
Volume 269, issues 1–2, 1 October 2010
Complete volume
ISSN 0378-5955
ELSEVIER
Hearing Research

Hearing Research 269 (2010) 34–41

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

What breaks a melody: Perceiving F0 and intensity sequences with a cochlear implant

Marion Cousineau [a,b,*], Laurent Demany [c], Bernard Meyer [d], Daniel Pressnitzer [a,b]

[a] Laboratoire Psychologie de la Perception, UMR 8158 CNRS – Université Paris Descartes, 45 rue des Saints-Pères, F-75006 Paris, France
[b] Département d'études cognitives, École normale supérieure, 29 rue d'Ulm, F-75005 Paris, France
[c] Laboratoire Mouvement, Adaptation, Cognition (UMR CNRS 5227), BP 63, Université de Bordeaux, 146 rue Leo Saignat, F-33076 Bordeaux, France
[d] APHP – Hopital Beaujon, Service d'Oto-Rhino-Laryngologie, 100, Bd du général Leclerc, F-92110 Clichy, France

ARTICLE INFO

Article history:
Received 17 March 2010
Received in revised form
16 July 2010
Accepted 20 July 2010
Available online 30 July 2010

ABSTRACT

Pitch perception has been extensively studied using discrimination tasks on pairs of single sounds. When comparing pitch discrimination performance for normal-hearing (NH) and cochlear implant (CI) listeners, it usually appears that CI users have relatively poor pitch discrimination. Tasks involving pitch sequences, such as melody perception or auditory scene analysis, are also usually difficult for CI users. However, it is unclear whether the issue with pitch sequences is a consequence of sound discriminability, or if an impairment exists for sequence processing *per se*. Here, we compared sequence processing abilities across stimulus dimensions (fundamental frequency and intensity) and listener groups (NH, CI, and NH listeners presented with noise-vocoded sequences). The sequence elements were firstly matched in discriminability, for each listener and dimension. Participants were then presented with pairs of sequences, constituted by up to four elements varying on a single dimension, and they performed a same/different task. In agreement with a previous study (Cousineau et al., 2009) fundamental frequency sequences were processed more accurately than intensity sequences by NH listeners. However, this was not the case for CI listeners, nor for NH listeners presented with noise-vocoded sequences. Intensity sequence processing was, nonetheless, equally accurate in the three groups. These results show that the reduced pitch cues received by CI listeners do not only elevate thresholds, as previously documented, but also affect pitch sequence processing above threshold. We suggest that efficient sequence processing for pitch requires the resolution of individual harmonics in the auditory periphery, which is not achieved with the current generation of implants.

1. Introduction

The cochlear implant, a surgically-implanted device that bypasses cochlear processing to directly stimulate the auditory nerve, has been used to restore auditory function in many individuals with profound deafness. The original aim of the implant design was, understandably, to enable speech intelligibility. But whereas speech intelligibility in quiet can be achieved with a coarse representation of acoustic information (Shannon et al., 1995), other auditory abilities may require acoustic cues that are not currently available to cochlear implant (CI) users. In particular, pitch perception seems to be impaired when using a cochlear implant (for reviews, see McDermott, 2004; Moore and Carlyon, 2005; Drennan and Rubinstein, 2008). This is problematic, as accurate processing of pitch patterns is essential for speech perception (intonation, tonal languages), music perception (melodies), or auditory scene analysis (streaming, speech in noise). A better understanding of which aspects of pitch patterns' perception are impaired in CI users is therefore of fundamental importance.

Pitch perception in CI users has often been assessed using pitch-ranking tasks with acoustically presented stimuli (Geurts and Wouters, 2001; Laneau et al., 2004; McDermott, 2004; Pressnitzer et al., 2005; Vandali et al., 2005; Sucher and McDermott, 2007; Looi et al., 2008a, 2008b). A similar procedure was used in all of these studies: two complex sounds (often sung vowels) that differed

Abbreviations: ANOVA, analysis of variance; CI, cochlear implant; dB, decibel; F0, fundamental frequency; F condition, sequences varying in F0; FSD, frequency-shift detector; I condition, sequences varying in intensity; N, number of elements in a sequence; NH, normal-hearing; NH-voc, normal-hearing listeners presented with noise-vocoded stimuli; ST, semitone.

* Corresponding author. Tel.: +33 1 4432 2673; fax: +33 1 4432 3610.
E-mail addresses: Marion.Cousineau@ens.fr (M. Cousineau), laurent.demany@u-bordeaux2.fr (L. Demany), bernard.meyer@sat.ap-hop-paris.fr (B. Meyer), Daniel.Pressnitzer@ens.fr (D. Pressnitzer).

0378-5955/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.heares.2010.07.007

in fundamental frequency (F0) were presented. Listeners had to rank them on the pitch dimension by indicating which sound was higher in pitch, and threshold was estimated as the smallest F0 difference producing a consistent ranking. Results showed average thresholds much poorer than those observed for normal-hearing (NH) listeners (e.g. around 10% of F0 for the best performers in the CI groups, compared to less than 0.5% for NH, Geurts and Wouters, 2001; Pressnitzer et al., 2005). Moreover, a prominent feature of all results is a large inter-subject variability: for a 50% difference in F0 (7 semitones), performance of CI users may range from chance to near perfect (McDermott, 2004).

Another type of measure has focused on the processing of pitch sequences that extend over time, because of their immediate relevance to music perception (Cooper et al., 2008; Fujita and Ito, 1999; Galvin et al., 2007; Gfeller et al., 2002; Kong et al., 2004; Looi et al., 2008a, 2008b; Pressnitzer et al., 2005; Singh et al., 2009). In some of these studies, familiar melodies were to be recognized from a closed set (e.g. Fujita and Ito, 1999) whereas in others, simple melodic contours were to be discriminated (e.g. Galvin et al., 2007). Results generally indicate that melody perception is poor in CI users, with, again, a large variability between subjects. Galvin et al. (2007), for instance, obtained results ranging from chance to near-NH performance in a 5-note melodic contour identification task.

There are many potential sources for the variability reported across these studies, from different processing strategies to various nerve survival rates in individual CI listeners (Moore and Carlyon, 2005). It is unclear, moreover, if the variability observed in the pitch sequence tasks is simply a reflection of the diverse pitch-ranking abilities of individual CI users. Obviously, a large pitch-ranking threshold should induce poor pitch sequence representation. Looi et al. (2004) found that subjects' ability to rank pitches was correlated with their ability to recognize melodies. However, it is also possible that, in addition, pitch sequence processing *per se* is impaired in CI users. For NH listeners, McFarland and Cacace (1992) suggested that sequences of pitch were more accurately processed than sequences of loudness, even though the discriminability between single elements of the sequences was approximately equated on each dimension. Cousineau et al. (2009), using a method that took into account the exact discriminability thresholds of individual listeners on each dimension, confirmed the advantage for pitch sequence processing for NH listeners. Pitch sequence discrimination performance was found to be superior to loudness sequence discrimination performance, presumably because of contour-encoding mechanisms available only for pitch (Demany and Ramos, 2005; Demany et al., 2009). In addition, it was found that the pitch sequence advantage was restricted to sounds made of resolved harmonics; pitch sequences made up of complex tones without any resolved harmonic were processed no more accurately than loudness sequences.

The latter finding leads to the prediction that CI users, being generally unable to resolve the individual harmonics of complex tones (Laneau et al., 2004), may suffer from a specific impairment in pitch sequence processing, independent of their pitch-ranking abilities. The following experiments were designed to test this hypothesis. We used the psychophysical method of Cousineau et al. (2009) to test the perception of sequences varying in either F0 or intensity in three groups of listeners: CI users, NH listeners, and NH listeners presented with noise-vocoded sequences (NH-voc). Importantly, the method aimed to uncouple sequence processing performance from any difference in terms of pitch discriminability that was expected between (and within) the different groups of listeners.

2. Methods

2.1. Subjects

2.1.1. CI group

The CI group consisted of five post-lingually deafened adult CI users ($M = 65.4$ years, $SD = 9.8$). These listeners used a variety of implanted devices and processing strategies, which were set to their recommended settings during the experiment. Some relevant details about individual listeners can be found in Table 1. All listeners had already participated in psychophysical experiments and were relatively good performers on consonant and vowel closed-set identification in silence (Poncet-Wallet et al., 2008; Table 1). All listeners were fully informed about the goal of the present study and provided written consent before their participation. The study was carried out in accordance with the declaration of Helsinki. Clinical investigation with CI listeners received prior approval of the French "Regional Ethics Committee" CPP Ile de France VI.

2.1.2. NH group

The NH group consisted of seven young normal-hearing listeners with no self-reported hearing deficit ($M = 25.7$ years; $SD = 2.3$). These listeners were tested using unprocessed stimuli.

2.1.3. NH-voc group

The NH-voc group consisted of five other young normal-hearing listeners with no self-reported hearing deficit ($M = 26.0$ years; $SD = 3.4$), who were tested using noise-vocoded stimuli.

2.2. Stimuli

2.2.1. Unprocessed

Sequences varying in F0 and sequences varying in intensity were used. For the F0 sequences, broadband complex tones were generated by filtering click trains between 0.1 kHz and 10 kHz. The F0 of the filtered click trains was close to 125 Hz, well within the voice-pitch range and with a periodicity that is appropriate for melody perception in CI listeners (Pijl and Schwarz, 1995). For the intensity sequences, pink noises were generated in the frequency domain with spectral energy between 0.1 and 10 kHz (−3 dB/oct).

Table 1
Information about the CI participants in the study.

Subject	A	B	C	D	E
Age	82	56	64	64	61
Etiology of Deafness	Otospongiose	Progressive	Otospongiose	Progressive	Family Deafness
Duration of deprivation (years)	1	13	37	6	1
Duration post implantation (month)	34	12	60	62	19
Implant brand	MXM	Cochlear	MXM	Medel	Medel
Processor type	Digi SP	Freedom	Digi BTE	Tempo +	Opus2
Processing strategy	MPIS	ACE MP1+2	MPIS	CIS +	FSP (fine structure)
Performance VCV in silence (% correct)	38.0	75.5	34.4	64.7	70.3
Performance CVC in silence (% correct)	67.4	54.6	52.3	26.5	47.0

All sound elements had a 200-ms duration, including 25-mc onset and offset cosine ramps. For the CI listeners, the average level of stimuli was adjusted to be comfortable, and was kept constant throughout the experiment. For the NH listeners, the level of the stimuli was set close to 65 dB SPL (sound pressure level).

Sequences were constructed by presenting several sound elements consecutively, with no temporal gap between them. For the F0 sequences (F0 condition), the F0 of the elements was varied and it could only take one of two possible values, 125 Hz and 125 Hz + ΔHz. ΔHz was adjusted for each listener (see below). For the intensity sequences (I condition), the sound pressure level was varied and, similarly, could only take one of two possible values: 65 dB and 65 dB + ΔdB. Examples of such "binary" sequences are given in Fig. 1.

2.2.1. Noise-vocoded

The sequences described above were processed by an 8-channel noise vocoder (Shannon et al., 1995). Sounds were first passed through a bank of overlapping sixth-order Butterworth bandpass filters. The filters covered the full range of the stimuli (from 0.1 to 10 kHz). The frequency cutoffs were computed according to Greenwood's formula (1990) so as to simulate equally spaced electrodes in the cochlea (-3 dB points: 100, 245, 477, 846, 1433, 2368, 3856, 6226, and 10000 Hz). The envelope in each frequency band was extracted using a second-order Butterworth filter with a cutoff frequency of 400 Hz, after half-wave rectification. For each channel, the envelope was imposed on Gaussian noise, and the modulated noise was then bandpass filtered by the filter corresponding to the analysis channel. All channels were summed together to produce the noise-vocoded stimuli. Finally, all stimuli were equalized in *rms* level. An illustration of the effect of noise-vocoding on the F0 stimuli is shown in Fig. 2.

2.2.2. CI listeners

Stimuli were presented acoustically to CI listeners. This was to approximate realistic listening conditions, and also to be able to test a variety of devices. An illustration of the typical electrical stimulation pattern for the F0 stimuli is provided in Fig. 2 (simulation of the Nucleus® Freedom™ with the ACE processing strategy, as worn by CI subject B). As expected, the information about the stimulus periodicity can only be seen in the temporal envelope of the stimulation for each electrode, which is clearly modulated at F0. There is no spectral cue, as all active channels (12 out of 22 in the mid-frequency range, due to the coding strategy and simulated microphone frequency response) have a similar stimulation rate.

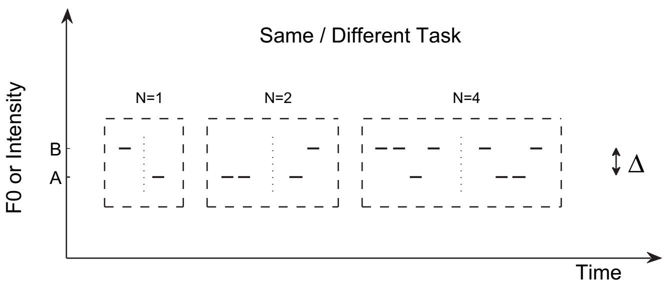

Fig. 1. Illustration of the sequence task. A same/different sequence task was used in the main part of the experiment, with different sequence lengths (N elements). Binary sequences were randomly constructed from only two sounds (A and B), differing in either F0 or SPL. The value of the difference, Δ, was adjusted for each listener and condition in order to yield equivalent discriminability for $N = 1$.

2.2.3. Apparatus

Each subject was tested individually, in a quiet room of the St Antoine hospital, Paris, for the CI listeners, and in a double-walled soundproof booth (Industrial Acoustics) for the NH and NH-voc listeners. Sound was played at a sampling rate of 44.1 kHz through an Echo Indigo sound card for CI listeners and an RME Fireface sound card for NH and NH-voc listeners. For consistency, sound was delivered diotically through Sennheiser HD250 linear II headphones for all groups of listeners. Two CI listeners (subjects A and D) had residual hearing in the non-implanted ear. They were asked to turn off their hearing aid during the experiment. Listeners provided their responses by means of a computer interface, in a self-paced manner, with no intervention of the experimenter for any of the experimental group.

2.3. Procedure

The experiment included a preliminary adjustment step, where differences in F0 or SPL were adjusted for each condition and listener so as to yield an equal level of discriminability across listeners and conditions. This helped to familiarize listeners with the experiment, even though no specific training was provided. Subsequently, in the main part of the experiment, sequence discrimination performance was measured. The whole experiment lasted for about 2.5 h and was run in two sessions separated by an average of 3.6 days.

2.3.1. Adjustment step

On each trial, the listener was presented with two successive sounds that were either identical or different; a same/different judgment had to be made. When different, the two sounds differed in F0 or SPL by a small value, Δ. The index of sensitivity d' (Green and Swets, 1966) was measured. Blocks of 50 trials were run with different Δ values until a Δ value was found that produced a stable level of performance of $d' = 2$. No feedback was provided. The number of blocks and the changes in Δ were chosen heuristically by the experimenter. Note that the accuracy of this initial adjustment step was independently verified in the main part of the experiment.

2.3.2. Sequence discrimination task

The values of Δ selected following the adjustment step described above were used to construct binary sequences of $N = 1, 2$ or 4 sounds. On each trial, two sequences separated by a 400-ms silent interval were presented (see Fig. 1). In the first sequence, each tone was, at random, either a reference stimulus, A, or another stimulus, B, differing from A by Δ Hz in F0 for the F0 condition, or by Δ dB in SPL for the I condition. The second sequence was equiprobably identical to the first sequence or different from it. In the latter case, a single, randomly chosen element was changed from A to B or vice versa. Listeners had to make a same/different judgment on the two sequences. For each listener, condition, and *N* value, four blocks of 50 trials were run. The ordering of conditions and *N* values was randomized within each group of subjects (CI, NH, NH-voc). No feedback was provided.

3. Results

3.1. Adjustment step

The mean Δ values obtained during the adjustment step are shown in Fig. 3, with the corresponding individual Δ values listed in Table 2. For the NH group, the mean Δ value was 0.39 semitones (ST) for the F0 condition, and 2.95 dB for the I condition. For the CI group, the mean Δ value was 7.80 ST for the F0 condition and

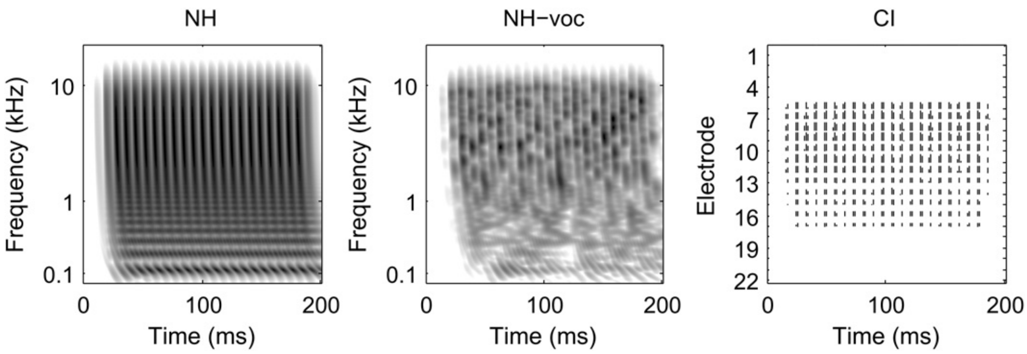

Fig. 2. Illustration of experimental stimuli. Left and Middle panels: The F0 stimuli were run through a gammatone filterbank (Patterson et al., 1995) to simulate peripheral auditory filtering. Left: The unprocessed F0 stimuli were broadband harmonic complex tones, containing both resolved harmonics (visible as horizontal stripes in the low-frequency channels) and unresolved harmonics (producing beats at the fundamental frequency F0 in the remaining channels). Middle: The processed stimuli were generated by applying an eight-channel noise vocoder, with an envelope cutoff frequency of 400 Hz. Envelope modulation at the fundamental frequency is still visible, without any obvious spectral cues. Right panel: Typical electrical stimulation pattern for the F0 stimuli after cochlear implant processing. The electrodogram was obtained with the Nucleus Matlab Toolbox provided by one of the reviewers, to simulate the Cochlear Nucleus® Freedom™ processor. Temporal cues to F0 are visible, with no spectral cue.

5.66 dB for the I condition. For the NH-voc group, the mean Δ value was 1.98 ST for the F0 condition and 3.18 dB for the I condition.

For statistical testing, we applied a log-transform to the data for the F0 condition, in order to make the standard deviations comparable across groups of listeners. We also applied a Bonferroni correction to all *t*-tests in the paper, so a Type I risk of 5% corresponds to $p = 0.017$. In the F0 condition, *t*-tests for independent samples showed that the Δ values for NH listeners were significantly smaller than for CI listeners ($p = 0.0001$) and for NH-voc listeners ($p = 0.0006$). There was no significant difference between the CI and NH-voc groups ($p = 0.026$) in spite of a trend for smaller Δ values on average for the NH-voc group. In the I condition, the *t*-tests for independent samples revealed no significant differences between the three groups.

It is important to keep in mind that stimuli were presented acoustically to CI listeners, who were using different devices with their own preferred settings. In particular, no attempt was made to equalize the effect of the automatic gain control during the intensity discrimination task. Therefore, it is likely that the intensity discrimination thresholds for CI listeners were influenced by both subject- and hardware-dependent factors. The adjustment step served to equate all of these factors for the main experiment.

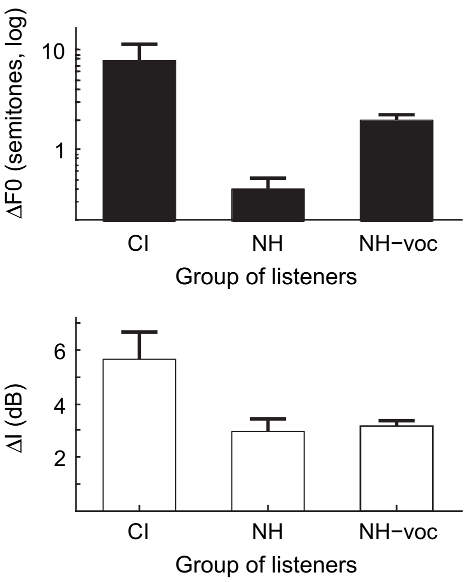

Fig. 3. Adjustment results. Δ values used to obtain equal discriminability in F0 or intensity, for the different groups of listeners. The mean Δ values are plotted as semitones for the F0 condition (black bars) and dB for the intensity condition (white bars). Error bars represent ± 1 standard error about the mean.

3.2. Sequence discrimination task

The mean results for the sequence discrimination task are shown in Fig. 4. The performance for $N = 1$, that is, "sequences" composed of a single element, represents a control for the accuracy of the adjustment step. Performance was similar for all groups of listeners and conditions for $N = 1$, which indicates that the adjustment step was on average successful in equating discriminability across listeners and conditions.

For the NH group, the effect of N was qualitatively different in the F0 and I conditions: in the I condition, performance decreased regularly when more elements were added to the sequences, whereas in the F0 condition, performance remained the same for sequences of $N = 1$, 2 or 4 elements. In contrast, for the CI and NH-voc listeners, results were similar in the F0 and I conditions: performance always decreased when more elements were added to the sequences. Finally, performance in the I condition was similar across the three groups.

Statistical analyses confirmed these observations. An ANOVA was performed on the d' values with two within-subject factors: Condition (F0 or I) and N (1, 2 or 4). The contrast between groups

Table 2
Individual Δ values used in the experiment for the three groups of subjects. These values are expressed in semitones for the F0 condition, and in dB for the intensity condition.

Group	Subject	Δ F0 semitones	Δ Intensity dB
CI	A	3.2	2.5
	B	5.7	8.2
	C	3.8	7.2
	D	4.7	4.5
	E	21.6	6.0
NH	F	0.2	3.0
	G	0.4	4.9
	H	0.1	1.7
	I	0.3	3.9
	J	0.2	1.7
	K	0.5	1.8
	L	1.1	3.6
NH-voc	M	1.5	2.5
	N	2.9	3.6
	O	2.1	3.6
	P	2.2	3.2
	Q	1.3	2.9

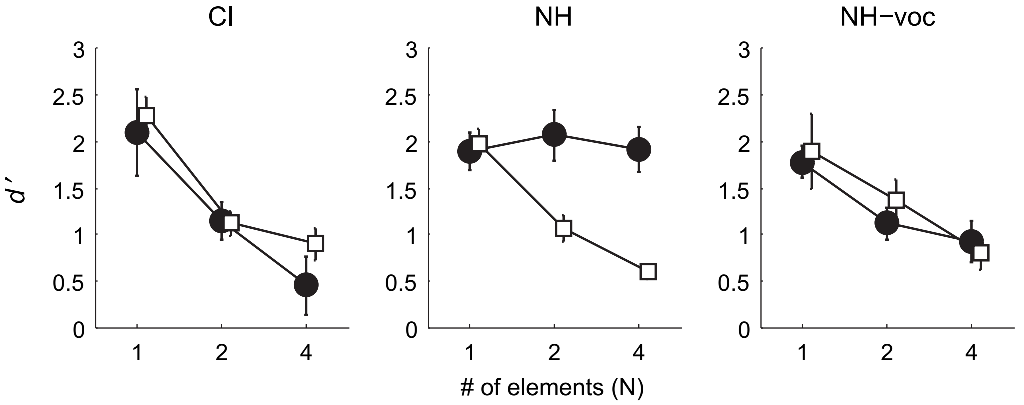

Fig. 4. Mean results for the sequence task. The mean d' values for the sequence task are plotted as a function of N. Error bars represent ± 1 standard error about the mean. For all groups and conditions, as a result of the adjustment phase and the different Δ values, mean performance for $N = 1$ is close to $d' = 2$. When N is greater than 1, NH listeners show an advantage for the processing of F0 sequences (filled circles) over intensity sequences (open squares). This advantage is not observed for CI listeners, nor for NH listeners presented with a noise-vocoded version of the stimuli (NH-voc).

(CI, NH, NH-voc) was an additional inter-subject factor. The only main effect found to be significant was that of N [$F(2,28) = 44.2$, $p < 0.0001$]. The interactions between Group and Condition [$F(2,14) = 5.1$, $p = 0.02$], Group and N [$F(4,28) = 2.9$, $p = 0.04$], and Condition and N [$F(2,28) = 6.6$, $p = 0.005$] were all significant. In addition, the interaction between the three factors was also significant [$F(4,28) = 5.6$, $p = 0.002$].

An additional statistic was computed to summarize the effect of N on performance. This statistic, termed the "*d′ slope*", was the slope of straight lines fitted to the individual data for the three values of N (N being scaled logarithmically). A smaller *d′ slope* indicates a slower decrease in performance when more elements are added to the sequences, i.e., better sequence processing. Fig. 5 shows, for each individual listener, the *d′ slope* for the F0 condition against the *d′ slope* for the I condition. All listeners from the NH group are positioned above the diagonal, indicating a greater sequence processing accuracy for F0 sequences compared to intensity sequences. Listeners from the NH-voc group are scattered around the diagonal, indicating similar sequence processing ability for F0 and I conditions. Listeners from the CI group were also scattered around the diagonal, with perhaps more variability, but no systematic pattern. One CI listener (listener E) appears to show a distinctive advantage of F0 sequences over intensity sequences. Note, however, that this subject displayed especially poor F0 discrimination capabilities (Table 2). A very large Δ value of 21.6 ST had to be used and it is not clear which cues were the basis of the perceptual judgments. Moreover, this subject was not well adjusted for $N = 1$ in the F0 condition ($d' = 1.35$ instead of the target value of 2). Having a low performance at $N = 1$ possibly prevented the drop of performance from being as large as for the other subjects.

An ANOVA was performed on the *d′ slope* data. It revealed significant main effects of both Group [$F(2,14) = 4.36$, $p = 0.03$] and Condition [$F(1,14) = 8.7$, $p = 0.01$] and, importantly, a significant interaction between the two factors [$F(2,14) = 7.0$, $p = 0.008$]. Paired *t*-tests were used to evaluate the effect of condition in each group. They revealed that there was a significant difference between *d′ slopes* for F0 and I conditions in the NH group ($p = 0.002$), but not in the CI group ($p = 0.62$) nor in the NH-voc group ($p = 0.37$). In the I condition, *t*-tests for independent samples revealed that the *d′ slopes* did not differ significantly between the NH and CI groups ($p = 0.99$), between the NH and NH-voc groups ($p = 0.37$) and between the CI and NH-voc group ($p = 0.27$).

4. Discussion

4.1. NH group

In NH listeners, we found that F0 sequences are processed very accurately: performance for discriminating two sequences of four elements was just as good as performance for discriminating two single sounds. These results replicate, with different listeners and stimuli, the findings of Cousineau et al. (2009). Cousineau et al.

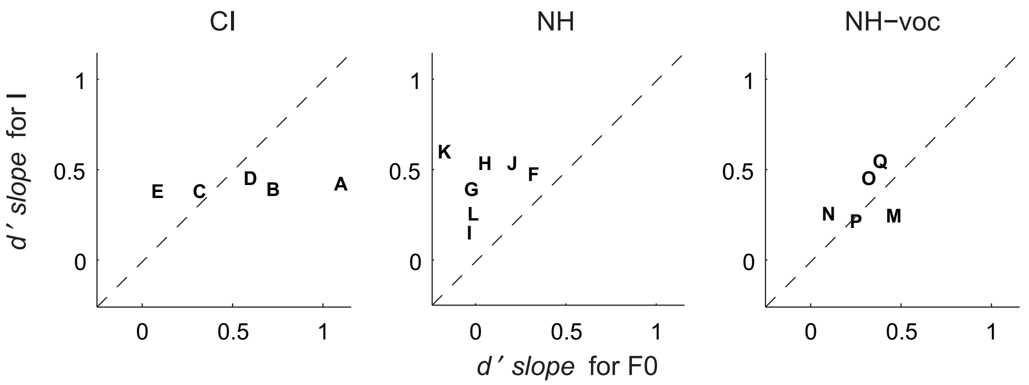

Fig. 5. Individual results for the sequence task. For each listener and condition, *d′ slopes* were obtained by linear regression on the d' data (see text for details). A high slope indicates a rapid degradation of performance as more elements are added to a sequence. The *d′ slopes* for F0 sequences are plotted against the *d′ slopes* for intensity sequences. Each letter represents an individual listener (see Tables 1 and 2). All listeners of the NH group show an advantage for F0 sequences over intensity sequences. For the other groups, CI and NH-voc, no systematic advantage is observed.

(2009) pointed out that such a pattern of results was inconsistent with the hypothesis that listeners break down each sequence into its individual elements and perform multiple element-to-element comparisons between sequences. Under this hypothesis, as the number of elements in the sequences increases, so should the number of comparisons and thus the probability of making an error. Instead, Cousineau et al. (2009) suggested that an additional mechanism may be available for NH listeners when processing pitch sequences. Frequency-shift detectors (FSDs), that specifically encode frequency-changes in addition to absolute frequency, could provide such a mechanism (Demany and Ramos, 2005; Demany et al., 2008, 2009). FSDs could also explain why bandpass filtered harmonic complexes without any resolved harmonic did not produce a pitch sequence advantage in the study of Cousineau et al. (2009), as FSDs are thought to detect spectral shifts. In the present study, the stimuli used were broadband sounds, contrary to the sounds used by Cousineau et al. (2009). A robust F0 sequence advantage was nevertheless observed, which shows that the pitch sequence advantage persists in NH listeners when both resolved and unresolved harmonics are available, as is the case with most natural sounds.

4.2. CI group

In contrast to the NH group, the CI group did not process F0 sequences any more accurately than intensity sequences. This was the case even though the elements of the sequences were equally discriminable in the F0 and I conditions, and just as discriminable for the CI group as for the NH group.

There is an age difference between the groups of NH listeners (NH and NH-voc) on the one hand, and the group of CI listeners on the other hand. It could thus be that aging caused the poorer performance of the CI group. When the effect of age on discrimination thresholds is investigated by comparing young and elderly NH listeners, discrimination thresholds are indeed found to become poorer with aging, both for pitch and loudness (He et al., 1998; Moore and Peters, 1992). Higher-order aspects of auditory processing might also be impaired in elderly subjects, although there is less consensus on the topic. Some studies find that when the contribution of thresholds is neutralized by using highly discriminable stimuli, no evidence is found for a decline of short-term memory capacity with aging (Murphy et al., 2000). Other studies find that elderly listeners show impairment in sequential processing of complex stimulus patterns (Fitzgibbons and Gordon-Salant, 1996).

However, three important aspects of the data make it highly unlikely that aging explained the lack of F0 sequence advantage for the CI group. First, our initial adjustment step factored out sound discriminability among groups of listeners, so any effect of aging on sound discriminability will have been compensated by the adjustment. Second, the same pattern of results was observed for the CI and NH-voc groups, the latter containing only young normal-hearing listeners, similar to the members of the NH group with respect to age. Third, performance for the intensity sequences was equivalent in all three groups, which is hard to reconcile with a putative impairment of higher auditory processing with age. A more parsimonious explanation of all the present findings is thus that the difference between the NH group on the one hand and the CI and NH-voc groups on the other hand is due to the nature of the available pitch cues for these three groups of listeners.

4.3. NH-voc group

The stimuli presented to the NH-voc group aimed at deteriorating the acoustic signal so that comparable cues were available to NH-voc and CI listeners. For the purposes of the present experiments, the main goal was to preserve the temporal cues to F0 while disrupting harmonic resolvability (see Fig. 2). Several variants of CI simulation are possible, so it is useful to consider which characteristics of the processing may have influenced the results.

We used noise-vocoding without any compression, which introduces some stochasticity in the temporal periodicity cues due to noise carriers and which does not simulate the enhancement in envelope modulation depth due to compression. For F0 discrimination tasks, such a noise-vocoding scheme sometimes underestimates the performance of CI listeners (Laneau et al., 2006). However, here discriminability was equated between CI and NH-voc in the main experiment. In addition, a previous study (Cousineau et al., 2009) used high-pass filtered click trains, which may be considered as another type of CI processing simulation without these issues (but limited to high-frequency channels). Results were similar, suggesting that the noise carriers and absence of compression were not the cause of poor F0 sequence processing for the NH-voc group.

Before vocoding, the envelope was extracted with a cutoff frequency of 400 Hz, in order to preserve F0 cues as far as possible. Shannon (1992) suggested that CI users have a poor ability to detect temporal envelope cues with a frequency higher than 300 Hz. Consequently, an envelope cutoff at 400 Hz is an optimistic simulation of the temporal pitch cues available to a typical CI user. This may explain why better Δ values were observed, on average, for the NH-voc group compared to the CI group (even though this difference was not significant). Still, the NH-voc group did not display any pitch sequence advantage. As for the spectral resolution of the simulations, eight analysis channels were used. More channels are available on current devices, but it has been shown that CI listeners may not benefit from having more than eight simultaneously active electrodes for speech in quiet, (Friesen et al., 2001). In addition, Cooper et al. (2008) showed that the number of simulated channels (from 4 to 16) had only a modest effect on melody perception with noise-vocoded stimuli. It would be interesting to investigate whether many more channels (e.g. 64 channels, as in Smith et al., 2002) would eventually restore a pitch sequence advantage, but it is unlikely that any currently plausible value of this parameter would do so.

4.4. Pitch sequence processing and harmonic resolvability

All available results using the sequence processing task with NH listeners (this study and Cousineau et al., 2009) can be summarized as follows. When sequences are formed of complex tones that contain resolvable harmonics, a pitch sequence advantage is observed. If there are no resolved harmonics, because of bandpass filtering (Cousineau et al., 2009) or noise-vocoding (NH-voc, this study), the pitch sequence advantage disappears. Therefore, we suggest that the absence of an F0 sequence advantage for CI users observed here is due to a lack of effective resolvability of individual harmonics, even though the FSDs may still be preserved for those listeners.

The definition and underlying mechanisms for harmonic resolvability are somewhat controversial. Resolvability could be determined by the number of harmonics per peripheral auditory channel (Shackleton and Carlyon, 1994), but it may also involve central factors (Bernstein and Oxenham, 2003), temporal fine structure cues to frequency (Srulovicz and Goldstein, 1983), or across-channel phase cues (de Cheveigné and Pressnitzer, 2006). There are therefore several possible reasons why users of current CI may not have access to resolved harmonics. Although high variability exists between individual CI users, there is for instance evidence for poorer frequency selectivity (Henry et al., 2005) and reduced sensitivity to fine structure cues (Kong et al., 2009; Zeng, 2002). Comprehensive reviews on this topic have been published recently (e.g. McDermott, 2004; Moore and Carlyon, 2005; Looi et al., 2008a). As our data do not provide any new suggestion as to which aspects of the electrical stimulation should be improved in priority to alleviate the problems related to harmonic resolvability, we will not repeat the arguments here. Our data do show, however, that harmonic resolvability does not only influence pitch-ranking abilities, but also determines whether or not a pitch sequence advantage can be observed.

4.5. Melody perception with an implant

The binary sequence method presented here has several interesting features as a measure of pitch pattern processing. The sequences are random and the intervals do not correspond to any musical scale, so previous familiarity with the experimental material is controlled. Interestingly, Galvin et al. (2007) did not observe any correlation between familiar melody recognition and novel melodies identification in CI listeners, suggesting that non-perceptual factors play an important role in tasks based on the identification of familiar melodies. The task of the listeners here is a same/different task. Thus, there is no need to instruct the listeners on the nature of the pitch cue they should be listening to. It is not even necessary to assume that they will only use pitch cues and not loudness or timbre cues, which may be correlated to F0 changes after implant processing. Whichever cue or cues were used by our CI listeners, they did not display as efficient sequence processing as NH listeners. Thus, when listening to melodies, CI listeners may not have access to specific pitch sequence mechanisms available for NH listeners.

It should be noted, nevertheless, that there are several important differences between the current task and realistic listening situations. First, the task is performed near the discrimination threshold. It is possible that sequence processing is different with larger steps resembling realistic musical scales (McDermott et al., 2008). This caveat is less relevant for CI listeners, as, even close to threshold, the differences on the pitch dimension are still larger than those most often used in Western tonal music. Second, from a pure methodological perspective, it should be acknowledged that the task is time-consuming. As a consequence, it is quite demanding for individual listeners, limiting its potential use in clinical situations. It would therefore be of interest in the future to derive an efficient measurement method, retaining the distinction between sound discriminability and sequence processing, but adapted to clinical testing.

In summary, the present data show that CI users are not able to process pitch sequences as accurately as NH listeners, even when the discriminability between individual elements of the sequences is factored out. This specific impairment is likely due to the nature of the pitch cues that are available with the current implants, which do not support harmonic resolvability. Any pitch sequence processing impairment would have an adverse effect on auditory scene analysis or music perception, two common issues for current CI users.

Acknowledgments

We would like to thank Brett Swanson for his helpful review and for his contribution in producing the electrodogram of Fig. 2.

References

Bernstein, J.G., Oxenham, A.J., 2003. Pitch discrimination of diotic and dichotic tone complexes: harmonic resolvability or harmonic number? J. Acoust. Soc. Am. 113, 3323–3334.

Cooper, W.B., Tobey, E., Loizou, P.C., 2008. Music perception by cochlear implant and normal hearing listeners as measured by the Montreal battery for evaluation of amusia. Ear. Hear. 29, 618–626.

Cousineau, M., Demany, L., Pressnitzer, D., 2009. What makes a melody: the perceptual singularity of pitch sequences. J. Acoust. Soc. Am. 126, 3179–3187.

de Cheveigné, A., Pressnitzer, D., 2006. The case of the missing delay lines: synthetic delays obtained by cross-channel phase interaction. J. Acoust. Soc. Am. 119, 3908–3918.

Demany, L., Ramos, C., 2005. On the binding of successive sounds: perceiving shifts in nonperceived pitches. J. Acoust. Soc. Am. 117, 833–841.

Demany, L., Pressnitzer, D., Semal, C., 2009. Tuning properties of the auditory frequency-shift detectors. J. Acoust. Soc. Am. 126, 1342–1348.

Demany, L., Trost, W., Serman, M., Semal, C., 2008. Auditory change detection: simple sounds are not memorized better than complex sounds. Psychol. Sci. 19, 85–91.

Drennan, W.R., Rubinstein, J.T., 2008. Music perception in cochlear implant users and its relationship with psychophysical capabilities. J. Rehabil. Res. Dev. 45, 779–789.

Fitzgibbons, P.J., Gordon-Salant, S., 1996. Auditory temporal processing in elderly listeners. J. Am. Acad. Audiol. 7, 183–189.

Friesen, L.M., Shannon, R.V., Baskent, D., Wang, X., 2001. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. J. Acoust. Soc. Am. 110, 1150–1163.

Fujita, S., Ito, J., 1999. Ability of nucleus cochlear implantees to recognize music. Ann. Otol. Rhinol. Laryngol. 108, 634–640.

Galvin 3rd, J.J., Fu, Q.J., Nogaki, G., 2007. Melodic contour identification by cochlear implant listeners. Ear Hear. 28, 302–319.

Geurts, L., Wouters, J., 2001. Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. J. Acoust. Soc. Am. 109, 713–726.

Gfeller, K., Turner, C., Mehr, M., Woodworth, G., Fearn, R., Knutson, J.F., Witt, S., Stordahl, J., 2002. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. Cochlear Implants Int. 3, 29–53.

Green, D.M., Swets, J.A., 1966. Signal Detection Theory and Psychophysics. Wiley, New York.

Greenwood, D.D., 1990. A cochlear frequency-position function for several species—29 years later. J. Acoust. Soc. Am. 87, 2592–2605.

He, N., Dubno, J.R., Mills, J.H., 1998. Frequency and intensity discrimination measured in a maximum-likelihood procedure from young and aged normal-hearing subjects. J. Acoust. Soc. Am. 103, 553–565.

Henry, B.A., Turner, C.W., Behrens, A., 2005. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. J. Acoust. Soc. Am. 118, 1111–1121.

Kong, Y.Y., Cruz, R., Jones, J.A., Zeng, F.G., 2004. Music perception with temporal cues in acoustic and electric hearing. Ear Hear. 25, 173–185.

Kong, Y.Y., Deeks, J.M., Axon, P.R., Carlyon, R.P., 2009. Limits of temporal pitch in cochlear implants. J. Acoust. Soc. Am. 125, 1649–1657.

Laneau, J., Wouters, J., Moonen, M., 2004. Relative contributions of temporal and place pitch cues to fundamental frequency discrimination in cochlear implantees. J. Acoust. Soc. Am. 116, 3606–3619.

Laneau, J., Moonen, M., Wouters, J., 2006. Factors affecting the use of noise-band vocoders as acoustic models for pitch perception in cochlear implants. J. Acoust. Soc. Am. 119, 491–506.

Looi, V., McDermott, H., McKay, C., Hickson, L., 2004. Pitch discrimination and melody recognition by cochlear implant users. In: Miyamoto, R. (Ed.), Cochlear Implants: Proceedings of the VIII International Cochlear Implant Conference. International Congress Series, 1273C. Elsevier, Amsterdam, pp. 197–200.

Looi, V., McDermott, H., McKay, C., Hickson, L., 2008a. Music perception of cochlear implant users compared with that of hearing aid users. Ear. Hear. 29, 421–434.

Looi, V., McDermott, H., McKay, C., Hickson, L., 2008b. The effect of cochlear implantation on music perception by adults with usable pre-operative acoustic hearing. Int. J. Audiol. 47, 257–268.

McDermott, H.J., 2004. Music perception with cochlear implants: a review. Trends Amplif. 8, 49–82.

McDermott, J.H., Lehr, A.J., Oxenham, A.J., 2008. Is relative pitch specific to pitch? Psychol. Sci. 19, 1263–1271.

McFarland, D.J., Cacace, A.T., 1992. Aspects of short-term acoustic recognition memory: modality and serial position effects. Audiology 31, 342–352.

Moore, B.C., Peters, R.W., 1992. Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity. J. Acoust. Soc. Am. 91, 2881–2893.

Moore, B.C.J., Carlyon, R.P., 2005. Perception of pitch by people with cohlear hearing loss and by cochlear implant users. In: Plack, C.J., Oxenham, A.J., Fay, R.R., Popper, A.N. (Eds.), Pitch, Neural Coding and Perception. Springer, New York, pp. 234–277.

Murphy, D.R., Craik, F.I., Li, K.Z., Schneider, B.A., 2000. Comparing the effects of aging and background noise on short-term memory performance. Psychol. Aging 15, 323–334.

Patterson, R.D., Allerhand, M.H., Giguere, C., 1995. Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. J. Acoust. Soc. Am. 98, 1890–1894.

Pijl, S., Schwarz, D.W., 1995. Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes. J. Acoust. Soc. Am. 98, 886–895.

Poncet-Wallet, C., Frachet, B., Meyer, B., Lancelin, D., Lorenzi, C., 2008. Identificaton de phonèmes isolés dans le silence et dans le bruit sous implant cochléaire. 115ème Congrès Natonal de la Société Française d'ORL, Paris, France.

Pressnitzer, D., Bestel, J., Fraysse, B., 2005. Music to electric ears: pitch and timbre perception by cochlear implant patients. Ann. N. Y. Acad. Sci. 1060, 343–345.

Shackleton, T.M., Carlyon, R.P., 1994. The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. J. Acoust. Soc. Am. 95, 3529–3540.

Shannon, R.V., 1992. Temporal modulation transfer functions in patients with cochlear implants. J. Acoust. Soc. Am. 91, 2156–2164.
Shannon, R.V., Zeng, F.G., Kamath, V., Wygonski, J., Ekelid, M., 1995. Speech recognition with primarily temporal cues. Science 270, 303–304.
Singh, S., Kong, Y.Y., Zeng, F.G., 2009. Cochlear implant melody recognition as a function of melody frequency range, harmonicity, and number of electrodes. Ear Hear. 30, 160–168.
Smith, Z.M., Delgutte, B., Oxenham, A.J., 2002. Chimaeric sounds reveal dichotomies in auditory perception. Nature 416, 87–90.
Srulovicz, P., Goldstein, J.L., 1983. A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. J. Acoust. Soc. Am. 73, 1266–1276.
Sucher, C.M., McDermott, H.J., 2007. Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. Hear. Res. 230, 80–87.
Vandali, A.E., Sucher, C., Tsang, D.J., McKay, C.M., Chew, J.W., McDermott, H.J., 2005. Pitch ranking ability of cochlear implant recipients: a comparison of sound-processing strategies. J. Acoust. Soc. Am. 117, 3126–3138.
Zeng, F.G., 2002. Temporal pitch in electric hearing. Hear. Res. 174, 101–106.