

Published in final edited form as:

J Acoust Soc Am. 2010 January ; 127(1): 339–349. doi:10.1121/1.3269042.

Temporal pitch percepts elicited by dual-channel stimulation of a cochlear implant

OLIVIER MACHEREY and ROBERT P. CARLYON

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

Abstract

McKay and McDermott [*J. Acoust. Soc. Am.* 100, 1081-1092 (1996)] found that, when two different amplitude modulated pulse trains are presented to two channels separated by < 1.5 mm, some cochlear implant (CI) patients perceive the aggregate temporal pattern. The present study attempted to extend this general finding and to test whether dual-electrode stimulation would increase the upper limit of temporal pitch perception in CIs. Six subjects were asked to rank twelve dual-channel stimuli differing in their rate (ranging from 92 to 516 pps on each individual channel) and in their inter-channel delay (pulses on the two channels being either nearly simultaneous or delayed by half the period). The data showed that, for an electrode separation of 0.75 or 1.1 mm, a) the perceived pitch was on average slightly higher for the long- than for the short-delay stimuli but never matched the pitch corresponding to the aggregate temporal pattern; b) the upper limit of temporal pitch did not increase using long-delay stimuli c) the pitch differences between short- and long-delay stimuli were largely insensitive to channel order and to electrode configuration. These results suggest that there may be more independence between CI channels than previously thought.

I. INTRODUCTION

In cochlear implants (CIs), sound information is transmitted through amplitude modulated high-rate electrical pulse trains delivered to multiple locations along the cochlea. Although existing CIs use carrier rates up to several thousands of pulses per second (pps) per channel, there is some evidence that most implanted subjects are insensitive to temporal fine structure at such high rates. For example, several reports showed little or no improvement in speech recognition for increases in pulse rate above 400-800 pps per channel (e.g. Friesen *et al.*, 2005). It has also been suggested that very high rates may increase non-simultaneous channel interactions and therefore impair the coding of temporal modulations (Middlebrooks, 2004).

The limitation in the extraction of temporal fine structure experienced by CI listeners may partly relate to their inability to discriminate between single-channel pulse trains of different rates when the standard rate is higher than approximately 300 pps, or to perceive an increase in pitch above such a rate. The reason for this upper limit of temporal pitch is unclear given that subjects show a great amount of variability with some of them being able to discriminate between rates as high as 800 pps (Townshend *et al.*, 1987; Kong *et al.*, 2009). Several possible reasons for this limitation have been put forward, including the absence of a place-rate match in the cochlea (e.g. Oxenham *et al.*, 2004), the fact that CIs do not stimulate very apical neurons (Middlebrooks and Snyder, 2009) or the possibility that more central processes have lost their ability to follow high rates due to prolonged deafness. Moreover,

normal hearing listeners show a higher upper limit of temporal pitch than the majority of CI listeners, being able to discriminate between unresolved harmonic complexes filtered in a fixed frequency region for rates as high as at least 700 pps (Carlyon and Deeks, 2002). Previous studies on the perception of temporal pitch by CI users have usually involved single channel stimulation. In this study, we will investigate the temporal pitch percepts elicited by unmodulated pulse trains presented to two neighboring channels and we will study how these percepts are altered by changing the inter-channel delay.

The sensitivity of CI users to inter-channel delay has mainly been investigated in discrimination studies (Tong and Clark, 1986; McKay and McDermott, 1996, 1999; Carlyon *et al.*, 2000). Tong and Clark (1986) presented 100-pps pulse trains to two bipolar channels and measured the just-discriminable inter-channel delay. Their reference stimulus had a 5-ms inter-channel delay (half-a-period) and that of the signal was varied in steps of plus or minus 0.1 ms. Delay difference limens (DLs) were measured for several distances of the two bipolar channels, including a reference condition where both pulse trains were presented to the same channel. They obtained the smallest DLs (less than 0.5 ms) in the “same channel” condition and found the DLs to increase as the distance between the two bipolar pairs increased.

Similarly, McKay and McDermott (1996) performed a discrimination experiment using 500-pps pulse trains, also presented to two bipolar channels. Each pulse train was amplitude modulated at 100 Hz and consisted of one “big” and four “smaller” pulses (with the amplitude of the small pulses reduced to 50% of the dynamic range). They found that subjects could detect the inter-channel delay when the two bipolar channels were separated by less than approximately 3-4 mm but not for larger separations.

The results of these two studies are consistent with the hypothesis of overlapping spreads of excitation: if the two stimulating channels are close to each other, some auditory nerve fibers will respond to both of them and will convey a temporal code highly dependent on the inter-channel delay. As the spatial distance between channels increases, each channel will excite a more discrete neural population and the amount of overlap will decrease. In another study, McKay and McDermott (1999) used dual-channel stimulation as a way to assess spatial spread of excitation as a function of pulse duration. They found that subjects could discriminate between stimuli with different inter-channel delays at larger channel separations for long (266 μ s) than for shorter (less than 100 μ s) phase durations, consistent with a broader spatial spread of excitation for the longer-phase stimuli.

In addition to the phenomenon of overlapping spreads of excitation, there is some evidence that another process is at work in CIs. Carlyon *et al.* (2000) showed that cochlear implantees could reliably discriminate between inter-channel delays of 0.1 and 2 ms when *remote* bipolar channels were stimulated. The fact that this sensitivity was still present when the pulse polarity was inverted or when an additional “masker” pulse train was presented on a channel in between the two target channels argued against their results being due to overlapping spreads. One possible interpretation was that a more central mechanism (e.g. co-incidence detectors) could detect fine timing differences between the nerve firings of distinct populations of auditory nerve fibers. It is also worth noting that normal-hearing listeners can detect asynchronies between pairs of concurrent pulse trains that have been bandpass filtered into separate frequency regions, even when potential within-channel interactions are masked by noise (Carlyon, 1992; Carlyon and Shamma, 2003).

Recently, much attention has been given to the effects of simultaneous or quasi-simultaneous dual-channel stimulation on place-pitch perception (e.g. Donaldson *et al.*, 2005). However, little is known about the temporal pitch information that is conveyed when

more than one channel is stimulated. To our knowledge, this has only been studied by McKay and McDermott (1996). They compared the pitches evoked by several dual-channel stimuli (pulse trains modulated at 100 Hz) in a pitch scaling experiment. For channel separations smaller than 1.5 mm, pitch increased with increases in inter-channel delay for three of the five subjects tested. The magnitude of this increase suggested that the aggregate temporal pattern rather than the individual channel patterns could be conveyed to higher levels of the auditory system. McKay and McDermott further proposed two alternative interpretations for this observation. First, subjects might have perceived the aggregate pattern because some auditory nerve fibers were responding to both stimulating channels and were therefore conveying a different temporal pattern when the inter-channel delay changed. Second, it was possible that each channel excited a different set of auditory nerve fibers but that a more central neural mechanism was able to combine information arising from these two intracochlear sites.

The present study aims to disentangle these two hypotheses and to quantify the pitch variations as a function of inter-channel delay in dual-channel stimulation. If the second hypothesis of McKay and McDermott is true, then, we may expect the upper limit of temporal pitch to *increase* when two channels are stimulated with a delay of half-a-period compared to when only one channel is stimulated. For example, if neural populations proximal to two CI electrodes are able to follow a rate of 300 pps but not higher due to a limitation at the auditory nerve level, more central neurons may receive inputs from these two populations and convey a temporal code of 600 pps. The existence of such a “summing” central process is also suggested by physiological studies of acoustic stimulation. When primary auditory neurons are excited by a single tone with a frequency less than approximately 4–6 kHz, their responses are phase-locked to the stimulus. However, a given neuron usually does not fire on every cycle of the stimulus, especially as the frequency is increased beyond 200 Hz (Rose *et al.*, 1967). The encoding of the frequency of the waveform therefore requires a combination of the firings of several neurons.

To our knowledge, there has not been a study investigating direct pitch comparisons between dual-channel pulse trains. Moreover, the pitch scaling experiment of McKay and McDermott (1996) was limited to a relatively low (100 Hz) modulation rate. In the present study, we will compare the pitches elicited by dual-channel stimuli differing in their inter-channel delay (quasi-simultaneous or delayed by half-a-period) for pulse rates ranging from 92 to 516 pps per channel. In Section II, we will show that the perceived pitch of a half-period delay dual-channel stimulus is only slightly higher than that of a quasi-simultaneous stimulus and that the upper limit of temporal pitch does not increase using half-period delay stimuli. In Section III, we will demonstrate that these results are robust to modifications of the electrode configuration (monopolar or bipolar) and of the channel order of stimulation (apical or basal-first). Finally, we will focus in Section IV on the discriminability of dual-channel stimuli with different inter-channel delays.

II. EXPERIMENT 1: Pitch ranking of dual-channel stimuli

A. Rationale and Methods

Subjects and Stimuli—Four Advanced Bionics CII/HiRes90k (S1-S4) and two Nucleus CI24 (S5-S6) subjects participated in Experiment 1. Their biographical and clinical details are shown in Table 1. The stimuli consisted of 400-ms duration, unmodulated pulse trains presented to two neighboring bipolar channels of the implant. The pulses were anodic-first, symmetric biphasic with a short inter-phase gap (0 or 8 μ s). Polarity refers to the most apical electrode of the bipolar pair. The twelve stimuli differed in their rate (92, 129, 184, 258, 368 or 516 pps per channel) and in their inter-channel delay as follows. For each rate, there was a short- and a long-delay condition (cf. Fig. 1A). The short delay was set to 388

μs and was constant across the different rate conditions. It allowed the pulses on both channels to be nearly synchronous while limiting interaction effects due to residual polarization (de Balthasar *et al.* 2003). The long delay was set so that the second channel would deliver a pulse delayed by half-a-period relative to the first channel and its value consequently co-varied with stimulation rate.

As we aimed to extend the results of McKay and McDermott (1996), we used similar parameters as in their study, i.e. a narrow bipolar separation (“BP+1” for S1-S4, and S6, and “BP+2” for S5), a phase duration ranging from 97 μs (for S1 to S4) to 150 μs (for S5 and S6) and the smallest distance possible between the two channels (one electrode separation). “BP+X” refers to the separation between the active and return electrodes of the bipolar pair (e.g., “BP+1” means that the two electrodes of the pair are separated by two electrode distances). The distance between two neighboring electrodes is 1.1 mm for the Advanced Bionics and 0.75 mm for the Nucleus. The electrodes tested for each subject are indicated in Table 1.

In this experiment, the leading channel was the most basal electrode pair. This was to avoid any temporal effect being confounded by a place of excitation cue. In the short-delay condition, we may indeed expect the first pulse (leading channel) to partially mask the effect of the second pulse (delayed channel). Consequently, if the first pulse is presented on the basal channel, we may expect the short-delay stimulus to have a more basal place of excitation, and therefore to elicit a higher pitch sensation than the long-delay stimulus if the pitch difference was only due to a place cue. We anticipate that, if the results show that the long-delay stimulus sounds higher in pitch than the short-delay one, then, this will strongly suggest that the pitch difference reflects a change in the temporal pattern of stimulation. Note that the effect of channel order will be specifically addressed in Section III.

All data were collected using the APEX experimental software platform (Laneau *et al.*, 2005), which acted as an interface to the BEDCS and NIC2 software routines provided by Advanced Bionics and Cochlear corporation, respectively. Impedances on all electrodes were checked prior to perform the experiments in order to ensure the implants were driven below compliance. This research was approved by the Cambridge Local Research Ethics Committee.

Loudness balancing—Prior to perform the pitch comparisons, the dual-channel stimuli were equated in loudness, using an adjustment paradigm. Each loudness-balancing trial consisted of two sounds presented consecutively with a 500-ms gap between them. The first sound was the reference and was fixed across the adjustment; the second sound was the one to match and subjects did so by pressing one of six virtual buttons labelled “–”, “– –”, “– – –” to make the sound softer and “+”, “+ +”, “+ + +” to make it louder. The different buttons corresponded to different current steps: 8, 16 and 24 μA for Advanced Bionics subjects and 1, 2 and 3 current units (approx 0.176 dB) for CI24 subjects. Each time a button was pressed, the two stimuli were presented again with the new level of the target and this was repeated until the subject perceived them as equally loud.

For each pair of sounds that had to be loudness-balanced, two adjustments were performed. In the first one, the reference was fixed to a specific level and the target was adjusted to match its loudness. In the second one, the initial target became the reference and its level was set at the balanced level obtained in the first adjustment. The initial reference stimulus became the target and was the one to adjust. This was done to counterbalance any potential response bias due to the order of presentation. Once these two adjustments were performed, the balanced level of the initial target was obtained by averaging the level differences (in

dB) obtained in the two adjustments and subtracting it from the level of the initial reference sound.

The loudness balancing of the twelve dual-channel stimuli was done in four steps:

- i. The most comfortable level (MCL) of the 92-pps pulse train presented on the basal channel was obtained by first increasing its level until the subject found it slightly too loud and then decreasing it to reach MCL. The 129-pps pulse train on the basal channel was further loudness balanced to the 92-pps stimulus at MCL. Similarly, the 258-pps stimulus was balanced to the 129 pps and the 516 pps to the 258 pps. The levels of the 184-pps and 368-pps stimuli were obtained by logarithmic interpolation of the balanced levels found for the other rates.
- ii. Each rate-stimulus on the apical channel was then loudness-balanced to the same rate-stimulus presented on the basal channel at MCL (as obtained in (i)).
- iii. The dual-channel, short-delay stimuli were then constructed and the MCL for the 92-pps stimulus was determined by increasing the levels on both channels at the same time, keeping their level difference in dB (as found in (ii)) constant. Let d be the difference in dB between the MCLs of the 92-pps stimulus when presented alone on the basal channel and when presented in the dual-channel, short-delay condition. The higher-rate, short-delay stimuli levels were either inferred from the 92-pps dual-channel stimulus by reducing, for each rate, the levels obtained in the single-channel case by d or, if time allowed, were formally loudness-balanced to the 92-pps short-delay stimulus. In the latter case, the results were consistent with d being constant at all rates. This is also consistent with the loudness model of McKay *et al.* (2003) which first performs a temporal integration on each individual channel and then sums the neural activity obtained at different cochlear locations. The value of d averaged across subjects was 1.5 dB.
- iv. Each of the six long-delay stimuli was finally balanced to the corresponding short-delay stimulus at the same rate.

The loudness-balancing results of S1 at the step (i) stage showed a large amount of variation. His results seemed to be particularly influenced by the starting level of the signal. The MCLs of all stimuli were therefore determined using a loudness chart and asking him to estimate the loudness of the sounds. The same four steps were repeated as for the other subjects except that there was no loudness balancing involved. It was noticed that this subject would rate an unusually wide range of levels as “most comfortable”.

Pitch Ranking—The aim of this experiment was to pitch rank the twelve dual-channel stimuli previously described. This was done using the midpoint comparison procedure developed by Long *et al.* (2005) and originally designed to optimise the fitting of auditory brainstem implants. This procedure consists of making pitch comparisons between pairs of sounds with the provisional pitch ordering being updated as more comparisons are made. To illustrate its functioning, assume a list of stimuli labelled A-Z that need to be ranked in pitch and that, at one point of the procedure, the provisional order is [F B G C A E D]. The next stimulus to be ranked (H) will first be compared to the middle-ranked stimulus of the list (C). If it is higher, the list will be bisected and H will be compared to the middle-ranked stimulus (E) of the higher part of the provisional list ([A E D]). Subsequently, H will be compared to D or to A depending on whether it was higher or lower than E. This will eventually lead to a new provisional ranking with more stimuli in the list.

The procedure was repeated at least 15 times with each subject and, for each of the twelve dual-channel stimuli, we calculated the mean and standard error of the rank. The order of presentation of the stimuli was randomized across the different repetitions.

B. Results

Fig. 2 shows, for each subject, the mean rank as a function of pulse rate for both short- (filled squares) and long- (open circles) delay stimuli. Although the data of the two delay conditions are connected using different lines and symbols, all stimuli were mixed within the same block. The short-delay function shows that pitch increased as a function of pulse rate up to about 258 pps for S2, S3 and S4 after which it started to asymptote. The three other subjects (S1, S5 and S6) perceived the pitch to increase up to the highest rate tested (516 pps).

It was initially expected that, at least at a low pulse rate (e.g. 92 pps, similar to the modulation rate used by McKay and McDermott (1996)), the pitch of the long-delay stimulus would be more similar to the pitch of a short-delay stimulus at twice the rate (184 pps), which would correspond to an exact doubling of the perceived pitch. At no instance did the results show this pattern. Overall, the pitch of the long-delay stimuli was only slightly higher than that of the short-delay stimuli when compared at the same rate. The data were analysed in a two-way repeated-measures ANOVA with rate and delay as treatment factors. The analysis shows the effects of pulse rate ($F(1,5)=99.9$, $p<0.001$) and delay ($F(1,5)=37.1$, $p=0.002$) to be highly significant. However, there was no significant interaction between the two factors ($F(1,5)=0.63$, $p=0.46$). The mean rank difference between short- and long-delay stimuli averaged across rate was relatively small, only 0.6 rank. As an example, the pitch of the long-delay stimulus at 92 pps was always lower than the pitch of the short-delay stimulus at 129 pps. Because consecutive rates are separated by approximately 40%, this means that the increase in pitch from short- to long-delay at 92 pps was always less than 40%.

The arrows in Fig. 2 point to the rates for which the rank difference between short- and long-delay stimuli was significant ($p<0.05$, Bonferroni-corrected). For the two highest rates tested (368 and 516 pps), the ranks were very similar for both delay conditions. Furthermore, the two functions (long- and short-delay) followed the same pattern. For example, S2, S3 and S4 showed a similar plateau as in the short-delay condition for rates higher than about 258 pps. These two observations suggest that dual-channel stimulation may not be useful to increase the upper limit of temporal pitch in CIs or, that if it is, this combination does not depend on inter-channel delay. The third, and somewhat puzzling, observation was that the individual rates where subjects showed a pitch difference between the two delay conditions were not necessarily the lowest rates. For example, S1 could not differentiate the pitches of the short- and long-delay stimuli at 92 pps but could do so at 258 pps. A similar “nonmonotonic” pattern was observed for S2, S3 and S5. Possible reasons for this trend will be further investigated in Section IV.

III. EXPERIMENT 2: Effects of electrode configuration and of channel order on the pitch of dual-channel stimuli

A. Rationale and Methods

The absence of any pitch doubling in the long-delay condition of Experiment 1 contrasted with the results of some of the subjects tested by McKay and McDermott (1996). Experiment 2 provides several additional conditions aimed to test the generality of our results.

First, it was checked that the results were not due to conflicting place and temporal cues due to the specific channel order that was used (basal-channel first). The same procedure was therefore repeated with a leading *apical* channel (cf. Fig. 1B). The levels on both channels were the same as in the basal-leading condition.

Second, we implicitly assumed in Section II that nearly simultaneous pulses would produce a temporal pattern of neural activity similar to truly simultaneous pulses. However, this needed to be confirmed. The same procedure was therefore repeated with both pulse trains presented on the same bipolar channel (cf. Fig. 1C). In this case, we would expect the second pulse of the short-delay stimulus either to fall within the absolute refractory period of the nerve due to the first pulse or to evoke spikes shortly after the first pulse. The pitch evoked by the long-delay stimulus should therefore be more similar to that of the short-delay stimulus at twice the rate. The twelve single-channel stimuli were loudness-balanced in a similar way as in Experiment 1. Five subjects performed this condition (all except S3). The electrodes tested were the same electrodes as those of the basal channel of Experiment 1 ([8, 10] for S1, S2 and S4; [12, 9] for S5; [12, 10] for S6).

Third, it is possible that the results of Experiment 1 were due to the two bipolar channels producing very focussed stimulation. The same procedure was therefore repeated in monopolar mode which (1) should theoretically produce a spatially broader excitation (e.g. Snyder *et al.*, 2008) and (2) is used in most contemporary CI speech-processing strategies. Four subjects (S1, S2, S5 and S6) performed this condition with a basal-leading and an apical-leading channel (cf. Fig. 1A, B). The levels were determined in the same way as described in Section II.A. The active basal and apical electrodes were number 8 and 7 for the Advanced Bionics subjects (S1-S2) and number 12 and 13 for the CI24 subjects (S5-S6), respectively. The return electrode was the extracochlear case electrode for the Advanced Bionics subjects and the two extracochlear reference electrodes for the CI24 subjects (so-called “MP 1+2 mode”).

B. Results

Apical-leading bipolar condition—Fig. 3 shows the results of the apical-leading condition in bipolar mode. The results were very similar to those of Experiment 1. A two-way repeated measures ANOVA showed the effects of rate ($F(1,5)=89.5$, $p<0.001$) and delay ($F(1,5)=13.9$, $p=0.014$) to be significant while the interaction factor was not ($F(1,5)=3.0$, $p=0.14$). To check for any interaction between channel order and delay, an additional three-way repeated measures ANOVA was performed on the mean data with pulse rate, delay and channel order as treatment factors. Of course, no main effect of channel order was expected as the two conditions (apical- and basal-leading) were performed in separate blocks and therefore had the same mean pitch rank. The analysis showed the effects of pulse rate ($F(1,5)=98.1$, $p<0.001$) and delay ($F(1,5)=28.9$, $p=0.003$) to remain highly significant. However, the interaction between delay and channel order ($F(1,5)=1.4$, $p=0.3$) was not significant, nor were any of the other interaction factors. This lack of an interaction between delay and channel order combined with the similarity of the results in the basal-leading and apical-leading conditions strongly suggest that the pitch difference between short- and long-delay stimuli reflects a difference in temporal rather than in place cues.

Single-channel condition—Fig. 4 shows the results of the single-channel condition. The long-delay stimulus is, here, equivalent to a regular, unmodulated pulse train (cf. Fig. 1C, bottom) at twice the nominal rate. The data of the long-delay condition are replotted as a function of the “true” pulse rate using open triangles (long-delay “shifted” function). For three subjects (S1, S5 and S6), the triangles and filled squares perfectly overlap, consistent with a doubling in the perceived pitch of the long-delay stimulus compared to that of the

short-delay stimulus. For S2, the trends of the short- and long-delay functions were similar (non-monotonic) although at 184 pps, the pitch of the long-delay stimulus was significantly higher than that of the short-delay condition (when compared at the same “true” rate). The results of S4 were very variable (showing large standard deviations) and the subject seemed to have been confused by the task. This may relate to the presence of several very high rate stimuli that were not discriminable. Another explanation could be that the short-delay stimuli had a different sound quality due to spikes being elicited by both of the pulses presented in each period and that this difference would prevent the subjects from comparing the stimuli on the basis of pitch. Overall, the large contrast between the single-channel (Fig. 4) and the dual-channel (Fig. 2 and 3) results reinforces our finding that two neighboring channels separated by only 0.75 or 1.1 mm do not convey the aggregate temporal code to higher levels of the auditory system.

Basal- and Apical-leading monopolar conditions—Fig. 5 shows the results of the basal-leading and apical-leading conditions in monopolar mode. Here again, the pitch differences between short- and long-delay stimuli were small. It is worth noting that the number of rates at which the pitches of the long- and short-delay stimuli differed significantly appeared to be larger for the CI24 (S5 and S6) than for the Advanced Bionics subjects. This may relate to the channel separation that was used (one electrode), which is smaller for the CI24 implant (0.75 vs. 1.1 mm.).

As in Experiment 1, two-way repeated-measures ANOVAs were performed separately for the basal-leading and apical-leading conditions. While the pulse rate still had a highly significant contribution on the perceived pitch, the effect of delay failed to reach significance in both cases ($F(1,3)=3.36$, $p=0.16$ for basal-first; $F(1,3)=3.68$, $p=0.15$ for apical-first). To check for any interaction effect between delay and electrode configuration, an additional four-way repeated measures ANOVA was performed on the mean data of the four subjects (S1, S2, S5 and S6) who took part in this sub-experiment. The treatment factors were the delay, the pulse rate, the stimulation mode (bipolar or monopolar) and the channel order (basal- or apical-first). Here again, no main effect of stimulation mode or channel order was expected as the four different conditions were run in different blocks and therefore had the same mean rank. The effects of pulse rate ($F(1,3)=266.2$, $p=0.001$) and delay ($F(1,3)=13.2$, $p=0.036$) were significant but there was no interaction between any of the treatment factors. Only the interaction between delay and channel order approached significance ($F(1,3)=7.9$, $p=0.067$). Furthermore, for a given subject, the rates showing a significant difference were often the same for the apical-first and basal-first conditions. These data suggest that the pitch difference between short- and long-delay stimuli is largely independent of the stimulation mode and of the channel order of stimulation.

IV. EXPERIMENT 3: Discrimination of inter-channel delay

A. Rationale and Methods

An intriguing result of Experiment 1 and 2 was that the long-delay stimuli were, for some subjects, perceived as higher in pitch than the short-delay ones only at intermediate rates. In this experiment, we investigated the discriminability of short- and long-delay stimuli using two different methods.

First, the stimuli of Experiment 1 (basal-leading, bipolar mode) were used in an odd-man out task. The main difference was that subjects could use any cues to perform the task, which consisted of a 3-interval, 2-alternative forced choice procedure. The first interval was fixed and always contained the short-delay stimulus. The second and third intervals were randomly assigned the short- or long-delay stimulus. Subjects were asked to indicate which of the second or third intervals was different from the other two, i.e. which one contained the

long-delay stimulus. Performance was measured at the two extreme rates used in the pitch ranking experiments (92 pps and 516 pps per channel) and at a third intermediate rate corresponding to the highest rate for which there was a significant pitch difference in Section II. The value of this intermediate rate differed across subjects (129 pps for S3, 184 pps for S2 and S6, and 258 pps for S1 and S5, none for S4). The experiment consisted of blocks of 20 trials of the same rate condition which were repeated between two and five times each. The different pulse rate conditions alternated from block to block.

Second, regarding the initial hypothesis that more central neurons are able to combine the temporal information conveyed by nearby auditory nerve fibers, a possibility would be that these central neurons can only do so at some specific rates. To investigate this, we measured inter-channel delay DLs at the highest rate for which there was a significant pitch difference in Experiment 1 and compared it to inter-pulse interval DLs using a single bipolar channel at the same pulse rate. The hypothesis was that if the DLs are smaller in the dual-channel than in the single-channel case, then this would necessarily imply either that a more central process is able to combine the inputs from two distinct populations of AN fibers in order to enhance the inter-channel delay sensitivity or, alternatively, that one channel conveys more accurate timing information than the other. However, if the DLs are smaller in the single-channel case, a sufficient explanation would be that a small subset of AN fibers is being stimulated by both channels. The task was a 3-interval, 2-alternative forced-choice, 3-down, 1-up adaptive procedure with feedback (Levitt, 1971). The procedure stopped after eight reversals and each estimate was the average of the last six reversals. The standard was the short-delay stimulus and the delay of the signal was varied in steps of $\pm 194 \mu\text{s}$. The signal had an initial delay of half-a-period (equivalent to the long-delay stimulus of Section II). Each adaptive procedure was repeated at least three times. Only four subjects participated (S1, S2, S5 and S6) because S3 decided to withdraw from the experiment and S4 did not show any pitch difference at any of the rates tested in Section II. The electrodes tested in this task corresponded to the electrodes of the basal channel of the dual-channel stimuli ([8, 10] for S1-S2; [12, 9] for S5; [12, 10] for S6).

B. Results

Discriminability of short- and long-delay stimuli—The results of the odd-man out task are shown in Fig. 6A for rates of 92 pps (black), 516 pps (white) and for an intermediate rate whose value differed across subjects (gray). Discrimination performance was better at 92 pps than at 516 pps for all subjects. At 92 pps, four subjects (S2, S4-S6) showed scores higher than 95%. Although this could have been expected for S5 and S6 who showed a significant pitch difference in Experiment 1, it is worth noting that S2 and S4 did not consistently pitch rank the two delay stimuli but could easily discriminate them. At 516 pps, most subjects could not discriminate between the two different delays, showing scores close to chance. In the intermediate rate condition, all subjects except S5 showed high levels of discrimination. An interesting observation was that S1 and S3 were better at the intermediate rate than at 92 pps. This result is consistent with their ranking data showing a larger pitch difference at the intermediate rate than at 92 pps.

Inter-channel difference limens—The results of the delay discrimination experiments are illustrated in Fig. 6B. Black and white bars show the smallest detectable delay for the dual-channel and for the single-channel case, respectively. For all subjects, this delay was either the same or smaller in the single-channel condition. This suggests that the sensitivity to long-delay stimuli in the dual-channel case can be reasonably explained by the same process that occurs in the single-channel case, i.e. that a subset of auditory nerve fibers is being stimulated by both channels.

V. DISCUSSION

A. Main findings

We have investigated the effect of inter-channel delay on temporal pitch perception for dual-channel stimuli with a channel separation of 0.75 or 1.1 mm. We showed that:

1. The perceived pitch was, on average, slightly but significantly higher for the long- than for the short-delay stimuli. Nevertheless, it never matched the pitch corresponding to the aggregate temporal pattern.
2. This result was independent of the mode of stimulation (bipolar or monopolar) and of the channel order (basal- or apical-first).
3. The upper limit of temporal pitch was not improved using long delays. In particular, subjects could not discriminate between short- and long-delay stimuli at the highest rate tested (516 pps).
4. The largest pitch differences between short- and long-delay were often obtained at intermediate and not at the lowest rate.
5. In an odd-man-out task, all subjects could discriminate between short- and long-delay stimuli at 92 pps, but most of them performed at chance at 516 pps.
6. Delay difference limens were smaller for single channel than for dual-channel stimuli when measured at rates for which there was a significant pitch difference between short- and long-delay stimuli.

The primary aim of this series of experiment was to evaluate the hypotheses of McKay and McDermott (1996) and to test whether subjects would perceive the aggregate temporal pattern (1) because of auditory nerve fibers responding to both channels or (2) because of a more central process combining inputs originating from two distinct populations of auditory nerve fibers. Our results are not consistent with either of these hypotheses as the perceived pitch never matched the pitch expected from the aggregate temporal pattern. The fact that the pitch was slightly higher for the long-delay than for the short-delay stimuli is consistent with most neurons conveying the single-channel temporal code and only a small subset responding to both channels and conveying the aggregate temporal code.

The fact that the pitch difference was sometimes larger at intermediate rates than at the lowest rate tested is intriguing. This may relate to the salience of the pitch percept being clearer at some rates than others. For example, if at a given rate, the aggregate pattern was conveying a more salient pitch cue than the individual channel patterns, this could make the subject “listen” preferentially to the neurons conveying the aggregate pattern. Incidentally, Kong *et al.* (2009) showed that rate discrimination by CI subjects was sometimes better at 200-300 pps than at 100 pps, suggesting that the transmission of temporal cues does not always degrade monotonically with increases in pulse rate.

B. Comparison to previous studies

Previous studies on dual-channel stimulation have been performed at a pulse rate or a modulation rate of 100 pps per channel (Tong and Clark, 1986; McKay and McDermott, 1996; Carlyon *et al.*, 2000). For small channel separations, these authors observed high levels of discrimination performance. The results of our discrimination experiment at 92 pps are consistent with these data and show that most of our subjects could reliably perceive a difference between short- and long-delay stimuli.

In their pitch scaling study, McKay and McDermott (1996) found that three out of five subjects could perceive a pitch consistent with the aggregate temporal pattern whereas this

was never the case in our experiment. There are, however, several differences between their study and ours. First, they compared the pitches of amplitude-modulated pulse trains having different inter-channel delays with the pitches of *unmodulated* pulse trains with a half-period delay (identical to our “long-delay” condition) having different rates. It is possible that these two groups of stimuli had different sound qualities and were therefore difficult to compare. In the present experiment, we only used unmodulated pulse trains and compared their pitches by varying both the inter-channel delay and the rate. Second, as suggested by the facts that our patients were implanted more recently than those tested by McKay and McDermott, and that criteria for implantation have become more relaxed over time, our subjects may have had better neural survival. If this is true, we would expect our subjects to show less overlap in the neural spreads of excitation produced by nearby channels. Third, they used a pitch scaling procedure where they asked subjects to give a numerical estimate of each stimulus individually. Such a method may be prone to several kinds of non-sensory response biases that can affect the accuracy of the pitch match (Poulton, 1979). In contrast, we performed direct pitch comparisons between the stimuli using a ranking task. Finally, and perhaps most importantly, all subjects tested by McKay and McDermott were users of the Mini System 22 implant. This device has banded electrode contacts assumed to lie along the outer wall of the scala tympani (so-called “straight” electrode). In contrast, all our subjects except S6 have an electrode array designed to have a perimodiolar placement (HiFocus electrode for S1-S4 and Contour electrode for S5). Perimodiolar electrodes presumably lie closer to the excitable neural elements and have been shown to yield lower thresholds and narrower forward-masking profiles (Cohen et al. 2006). Therefore, it is possible that our long-delay stimuli did not correspond to the aggregate temporal pattern because each individual channel produced a more spatially-selective stimulation than in the McKay and McDermott study. Incidentally, S6, who has a straight (outer-wall placement) electrode was the subject for whom the pitch difference between short- and long-delay stimuli was the largest at several rates.

C. Implications for cochlear implants

One of the most commonly-cited reasons for the limitations experienced by cochlear implant (CI) subjects in a range of tasks is the spread of current between neighboring electrode channels. Large current spreads have been reported in physiological studies for both monopolar and bipolar configurations (Snyder *et al.*, 2008). In human CI experiments, performance on speech recognition tasks usually does not improve when the number of channels is increased above approximately eight (Friesen *et al.*, 2001), suggesting that the number of functional channels is smaller than the number of implanted intracochlear electrodes. The fact that the perceived pitch of our dual-channel stimuli did not correspond to the pitch of the aggregate temporal pattern suggests that the temporal codes conveyed by neighboring channels are largely independent. In other terms, although the neural populations excited by neighboring channels may be overlapping, it is possible that the neurons conveying the meaningful temporal codes of each channel are located in spatially-restricted regions that do not overlap much between channels. The fact that several subjects were able to discriminate between long- and short-delay stimuli but were not able to pitch-rank them would support such a hypothesis. Another alternative explanation could be that the response to one or to both of the channels spreads away, leading to “off-place listening” (cf. Dingemans *et al.*, 2006).

Acknowledgments

We thank Colette McKay for helpful discussions. This study was supported by the Wellcome trust, grant number 080216.

REFERENCES

- Carlyon, RP. Detecting F0 differences and pitch-pulse asynchronies. In: Schouten, B., editor. *The Auditory Processing of Speech. From Sounds to Words*. Mouton-De Gruyter; Berlin: 1992. p. 149-156.
- Carlyon RP, Geurts L, Wouters J. Detection of small across-channel timing differences by cochlear implantees. *Hear. Res.* 2000; 141:140–154. [PubMed: 10713502]
- Carlyon RP, Deeks JM. Limitations on rate discrimination. *J. Acoust. Soc. Am.* 2002; 112:1009–1025. [PubMed: 12243150]
- Carlyon RP, Shamma S. An account of monaural phase sensitivity. *J. Acoust. Soc. Am.* 2003; 114:333–348. [PubMed: 12880045]
- Cohen LT, Saunders E, Knight MR, Cowan RS. Psychophysical measures in patients fitted with Contour and straight Nucleus electrode arrays. *Hear. Res.* 2006; 212:160–175. [PubMed: 16403611]
- de Balthasar C, Boéx C, Cosendai G, Valentini G, Sigrist A, Pelizzone M. Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects. *Hear. Res.* 2003; 182:77–87. [PubMed: 12948604]
- Dingemans JG, Frijns JH, Briare JJ. Psychophysical assessment of spatial spread of excitation in electrical hearing with single and dual electrode contact maskers. *Ear. Hear.* 2006; 27:645–657. [PubMed: 17086076]
- Donaldson GS, Kreft HA, Litvak L. Place-pitch discrimination of single- versus dual-electrode stimuli by cochlear implant users. *J. Acoust. Soc. Am.* 2005; 118:623–626. [PubMed: 16158620]
- Friesen LM, Shannon RV, Baskent D, Wang X. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J. Acoust. Soc. Am.* 2001; 110:1150–1163. [PubMed: 11519582]
- Friesen LM, Shannon RV, Cruz RJ. Effects of stimulation rate on speech recognition with cochlear implants. *Audiol. Neurootol.* 2005; 10:169–184. [PubMed: 15724088]
- Kong YY, Deeks JM, Axon PR, Carlyon RP. Limits of temporal pitch in cochlear implants. *J. Acoust. Soc. Am.* 2009; 125:1649–1657. [PubMed: 19275322]
- Laneau J, Boets B, Moonen M, van Wieringen A, Wouters J. A flexible auditory research platform using acoustic or electric stimuli for adults and young children. *J. Neurosci. Methods.* 2005; 142:131–136. [PubMed: 15652626]
- Levitt H. Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 1971; 49:467–477. [PubMed: 5541744]
- Long CJ, Nimmo-Smith I, Baguley DM, O'Driscoll M, Ramsden R, Otto SR, Axon PR, Carlyon RP. Optimizing the clinical fit of auditory brain stem implants. *Ear. Hear.* 2005; 26:251–262. [PubMed: 15937407]
- McKay CM, McDermott HJ. The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. *J. Acoust. Soc. Am.* 1996; 100:1081–1092. [PubMed: 8759961]
- McKay CM, McDermott HJ. The perceptual effects of current pulse duration in electrical stimulation of the auditory nerve. *J. Acoust. Soc. Am.* 1999; 106:998–1009. [PubMed: 10462805]
- McKay CM, Henshall KR, Farrell RJ, McDermott HJ. A practical method of predicting the loudness of complex electrical stimuli. *J. Acoust. Soc. Am.* 2003; 113:2054–2063. [PubMed: 12703716]
- Middlebrooks JC. Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *J. Acoust. Soc. Am.* 2004; 116:452–468. [PubMed: 15296005]
- Middlebrooks, JC.; Snyder, RL. Enhanced Transmission of Temporal Fine Structure Using Penetrating Auditory Nerve Electrodes. Association for Research in Otolaryngology, 32nd Midwinter Research Meeting; Baltimore, MD, USA. 2009.
- Oxenham AJ, Bernstein JG, Penagos H. Correct tonotopic representation is necessary for complex pitch perception. *Proc. Natl. Acad. Sci. USA.* 2004; 101:1421–1425. [PubMed: 14718671]
- Poulton EC. Models for biases in judging sensory magnitude. *Psychological Bull.* 1979; 86:777–803.
- Rose JE, Brugge JF, Anderson DJ, Hind JE. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J. Neurophysiol.* 1967; 30:769–93. [PubMed: 4962851]

- Snyder RL, Middlebrooks JC, Bonham BH. Cochlear implant electrode configuration effects on activation threshold and tonotopic selectivity. *Hear. Res.* 2008; 235:23–38. [PubMed: 18037252]
- Tong YC, Clark GM. Loudness summation, masking, and temporal interaction for sensations produced by electric stimulation of two sites in the human cochlea. *J. Acoust. Soc. Am.* 1986; 79:1958–1966. [PubMed: 3722606]
- Townshend B, Cotter N, Van Compernelle D, White RL. Pitch perception by cochlear implant subjects. *J. Acoust. Soc. Am.* 1987; 82:106–115. [PubMed: 3624633]

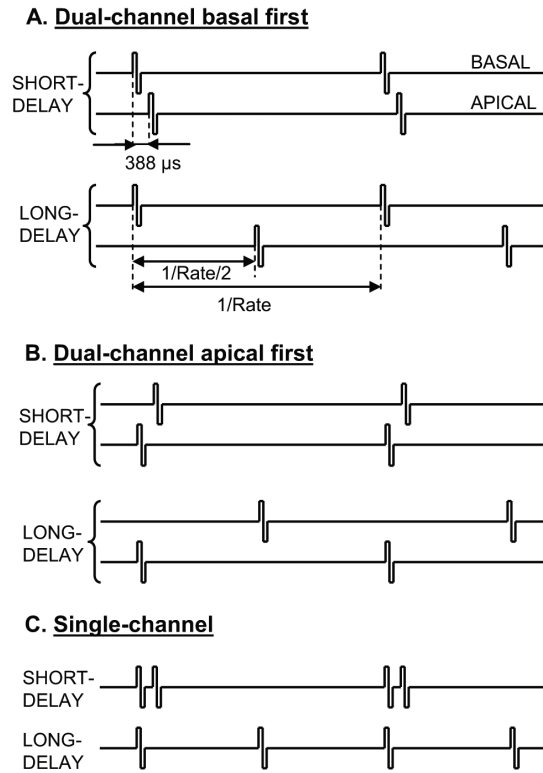


Figure 1. Schematic representation of the stimuli used in the study. **A**, Dual-channel stimuli with a basal-leading channel for the short- (388 μ s) and long-delay (half-a-period) conditions. **B**, Same as A with an apical-leading channel. **C**, Same as A and B for single channel stimulation. The two pulse trains are merged on the same channel.

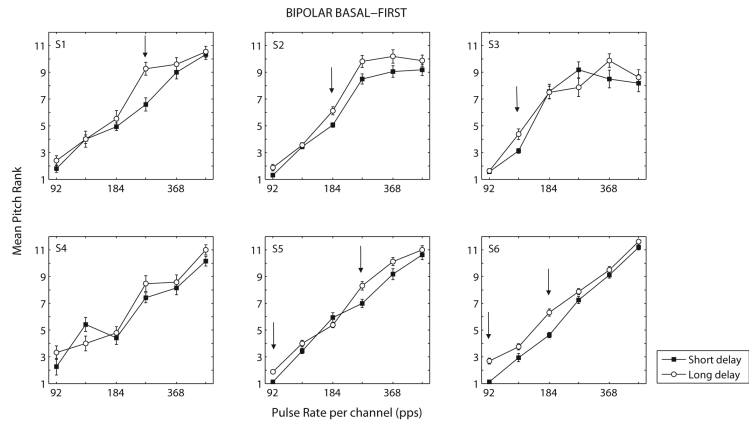


Figure 2. Results of Experiment 1 (bipolar, basal-leading dual-channel condition). Mean pitch rank as a function of pulse rate per channel for the six subjects. Errorbars indicate \pm standard error.

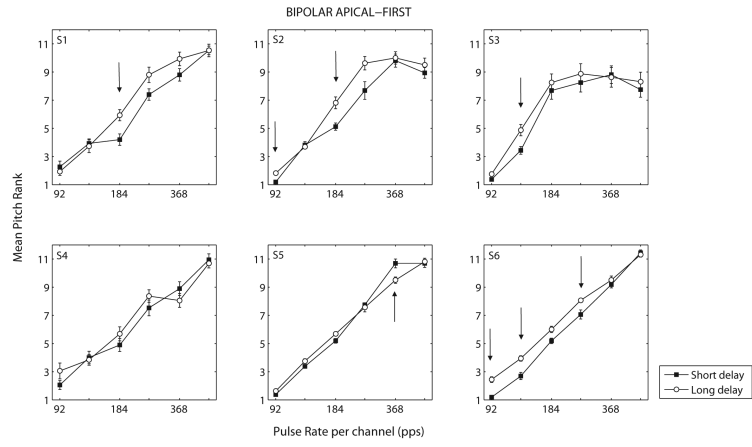


Figure 3. Results of Experiment 2 for the bipolar, apical-leading dual-channel condition. Mean pitch rank as a function of pulse rate per channel for the six subjects. Errorbars indicate +/- standard error.

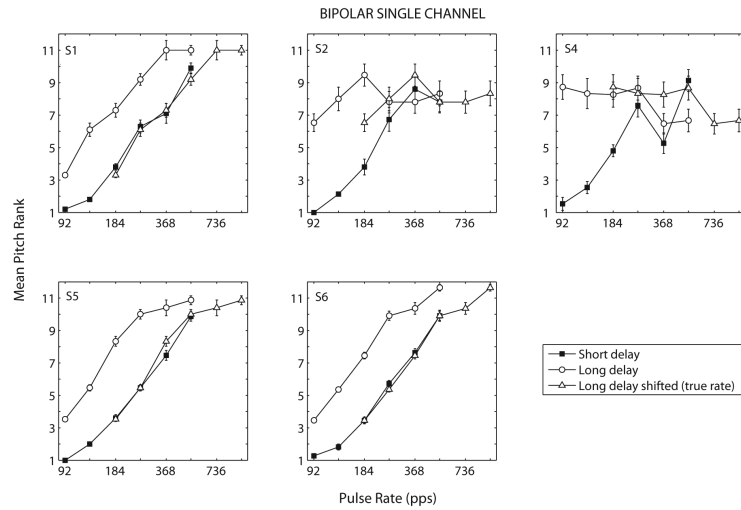


Figure 4. Results of Experiment 2 for the single-channel condition. Mean pitch rank as a function of pulse rate for five subjects. Errorbars indicate \pm standard error. The open circles show the rank of the long-delay stimuli as a function of rate as defined in Fig. 1A. Open triangles show the same data plotted as a function of the “true” rate on the single channel.

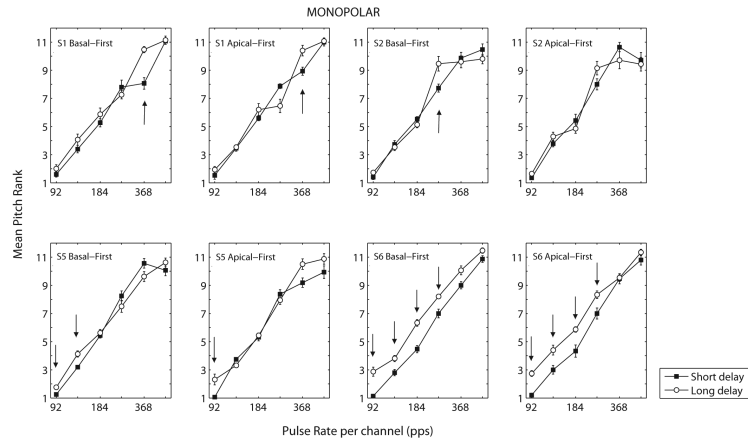


Figure 5. Results of Experiment 2 for the monopolar, basal- and apical-leading dual-channel conditions. Mean pitch rank as a function of pulse rate for four subjects. Errorbars indicate +/- standard error.

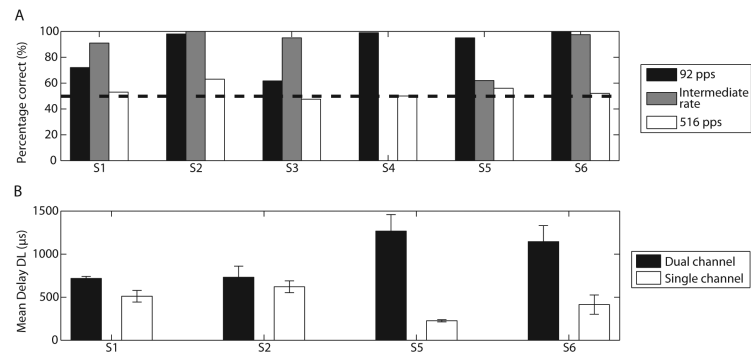


Figure 6. Results of Experiment 3. **A**, Percentage correct in the odd-man out task at three different rates for the six subjects. Only S4 did not perform the intermediate rate condition because he could not reliably pitch rank short- and long-delay stimuli at any of the rates (cf. Fig. 2). **B**, Smallest detectable delay in μs at the intermediate rate for dual- and single-channel stimuli.

Table 1

Biographical and clinical details of the subjects who took part in the study. The last three columns show the stimulus parameters used in the bipolar dual-channel conditions.

Subject	Age (years)	Etiology	Duration of deafness	Duration of implant use	Implant type	Phase duration (μ s)	Basal channel	Apical channel
S1	64	Progressive	> 12 years	6 months	HR90K	97	(8,10)	(7,9)
S2	73	Unknown sudden	< 2 years	6 years	CII	97	(8,10)	(7,9)
S3	61	Unknown progressive	34 years	5 years	HR90K	97	(8,10)	(7,9)
S4	54	Meningitis	1 year	6 years	CII	97	(8,10)	(7,9)
S5	60	Sudden viral	> 30 years	3 years	CI24R (CA)	150	(12,9)	(13,10)
S6	63	CSOM	10 years	10 years	CI24M	150	(12,10)	(13,11)