Combined neural and behavioural measures of temporal pitch perception in cochlear implant users

Robert P. Carlyon and John M. Deeks

Citation: The Journal of the Acoustical Society of America **138**, 2885 (2015); doi: 10.1121/1.4934275
View online: https://doi.org/10.1121/1.4934275
View Table of Contents: https://asa.scitation.org/toc/jas/138/5
Published by the Acoustical Society of America

Combined neural and behavioural measures of temporal pitch perception in cochlear implant users[a)]

Robert P. Carlyon[b)] and John M. Deeks
MRC Cognition & Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, England

(Received 6 May 2015; revised 15 September 2015; accepted 5 October 2015; published online 9 November 2015)

Four experiments measured the perceptual and neural correlates of the temporal pattern of electrical stimulation applied to one cochlear-implant (CI) electrode, for several subjects. Neural effects were estimated from the electrically evoked compound action potential (ECAP) to each pulse. Experiment 1 attenuated every second pulse of a 200-pps pulse train. Increasing attenuation caused pitch to drop and the ECAP to become amplitude modulated, thereby providing an estimate of the relationship between neural modulation and pitch. Experiment 2 showed that the pitch of a 200-pps pulse train can be reduced by delaying every second pulse, so that the inter-pulse-intervals alternate between longer and shorter intervals. This caused the ECAP to become amplitude modulated, but not by enough to account for the change in pitch. Experiment 3 replicated the finding that rate discrimination deteriorates with increases in baseline rate. This was accompanied by an increase in ECAP modulation, but by an amount that produced only a small effect on pitch in experiment 1. Experiment 4 showed that preceding a pulse train with a carefully selected "pre-pulse" could reduce ECAP modulation, but did not improve rate discrimination. Implications for theories of pitch and for limitations of pitch perception in CI users are discussed.
 [http://dx.doi.org/10.1121/1.4934275]

[ELP] Pages: 2885–2905

I. INTRODUCTION

Modern cochlear implants (CIs) convey the fundamental frequency (F0) of complex sounds by the temporal pattern of electrical stimulation. For example, in the continuous interleaved sampling (CIS) strategy (Wilson *et al.*, 1991), the envelope fluctuation in each frequency channel is superimposed upon a fixed high-rate pulse train that is presented to one electrode. For a harmonic sound, such as a vowel, the envelope repeats at a rate equal to F0, which is therefore reflected in the electrical waveform. Another type of strategy is based on the so-called "n of m" approach, such as the ACE and SPEAK strategies (McDermott *et al.*, 1992) implemented in devices manufactured by the Cochlear corporation. In this case, the pulse rate on each electrode is also fixed, and is also modulated by the envelope in a given frequency band, but only a subset ("n") of the total number of electrodes ("m") is selected for stimulation during each time frame. Unfortunately, pitch perception by CI users is typically poor (Moore and Carlyon, 2005). A number of modifications and alternatives to these strategies have been implemented, including the enhancement and alignment of the F0-related modulations across channels (Vandali *et al.*, 2005; Laneau *et al.*, 2006), and the "fine structure" coding strategy, in which short bursts of pulses are presented at the zero-crossings of the three or four lowest-frequency bandpass filters (Riss *et al.*, 2008). An important constraint on the potential success of such algorithms is that, even when the speech processor is bypassed and the listener presented with simple pulse trains, perception by CI users typically falls far short of that obtained for most periodic sounds by normal hearing (NH) listeners (Shannon, 1983; Tong *et al.*, 1983; Townshend *et al.*, 1987; Kong *et al.*, 2009). Hence there are some basic limitations inherent in the ability of the auditory system to process temporal patterns of electrical stimulation, at least for deafened individuals presented with the fairly simple stimuli used by experimenters to date. It is therefore of some interest to explore the physiological basis for these limitations. The present article does so by comparing measures of the auditory nerve (AN) response with behavioural data obtained with the same subjects and almost identical stimuli.

One well-known and important finding concerns the upper limit of temporal pitch, which has typically been studied by stimulating a single electrode with isochronous pulse trains, in which the intervals between all adjacent pulses are equal. When an isochronous 100-pps pulse train is presented to one electrode of a CI, increasing the pulse rate causes an increase in perceived pitch. However, the majority of CI users are unable to detect increases in pulse rate beyond about 300 pps, although there is substantial across-user variability, with a minority of users being able to detect increases up to at least 700 pps (Hochmair-Desoyer *et al.*, 1983; Kong and Carlyon, 2009).

[a)]Portions of the work described here were presented at the *2012 International Symposium on Hearing* and were reported in summary form by Carlyon and Deeks [Carlyon, R. P., and Deeks, J. M. (**2013**). "Relationships between auditory nerve activity and temporal pitch perception in cochlear implant users," in *Basic Aspects of Hearing: Physiology and Perception*, edited by B. C. J. Moore, R. D. Patterson, I. M. Winter, R. P. Carlyon, and H. E. Gockel (Springer, New York), pp. 363–372].

[b)]Electronic mail: bob.carlyon@mrc-cbu.cam.ac.uk

Despite the obvious importance of the upper limit for pitch perception in CIs, its neural basis remains elusive. Several findings argue for a peripherally based limitation. For a given subject, the upper limit can differ between electrodes, and the limit is lower for those sites of stimulation that require the least current to achieve a comfortable loudness (Kong *et al.*, 2009; Macherey *et al.*, 2011). This latter finding has been interpreted in terms of variations in local neural survival. Evidence for a peripheral limitation at high rates also comes from experiments investigating whether, in patients with two CIs, discrimination between a lower and a higher-pulse rate, presented to one ear, could be improved by presenting a copy of the lower-rate pulse train to an electrode in the opposite ear, in both intervals of each trial. A benefit occurred at low rates by providing the listener with a binaural cue; the lower-rate trains sounds fused in the center of the head, whereas the higher-rate train sounds diffuse. However, this did not help discrimination when the lower rate was 300 pps, showing that the "upper limit" is not restricted to situations where the listener must derive a pitch (van Hoesel and Clark, 1997; van Hoesel, 2007; Carlyon *et al.*, 2008a). Further evidence comes from experiments with acoustic pulse trains filtered into high frequency regions. These stimuli present normal-hearing (NH) listeners with purely temporal rate information, in the absence of any place-of-excitation cues, and have been successfully used as simulations of electric pulse trains presented to CIs (McKay and Carlyon, 1999; Carlyon *et al.*, 2002; van Wieringen *et al.*, 2003; Carlyon *et al.*, 2008a; Carlyon *et al.*, 2008b; Carlyon *et al.*, 2011). NH listeners can detect differences in pulse rate up to about 700 pps (Carlyon and Deeks, 2002; Macherey and Carlyon, 2014), which is higher than the 300 pps observed for the majority of CI users, and similar to that of the best CI users, consistent with most patients having a limitation arising from damage to, or degradation of, the peripheral auditory system. On the other hand, the upper limit is immune to several manipulations that would be expected to influence the AN response to a pulse train: these include the introduction of slow onset ramps, the addition of continuous low-level "desychronising" pulses (Rubinstein *et al.*, 1999), and the use of modulated high-rate pulse trains instead of low-rate pulse trains having a single pulse per period (Carlyon *et al.*, 2010). In addition, although applying the standard and signal pulse rates to multiple electrodes rather than to single electrodes can improve discrimination in odd-man-out type tasks (Carlyon *et al.*, 2010; Venter and Hanekom, 2014), there is evidence that this improvement does not arise from an improvement in pitch perception (Carlyon *et al.*, 2010). To summarise, the upper limit in rate discrimination for pulse trains is usually but not universally lower in CI than in NH listeners, and is not specific to pitch perception, but has not been associated with any particular pattern of AN activity.

A possible mechanism for the upper limit is illustrated in Figs. 1(a) and 1(b). Measures of the electrically evoked compound action potential (ECAP) to isochronous pulse trains have shown that the response becomes amplitude modulated at high rates, with a larger response to odd-numbered than to even-numbered pulses (Wilson, 1997). This could cause some intervals equal to twice the true period to be conveyed to the more central stages of the auditory system, thereby lowering pitch. As the depth of the modulation increases with pulse rate, it could potentially counteract the increase in pitch that there would otherwise have been.

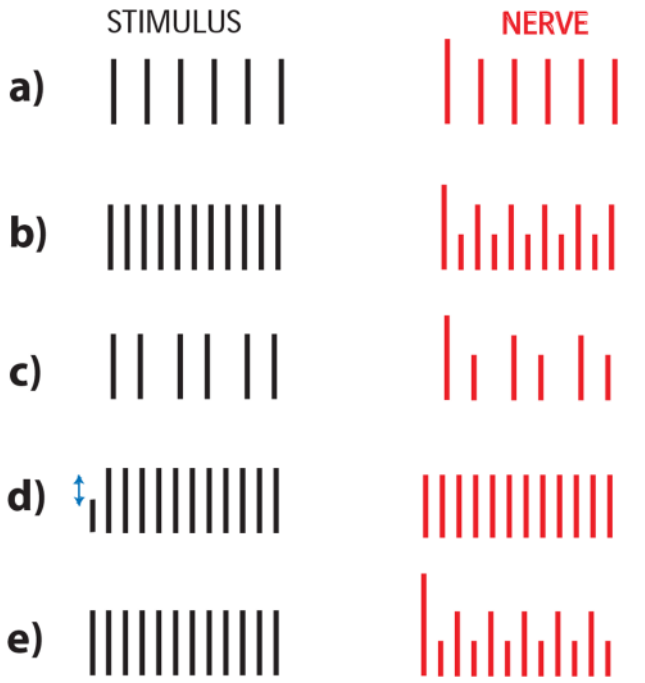

FIG. 1. (Color online) Schematic of pulse trains (left column) and neural response (right column). (a), (b) Unmodulated isochronous pulse trains of a low and a high rate, respectively. (c) Alternating-interval pulse train. (d) High-rate isochronous pulse train with a pre-pulse whose level is adjusted to minimise ECAP modulation. In (e), the pre-pulse is the same amplitude as the others, and so the alternating-amplitude pattern in the neural response returns.

Modulation of the AN response has also been proposed as an explanation for another phenomenon, illustrated in Fig. 1(c). When the inter-pulse-interval (IPI) alternates between two values, such as 4 and 6 ms, the perceived pitch corresponds neither to the true period (10 ms) nor to the mean IPI of the stimulus (5 ms). This finding has been observed both for electrical stimulation of individual CI electrodes, and for bandpass filtered acoustic pulse trains presented to NH listeners (Carlyon *et al.*, 2002; van Wieringen *et al.*, 2003; Carlyon *et al.*, 2008b). For acoustic stimulation, measures of the compound action potential (CAP) show a smaller response for pulses occurring after the shorter, compared to the longer, IPI [Carlyon *et al.*, 2008b; cf. Fig. 1(c)]. This is consistent with refractory effects being stronger after shorter IPIs, and leads to the CAP being modulated. It has been suggested that, as a result of this modulation, some neurons central to the AN may code only the intervals between the larger responses (e.g., 10 ms for a "4–6" pulse train), and that pitch is derived from a weighted sum of these longer intervals and of the shorter intervals (e.g., 4 and 6 ms) corresponding to the IPIs in the stimulus (Carlyon *et al.*, 2008b). Consistent with this interpretation, there is some evidence that the pitch of 4–6 pulse trains is lower when the pulses occurring after the shorter IPI are attenuated than when the same attenuation is applied to the pulses occurring after the

longer IPI (van Wieringen *et al*., 2003). This would occur if the modulation in the neural response were exaggerated in the former case and attenuated in the latter. This finding was significant across the four NH listeners tested by van Wieringen *et al*. (2003), and was significant for three out of their four CI listeners.

The experiments described here investigate the potential role of amplitude modulation in the AN response, as measured by the ECAP, on temporal pitch perception. Experiment 1 compared modulation in the ECAP response and pitch perception for physically amplitude modulated pulse trains. It showed that ECAP modulation, rather than the physical modulation of the stimulus, could predict the variation in pitch perception across subjects. Experiment 2 obtained both of these measures for unmodulated alternating-interval pulse trains, and found that modulation in the ECAP response was too small to account for the pitch of those stimuli. Experiment 3 measured rate discrimination and ECAP modulation as a function of baseline rate, and found that the upper limit of rate discrimination could not be explained by ECAP modulation. This conclusion was confirmed by our final experiment, which introduced a manipulation that greatly reduced modulation in the neural response, but that did not improve rate discrimination performance. We argue that the results constrain explanations for the poor pitch perception experienced by CI users, and discuss the implications of the results for theories of pitch perception.

II. EXPERIMENT 1: RELATING ECAP MODULATION TO PITCH CHANGES

A. Rationale

A problem with interpreting psychophysical data in terms of physiological mechanisms is that the former are usually obtained from awake humans, whereas the most accurate physiological recordings are typically obtained from anaesthetised animals. Here we obtain ECAPs and psychophysical measures from the same (human) CI users and the same stimuli. Another problem remains, however: the dependent variables arising from behavioural and physiological measures are necessarily different, and so a direct comparison between them is not straightforward. To address this, experiment 1 physically attenuated every second pulse of an isochronous pulse train, and measured the effect of this physical modulation on both the perceived pitch and on the ECAP to the pulse train. This provided us with an estimate of the pitch change corresponding to a given amount of modulation in the ECAP to a pulse train. This information was then used in experiments 2 and 3 to evaluate the possible effect of ECAP modulation both on the upper limit of rate discrimination, and on the pitch of alternating-interval pulse trains.

B. Subjects and stimuli

Seven users (S1 to S7) of the Cochlear Corporation CI24RE implant took part. Details of their age, duration of implant use, duration of profound hearing loss, and electrode used in the experiment are shown in Table I, along with those of three other subjects (S8–S10) who participated in experiments 2–4. All stimuli consisted of trains of symmetric biphasic pulses, presented in monopolar ("MP1 + 2") mode to an active electrode near the middle of the array, except S9 who used a more apical electrode [Table I]. Active electrode selection was guided by preliminary measures of ECAP responses, with non-responsive electrodes discarded in favour of more responsive electrodes. Each pulse consisted of two 25-μs phases of opposite polarity, separated by an inter-phase gap of 8 μs. The total duration of all pulse trains was 100 ms. Stimuli were presented via the NIC2 interface provided by Cochlear Corp and by research processors also provided by Cochlear; these were the L34 for the behavioural experiments and the SP12 for the ECAP recordings. The same stimulus levels were used in both parts of the experiment, and were selected using a loudness-balancing procedure, described in Sec. II C.

C. Behavioural experiment

1. Methods

Each trial consisted of a 100-ms 200-pps modulated pulse train (the "standard") and an unmodulated pulse train (the "comparison") presented in random order. For the standard, the amplitude of the even-numbered pulses could be attenuated by either 1, 4, or 8 clinical current units ("CUs"), corresponding to about 0.17, 0.68, and 1.36 dB, respectively; five of the seven subjects were tested only at the two smaller modulation depths. The rate of the unmodulated pulse train was selected from 100, 119, 141, 168, 200, and 238 pps. The subject's task was always to judge which interval contained the stimulus with the higher pitch. No feedback was provided. Each block of trials consisted of the same modulated pulse train (0.17, 0.64, or 1.36 dB depth) paired with each of the unmodulated pulse trains 10 times, leading to a total of 60 trials per block. Between five and eight blocks were run, depending on the subject's time and availability, so each data point was based on between 50 and 80 trials. The psychometric functions (proportion of trials in which the unmodulated sound was judged higher, as a function of its pulse rate) for each subject and modulation depth were then subjected to a probit analysis (SPSS, 2012). The point of

TABLE I. Details of the subjects who took part in the experiments.

Subject	Age in years	Active electrode	Years of implant use	Duration of profound deafness
S1	69.6	11	3.5	20
S2	60.9	11	4.5	5
S3	64.3	11	5.6	3
S4	69.9	17	5.6	5
S5	52.5	11	1.2	> 45
S6	54.8	15	4.5	> 5
S7	67.2	11	4.9	9
S8	70.9	16	0.9	> 10
S9	35.7	20	0.6	> 5
S10	61.2	11	1.3	15

subjective equality ("PSE") was then defined as the pulse rate corresponding to the 50% point on the function.

Stimuli were loudness balanced so as to minimise any effects of loudness on pitch judgements. The results of the loudness balancing was used to set the levels for both behavioural and ECAP measures. Subjects first adjusted the level of all modulated stimuli so that the loudness corresponded to point 6, termed "most comfortable level (MCL)," of an eleven-point scale. A note of the 7 ("loud, but comfortable") level was also made. Each modulated pulse train, at its MCL, was then used as a reference to which an unmodulated 200-pps pulse train was balanced, using a method described by Landsberger and McKay (2005). Briefly, the modulated and unmodulated stimuli were presented in pairs, and subjects adjusted the level of the unmodulated stimulus to match the loudness of the modulated stimulus. The average level obtained from four repeats of this procedure was then used to set the level of the unmodulated stimulus, and the procedure was then repeated but this time varying the level of the modulated stimulus. The adjusted level of the 200 pps unmodulated stimulus was calculated using the average from these two stages. Next, in order to achieve equal loudness of the range of pulse rates of unmodulated pulse trains, the 200 pps stimulus was loudness balanced with both the 237- and 100-pps stimuli in a similar fashion. The levels of the 168-, 141-, and 118-pps stimuli were taken from a linear interpolation (in CUs) of the levels at 200 and 100 pps.

2. Results

Psychometric functions for each modulation depth are shown in Fig. 2. Not surprisingly, as the pulse rate of the unmodulated pulse train increased, so did the proportion of trials on which it was judged to have a higher pitch than the standard. As the modulation depth of the standard increased, these curves shifted upwards and to the left, consistent with a decrease in the pitch of the standard. The PSE for each subject and modulation depth was calculated, converted into a percentage shift re 200 pps, and presented in the first three columns in bold type in Table II. Generally speaking, these shifts were very small for all subjects at a modulation depth of 0.17 dB, and for four subjects (S2, S4, S5, and S7) at a modulation depth of 0.68 dB. For the five subjects tested with a modulation depth of 1.36 dB, three (S2, S5, and S7) show very small pitch shifts, whereas S1 and S4 show shifts of 37 and 38%, respectively.

D. ECAP measurements

1. Methods

ECAPs provide a measure of the compound auditory nerve response to an electrical stimulus, and can be measured by stimulating one intra-cochlear electrode and recording from another intra-cochlear electrode. A number of methods have been proposed to reduce the effect of the electrical artefact; here we use the alternating-polarity method in which the responses to opposite-polarity pulse trains are summed. We also imposed a short "recording delay" between the end of the stimulus and the time at which we started to record the response.

ECAP measurements were obtained for 200-pps pulse trains with modulation depths of 0, 0.16, 0.64, and 1.36 dB, using the same stimulus levels and subjects as in the behavioural experiment. Technical reasons (specifically, the memory of the SP12 device) limit the number of ECAP measurements to ten per stimulus. Therefore, the majority of the measurements were obtained using 10-pulse (50-ms) stimuli. As the standards used in the behavioural experiment consisted of 20 pulses (100-ms duration), we needed to

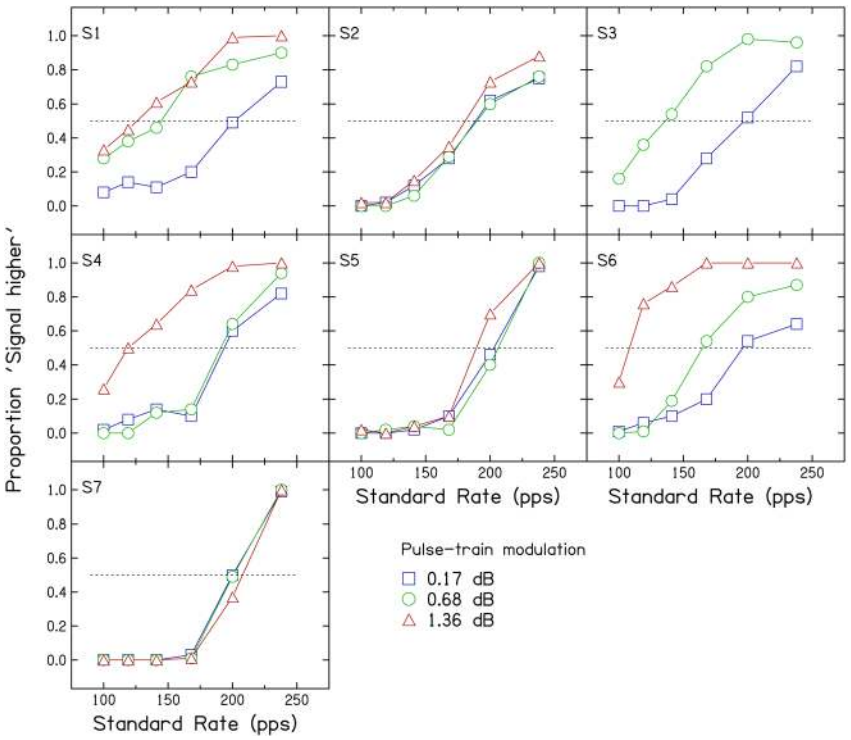

FIG. 2. (Color online) Results of experiment 1. Each panel shows psychometric functions for one subject and for 200-pps standard stimuli amplitude modulated by amounts selected from 0.17, 0.68, and 1.36 dB. The abscissa shows the pulse rate of the isochronous comparison sounds. The horizontal dashed lines show the point at which the comparison is equally likely to be judged as higher or lower in pitch than the standard.

TABLE II. ECAP modulations and pitch shifts for experiment 1 and for the 4-6 and 3-7 stimuli of experiment 2. The last three rows show the means across 5, 7, and 9 listeners, respectively; this is done so as to aid comparisons across conditions or experiments involving different numbers of subjects.

		0.17 dB		0.68 dB		1.36 dB		4-6		3-7	
Subject	0 dB ECAP mod	ECAP mod	Pitch Shift (%)	ECAP mod	Pitch Shift (%)	ECAP mod	Pitch Shift (%)	ECAP mod	Pitch Shift (%)	ECAP mod	Pitch Shift (%)
S1	3.5	13.7	**−2.8**	48.9	**30.6**	74.3	**37**	−0.8	**19.1**	34.5	**27.6**
S2	1.4	7.4	**1.5**	17.3	**0.6**	31.8	**8.3**	10.7	**11.5**	16.2	**33**
S3	12.6	32	**0.5**	56.3	**32.2**	—	—	16	**15.7**	31.5	**39.6**
S4	0.7	15.3	**1.3**	42.8	**4.6**	67.2	**38**	9.2	**22**	21.7	**36.3**
S5	1	3.8	**0.7**	11.2	**−0.3**	20.1	**5.8**	−3.8	**−7**	4.3	**36.4**
S6	0.4	8.3	**−4.7**	26.8	**11.9**	—	—	1.5	**17.8**	5.4	**33.1**
S7	0.2	7	**−0.1**	24.9	**−0.1**	43.8	**−1.9**	5.7	**−6.4**	11.7	38.4
S8	—	—	—	—	—	—	—	22.9	**11.7**	30.3	**29.6**
S9	—	—	—	—	—	—	—	7.3	**18.9**	16.6	**30.4**
Mean (N = 5)	*1.3*	*9.5*	***0.1***	*29*	*7.1*	*47.5*	***17.4***	*4.2*	*7.8*	*17.7*	*34.3*
Mean (N = 7)	*2.8*	*12.5*	***−0.5***	*32.6*	***14***	—	—	*5.5*	***10.4***	*17.9*	***34.9***
Mean (N = 9)	—	—	—	—	—	—	—	7.6	**11.5**	19.1	**33.8**

check that any amplitude modulation of the ECAP response did not change markedly between the first and second 10 pulses of the stimulus. To do this, for a subset of subjects and stimuli, we used 20-pulse stimuli in which ECAPs were obtained to the last 10 pulses. These were then compared to the ECAPs for the corresponding 10-pulse stimulus.

The ECAP for each pulse train was obtained by averaging the responses to 100 pairs of 100-ms stimuli, separated by 205 ms, where the pulses in the first stimulus in each pair were cathodic-leading, and where those in the second were anodic-leading. The interval between pairs was also 205 ms. ECAPs were defined as the amplitude between the first negative and positive peak of the response (N1 and P1, respectively). Amplifier gain and recording delay parameters were set to give optimized recordings for each individual subject. The recording sample rate was 20.0 kHz. The active electrode was the same electrode as used in the behavioural task for each subject, and, as in the behavioural task, current was returned using the ball (MP1) and case (MP2) electrodes in parallel ("MP1 + 2" mode). The recording electrode was two electrodes apical to the active electrode; the recording return electrode was MP2. All responses were lowpass filtered (4th-order Butterworth) at 8.6 kHz. For responses to pulses 2 through to 9, the values of N1 and P1 were identified using an algorithm in which N1 was identified as the minimum sample value over the first 400 μs of the response, and P1 as the maximum sample value from the time point of N1 until a further 800 μs.

To measure the ECAPs to multiple pulses within a pulse train, the SP12 processor issued a command for each pulse both to present that pulse and to record an ECAP. This is more efficient than the method used in most previous studies, which was to measure the ECAP to the final pulse of pulse trains of different lengths (e.g., Hay-McCutcheon *et al.*, 2005). A technical issue arose from the fact that, when the radio frequency (RF) coil of the Cochlear Corp. implant sends a command to transmit a pulse, it also transmits power to the internal part of the device. When the interval between pulses is relatively long, as in the present experiment, it is necessary to transmit "power up" pulses in the gaps between the stimulus pulses. We used power-up pulses with a rate of 5000 pps, a duration of 25 μs/phase, and a nominal level of 1 current unit (about 18 μA, well below detection threshold). However, it is not possible to record an ECAP during the time that any stimulation (including power up pulses) is taking place. Our solution was, for the ECAP experiments, to leave a gap of 2 ms after every stimulus pulse, during which time the ECAP was recorded and no power-up pulses were presented. The power-up pulses resumed immediately after the gap and continued until the next pulse. All 10-pulse sequences were preceded by 1200 power-up pulses. Inspection of the stimulus using a test implant and digital storage oscilloscope confirmed that the amplitude of the pulses was unaffected by this manipulation. The calibration was performed with a piece of Perspex, having a thickness of 8 mm, inserted between the coil and test implant, to simulate the situation of a patient having a thick skin flap. A 10.8-kΩ resistor was placed between the active and return electrodes of the test implant. This method simulates a demanding condition in which the high impedance and large distance between the transmitter and receiver lead to high voltage requirements. It was used for all ECAP measurements reported here. The stimuli in the behavioural experiments, which did not have these power-up free periods, were also checked in a similar way.

To obtain a summary measure of the depth of any ECAP modulation, we analysed the ECAP amplitude to the 2nd to 9th pulse of each pulse train, removed linear trends using the MATLAB function detrend.m and then divided the mean ECAP to pulses (3, 5, 7, 9) by that to pulses (2, 4, 6, 8). Linear trends were removed because we would not expect pitch to be affected by a monotonic change in the size of the neural response as a function of pulse rate. (Where we describe the analysis of the responses to the second half of 20-pulse stimuli, the pulse numbers should each be increased by 10.) For most plots and analyses, this ECAP modulation ratio r was converted to a percentage modulation depth using the formula $100 - (100/r)$.

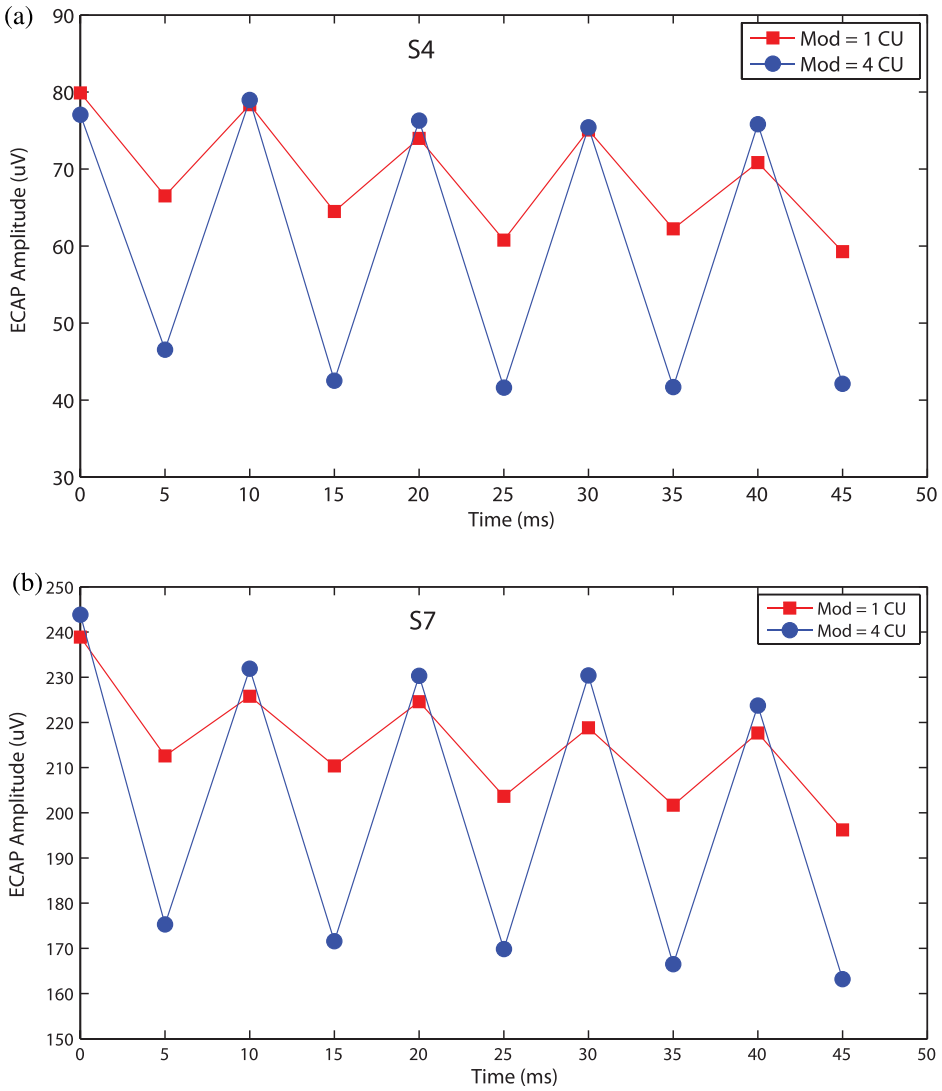

FIG. 3. (Color online) Each panel shows ECAP amplitude as a function of pulse position for two modulation depths applied to a 200-pps stimulus. Data from two subjects are shown.

2. Results

Figures 3(a) and 3(b) shows ECAP amplitudes for 200-pps pulse trains at two stimulus modulation depths, for two representative subjects (S4 and S7). As expected, there is a large ECAP to the first pulse followed by a modulated pattern, and the depth of the modulation on this neural response increases with that of the stimulus. This ECAP modulation depth, measured over the 2nd to 9th pulse in each pulse train, is shown for each subject and stimulus by the first four columns in faint type in Table II. The first of these columns shows the ECAP modulation to the physically unmodulated 200-pps pulse train. At this modulation rate the ECAP modulation depth is close to zero for all subjects except S3. All subjects showed the same general pattern as in Fig. 2, with greater ECAP modulation for stimuli having a larger modulation depth.

ECAPs were also obtained for three subjects for the 12th to 19th pulses of 100-ms 200-pps pulse trains. These measurements were obtained in different testing sessions from the measurements summarised for the 2nd to 9th pulses shown in Table II, and so the measures for the 2nd to 9th pulses were repeated. This resulted in very similar responses for S6 and S7, but the unexpectedly large modulation noted above for subject S3 had dropped to 2.8% on re-testing. Table III shows that the modulation in the ECAP was very similar for the 2nd to 9th and 12th to 19th pulses. A univariate analysis of variance (ANOVA), with physical modulation depth, subjects, and pulse range (2nd to 9th vs 12th to 19th pulses) as fixed factors, ECAP modulation depth as dependent variable, and

TABLE III. ECAP modulation depths measured from the 2nd–9th and 12th–19th pulse trains in experiment 1.

Subject	Stimulus Modulation (dB)	ECAP modulation (%) 2nd–9th	12th–19th	Difference
S3	0	2.8	1.4	−1.4
	0.17	32	29	−3
	0.68	65.6	60.1	−5.5
S6	0	0.4	1.4	1
	0.17	9.1	9.2	0.1
	0.68	30.3	27.4	−2.9
	1.36	46.4	43.6	−2.8
S7	0	0.1	-7.5	−7.6
	0.17	9.1	7.8	−1.3
	0.68	27.3	30.6	3.3
	1.36	45.7	50.5	4.8

that modelled only main effects, revealed significant effects of subject [$F(2,15) = 16.2$, $p < 0.001$] and physical modulation [$F(3,15) = 59.1$, $p < 0.001$], but not of pulse range [$F(1,15) = 0.2$, $p = 0.66$]. This is evidence that the ECAP modulation measured for the 10-pulse stimuli is representative of that occurring during the 20-pulse stimuli used for the behavioural measures.

E. Comparison of ECAP modulation and pitch shifts

The relationship between ECAP modulation and pitch shift is summarised by the circles in Fig. 4. Each panel shows the data for one subject. For each line, the horizontal displacement of the symbols shows the ECAP modulation, with the left-most circle representing the unmodulated 200-pps stimulus, and with increasing physical modulation depth shown by progressively more right-ward circles. The percentage pitch shift re 200 pps is indicated by the vertical position of each circle; for the unmodulated 200-pps pulse train the shift was defined as zero. Generally speaking, ECAP modulation depths below 20% corresponded to negligible pitch shifts, and those below 30% corresponded to shifts smaller than 10%. Although the pattern of results differed across subjects, it is worth noting that there were some cases—indicated by points with a large horizontal and a small vertical displacement—where substantial ECAP modulations corresponded to small or absent pitch shifts. For example, subjects S3, S4, and S7 all show instances where the ECAP modulation depth is greater than 30% and where the pitch shift is smaller than 5%. Nevertheless, it was generally true that, for a given physical modulation depth, subjects who show a larger ECAP modulation tended to also show a larger pitch shift. This was demonstrated by a univariate ANOVA performed on the data for the two modulation depths at which all subjects were tested—0.34 and 0.68 dB—with pitch shift as the dependent variable, ECAP modulation as the covariate, and physical modulation depth as a fixed factor. The analysis revealed a significant correlation between ECAP modulation and pitch shift ($r = 0.74$, $df = 10$, $p < 0.01$). Furthermore, the effect of physical modulation depth was not significant [$F(1,11) = 0.02$, $p = 0.89$]. A likely explanation for the significant correlation observed is that the slope of the relationship between stimulus level and neural response differs across subjects, and that the ECAP measure captures this difference. The correlation is important as it demonstrates that the ECAP captures some aspects of the neural response that are important for pitch perception; it is the ECAP, rather than the physical modulation depth, that successfully predicts the size of the pitch shift for different subjects. Nevertheless, as noted above, the data of some subjects such as S7 show small pitch shifts even for substantial ECAP modulations, suggesting that subjects may differ in the extent to which they "weight" the higher-amplitude and lower-amplitude neural responses.

III. EXPERIMENT 2: ALTERNATING-INTERVAL PULSE TRAINS

A. Method

Experiment 2 obtained both behavioural and ECAP responses to 100-ms pulse trains in which the IPI alternated between 4 and 6 ms, 3 and 7 ms, or 8 and 12 ms (termed "4–6," "3–7," and "8–12" stimuli, respectively). The method used for both sets of measurements was identical to that in experiment 1, with the following exceptions.

1. Behavioural experiment

The behavioural experiment measured the rate of an isochronous pulse train having a pitch equal to that of an alternating-interval pulse train. This was achieved using the same method of constant stimuli as in experiment 1. In each trial the alternating-interval pulse train (the standard) and an isochronous pulse train were presented in random order and the subject indicated which had the higher pitch. No feedback was given.

Nine subjects [S1–S9; Table I] took part in conditions involving 4–6 and 3–7 standards. In these two conditions the isochronous comparison stimulus could have an IPI of 2, 3, 4, 5, 6, 7, or 8 ms. For subject S7, all of these comparison stimuli were judged higher in pitch than the 3–7 stimulus and so, for this subject, that condition was re-run with comparison IPIs of 5, 6, 7, 8, 9, 10, and 11 ms and these new data are reported and analysed. In each block of trials the

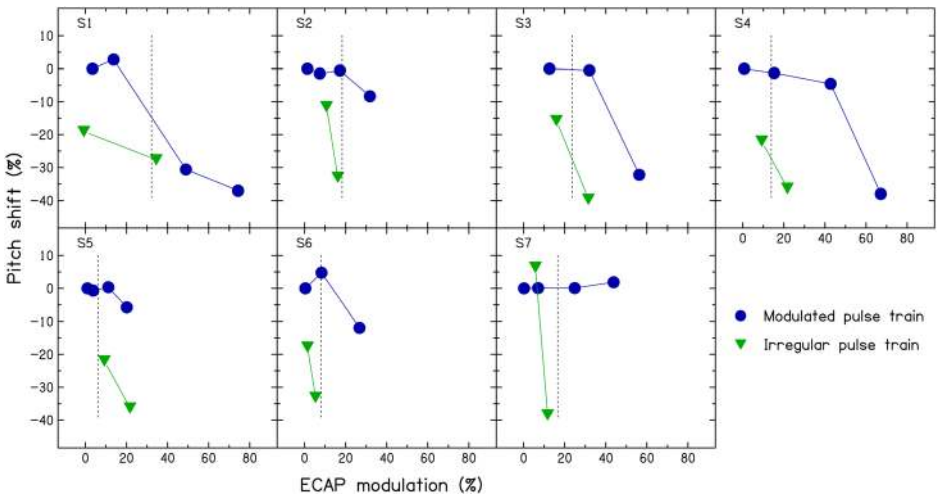

FIG. 4. (Color online) Each panel shows results for one subject. Each circle shows the ECAP modulation depth (abscissa) and pitch shift (ordinate) for one 200-pps stimulus from experiment 1. In each panel the stimulus modulation increases from left to right, so that the left-most circle always corresponds to an unmodulated pulse train. Each triangle shows the ECAP modulation and pitch shift obtained with the alternating-interval stimuli of experiment 2: the left-most and right-most triangles in each panel are for the 4-6 and 3-7 stimuli, respectively.

standard was paired with each comparison 10 times, so there were 70 trials per block. At least four blocks were run for each subject, leading to a minimum of 40 trials per point. The PSE was then calculated as in experiment 1. Each condition was run in separate blocks, alternating between 4–6 and 3–7 standards. In addition, five subjects [S1, S2, S6, S7, and S10; Table I] took part in a condition with an 8–12 standard, and with comparison stimuli having IPIs of 4, 6, 8, 10, 12, 14, and 16 ms.

Stimuli were loudness balanced using the method described previously in experiment 1. Here, MCL was obtained for the 4–6, 3–7, and 8–12 stimuli. For the 4–6 and 3–7 stimuli, the loudness matching procedure was carried out for a stimulus with IPI = 5 ms. This loudness adjusted stimulus was then used to adjust the loudness of both IPI = 2 ms and IPI = 8 ms stimuli. The levels of the IPI = 3, 4, 6, and 7 ms stimuli were taken from a linear interpolation (in CUs) of the adjusted levels at IPI = 2, 5, and 8 ms. Similarly, an IPI = 10 ms stimulus was loudness adjusted with the 8–12 stimulus, and then this IPI = 10 ms stimulus used to adjust IPI = 4 and 16 ms stimuli, again with linear interpolation of remaining rates.

2. ECAP measurements

ECAP measurements were obtained to the same standard stimuli as used in the behavioural experiment. They were also obtained to isochronous comparison stimuli having IPIs close to, and slightly shorter than, the PSEs expected in the behavioural experiment. This was done because we used the isochronous stimuli as a method for estimating the perceived pitch, and we wanted to check that those estimates were not unduly affected by modulation in the neural response. The IPIs tested always included 4, 5, and 6 ms except for subjects S4 and S7 for whom only an IPI of 5 ms was tested. For these IPIs the ECAP modulation was nearly always much less than 5%, and in no case did it exceed the value of 10% shown in experiment 1 to produce negligible pitch shifts.

Two further sets of measurements were obtained. For any alternating-interval stimulus, the IPI preceding any pulse is confounded by the position of the pulse in the sequence. For example, for the 4–6 pulse train, all even numbered pulses occur after an IPI of 4 ms and all odd-numbered pulses occur after an IPI of 6 ms. To distinguish between the effects of IPI and of pulse position, we also measured ECAPs to pulse trains that started with the longer IPI—e.g., 6–4, 7–3, and 12–8 stimuli, and to isochronous stimuli having IPIs of 5 and 10 ms. ECAP modulation was calculated as a percentage using the same formula as in experiment 1.

The alternating-interval stimuli used to obtain the ECAPs and to obtain the behavioural measures had the same levels as each other. The levels were also generally very similar to those used in the calibration experiment. The first two columns of Table IV list the level of the 5-ms-IPI isochronous comparison stimulus in experiment 2 and the (otherwise identical) unmodulated 200-pps stimulus used in the calibration experiment. These levels are within 3 CUs (0.51 dB) of each other for all subjects except S1, for whom the level was increased for experiment 2. At the start of experiment 2 this subject was able to tolerate a higher sound level than at the start of experiment 1. It was decided to use the higher level in order to improve the signal-to-noise ratio in the ECAP measurements. To check the effect of this level difference on the neural response we measured ECAPs to the 3–7 stimulus at a range of levels. The increase in level between experiments 1 and 2 produced only a modest increase in ECAP modulation depth, from 16.7% to 22.3%.

TABLE IV. Stimulus levels (in CUs) for each subject in experiments 1, 2, and 3. One CU is equal to approximately 0.17 dB.

	Experiment 1	Experiment 2				Experiment 3				
	Unmodulated	5-5	4-6	3-7	8-12	100	200	300	400	500
S1	186	193	200	200	196	200	198	196	196	196
S2	225	226	225	227	227	208	205	207	207	208
S3	196	193	193	193	—	195	193	192	192	191
S4	197	197	186	191	—	198	197	197	197	196
S5	237	239	237	237	—	240	237	235	229	229
S6	220	219	219	219	225	219	219	219	219	218
S7	201	203	204	208	212	195	193	191	190	190
S8	—	—	166	166	—	168	166	165	165	164
S9	—	—	203	203	—	201	199	199	199	198
S10	—	—	—	—	—	181	178	178	177	177

3. Supplementary experiment: 3–7 and modulated 4–6 pulse trains

Several months after the above experiments were completed, listeners S1 and S3 took part in a supplementary experiment. The aim was to amplitude-modulate a 4–6 pulse train so that it produced the same ECAP modulation as a 3–7 pulse train, and to determine whether the two stimuli would then have the same pitch. The methods and stimuli were the same as in the main experiments with the following exceptions. First, ECAPs were recorded in response to a 3–7 pulse train and to 4–6 pulse train in which the pulses occurring after the 4-ms intervals were physically attenuated by various amounts. Five repeats were averaged in order to obtain the ECAP modulation depth in response to each stimulus. The physically modulated 4–6 stimulus that produced an ECAP modulation closest to that produced by the (unmodulated) 3–7 stimulus was then selected for behavioural testing. For both subjects this physical modulation depth was 1 CU.

For the behavioural part of the experiment, both the modulated 4–6 and the unmodulated 3–7 stimulus were compared to a range of isochronous comparison pulse trains, and the point of subjective equality was estimated from probit fits to the psychometric functions, as in the main experiment. For subject S1 we initially used isochronous sounds with periods of 3, 4, 5, 6, 7, 8, and 9 ms for the modulated 4–6 stimulus and 4, 5, 6, 7, 8, 9, and 10 ms for the 3–7 stimulus, and with 70 trials per point. To control for any possible effect of using different ranges for the two stimuli, we repeated the 3–7 measurements with 40 trials per point with the same isochronous comparison stimuli as had been used for the modulated 4–6 pulse trains. No significant effect of range was observed. Nevertheless, only the results obtained with the same range of comparison stimuli are analysed and

presented here. For subject S3 the same isochronous stimuli (periods of 4, 5, 6, 7, 8, 9, and 10 ms) were used for both the 4–6 and 3–7 pulse trains.

B. Results

1. 4–6 and 3–7 stimuli

The PSEs for the 4–6 and 3–7 pulse trains are shown, in milliseconds, by the first two columns of Table V. The mean value of 5.7 ms for the 4–6 stimulus is in good agreement with that observed previously, both for CI users and for NH listeners presented with filtered acoustic pulse trains (Carlyon *et al.*, 2002; Carlyon *et al.*, 2008b). The PSE for the 3–7 stimulus was, on average, 7.6 ms. This was significantly longer than for the 4–6 stimulus (t-test, $df = 8$, $p < 0.001$). Perhaps more strikingly, it was significantly longer than 7 ms, which was the longest IPI present in the stimulus (95% confidence intervals +/− 0.3 ms). This is consistent with the idea that the neural response to the pulses occurring after the 3-ms intervals are reduced, leading to some intervals equal to 10 ms being conveyed to the brain (Carlyon *et al.*, 2008b). It is not consistent with pitch being derived from a weighted sum of all first-order intervals in the stimulus (Carlyon *et al.*, 2002), as such a scheme would necessarily predict a PSE somewhere between the longest and shortest IPI in the stimulus. The electrophysiological measurements, discussed next, shed light on whether this reduction can entirely be explained by the compound response of the auditory nerve.

ECAP modulations for the 4–6, 6–4, 3–7, and 7–3 stimuli are shown by the first four columns of Table VI. The values for the 6–4 and 7–3 stimuli are negative, reflecting a larger response to the even-numbered than to the odd-numbered pulses. This occurred because, for these stimuli, the even-numbered pulses occurred after the longer IPI. The fact that the values were negative means that the effect of IPI outweighed any possible effect of pulse position. To analyse these data we inverted the sign for the 6–4 and 7–3 stimuli, and then performed a two-way repeated-measures ANOVA with factors "IPI order" (short-long vs long-short) and "IPI type" (4 and 6 ms vs 3 and 7 ms). The effect of IPI type was significant, reflecting greater modulation for the 3–7 than for the 4–6 pulse train [$F(1,8) = 21.27$, $p < 0.01$]. However, it did not matter in which order the long and short IPIs occurred, as reflected by a non-significant effect of IPI order [$F(1,8) = 0.63$, $p > 0.05$]. The interaction was not significant [$F(1,8) = 0.26$, $p > 0.05$]. The ECAP modulation for the 3–7 and 7–3 stimuli correlated significantly across listeners ($df = 7$, $r = 0.86$, $p < 0.01$, two-tailed). The correlation between ECAP modulation for the 4–6 and 6–4 stimuli was of borderline significance ($df = 7$, $r = 0.66$, $p = 0.05$, two-tailed).

TABLE V. The PSE and corresponding pitch shift re the mean IPI in each stimulus for experiment 2.

	PSE, ms			Downward pitch shift, %		
Subject	4-6	3-7	8-12	4-6	3-7	8-12
S1	6.2	6.9	12.8	19.1	27.6	21.7
S2	5.7	7.5	11.2	11.5	33.0	10.8
S3	5.9	8.3	—	15.7	39.6	—
S4	6.4	7.9	—	22.0	36.3	—
S5	4.7	7.9	—	−7.0	36.4	—
S6	6.1	7.5	12.1	17.8	33.1	17.6
S7	4.7	8.1	8.8	−6.4	38.4	−13.7
S8	5.7	7.1	—	11.7	29.6	—
S9	6.2	7.2	—	18.9	30.4	—
S10	—	—	9.3	—	—	−7.5
Mean	**5.7**	**7.6**	**10.8**	**11.5**	**33.8**	**5.8**
95%	0.4	0.3	1.5	7.1	2.7	13.7

TABLE VI. ECAP modulation for alternating-interval pulse trains, including conditions where the initial IPI was longer or shorter than the mean (e.g., 4-6 vs.6-4 pulse trains).

	ECAP modulation, %							
Subject	4-6	6-4	3-7	7-3	5-5	8-12	12-8	10-10
S1	−0.8	−8.7	34.5	−29.4	−1.7	−1.3	−0.1	0.5
S2	10.7	−10.6	16.2	−13.3	2.0	4.0	−1.6	0.9
S3	16.0	−7.2	31.5	−23.1	4.8	—	—	—
S4	9.2	−5.8	21.7	−23.5	2.0	—	—	—
S5	−3.8	−1.4	4.3	−3.0	−0.1	—	—	—
S6	1.5	−0.4	5.4	−4.9	1.2	0.0	−0.0	0.1
S7	5.7	2.4	11.7	−12.8	−1.5	−0.6	0.8	0.4
S8	22.9	−14.5	30.3	−47.0	4.0	—	—	—
S9	7.3	−5.1	16.6	−14.7	1.3	—	—	—
S10	—	—	—	—	—	−1.9	2.7	0.1
Mean	**7.6**	**−5.7**	**19.1**	**−19.1**	**1.3**	**0.0**	**0.4**	**0.4**
95% conf.	5.5	3.5	7.30	8.9	1.4	2.0	1.4	0.3

Both the pitch shifts observed behaviourally and the ECAP modulation measured electrophysiologically were greater for the 3–7 than for the 4–6 pulse trains. To assess whether the ECAP modulation was sufficient to account for the size of the pitch shifts, on the assumption that the relationship between ECAP modulation and pitch shift was the same in experiments 1 and 2, we plotted the pitch shift as a function of ECAP modulation as triangles in Fig. 4. In each panel, the left-most and right-most triangles show the data for the 4–6 and 3–7 pulse trains, respectively. As for the modulated stimuli of experiment 1 (circles), the horizontal position shows the ECAP modulation depth and the vertical position shows the pitch shift. If pitch shifts were determined solely by ECAP modulation, the two sets of lines should overlap. However, it can be seen that the triangles mostly lie well below the circles. In other words, the pitch shifts for alternating-interval pulse trains are larger than those produced by amplitude modulated pulse trains that produce a similar amount of ECAP modulation. In many cases these differences are substantial. To take one example, subject S7 showed a larger ECAP modulation for a 200-pps pulse train that was amplitude modulated by 0.68 dB than for a 3–7 pulse train (24.9% vs 11.7%), whereas the pitch shift for the modulated pulse train only was −1.5%, compared to 38.4% for the 3–7 stimulus (compare the second circle from the right and the right-most triangle in the bottom right panel of Fig. 4).

2. 8–12 stimulus

The PSEs obtained in the behavioural experiment for the 8–12 stimulus are shown in more detail in the fourth column of Table V. It is interesting to compare the results between the 4–6 and 8–12 stimuli; both stimuli consist of alternating IPIs that differ by a factor of 1.5, with the overall value of the IPI being twice as long for the 8–12 than for the 4–6 stimulus. The PSEs were highly correlated across the four subjects who performed the experiment with both stimuli ($df = 2$, $r = 0.996$, $p < 0.01$, two-tailed), although it is acknowledged that the number of subjects is very small. For these four subjects, the mean PSE was 10.8 ms for the 8–12 stimulus—not far from 11.4 ms, which is double the PSE of 5.7 ms obtained for the 4–6 stimulus. However, it is worth noting that one subject, S7, showed a somewhat unusual result in that her PSE of 8.8 ms for the 8–12 stimulus was smaller than the mean IPI of 10 ms; this was also true to a lesser extent for subject S10.

The ECAP modulations for the 8–12 and 12–8 stimuli are shown in Table VI. The amount of modulation was generally small, and did not differ significantly between the two stimuli. To compare the ECAP modulation to that observed for the 4–6 and 6–4 stimuli, we inverted the sign of the ECAP modulation for conditions that started with the longer IPI (6–4 and 12–8), and entered the data into a two-way repeated-measures ANOVA with factors "IPI order" (short-long vs long-short) and "IPI type" (4 and 6 ms vs 8 and 12 ms). The effect of IPI type reached only borderline significance [$F(1,3) = 6.9$, $p = 0.078$]. There was no effect of pulse order [$F(1,3) = 0.0$, $p = 0.954$]. There was a positive across-subject correlation between the ECAP modulations for the 8–12 and 12–8 stimuli, but this failed to reach significance [$df = 3$, $r = 0.82$, $p = 0.09$, two-tailed], perhaps due to the small number of subjects.

3. 3–7 and modulated 4–6 pulse trains

Table VII shows the pulse levels and ECAP modulation depths for the experiment with 3–7 and modulated 4–6 pulse trains. It can be seen that, for each subject, the ECAP modulation depths were very similar for the two stimuli. Despite this, the PSEs were still longer for the 3–7 stimulus than for the modulated 4–6 stimulus (S1: 7.6 vs 6.3 ms, S3: 8.0 vs 6.9 ms). To evaluate the statistical significance of these differences, we entered the data for each subject into a univariate ANOVA, with the period of the standard and the condition (3–7 vs 4–6) as fixed factors, and the scores for each block of 10 trials as dependent variables. For both subjects there was a main effect of the comparison IPI [S1: $F(6,63) = 67.2$; $p < 0.001$, S3: $F(6,42) = 149.2$; $p < 0.001$], and, importantly, a main effect of condition [S1: $F(1,63) = 28.4$, $p < 0.001$. S3: $F(1,42) = 41.5$, $p < 0.001$]. The interaction was borderline for subject S1 [$F(6,63) = 2.4$, $p = 0.065$] and significant for S3 [$F(6,42) = 5.1$, $p = 0.001$]. The difference in PSEs for the 4–6 and 3–7 stimuli, shows that one cannot account for the lower pitch of 3–7 compared to (unmodulated) 4–6 stimuli solely in terms of the greater modulation measured at the level of the auditory nerve (cf. Carlyon *et al.*, 2008b)

TABLE VII. Stimulus levels and results for the experiment comparing unmodulated 3-7 and modulated 4-6 pulse trains. The row labeled "Level, CUs" shows the levels of the higher- and lower-amplitude pulses used for the 4-6 stimulus, and the level used for all pulses in the 3-7 stimulus. The next two rows show the measured ECAP modulation (with the standard error across measurements in parentheses), followed by the PSE obtained in the behavioural part of the experiment.

	S1		S3	
	4-6 (modulated)	3-7	4-6 (modulated)	3-7
Level, CUs	200,199	200,200	198,197	198,198
ECAP modulation (s.e.), %.	19.1 (0.9)	19.5 (0.7)	26.3 (1.0)	25.0 (2.6)
PSE, ms	6.3	7.6	6.9	8.0

IV. EXPERIMENT 3: THE UPPER LIMIT OF TEMPORAL PITCH

A. Method

1. Behavioural experiment

Eight users of the Nucleus Freedom cochlear implant, seven of whom also participated in experiments 1 and 2, took part; their details are shown in Table I. All stimuli were isochronous pulse trains presented in MP1 + 2 mode, with a phase duration of 25 μs and an inter-phase gap of 8 μs. On each two-interval forced-choice trial the listener heard a standard and a signal pulse train in random order and indicated which one had the higher pitch. Correct-answer feedback was provided after every trial. The standard pulse rate was 100, 200, 300, 400, or 500 pps, and the signal had a rate 30% higher than that of the corresponding standard. Answers were scored as correct when the subject identified the signal as having the higher pitch. The standard stimulus could consist of either 10 or 50 pulses; the corresponding signal had either 13 or 65 pulses, so that its duration was the same as that of the standard. We used 10-pulse standards so as to facilitate comparison with ECAP measures, described in the next subsection. We expected rate discrimination to be quite hard for these short stimuli, and so we included the 50-pulse standards partly to help keep the subjects "on task", and partly to provide a comparison with data obtained in other studies that typically used rather longer stimuli. Each block of trials consisted of 10 repetitions of each combination of standard rate and number of pulses, leading to a total of 100 trials per block. At least five, and usually between seven and ten blocks were run for each subject, so there were a minimum of 50 trials per data point.

Prior to the main part of the experiment all stimuli were loudness balanced using the same procedure described for experiments 1 and 2. Loudness balancing was carried out using the 50-pulse stimuli. Initially, subjects indicated MCL for the 100-pps stimulus. This rate, at MCL level, was used to adjust the loudness of the 300- and 500-pps stimuli. The MCL of the 100-pps stimuli, together with the loudness adjusted levels of the 300- and 500-pps stimuli, were then used to interpolate levels for the standard rates of 200- and 400-pps. A further control of possible loudness cues in the

discrimination task was made by loudness balancing the signal rate with each standard rate. For example, the 130-pps signal was loudness balanced with the 100-pps standard; the 260-pps signal with the 200-pps standard, and so on.

2. ECAP measurements

ECAPs to the same stimuli as used in the behavioural experiment were measured and analysed with the methods employed for experiments 1 and 2. An exception was that, for technical reasons, the highest pulse rate from which we could obtain an ECAP to every pulse was 478 pps. This rate was therefore used in lieu of the 520-pps signal that was paired with the 400-pps standard in the behavioural experiment. The level of this 478-pps stimulus was the same as used for the 520-pps stimulus in the behavioural task. No ECAPs were measured for the 500-pps standard or signal, which were therefore absent from statistical comparisons between the behavioural and ECAP measures.

The levels of the stimuli were the same for the behavioural and electrophysiological measurements, and are shown in the last five columns of Table IV. For five subjects, the level of the 200-pps stimulus was within 3 CUs of that of the unmodulated 200-pps stimulus in the calibration experiment. The level was higher in the present experiment than in the calibration experiment for subject S1, and lower for S7. For subject S7, we subsequently re-measured ECAPs for the 0.17-dB modulated stimulus of experiment 1 at the level used in that experiment and at the lower level used in experiment 3. The ECAP modulation at the lower level used here was 13.3%, compared to 8.9% at the higher level used previously. Note that for this listener, experiment 1 revealed very small pitch shifts even for an ECAP modulation larger than 40% (Fig. 4).

A potential limitation on the directness of our behavioural-neural comparison arises from the inter-stimulus-interval (ISI) of 205 ms between the pulse trains used to obtain our ECAP measures. Although recovery times to single pulses are typically less than 10 ms, ECAP recordings from the rat (Haenggeli *et al.*, 1998) to 200-ms pulse trains separated by 200 ms have shown that the response to the first pulse in each train decreased with increases in pulse rate. This can only have been due to the ECAP to the first pulse in each train being affected by preceding pulse trains. Because the ISI in our behavioural experiments was longer (400 ms between intervals, average of 1700 ms between trials) than the 205 ms gap between pulse trains for the ECAP measures, we re-tested the ECAPs for a subset of subjects (S1, S2, S7, S8, and S9) at 478 pps—the highest rate used here, with inter-stimulus intervals similar to those used in the behavioural experiment (400 ms between each pair of opposite-polarity pulse trains, and 1700 ms between pairs). These additional measurements were obtained several months after the main experiment, and so we also repeated the measures with the standard 205-ms intervals between all pulse trains. Four repeat measurements were obtained per condition for subjects S1, S2, and S7; at least six measurements per condition were obtained for subjects S8 and S9, whose ECAPs were smaller overall than for the other subjects.

B. Results

1. Behavioural experiment

Rate discrimination (Rau-transformed percent correct) is shown as a function of the standard pulse rate in Fig. 5; parts (a) and (b) of the figure show data for the 10-pulse and 50-pulse standards, respectively. Mean data, shown by the thick line with filled square symbols, show trends that were confirmed by main effects in a two-way RM-ANOVA: performance was better for 50-pulse than for 10-pulse standards [$F(1,7) = 16.1$; $p < 0.01$], and, as has been found previously, varied as a function of pulse rate [$F(4,28) = 5.6$, $p < 0.02$]. The interaction was not significant [$F(4,28) = 1.4$, $p = 0.28$]. There was also substantial variability in the pattern of results across listeners, as shown by the individual data plotted as fainter lines with symbols. Some listeners, such as S4, show the "classic" pattern of performance, being good at 100 and 200 pps but close to chance at higher rates. Others, such as S2, show best performance at intermediate rates—a pattern that has also been observed in some previous studies (Kong and Carlyon, 2009; Kong *et al.*, 2009).

2. ECAP measurements

Figure 6 shows ECAP amplitudes, plotted as a function of pulse position and at four pulse rates, for two example

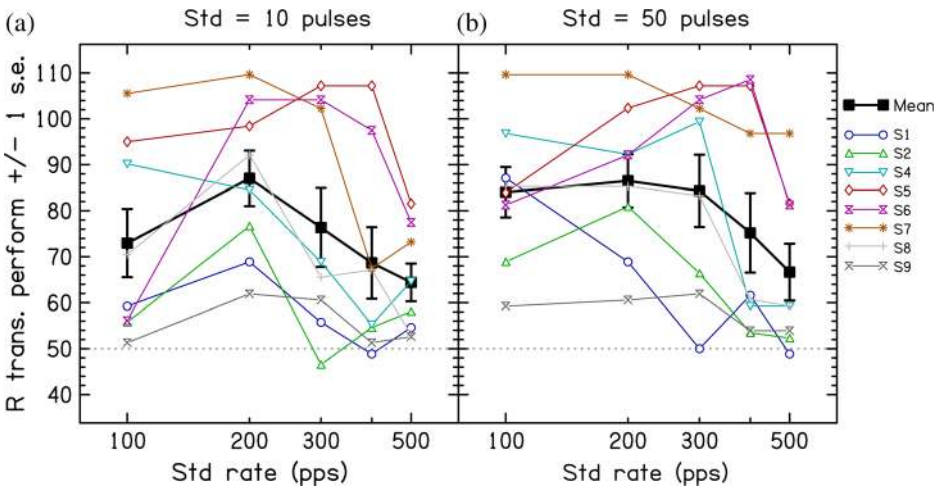

FIG. 5. (Color online) Rau-transformed % correct in the rate discrimination task of experiment 3, for stimuli where the standard consisted of 10 pulses (left panel) or 50 pulses (right panel). The large black symbols joined by heavy lines show the mean data across subjects, with error bars showing +/− one standard deviation. The fainter colored symbols show data for individual subjects as shown in the legend.

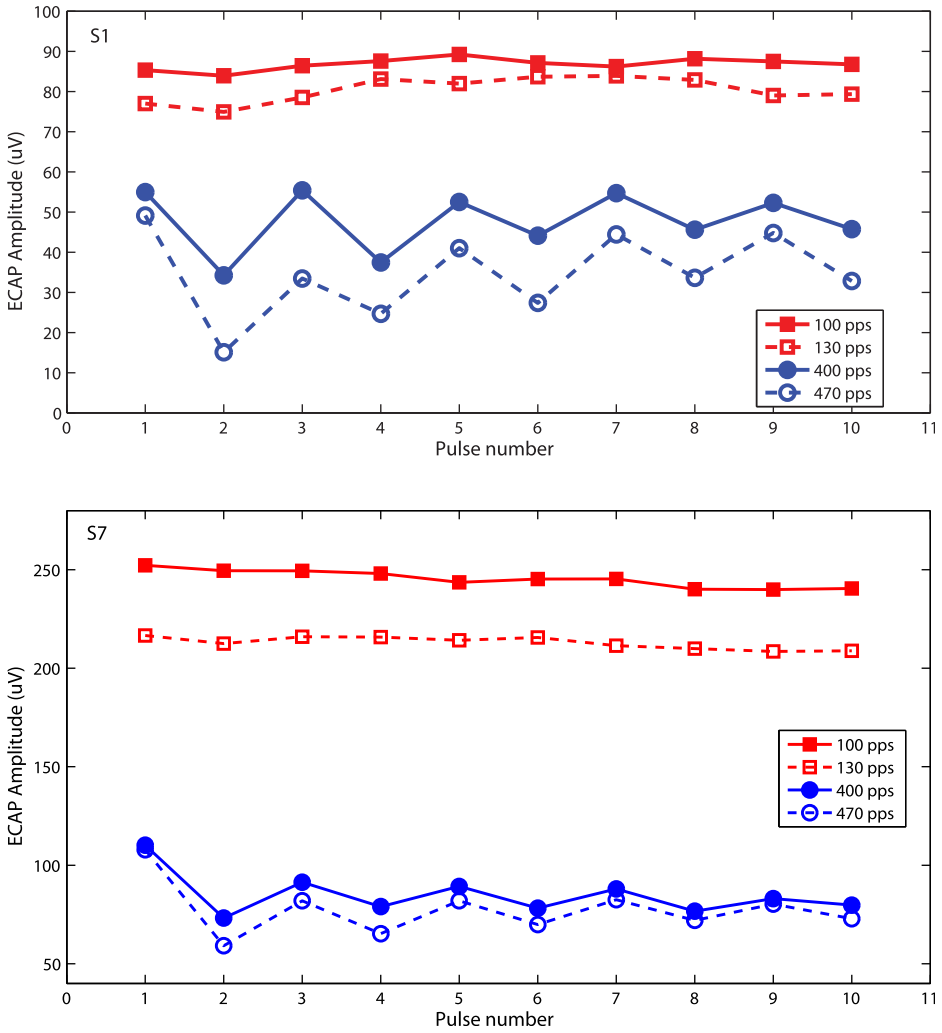

FIG. 6. (Color online) ECAP amplitudes as a function of pulse position for pulse rates of 100, 130, 400, and 478 pps. Data are shown for subjects S1 and S7.

subjects, S1 and S7. At 100 and 130 pps the ECAPs are roughly constant as a function of pulse position. In contrast, the ECAPs observed at 400 and 478 pps are larger for the odd-numbered than for the even-numbered pulses. Figure 7 shows the percentage ECAP modulation, for all subjects as a function of pulse rate, with the mean data shown by the thick bold line with filled square symbols. A one-way RM-ANOVA revealed a highly significant effect of pulse rate [$F(7,49) = 18.0$, $p = 0.001$], reflecting the increase in modulation depth with increasing rate.

The effect of changing the inter-stimulus interval on the ECAP is shown in Table VIII. When the average data for each subject were analysed using t-tests, there was no

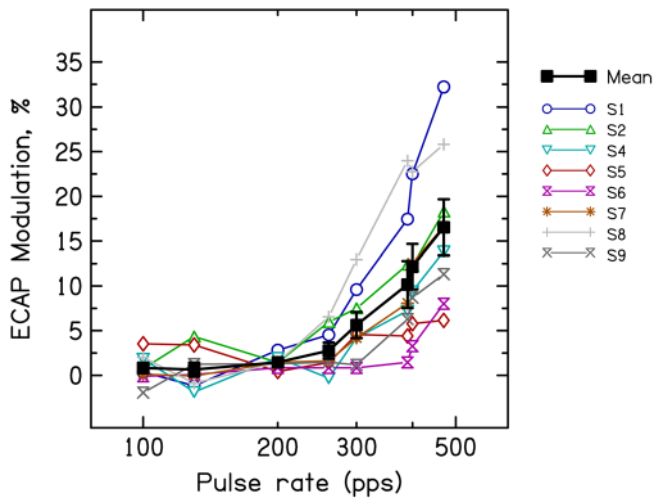

FIG. 7. (Color online) ECAP modulation depth as a function of stimulus rate, for all subjects who took part in experiment 3. Mean data are shown by the thick black line with large black square symbols. Fainter lines and smaller symbols are for individual subjects, as in Fig. 5.

TABLE VIII. Comparison of ECAPs obtained with short vs long ISIs (see text for details). The first two columns show the amplitude of the ECAP to the first pulse, in μV. The last two columns show the ECAP modulation depth in percent.

	1st pulse amplitude		ECAP modulation	
	Short ISI	Long ISI	Short ISI	Long ISI
S1	41.2	43.1	27.2	26.9
S2	57.2[b]	61.8[b]	18.7[b]	24.5[b]
S7	125.8	134.5	12.9	14.5
S8	15.4	13.2	36.9	45.6
S9	28.9	26.6	18.5[a]	24.1[a]

[a]Differences that were significant at the 5% level.
[b]Differences that were significant at the 1% level.

significant effect of condition on either the amplitude of the ECAP to the first pulse of each train, or on the amount of ECAP modulation. For individual subjects, there were instances where either the ECAP to the first pulse or the ECAP modulation differed between conditions. For subject S2 the first-pulse ECAP was slightly but significantly larger at the longer inter-stimulus interval (61.8 vs 71.2 μs, $t(8) = 4.18$, $p < 0.01$). This subject also showed greater ECAP modulation at the longer inter-stimulus interval (24.5% vs 18.7%, $t(8) = 4.32$, $p < 0.01$), as did subject S9 (24.1% vs 18.5%, $t(12) = 2.31$, $p < 0.05$). The differences observed for subject S2 would survive Bonferroni correction for multiple comparisons; the difference for S9 would not. Overall we consider these differences to be modest and do not believe that they have large implications for the conclusions drawn here. They do however extend previous findings from animal experiments to show that quite long-lasting effects can be observed in the human auditory nerve in response to electrical stimulation (Haenggeli *et al.*, 1998).

3. Comparison of ECAP and behavioural measures

Figures 5 and 7 show that, averaged across listeners, an increase in pulse rate was accompanied both by a deterioration in discrimination performance and by an increase in ECAP modulation. These main effects were to be expected from the literature, and the fact that discrimination and physiology co-vary with rate does not mean that they are necessarily related. To determine whether ECAP modulation and performance are related, it is therefore worthwhile to determine whether these two dependent variables co-vary when the effects of pulse rate are partialed out. Another consideration comes from the fact that any perceptual effect of modulation in the neural response should not depend on whether the larger response occurs to the odd-numbered pulses—leading to a modulation ratio greater than one, or to the even-numbered pulses—leading to a ratio smaller than 1. In order to treat any ECAP modulation ratio as being equivalent to its reciprocal, we calculated the absolute value of the logarithm of the modulation ratios for each signal and standard. According to the reasoning described in the Introduction and Fig. 1, performance should be poor when the ECAP to the signal was more modulated than that to the standard, thereby reducing the pitch difference between them. We therefore entered the difference between the transformed signal and standard modulation values, for each combination of subject and standard rate, into a univariate ANCOVA with performance as the dependent variable and rate as a fixed factor. This revealed a significant effect of the covariate [$F(1,27) = 7.8$, $p = 0.01$], which accounted for 22.4% of the variance in the data; the correlation is shown in Fig. 8. We also performed the analysis using the log of the modulation ratio, averaged across each standard stimulus and its corresponding signal; this was also significant [$F(1,27 = 6.4$, $p < 0.02$]. These analyses show that there is indeed a significant relationship between rate discrimination and both the average amount of modulation in the ECAP to the standard and signal, and the difference in the ECAP modulation between signal and standard. Specifically, subjects showing larger ECAP modulations, and larger modulation differences between the standard and signal, performed worse on the rate discrimination task.

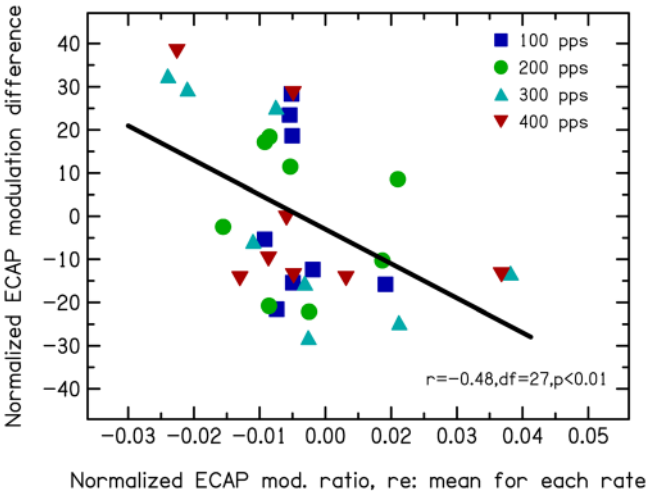

FIG. 8. (Color online) Normalised correlation obtained from experiment 3. The abscissa shows the difference between the absolute values of the log of the ECAP modulation obtained for the standard and signal pulse rates. The ordinate shows performance on the rate discrimination task. Both measures were normalised so as to remove main effects of pulse rate, so as to leave only between-subject differences. Each symbol shape and color shows data obtained for one pulse rate and eight subjects.

Despite the significant across-subject correlation between ECAP modulation and performance, the small size of the ECAP modulations observed here suggests that any such relationship is unlikely to be causal. Specifically, Fig. 7 shows that the ECAP modulation in response to the isochronous pulse trains used in experiment 3 are almost universally below 30%, even at the highest pulse rate tested. Experiment 1 revealed that ECAP modulations of this size corresponded to rather small pitch shifts. To illustrate this, the vertical dashed lines in Fig. 4 show the ECAP modulation depth for the 478-pps pulse trains of experiment 3, for the subjects who took part in both experiments 1 and 3. The intersection of these lines with the curves connecting the circles show the pitch shift that would be expected for a physically modulated 200-pps pulse train that produced the same ECAP modulation as the unmodulated 478-pps pulse trains of experiment 3. These predicted pitch shifts are below 5% for all subjects except S1, for whom the predicted shift is about 15%. We think that these pitch shifts would be unlikely to affect the discrimination between two pulse rates that differed by 30%. Furthermore, there are instances where rate discrimination performance is poor at lower rates and where the ECAP modulation is even smaller. For example, subject S2 scored only 46.6% at rate discrimination with a 300-pps standard, and the ECAP modulations for this standard and the corresponding signal were 8% and 12%, respectively. These were smaller than the 17% ECAP modulation observed in experiment 1 by attenuating every other pulse by 0.68 dB, a manipulation that produced no change in perceived pitch (Fig. 4, subject S2, third circle from the left).

A *caveat* with the above comparisons is that their validity rests on the assumption that a given modulation depth will have a similar effect on pitch for the 200-pps stimuli of experiment 1 as for the higher-rate stimuli from

experiment 3. However we should also note that ECAP modulation should impair rate discrimination only to the extent that it is greater in the signal than in the standard interval. Although this was generally true, these differences were much smaller than the total ECAP modulation discussed here and shown in Fig. 6.

V. EXPERIMENT 4: EFFECT OF REDUCING ECAP MODULATION

A. Rationale and method

As argued above, a comparison of the results of experiments 1 and 3 suggest that the relationship between ECAP modulation and rate discrimination is unlikely to be causal. However, we also noted that this conclusion depends on the assumption that the relationship between ECAP modulation and pitch is the same across the range of rates tested in the two experiments. Experiment 4 tested the possible causal relationship between ECAP modulation and rate discrimination, using a method that does not require this assumption. To do so, we added a "pre-pulse" before the standard and signal pulse trains, and adjusted its level so as to minimise the modulation in the ECAPs to all the pulses in the pulse train. The reasoning was that, by a judicious choice of the level of the pre-pulse, we could reduce the response to the first pulse by just enough so that the ECAPs to the pre-pulse and first pulse were equal, as illustrated in Fig. 1(d). This would avoid having the majority of neurons responding to any one pulse, and might in turn reduce modulation in the ECAPs to subsequent pulses. The level of this pre-pulse was selected by performing a set of ECAP recordings at which several different pre-pulse levels were tried, and the ECAP modulation of the remaining pulses (2 through 9) measured in each case. The selected pre-pulse level was then checked by repeating the ECAP measurement at least four times, and the mean of the ECAP modulation ratios was used to calculate the ECAP modulation for this stimulus. This procedure was carried out for the two standard rates and corresponding signal rates used in the behavioural part of the experiment, described below. We also measured ECAP modulation with a pre-pulse whose amplitude was equal to that of the other pulses [Fig. 1(e)]; this left the ECAP modulation depth largely unchanged, and provided a control condition for the behavioural experiment, in which the total durations of the stimuli were the same as with the "optimised" pre-pulse levels.

In the behavioural part of the experiment, rate discrimination was measured using the same method as in the previous two experiments, except that only two standard rates were used. For most subjects, one of these rates was 368 pps, selected so that we could use a signal with a rate 30% higher and equal to 478 pps, which, for technical reasons, was the highest rate at which we could measure an ECAP to every pulse. The other rate was either 100 or 200 pps and was selected for each subject to be one at which rate discrimination was good in experiment 3; this was 100 pps for subjects S1 and S7, and 200 pps for S2 and S8. For subject S7, whose rate discrimination at low rates was close to ceiling when the standard rate was 30% higher than that of the standard, the signal rate was set to 10% higher than the standard in this experiment, for both standard rates. For this subject only, the higher-rate standard had a rate of 400 pps. Each standard stimulus consisted of 10 "main" pulses, each preceded by a pre-pulse; this was the same as in the ECAP measurements. Two levels of pre-pulse were used. One of these was the "optimal" level selected during the preliminary experiment, whereas the other had the same level as the other pulses in the pulse train. The level of the two standard and corresponding signal pulse trains were interpolated from the loudness-balanced levels for each rate determined for each subject in experiment 3.

B. Results

For the lower-rate standards and signals there was generally very little ECAP modulation regardless of the pre-pulse amplitude. Figure 9(a) shows the ECAP modulations obtained for the 368-pps (400 pps for S7) standard and corresponding 478-pps (440 pps for S7) signal, with the "optimal" pre-pulse level (solid bars) and for the condition where the pre-pulse had the same level as the "main" pulses (cross-hatched bars). It can be seen that the ECAP modulation was greatly reduced by the optimal pre-pulse. An illustration of the ECAP amplitudes with the two pre-pulse amplitudes is shown for subject S7 in Fig. 9(b); the ECAP to the pre-pulse itself is not shown. With the full-level pre-pulse, ECAPs are smaller to odd-numbered than to even-numbered pulses; this is because the pre-pulse has shifted each pulse number by one (e.g., the first main pulse is really the second pulse in an equal-amplitude pulse train). In contrast, the optimal pre-pulse results in an ECAP pattern that is mostly flat throughout each pulse train.

Figure 9(c) shows that, despite the different effects of the two pre-pulse amplitudes on the pattern of ECAP modulation, performance on the rate discrimination task was essentially unchanged. This argues strongly against a causative effect of ECAP modulation on the upper limit of rate discrimination.

VI. DISCUSSION

A. Comparison of physiological measures with behavioural responses

Understanding the neural basis for sensory acuity and for the limitations of human performance has been the goal of auditory scientists for many decades. Each approach has the potential to provide important insights, but, perhaps inevitably, the interpretation of the results is not straightforward. For example, a number of authors have obtained single-unit recordings from anaesthetised animals and compared the ideal performance that would be predicted, based on different decoding mechanisms, to that obtained in human psychophysical experiments. A classic example comes from studies comparing information on place-of-excitation, phase-locking, and combined codes on pure-tone frequency in cat to the frequency DLs observed in human (e.g., Siebert, 1970; Heinz *et al.*, 2001). Those studies provide important information on the potential limits of frequency encoding, but often rely on additional assumptions concerning the frequency

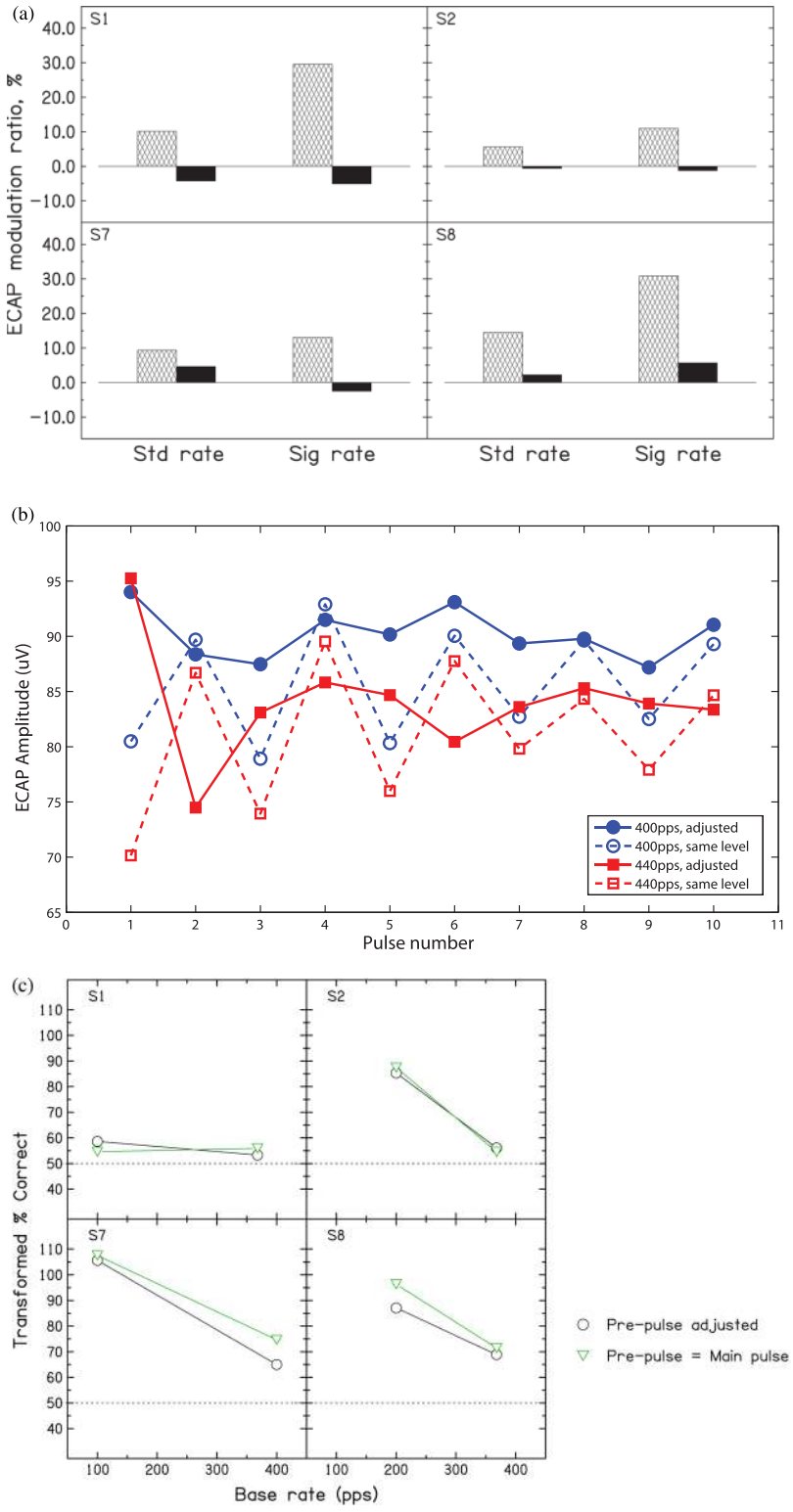

FIG. 9. (Color online) (a) ECAP modulation for each subject and with the pre-pulse amplitude set either to the "optimal" level (solid bars) or to the same amplitude as the other pulses (cross-hatched bars). Data are shown for the "high rate" standard and signal rates. (b) Example of ECAP amplitude as a function of pulse position for a 400-pps (circle symbols) and 440-pps (square symbols) pulse train for subject S7. Dashed lines and open symbols are for the condition with the pre-pulse at the same level as the main pulses. Solid lines and filled symbols are for the condition with the pre-pulse set to the optimal level. (c) Rau-transformed percent correct for conditions with the pre-pulse set to the optimal level (circles), or set to the same amplitude as the other pulses (inverted triangles).

selectivity of the auditory system, and compare physiology in one species with the psychophysics of another. We believe that our approach has some unique advantages, which we describe first, before proceeding to describe its limitations and underlying assumptions.

Although it is widely accepted that temporal processing is important for pitch perception, several important issues remain. These include the potential role of timing differences between the responses of neurons tuned to different characteristic frequencies (CFs) (Shamma, 1985; Moore and Carlyon, 2005; Cedolin and Delgutte, 2010; Carlyon *et al.*, 2012), and whether place-of-excitation cues provide an additional mechanism for encoding musical pitch (Oxenham *et al.*, 2011). The analysis of the pitch perceived when such cues are absent provides a way of studying within-channel temporal processing in isolation, and may provide insight both into the nature of that processing and of more general accounts of pitch. Electric pulse trains presented to CI users

provide a very effective way of removing place-of-excitation and across-channel-timing cues. Hence, both the discrimination of pulse rate and the judgement of the pitch of these stimuli must be based on the temporal information present in the auditory nerve, and, furthermore, it is not necessary to consider the relative timing of the responses of different AN fibers. By recording the whole-nerve response to electrical pulse trains, and by comparing these responses to behavioural measurements obtained with the same subjects and stimuli, one can obtain a relatively direct comparison of the AN response to a stimulus and of the information that is extracted and encoded by the listener.

It is worth noting that the size of the ECAP can be affected by several factors including the number of neurons responding, spike amplitude, and the variance in the firing rate across neurons. Our approach was to compare the ECAP modulation arising from different manipulations, including physically amplitude-modulating the pulse train, using alternating-interval pulse trains, and increasing the pulse rate. An implicit assumption is therefore that either the different manipulations influence ECAP amplitude in the same way (for example, by primarily changing the number of neurons firing at any one time), or that, if they do differ, the effect on the responses at the next stage of processing (the cochlear nucleus) are similar. For example, Miller *et al.* (2001) measured the response to pairs of electrical pulses in the auditory nerve of the cat, and showed that the spike amplitude to the second pulse decreased for IPIs less than 1 ms. They argued that this would decrease the response at the level of the cochlear nucleus and at more central sites; this would also be the case for a decrease in the number of AN fibers responding.

The ECAP analyses that we have presented were applied to the averages of 100 pairs of opposite-polarity pulse trains. Such averaging is ubiquitous in electrophysiological research, and allowed us to obtain accurate measurements of, for example, the average amount by which the neural response to odd-numbered pulses exceeded that to even-numbered pulses. However, it should be remembered that, in each trial of a psychophysical task, listeners do not have access to 200 presentations of the stimuli to be discriminated. One extreme way in which the average of multiple presentations could provide a misleading representation of the neural response on each trial would be if the response to even-numbered pulses exceeded that to odd-numbered pulses on some proportion of trials, with the opposite effect occurring for the remainder of trials; these neural modulations would then, to some extent, cancel out in the average. Fortunately, we were able to repeat the ECAP measures for subjects S1 and S2 using a slightly modified method that gave us access to the traces from individual presentations. The stimuli tested corresponded to the standard rates in experiment 3. In every case the ECAP modulation ratio was unimodally distributed across sweeps, inconsistent with the ratio of the amplitudes of the odd- vs even-numbered responses being bimodally distributed across presentations. However it is still possible that random pulse-to-pulse fluctuations in the amplitude of the neural response could have affected pitch judgements. For this to influence our conclusions the perceptual effect of those idiosyncratic fluctuations would have to be comparable to, or larger than, that of the alternating-amplitude fluctuations present in the averaged recordings. To account for the results of experiment 2, this effect would also have to be larger for alternating-interval (4–6 and 3–7) pulse trains than for physically modulated pulse trains (experiment 1). To account for the deterioration in rate discrimination observed at high rates in experiment 3, the effect would also have to be greater for high-rate equal-amplitude pulse trains than for the physically modulated 200-pps pulse trains of experiment 1.

Finally, although this article is primarily concerned with the effects of variations in the *amplitude* of the auditory nerve response on pitch perception, it is worth briefly considering whether variability in *temporal* aspects of the response could limit rate discrimination. Each panel of Fig. 10 shows the ECAP in response to the fourth pulse of a pulse train. The ten traces in each panel show the ECAP averaged from 10 different "sweeps," where each sweep consists of a single presentation of two opposite-polarity pulse trains. The top row shows ECAPs for a 300-pps pulse train and the bottom row shows ECAPs to a 400-pps pulse train—close to the signal rate paired with a 300-pps standard in experiment 3. The left column is for subject S1, for whom performance in experiment 3 had dropped to 55.8% for the 300-pps standard. The right column is for subject S2, for whom performance had dropped to 46.6% at 300 pps. In all cases the timing of the ECAP is consistent across different sets of 10 sweeps, and the variation in the timing of N1 is considerably less than 100 μs. The thick bar at the bottom of each plot shows the 769–μs difference in inter-pulse interval between the 300-pps standard and 390-pps signal in experiment 3. It appears that this difference is considerably larger than the variation in the timing of the ECAP shown in Fig. 10. Hence although we have not performed a quantitative analysis of the variation in the neural response, it appears that such variation is unlikely to account for the limitations on rate discrimination observed at high rates in experiment 3.

B. Alternating-interval pulse trains

Three previous attempts have been made to model data on the pitch of alternating-interval pulse trains. Those models were applied to the pitch heard by normal-hearing listeners when presented with acoustic pulse trains that had been bandpass filtered so as to contain only unresolved harmonics. As noted in the Introduction, such stimuli produce pitches that are very similar to those obtained with pulse trains presented to a CI electrode, both in absolute terms and in their variation with changes in stimulus parameters (McKay and Carlyon, 1999; Carlyon *et al.*, 2002; van Wieringen *et al.*, 2003; Carlyon *et al.*, 2008a; Carlyon *et al.*, 2008b; Carlyon *et al.*, 2011).

Carlyon *et al.* (2002) pointed out that the autocorrelation of a 4–6 pulse train contains peaks at 4, 6, and 10 ms. They showed that a popular autocorrelogram model (Meddis and O'Mard, 1997) produced peaks at these values. That model used a metric for predicting pitch matches based on the squared Euclidean distance between the "summary autocorrelograms"

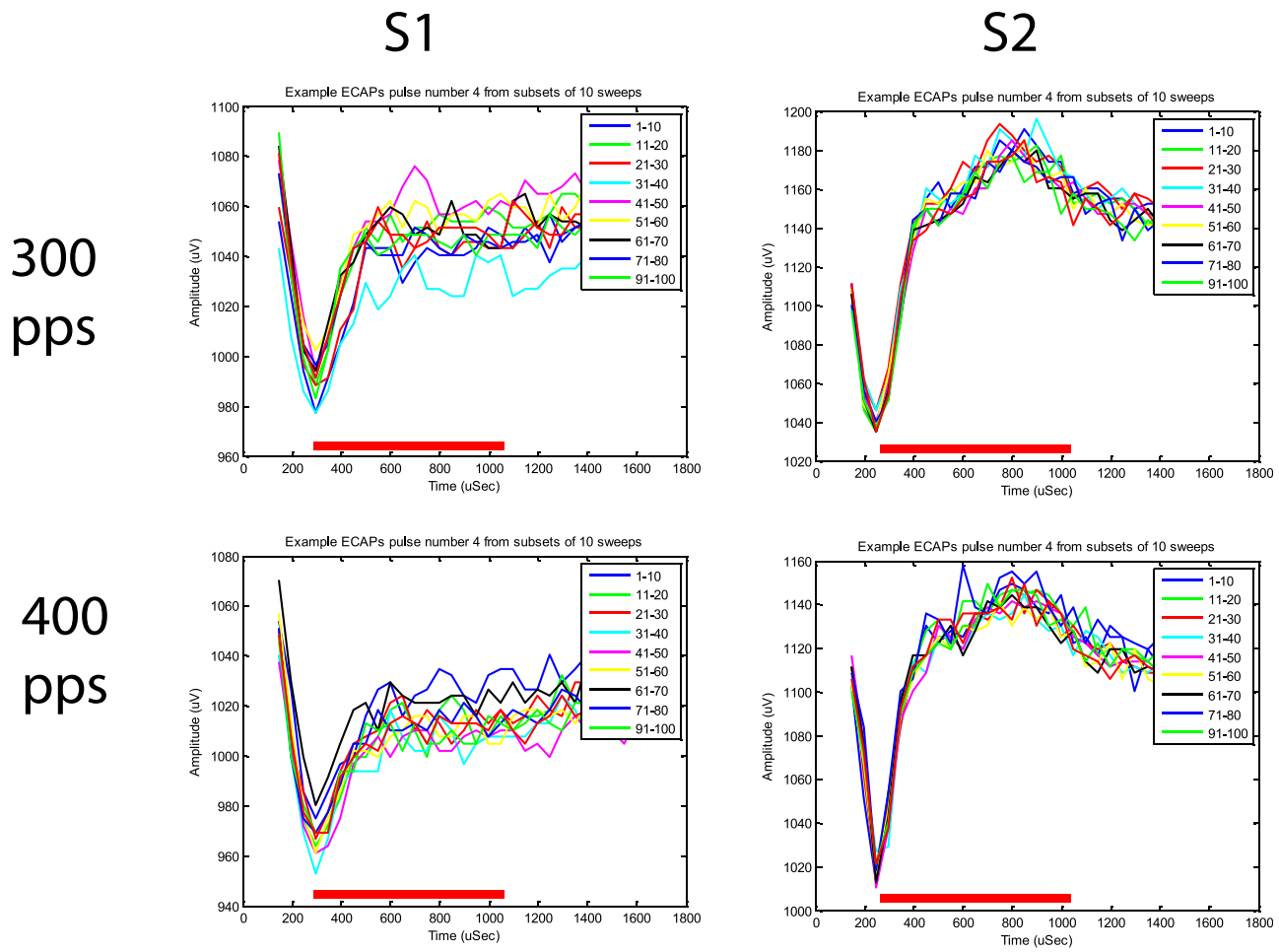

FIG. 10. (Color online) ECAPs obtained from subsets of 10 sweeps for pulse number 4 of pulse trains having rates of 300 pps (top row) and 400 pps (bottom row). Data for subjects S1 and S2 are shown in the left- and right-hand columns, respectively.

(SACFs) corresponding to the matching and reference stimuli. For the 4–6 stimuli used by Carlyon *et al.* (2002), the model predicted that the best-matching isochronous stimulus had an IPI of 5 ms; its SACFs contained a peak at 5 ms, which did a reasonable job of matching the 4- and 6-ms peaks of the SACF of the 4–6 stimulus, and a peak at 10 ms that had an exact counterpart in the 4–6 SACF. This prediction contrasted with the PSE of about 5.7 ms obtained both for acoustic and electric pulse trains used by Carlyon *et al.* (2002). Because of this, and of the failure of Meddis and O'Mard's model to account for some other data obtained using filtered pulse trains (Plack and White, 2000), Carlyon *et al.* proposed an alternative model in which pitch was derived from a weighted sum of first-order IPIs in the stimulus. The model applied weights that increased with increasing IPI, and could account not only for the pitch of the "4–6" pulse trains but also of other stimuli studied in that article and by Plack and White (2000).

Subsequently, Carlyon *et al.* (2008b) suggested that one may be able to dispense with the central weighting function by assuming that, as had been pointed out in a different approach by Pressnitzer *et al.* (2004), higher-order intervals in the stimulus may be converted to 1st-order intervals in the AN *response*. Carlyon *et al.* showed that the CAP to alternating-interval pulse acoustic trains was amplitude modulated, and suggested that some "more central" neurons might respond only to the pulses that elicit the higher-amplitude CAPs—effectively transforming a 2nd-order stimulus IPI into a 1st-order neural IPI [cf. Fig. 1(b)]. As discussed in our Introduction, they suggested that, e.g., for a 4–6 stimulus, pitch is derived from a weighted sum of the 4, 6, and 10 ms intervals, with the weight given to the 10-ms intervals being proportional to the modulation depth of the neural response.

Our experiment 2 showed that the pitch of a 3–7 stimulus was judged equal to that of an isochronous pulse train having an ISI of 7.6 ms—longer than the longest IPI present in the stimulus. This means that any account based on 1st-order intervals must indeed do so with respect to the intervals in the neural response (Carlyon *et al.*, 2008b), rather than to those in the stimulus (Carlyon *et al.*, 2002). The experiment also shows that the simple model proposed by Carlyon *et al.* (2008b) cannot be based on modulation in the neural response at the level of the AN. That model would predict the same pitch for a (physically) modulated 200-pps pulse train and an unmodulated 4–6 or 3–7 pulse train, whenever the stimuli elicit the same depth of modulation in the neural response. Figure 4 shows clearly that this is not the case. Hence the model would require an additional source of refractoriness, central to the AN, that had a substantial impact on pitch judgements. This conclusion is further supported by our finding that the pitch of a 3–7 stimulus remains lower than that of a 4–6 stimulus, even when the

latter is amplitude modulated such that the two sounds produce similar amounts of ECAP modulation. As noted in the Introduction, evidence for an effect of refractoriness on pitch comes from the finding of van Wieringen *et al.* (2003) that the pitch of 4–6 pulse trains is lower when the amplitudes of pulses occurring after 4-ms intervals are attenuated, compared to when the attenuation is applied to pulses occurring after the 6-ms intervals. Taken together, the results presented here and by van Wieringen *et al.* (2003) are consistent with an effect on pitch of refractoriness at a site central to the auditory nerve.

Ballester *et al.* (2008) proposed a modification to the autocorrelogram model of Meddis and O'Mard (1997), which, they argued, allowed it to account for the pitch of 4–6 pulse trains in acoustic hearing. The revised model produced an estimated PSE of 5.9 ms, not far from the 5.7 ms observed experimentally, although it incorrectly predicted a bimodal pattern of pitch matches, unlike the unimodal distribution reported by Carlyon *et al.* (2002). Ballester *et al.* were not sure why the fit to the mean data was better for the new model, but assumed that the reason was related to the updated modeling of the auditory periphery. We think, however, that the nature of the peripheral transduction is unlikely to be important, both because for NH listeners the PSE is unaffected by a 24-dB difference in input level (Carlyon *et al.*, 2002; Carlyon *et al.*, 2008b), and, more importantly, because, as noted above, the PSEs obtained with acoustic and electric stimulation are very similar both in absolute terms and in their variation with changes in stimulus parameters (Carlyon *et al.*, 2002; van Wieringen *et al.*, 2003). Our data provide a more direct way of accounting for the effects of the auditory periphery by providing measurements of the compound AN response.

To illustrate the predictions of a simple autocorrelation-based model to our electric pulse trains, we measured the autocorrelation to trains of pulses that were amplitude modulated in the same way as the ECAPs measured in experiments 1 and 2. That is, the input to the autocorrelation was not, as is common in studies of acoustic hearing, the output of a model of peripheral processing, but was instead an estimate of the auditory nerve response obtained from our recordings. Specifically, we measured the autocorrelation of an unmodulated isochronous pulse train with an ISI of 5 ms, and of 4–6 and 3–7 pulse trains amplitude modulated by 7.6 and 19.1%, respectively—these amounts being the average ECAP modulation observed for the two stimuli (cf. Table VI). Two features of these autocorrelation functions (Fig. 11) are worth noting. First, there are no peaks in the 4–6 and 3–7 functions corresponding to the obtained PSEs (periods of 5.7 and 7.6 ms, respectively). As noted above, this was previously observed for acoustic 4–6 stimuli passed through the autocorrelogram model of Meddis and O'Mard (Carlyon *et al.*, 2002); as in that case, a model (not shown) based on squared Euclidean distances between the autocorrelation of the test and a range of isochronous comparison stimuli failed to predict the observed pitches. Second, the different depths of amplitude modulation applied to the 4–6 and 3–7 simulated neural responses did not have a marked effect on the heights of the peaks in the autocorrelation; in particular, the peaks at 4 and 6 ms have very similar heights to those at 3 and 7 ms [Figs. 11(b) and 11(c)]. This is consistent with our findings that modulation in the auditory nerve response, of the amounts observed in our experiments, did not have a large effect on pitch.

Although our (admittedly preliminary) analyses suggest that existing autocorrelation models do not, in their current form, correctly predict the pitches of alternating-interval pulse trains, it is possible that two fairly simple modifications may allow them to do so. For electric pulse trains, and for complex tones consisting of unresolved harmonics, it is likely that there are no consistent differences in the pattern of the neural response across the auditory nerve fiber array. It is also known that, when inharmonically related mixtures of such stimuli having different rates are applied to the same electrode or frequency region, listeners are unable to "hear out" the rates corresponding to each stimulus, but instead hear a composite sound that, at least in the acoustic case, sounds like a "crackle" (Carlyon, 1996; Micheyl *et al.*, 2006; Micheyl *et al.*, 2010; Wang *et al.*, 2012). It is therefore possible that listeners do not "hear out" the individual peaks in the autocorrelogram, except when they occur in different frequency channels, but instead judge a single pitch that is derived from a weighted combination of those individual peaks. Two simple rules would be to apply larger weights to larger peaks, and/or, when peaks occur at lags that are integer multiples, to apply smaller weights to the higher multiples in order to reduce octave errors. This would mean that, for an isochronous pulse train, the pitch would be dominated by the first peak, which is at the true period of the pulse train. For a 4–6 pulse train there would be some trade-off between the peaks at 4 and 6 ms and that at 10 ms—the former two being shorter than, but also smaller than, the latter (cf. Figure 11). This would lead to a larger weight being applied to the 10-ms peak than is the case for an isochronous pulse train having a 5-ms IPI, and could potentially account for the fact that the 4–6 stimulus has the lower pitch. An additional factor would have to be invoked to explain why the 3–7 stimulus has an even lower pitch than the 4–6 stimulus. One possibility is that the peaks in the autocorrelation are smoothed by some "lag window," such that peaks that are close together become combined into one larger peak. If so then the degree of summation would depend on the separation between the peaks, and so the peaks at 4 and 6 ms would produce a larger combined peak than those at 3 and 7 ms.

To summarise, experiments 1 and 2 do not unambiguously identify the processes involved in purely temporal pitch perception but do demonstrate that modification of existing models is necessary. Specifically, (i) any analysis of 1st-order intervals must account for the fact that neural refractoriness converts 2nd-order intervals in the stimulus into 1st-order intervals in the response, (ii) a simple version of that model would require that refractoriness central to the AN has a substantial effect on pitch (cf. van Wieringen *et al.*, 2003), (iii) existing autocorrelation-based models are unlikely to be able to account for the pitch of alternating-interval stimuli purely by invoking nonlinearities in the response of the peripheral auditory system to acoustic pulse trains, and (iv)

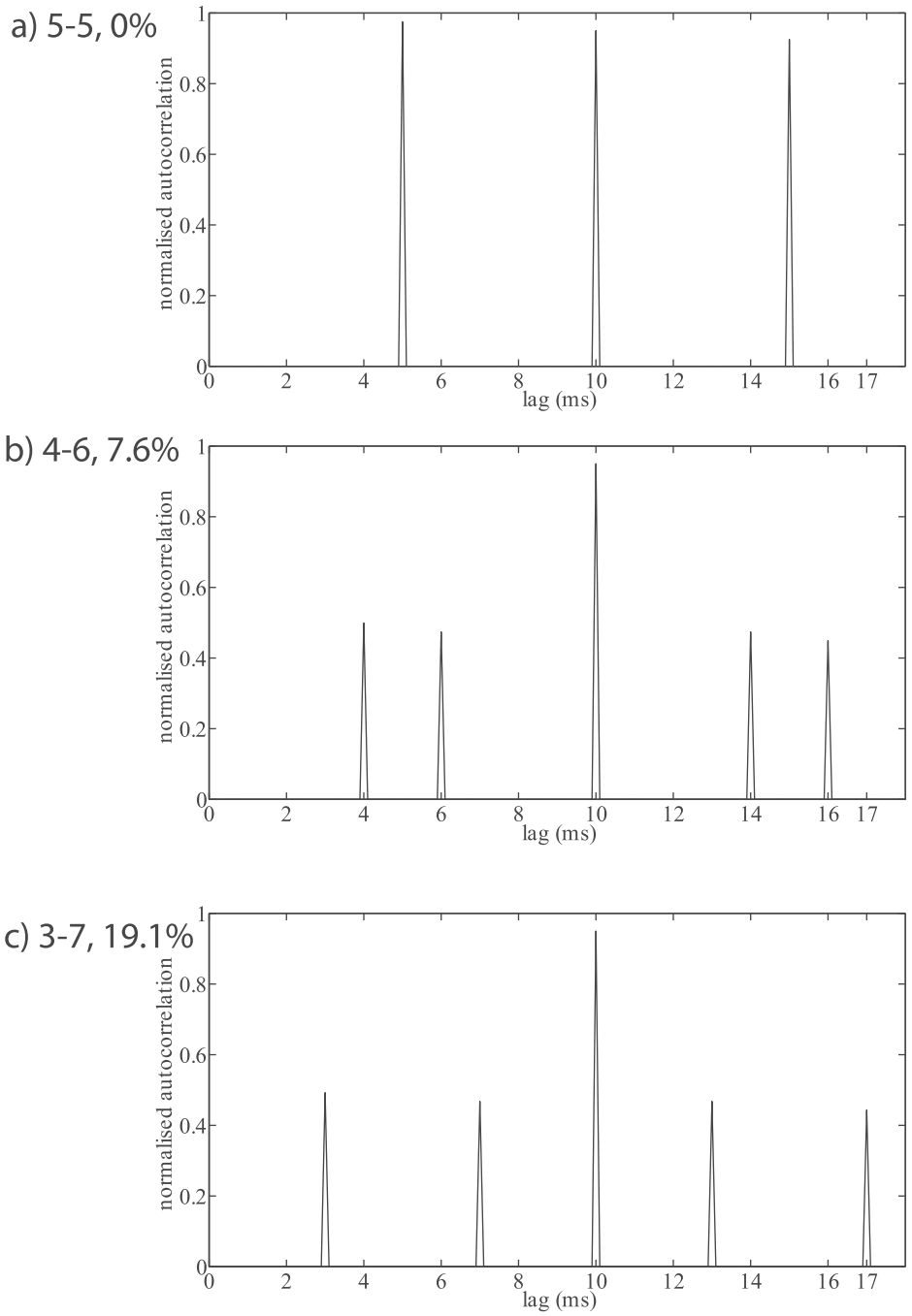

FIG. 11. Normalised autocorrelations of pulse trains, where panel (a) is for an unmodulated isochronous stimulus having a period of 5 ms. Panels (b) and (c) are for 4-6 and 3-7 pulse trains amplitude modulated by the same amount as the ECAPs in experiment 3.

autocorrelation-based models may be able to account for the data by using a different decision metric, that invokes a weighted sum of the different periodicities, and where smaller weights are applied to longer periodicities and, possibly, to those that are integer multiples of shorter periodicities present.

C. Rate discrimination as a function of rate

The results of experiments 3 and 4 provide convincing evidence that modulations in the auditory nerve response, as measured by the ECAP, are not responsible for the deterioration in rate discrimination with increasing baseline rate. Not only were those modulations smaller than those needed to produce a marked effect on pitch in experiment 1, but a manipulation that reduced ECAP modulation substantially did not consistently improve rate discrimination. Preliminary analyses of variability in both the amplitude and timing of the neural response, derived from small subsets of presentations, also failed to find any evidence for an auditory-nerve based limitation.

It is worth noting that, although there may be enough information present at the level of the AN for an ideal observer to perform the rate discrimination task, this does not necessarily mean that variations in auditory nerve survival do not affect performance. For example, extensive neural loss might require that a comfortably-loud stimulus excite neurons with a wide range of CFs, and, if more central structures "corrected" for the traveling-wave delay that would occur in acoustic hearing, this might cause the representation of the electric pulse train, which is synchronous across auditory nerve fibers, to become "blurred" at more central levels of

processing. This might reconcile the fact that rate discrimination correlates with ECAP modulation across subjects (experiment 2) and MCL across sites (Macherey *et al.*, 2011) with our evidence that the AN conveys enough information about pulse rate even at rates where performance is at chance. That is, ECAP modulation and MCL might co-vary with peripheral neural survival, which in turn has an indirect effect on the decoding of auditory nerve information by the brain. However, it is also worth remembering that it is well known that the upper limit of phase locking decreases at higher levels of processing, and that physiological experiments have begun to characterise the biological basis of some of these limitations, such as the transition from sustained to onset-only responses in neurons of the inferior colliculus to binaural stimulation (Smith and Delgutte, 2007; Hancock *et al.*, 2012). Hence although we believe that our results impose some important constraints on the neural basis of the upper limit of rate discrimination, the precise nature of those limitations remains to be determined.

VII. SUMMARY

(i) Modulation in the ECAP provides a measure of auditory nerve activity that partially accounts for across-subject differences in the effect of physical amplitude modulation on the pitch of a 200-pps pulse train.

(ii) ECAP modulation does not account for the reduction in pitch caused by delaying alternate pulses in an otherwise isochronous pulse train. Previous evidence suggests that the pitch of such "alternating interval" stimuli are affected by refractoriness; the present results suggest that any such refractoriness has a component central to the auditory nerve.

(iii) The pitch of alternating-interval pulse trains cannot be captured by a simple autocorrelation of the auditory nerve (ECAP) response. However, a number of modifications are suggested that may allow an autocorrelation-based approach to account for the data presented here and elsewhere.

(iv) Subjects who show large ECAP modulations tend to be worse at rate discrimination. However, there is not a causal link between modulation in the neural response and the upper limit of rate discrimination.

(v) Our analyses of the pattern of ECAP amplitudes, combined with preliminary analyses of the variability in the amplitude and timing of the ECAPs, indicate that the limitation on rate discrimination lies more centrally than the auditory nerve.

ACKNOWLEDGMENTS

We thank Mark White and Hedwig Gockel for very helpful discussions, and Alex Billig for comments on a previous version of this article.

Ballester, E. B., Denham, S., and Meddis, R. (**2008**). "A cascade autocorrelation model of pitch perception," J. Acoust. Soc. Am. **124**, 2186–2197.

Carlyon, R. P. (**1996**). "Encoding the fundamental frequency of a complex tone in the presence of a spectrally overlapping masker," J. Acoust. Soc. Am. **99**, 517–524.

Carlyon, R. P., and Deeks, J. M. (**2002**). "Limitations on rate discrimination," J. Acoust. Soc. Am. **112**, 1009–1025.

Carlyon, R. P., Deeks, J. M., and McKay, C. M. (**2010**). "The upper limit of temporal pitch: Stimulus duration, conditioner pulses, and the number of electrodes stimulated," J. Acoust. Soc. Am. **127**, 1469–1478.

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., Long, C. J., and Micheyl, C. (**2012**). "Across-channel timing differences as a potential code for the frequency of pure tones," J. Assoc. Res. Otolaryngol. **13**, 159–171.

Carlyon, R. P., Macherey, O., Frijns, J. H. M., Axon, P. R., Kalkman, R. K., Boyle, P., Baguley, D. M., Briggs, J., Deeks, J. M., Briaire, J. J., Barreau, X., and Dauman, R. (**2011**). "Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear," J. Assoc. Res. Otolaryngol. **11**, 625–640.

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. (**2010**). "Spatio-temporal representation of the pitch of harmonic complex tones in the auditory nerve," J. Neuroscience **30**, 12712–12724.

Haenggeli, A., Zhang, J. S., Vischer, M. W., Pelizzone, M., and Rouiller, E. M. (**1998**). "Electrically evoked compound action potential (ECAP) of the cochlear nerve in response to pulsatile electrical stimulation of the cochlea in the rat: Effects of stimulation at high rates," Audiology **37**, 353–371.

Hancock, K. E., Chung, Y., and Delgutte, B. (**2012**). "Neural ITD coding with bilateral cochlear implants: effect of binaurally coherent jitter," J. Neurophysiol. **108**, 714–728.

Hay-McCutcheon, M. J., Brown, C. J., and Abbas, P. J. (**2005**). "An analysis of the impact of auditory-nerve adaptation on behavioral measures of temporal integration in cochlear implant recipients," J. Acoust. Soc. Am. **118**, 2444–2457.

Heinz, M. G., Colburn, H. S., and Carney, L. H. (**2001**). "Evaluating auditory performance limits: I. One-parameter discrimination using a computational model for the auditory nerve," Neural Comput. **13**, 2273–2316.

Hochmair-Desoyer, I. J., Hochmair, E. S., Burian, K., and Stiglbrunner, H. K. (**1983**). "Percepts from the Vienna cochlear prosthesis," Ann. N. Y. Acad. Sci. **405**, 295–306.

Kong, Y.-Y., and Carlyon, R. P. (**2009**). "Temporal pitch perception at high rates in cochlear implants," J. Acoust. Soc. Am. **127**, 3114–3123.

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 differences between low and high rates of stimulation on single electrodes for cochlear implantees," J. Acoust. Soc. Am. **117**, 319–327.

Laneau, J., Wouters, J., and Moonen, M. (**2006**). "Improved music perception with explicit pitch coding in cochlear implants," Audiol. Neuro-Otol. **11**, 38–52.

Macherey, O., and Carlyon, R. P. (**2014**). "Re-examining the upper limit of temporal pitch," J. Acoust. Soc. Am. **136**, 3186–3199.

Macherey, O., Deeks, J. M., and Carlyon, R. P. (**2011**). "Extending the limits of place and temporal pitch perception in cochlear implant users," J. Assoc. Res. Otolaryngol. **12**, 233–251.

McDermott, H. J., McKay, C. M., and Vandali, A. E. (**1992**). "A new portable sound processor for the University-of-Melbourne nucleus limited multielectrode cochlear implant," J. Acoust. Soc. Am. **91**, 3367–3371.

McKay, C. M., and Carlyon, R. P. (**1999**). "Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains," J. Acoust. Soc. Am. **105**, 347–357.

Meddis, R., and O'Mard, L. (**1997**). "A unitary model of pitch perception," J. Acoust. Soc. Am. **102**, 1811–1820.

Micheyl, C., Bernstein, J. G. W., and Oxenham, A. J. (**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, M. V., and Oxenham, A. J. (**2010**). "Pitch perception for mixtures of spectrally overlapping harmonic complex tones," J. Acoust. Soc. Am. **128**, 257–269.

Miller, C. A., Abbas, P. J., and Robinson, B. K. (**2001**). "Response properties of the refractory auditory nerve fiber," J. Assoc. Otolaryngol. **2**, 216–232.
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.
Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A., and Santurette, S. (**2011**). "Pitch perception beyond the traditional existence region of pitch," Proc. Natl. Acad. Sci. U.S.A. **108**, 7629–7634.
Plack, C. J., and White, L. J. (**2000**). "Pitch matches between unresolved complex tones differing by a single interpulse interval," J. Acoust. Soc. Am. **108**, 696–705.
Pressnitzer, D., de Cheveigne, A., and Winter, I. M. (**2004**). "Physiological correlates of the perceptual pitch shift for sounds with similar waveform autocorrelation," Acoust. Res. Lett. Online **5**, 1–6.
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.
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.
Siebert, W. M. (**1970**). "Frequency discrimination in the auditory system: Place or periodicity mechanisms," Proc. IEEE **58**, 723–730.
SPSS. (**2012**). "IBM SPSS Statistics version 21."
Smith, Z. M., and Delgutte, B. (**2007**). "Sensitivity to interaural time differences in the inferior colliculus with bilateral cochlear implants," J. Neuroscience **27**, 6740–6750.
Tong, Y. C., Blamey, P. J., Dowell, R. C., and Clark, G. M. (**1983**). "Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple-channel cochlear implant," J. Acoust. Soc. Am. **74**, 73–80.
Townshend, B., Cotter, N., van Compernolle, D., and White, R. L. (**1987**). "Pitch perception by cochlear implant subjects," J. Acoust. Soc. Am. **82**, 106–115.
Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., and McDermott, H. J. (**2005**). "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies," J. Acoust. Soc. Am. **117**, 3126–3138.
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 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.
Venter, P. J., and Hanekom, J. J. (**2014**). "Is there a fundamental 300 Hz limit to pulse rate discrimination in cochlear implants?," J. Assoc. Res. Otolaryngol. **15**, 849–866.
Wang, J., Baer, T., Glasberg, B. R., Stone, M. A., Ye, D., and Moore, B. C. J. (**2012**). "Pitch perception of concurrent harmonic tones with overlapping spectra," J. Acoust. Soc. Am. **132**, 339–356.
Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., and Rabinowitz, W. M. (**1991**). "Better speech recognition with cochlear implants," Nature **352**, 236–238.
Wilson, B. S. (**1997**). "The future of cochlear implants," Brit. J. Audiol. **31**, 205–225.