (1983). "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics," *Hear. Res.* **11**, 157–189.
- Shepherd, R. K., and Javel, E. (1997). "Electric stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status," *Hear. Res.* **108**, 112–144.
- Snyder, R., Leake, P. A., Rebscher, S. J., and Beitel, R. (1995). "Temporal resolution of neurons in cat inferior colliculus to intracochlear electrical stimulation: Effects of neonatal deafening and chronic stimulation," *J. Neurophysiol.* **73**, 449–466.
- Stickney, G. S., Nie, K. B., and Zeng, F. G. (2005). "Contribution of frequency modulation to speech recognition in noise," *J. Acoust. Soc. Am.* **118**, 2412–2420.
- Townshend, B., Cotter, N., van Compernelle, D., and White, R. L. (1987). "Pitch perception by cochlear implant subjects," *J. Acoust. Soc. Am.* **82**, 106–115.
- van Hoesel, R. J. M. (2007). "Sensitivity to timing in bilateral cochlear implant users," *J. Acoust. Soc. Am.* **121**, 2192–2206.
- van Hoesel, R. J. M., and Clark, G. M. (1997). "Psychophysical studies with two binaural cochlear implant subjects," *J. Acoust. Soc. Am.* **102**, 495–507.
- van Hoesel, R. J. M., and Tyler, R. S. (2003). "Speech perception, localization, and lateralization with bilateral cochlear implants," *J. Acoust. Soc. Am.* **113**, 1617–1630.
- van Wieringen, A., Carlyon, R. P., Long, C. J., and Wouters, J. (2003). "Pitch of amplitude-modulated irregular-rate stimuli in electric and acoustic hearing," *J. Acoust. Soc. Am.* **114**, 1516–1528.
- Wilson, B., Finley, C., Lawson, D., and Zerbi, M. (1997a). "Temporal representations with cochlear implants," *Am. J. Otol.* **18**, s30–s34.
- Wilson, B., Zerbi, M., Finley, C., Lawson, D., and Honert, C. V. D. (1997b). "Speech processors for auditory prostheses (Eighth Quarterly Progress Report)," NIH Project No. N01-DC-5-2103.
- Zeng, F.-G. (2002). "Temporal pitch in electric hearing," *Hear. Res.* **174**, 101–106.
- Zeng, F. G., Fu, Q. J., and Morse, R. (2000). "Human hearing enhanced by noise," *Brain Res.* **869**, 251–255.
- Zhang, F., Miller, C. A., Robinson, B. K., Abbas, P. J., and Hu, N. (2007). "Changes across time in spike rate and spike amplitude of auditory nerve fibers stimulated by electric pulse trains," *J. Assoc. Res. Otolaryngol.* **8**, 356–372.