DTU Library

Pitch matching in bimodal cochlear implant patients: Effects of frequency, spectral envelope, and level

Maarefvand, Mohammad; Blamey, Peter J.; Marozeau, Jeremy

Link back to DTU Orbit

Pitch matching in bimodal cochlear implant patients: Effects of frequency, spectral envelope, and level[a)]

Mohammad Maarefvand,[1,b)] Peter J. Blamey,[2] and Jeremy Marozeau[3]

[1]*The Bionics Institute, Melbourne, Victoria, Australia*
[2]*Department of Medical Bionics, the University of Melbourne, Melbourne, Victoria, Australia*
[3]*Hearing Systems Group, Technical University of Denmark, Lyngby, Denmark*

This study systematically investigated the effects of frequency, level, and spectral envelope on pitch matching in twelve bimodal cochlear implant (CI) users. The participants were asked to vary the frequency and level of a pure or complex tone (adjustable sounds) presented in the non-implanted ear to match the pitch and loudness of different reference stimuli presented to the implanted ear. Three reference sounds were used: single electrode pulse trains, pure tones, and piano notes. The data showed a significant effect of the frequency and complexity of the reference sounds. No significant effect of the level of the reference sounds was found. The magnitude of effect of frequency was compressed in the implanted ear: on average a difference of seven semitones in the non-implanted ear induced the same pitch change as a difference of 19 to 24 semitones for a stimulus presented to the implanted ear. The spectral envelope of the adjustable sound presented to the non-implanted ear also had a significant effect. The matched frequencies were higher by an average of six semitones for the pure tone compared to a complex tone. Overall, the CI listeners might have matched the stimuli based on timbre characteristics such as brightness.
© 2017 Acoustical Society of America. https://doi.org/10.1121/1.5009443

[VMR]   Pages: 2854–2865

I. INTRODUCTION

The sound processing algorithms and neural stimulation strategies used in cochlear implants have primarily been designed to optimise the perception of speech and have been especially successful in quiet environments (Marozeau *et al.*, 2014). When cochlear implant (CI) sound processing algorithms and stimulation strategies are used for music instead of speech, they are much less successful and it has been widely reported that most CI users are not satisfied with music perception and appreciation as a result (Kong *et al.*, 2004; Nimmons *et al.*, 2008).

Appreciation of most types of music requires perception of rhythm, melody, and timbre. Present day CIs are effective at delivering the coarse temporal contrasts required for rhythm perception, but less effective for the melodic contrasts and timbres of music (Kong *et al.*, 2004; Nimmons *et al.*, 2008). Thus, it is most likely that improved musical appreciation will require improved perception of the pitch and timbre of musical notes. For this reason, the research reported here focused on the factors influencing perception of musical sounds with different pitch and timbre presented to cochlear implant users through their sound processors.

In spite of generally good speech perception with CIs, there has not been an obvious improvement in musical appreciation and pitch perception. One of the approaches in improving speech perception was the systematic analysis of the main stages of the sound processing chain between the speaker and the perception of the message (Millar *et al.*, 1984) and the repeated evaluation of factors affecting overall performance (Blamey *et al.*, 1985; Blamey *et al.*, 1996). A similar factorial approach was adopted here to study pitch perception in CI users.

For normal hearing (NH) listeners, the precepts pitch, loudness, timbre, and rhythm correspond to the physical characteristics frequency, intensity, and spectral and temporal envelope. This correspondence does not apply as straightforwardly for CI patients. For instance, a frequency change at the input to a CI sound processor may change the perceived pitch, loudness, and timbre all at once. Indeed, four stages may influence the perception of pitch in CI patients: (1) the input signal (e.g., its fundamental frequency, spectral envelope and intensity); (2) the sound coding (e.g., the strategy, the frequency-to-electrode map, the intensity-to-stimulus-level map, and the number and timing of electrical pulse); (3) the electro-neural interface (e.g., insertion depth of the electrodes, the spread of current delivered to the electrodes, and the amount and pattern of neural survival); (4) the central processing (e.g., the neural firing patterns, musical experience pre- and post-operatively and plasticity).

It has been clear since the early work of Fletcher (1935) that frequency is the major physical determinant of pitch for people with normal hearing. Thus, a CI sound processor that aims to replicate the pitch perception of listeners with normal hearing (NH) should employ sound coding algorithms so that the effects of spectral shape and intensity do not interfere with the perception of pitch. Similarly, the sound coding algorithms should minimise the effects of the electro-

neural interface by compensating for factors such as depth of electrode insertion, spread of electrical current in the cochlea, and neural survival patterns for individual CI patients.

Intensity has a small effect on perceived pitch for people with normal hearing: less than 1% for pure tones below 1 kHz and less than 5% between 1 and 2 kHz (Verschuure and Meeteren, 1975). In contrast to NH, Carlyon et al. (2010a) reported a large pitch perception change with increase in level (2.3 semitones or 16%). Shannon (1983) and Allen and Oxenham (2014) reported that pitch increased with increasing level while Townshend et al. (1987) and Pijl (1997) reported a significant decrease in pitch with increase in level. Carlyon et al. (2010a) showed both increase and decrease in pitch with level which varied from one electrode to another, which was in line with the results of the pitch matching experiment of Green et al. (2012) in bimodal CI users.

The factors that affect pitch perception at different stages are not independent of one another. High intensity can lead to greater current spread in the cochlea compared to low level stimulation. Increasing input intensity to the sound processor can also lead to stimulation of neighbouring electrodes (electrode activation spread), even for a pure tone input. Both current spread and electrode activation spread lead to excitation of larger numbers of neurons which can affect pitch perception. For a harmonic complex, both current spread and electrode activation spread can happen around each harmonic. The extent and direction of spread of neural excitation may depend on factors such as the pattern of neural survival close to the electrode and the proximity of the stimulated electrode to the surviving neurons. Neural survival and positions of electrodes relative to the modiolus in the cochlea are expected to vary considerably between CI patients and may thus account for the inconsistent and variable effects of level observed in published pitch matching studies.

Under some circumstances pitch perception in NH listeners is known to be influenced by timbre and quality of sound (Zarate et al., 2013). The quality of sounds was reported to be different in the implanted and non-implanted ears in an early study by Eddington et al. (1978) who tested pitch matching with a pure tone in one CI user with normal hearing up to 1000 Hz in the non-implanted ear. They reported that pitch matching was difficult and suggested that the quality of sound in the implanted ear might be revealed with variations in the spectral characteristics of complex acoustic stimuli presented to the non-implanted ear. This suggestion was tested with five bimodal CI users by Lazard et al. (2012). They asked their participants to match different sounds in the non-implanted ear to a pulse train presented to the most apical electrode in the implanted ear. They reported that in terms of quality, the perceived pitch in the implanted ear was similar to inharmonic complex sounds.

Hearing loss may also change the perception of pitch in impaired ears. Listeners with significant residual hearing in the non-implanted ear (bimodal listeners) are interesting cases to consider because it is likely that they hear different pitches in the two ears, arising primarily from the stimulus properties such as spectral shape and intensity, from the CI sound coding stage, and the electro-neural interface stage. For example,

variations in electrode insertion depth can give large differences in pitch perception between the two ears (Blamey et al., 1996). Additionally, substantial changes in pitch perception with increasing level may give rise to a mismatch between the perceived pitch in acoustic and electric stimulation which can be an obstacle to music perception in bimodal listeners. Some researchers have argued that the bimodal pitch mismatch between acoustic and electric stimuli may be reduced over time by post-operative perceptual adaptation (Reiss et al., 2007) but others have not observed adaptation (Carlyon et al., 2010b; Baumann et al., 2011). It seems that such adaptation does not happen to all CI users or for all sounds and the magnitude of the adaptation may vary across people.

Bimodal pitch matching of electric and acoustic stimuli in opposite ears has been used in many studies that have yielded diverse results with large individual variability (Eddington et al., 1978; Blamey et al., 1996; Dorman et al., 2007; Baumann et al., 2011). The large individual differences in these experiments might arise from the interplay of the numerous factors affecting pitch perception at different stages of the perception process, and the different ways in which these factors have been controlled in the previous research. Much bimodal pitch matching research has focused on the relationship between the pitch percept produced by stimulating a single electrode at a fixed place in the cochlea and the characteristic frequency of the hair cells found at that position in a normal cochlea as modelled by Greenwood (1961, 1990). Although Eddington et al. (1978) reported that acoustic-electric pitch matches were similar to those expected from Greenwood's function, others noticed that the acoustic matches to electrical stimulation were lower than predicted from the position of the electrodes in the cochlea (Blamey et al., 1996; Baumann et al., 2011).

The current study was motivated by the need for a better understanding of the factors affecting the perception of melodic pitch for CI users, including bimodal listeners, in the context of listening to musical sounds through their sound processor. Due to time and technical constraints, it was not possible to investigate the effect of all the previously mentioned factors. As explained above, some of the factors are not easily controllable, such as neural survival patterns for example. It was decided to keep some factors fixed across participants (such as musical experience, sound processor, and frequency-to-electrode map) and to concentrate on the effects of the stimulus parameters: frequency, spectral envelope, and intensity. Pure tone, complex tone, and single electrode stimuli were included to facilitate comparison of the data with published studies. Overall, this study tests the hypothesis that the dependency of the frequency, intensity and spectral envelope on pitch is significantly larger in CI listeners compared to normal-hearing listeners. A multifactorial experiment was designed to explore these factors in a systematic manner.

II. METHOD

A. Participants

The inclusion criteria for the study were: post-lingual deafness; at least 1 year of experience with both hearing aid

J. Acoust. Soc. Am. **142** (5), November 2017

Maarefvand et al.   2855

(HA) and CI together in opposite ears; use of HA and CI together for at least 70% of the time in everyday life. All of the participants were fitted with the Cochlear CI device with Contour electrode and all of them used the advanced combination encoder sound processing strategy. They used different brands of HAs in the non-implanted ear, so their HAs were not used for presenting sound to their non-implanted ear. Only people with measurable hearing thresholds in the non-implanted ear lower than or equal to 90 dB hearing level (HL) at 250 and 750 Hz or 1 KHz were included in this study. The practical constraints of finding a sufficient number of bimodal CI users did not allow us to apply stricter criteria for their inclusion in the study. The threshold of 90 dB HL in the non-implanted ear was considered adequate to allow pitch matching at these frequencies. The bimodal usage criteria were designed to provide adequate opportunity to develop a bimodal binaural benefit (Vermeire and Van de Heyning, 2009). A criterion for no musical training was adopted to avoid outliers who performed differently from the main body of bimodal listeners and CI users, such as the exceptional patient reported by Maarefvand *et al.* (2013).

Twelve bimodal cochlear implant users (4 females and 8 males) participated. Their age ranged from 61 to 87 years with an average of 73 years. Four had unknown etiology, three had noise-induced hearing loss, two had presbycusis, two had hearing loss of a hereditary nature, and one person had otosclerosis. They had an average duration of deafness of 12.6 years before implantation with a range of 2 to 36 years. They had used their HAs in the non-implanted ears for an average of 18.3 years with a range of 2 to 55 years at the time of the study. All of the participants were accustomed to their CIs, with average experience of 3.4 years and a range of 3 to 5 years. None of them had been a musician or had special music training before and they were not able to play any type of instrument. Figure 1 shows the audiograms of the participants in the non-implanted ear. Only S10 and S12 had thresholds worse than 90 dB HL at 1000 Hz. The recruitment was conducted through the Cochlear Implant Clinic at the Royal Victorian Eye and Ear Hospital, Melbourne,

Australia. All the participants gave written informed consent and were compensated for their travel expenses. This project conformed to The Code of Ethics of the World Medical Association (Declaration of Helsinki), and was approved by the Royal Victorian Eye and Ear Hospital Human Research Ethics Committee (Project 11/1040H, Improving music perception in bimodal cochlear implant users).

B. Stimuli

The method of pitch measurement chosen for this research was the adjustment of the frequency and level of an acoustic adjustable sound presented to a non-implanted ear to match the pitch and loudness of a reference sound presented to the opposite implanted ear via the cochlear implant sound processor. The experiment was composed of six conditions based on the combination of three reference and two adjustable sounds as shown in Fig. 2. All the sounds were synthesised and presented using MAX/MSP 6 software (Cycling 74 Co.).

1. Reference sounds

Three reference sounds were used: single electrode pulse trains, pure tones, and piano notes. For each reference sound there was a low frequency and a high frequency and each stimulus was presented at a low and a high intensity. The reference stimuli were presented through the Cochlear Ltd. approved direct audio input (DAI) accessory cable to a Freedom CI sound processor which was fitted with the participant's everyday map. The single electrode pulse trains had a duration 600 with 10 ms rise time, 100 ms fall time and were presented to the most apical electrode (number 22 which corresponded to 250 Hz in the CI sound processor frequency-to-electrode map) in the low frequency condition or to a more basal electrode in the high frequency condition. The basal electrode was electrode 16 for participants with hearing threshold better than 90 dB HL at 1000 Hz in the non-implanted ear. Otherwise electrode 18 was chosen, corresponding to 750 Hz in the frequency-to-electrode map.
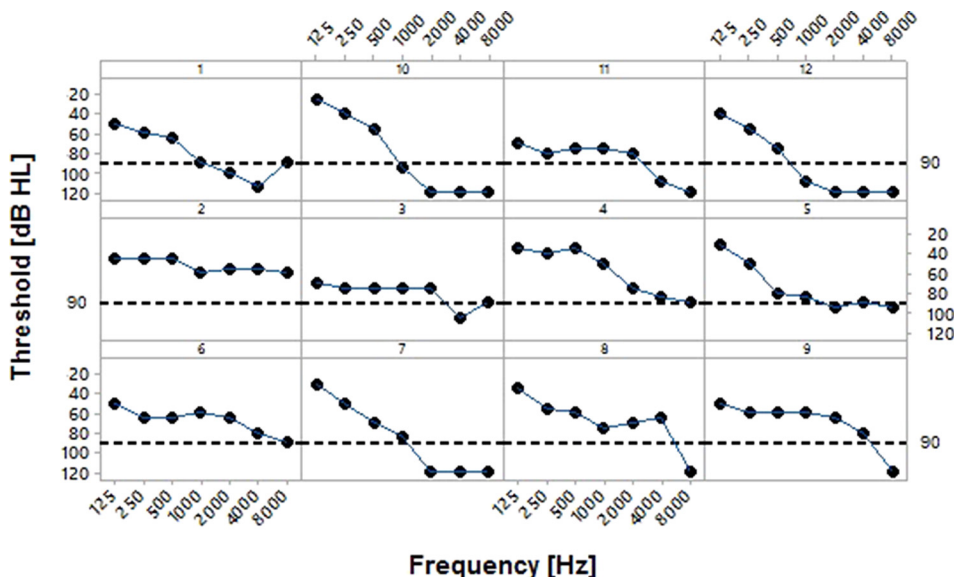

FIG. 1. (Color online) Participants' audiometric thresholds in the non-implanted ear. The dashed line indicates 90 dB HL, which was the benchmark for selection of the high frequency sounds.

Implanted ear   Non-Implanted ear

Reference sounds

Single electrode
Pure tone ←→ Pure tone
Piano note

Single electrode
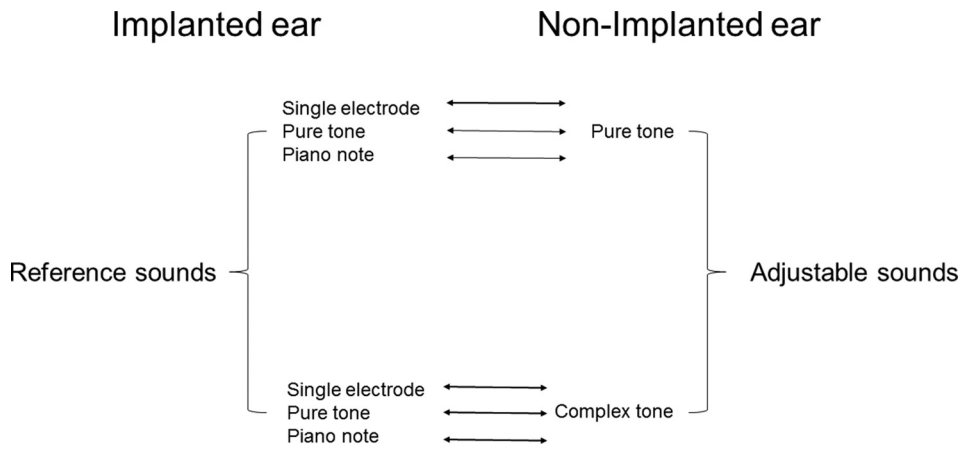Pure tone ←→ Complex tone
Piano note

Adjustable sounds

FIG. 2. Reference and adjustable sounds for each experimental condition.

These electrode choices ensured that every participant had hearing threshold better than 90 dB HL at the corresponding frequency in the non-implanted ear. The stimulation rate was set as the same as in the participant everyday map (in a range from 500 to 1200 pps). In order to create a single electrode pulse train from a pure tone at the direct audio input to the sound processor, the CI map was modified to enable only one electrode to be activated.

The pure tone reference sounds were tones with 10 ms rise time, 100 ms fall time, and total duration 600 ms presented to a sound processor which was fitted with the participant's everyday map with all the electrodes enabled. Due to the overlap of the band pass filters assigned to each electrode, this stimulus could activate one or more electrodes depending on the input intensity of the pure tone (Fig. 3). In the low frequency condition, the frequency was set to 250 Hz. In the high frequency condition, the frequency was set to 750 Hz (S10 and S12) or 1000 Hz (all other participants).

The piano note reference sounds were piano note samples extracted from Ableton Live sequencer software. The fundamental frequencies (F0) of the stimuli were the closest musical note to 250 Hz (246 Hz) for the lower frequency trials and either 750 (740 Hz) or 1000 Hz (987 Hz) for the higher frequency trials. The samples were limited in duration to 1.7 s.

Figure 4 shows the simulated electrodogram of a piano note as processed by a typical CI, at low frequency, 250 Hz, and low level (A); at low frequency and high level (B); at high frequency and low level (C); and at high frequency and high level (D). The electrodogram is analogous to a conventional spectrogram which shows electric current level for each electrode on the vertical axis as a function of time.

Prior to starting the pitch matching trials, a seven-point loudness scale from inaudible to too loud (Blamey and Martin, 2009) was used to find intensities for each reference stimulus that evoked "soft" and "loud but comfortable" responses for each participant. The two intensities were combined with the two frequencies and three stimulus types to give twelve reference stimuli in total for each participant.

2. Adjustable sounds

The adjustable sounds were presented to the non-implanted ear though a 3A E.A.R. gold insert phone (Etymotic

Research, Inc., USA) with a flat frequency response up to 10 KHz. The adjustable sounds were either pure tones, or complex tones. The complex tones were based on the parameters that gave the smallest timbre difference between acoustic and pulse train electric stimuli in a study by Lazard et al. (2012). They were composed of 11 harmonics, passed through a fourth-order Butterworth filter with centre frequency set to 1.64 times of the fundamental frequency and a Q factor of 13.4 (Fig. 5). The temporal envelope of the adjustable sounds was similar to the envelope of the single electrode pulse train and pure tone reference sounds. It was assumed that different adjustable sounds in the non-implanted ear induce different quality of sounds or timbre. "Timbre" and "quality of sounds" are used interchangeably in this article.

Since different participants had different patterns of hearing loss across frequency (Fig. 1), the adjustable sounds were amplified with a customised filter before presentation to the non-implanted ear. The filter response at each frequency was calculated and customised using the NAL-RP
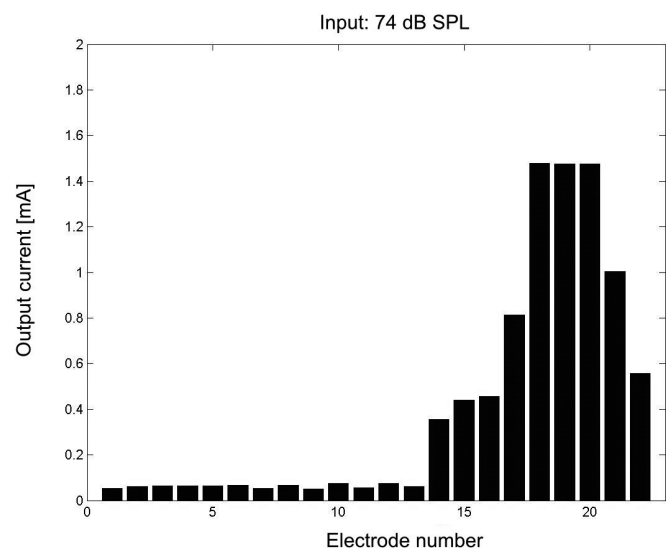

FIG. 3. Electrode activation pattern at 74 dB sound pressure level for a 590 Hz pure tone put into a sound processor. The horizontal axis shows electrode number 1 to 22. The vertical bars indicate the amount of activation on each electrode in milliamps. At this input level, the neighbouring electrodes to electrode 19 (which is the electrode corresponding to 590 Hz) also receive activation.
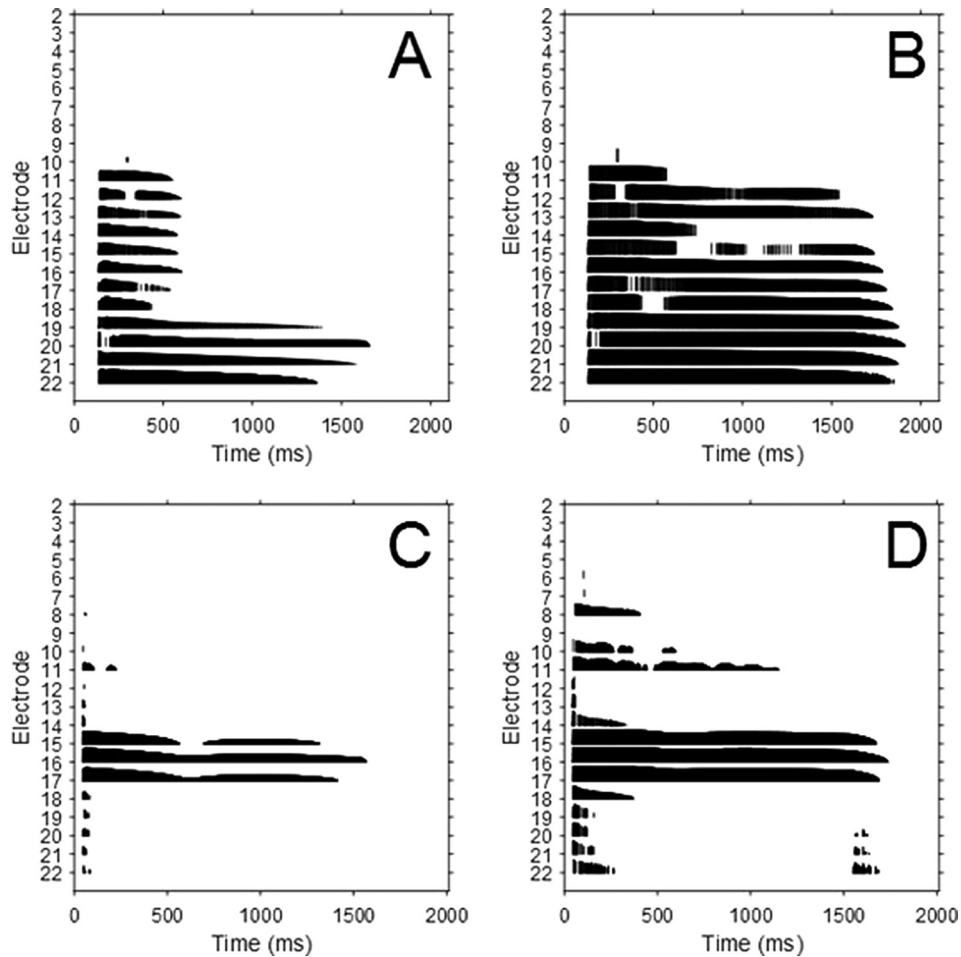
FIG. 4. Electrodogram showing electrical stimulation for a piano note with low frequency presented at a soft (A) and a loud (B) loudness levels and high frequency presented at soft (C) and aloud (D) loudness levels. The amplitude and number of pulses is greater for loud stimuli than for soft stimuli and extends further along the horizontal axis.

linear HA fitting formula. Thus, the participants' different brands of HAs were not used and were not a potential source of individual differences in the experiment.

C. Procedure

There were 12 reference sounds and 12 participants and the order of testing was determined by a Latin square table to balance out order effects across the study. Three pitch matches were obtained for each reference and adjustable sound condition, making a total of 3 repeats × 12 reference sounds × 2 different adjustable sound = 72 trials per participant.

During each trial, the participant listened to the reference and adjustable sounds alternately, controlled by the MAX/MSP 6 software. Two unmarked knobs controlled the intensity and the frequency of the acoustic adjustable sound (PowerMate 3.0, Griffin Technology, USA). The participant was first asked to vary the intensity of the adjustable sound to match the loudness of the reference sound. The intensity of the adjustable sound could be changed in steps of 0.5 dB. Upon matching the loudness, the participant started changing the second knob to match the pitch of the adjustable sound to the reference sound. A wide range of frequencies was made available to the participant for each trial in order to avoid potential bias (Carlyon *et al.*, 2010b; Green *et al.*, 2012). The knob changed the frequency or F0 of the sounds in a range from 80 to 2000 Hz on a logarithmic scale. The starting phases of the harmonics of the complex tone were randomised in each test trial

and the starting frequency was chosen randomly to be higher or lower than the frequency of the reference sounds. The final step was to adjust both level and frequency knobs together to achieve the best possible pitch match.

D. Similarity rating

Although the participants were asked to spend enough time to reach the best possible pitch match, it was still
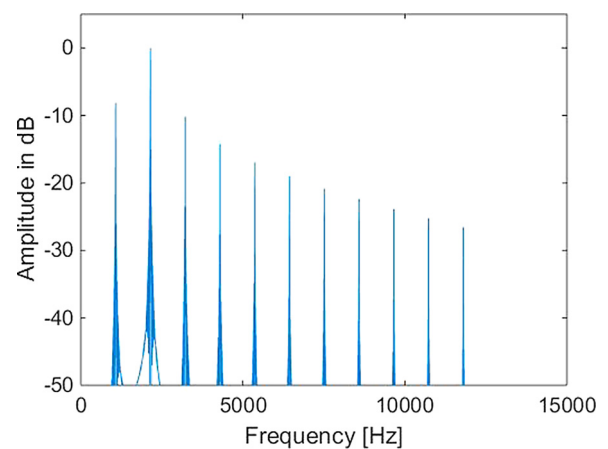

FIG. 5. (Color online) Spectral representation of a typical complex adjustable sound. The horizontal axis shows the frequency of each harmonic on a logarithmic scale. The vertical axis shows the amplitude of each harmonic in dB of attenuation from the maximum output of the sound card.

possible that the two stimuli differed in sound quality. Therefore, after each trial, the participant was asked to rate the similarity of the matched reference and adjustable sounds. To quantify the degree of similarity between the reference and adjustable sounds, a continuous rating scale from very similar at one end (100%) to very dissimilar at the other end (0%) was made available to the participant after each match.

E. Statistical analysis

The data were analysed through a linear mixed-model with the following fixed effects: the complexity of the reference (single electrode, pure tone, or piano note), frequency of the reference (low, high), level of the reference (soft, loud), the spectral envelope of adjustable sound (pure tone, and complex tone), and all their interactions. The participants and their interactions with all the fixed effects were included as random effects. The dependent factor was the matched frequency from each trial. The frequency of the reference and adjustable sounds referred to the fundamental frequency for complex sounds and frequency for pure tones. The matched frequency was transformed to semitones [musical instrument digital interface (MIDI) notes] using the following equation:

$$MIDI\ note\ number = 12\,log_2(frequency/440) + 69.$$

The logarithmic transformation was required to meet the normality requirement of the residual errors in the analysis of variance (ANOVA) analyses that were used. It also had the added advantage that the MIDI note or semitone scale is meaningful for interpreting the results in a musical context where 12 semitones difference is an octave. The reference frequencies of 250, 750, and 1000 Hz are equivalent to MIDI numbers of 59.2, 78.2, and 83.2, respectively. In order to simplify the analysis the fact that the high frequency could be either 750 or 1000 Hz (MIDI note 79.2 or 83.2, respectively) was ignored.

III. RESULTS

A. Pitch matching results

Figure 6 shows the matched frequency as a function of the four main fixed effects. In the analysis shown in Table I, only the main effects of the reference frequency and spectral envelope of the adjustable sound (pure tone and complex tone) were statistically significant at the $p < 0.05$ level. There was also a significant interaction between the complexity of the reference (single electrode, pure tone, or piano note) and frequency. All other interactions between the four independent factors were not statistically significant.

The effect of the reference frequency was significant [$F(1,11) = 19.0718$; $p = 0.0011$]. The matched pitch increased with the frequency of the reference sounds as expected. Figure 6 shows that with the sole exception of the soft high frequency electrode reference paired with the adjustable pure tone, the averaged matched pitches for high frequency stimuli were all lower than the MIDI numbers (83) used to generate the stimuli for most cases. All the averaged matched pitches for low frequency stimuli were higher than the MIDI number (59) used to generate the stimuli. The mean difference between the matched pitches for high and low reference frequency (7.1 semitones) was less than the expected value of 19 or 24 semitones between the frequencies of the low and high reference signals for participants S10 and S12 or the other ten participants, respectively.

The main effect of spectral envelope of the adjustable sound on matched pitch was statistically significant [$F(1,11) = 43.07$; $p < 0.0001$]. The matched frequencies for the pure tone were always higher than those for the

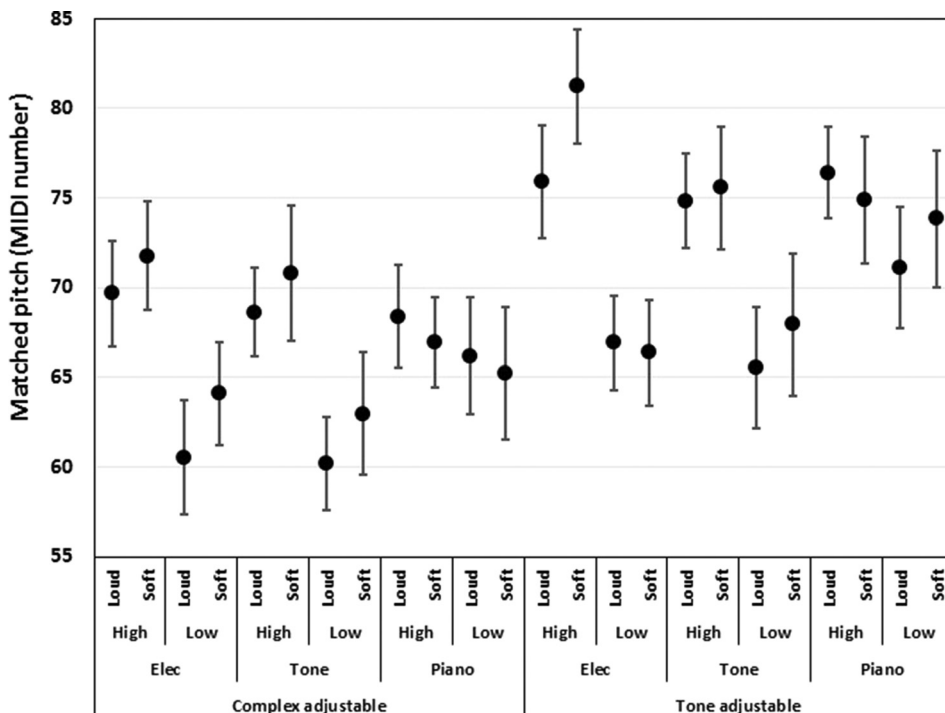FIG. 6. Mean and 95% confidence intervals for matched frequency on the vertical axis as a function of level (low and high), reference frequency (low and high), reference complexity (single electrode, pure tone, and piano note), and adjustable timbre spectral envelope (pure tone and complex tone) on the horizontal axis. The MIDI numbers for low and high frequency reference sounds were 59 and 83, respectively.

TABLE I. ANOVA for the pitch matching results. The main effects and interactions with p < 0.05 are shown in bold. "Reference" refers to the complexity of the reference sounds (single electrode, pure tone and piano note), "frequency" to the frequency of the reference sound (low and high), "level" to the level of the reference sound (low and high), and "adjustable" to the spectral envelope of the adjustable sound. Pure tone and complex tone are the other independent factors. The dependent variable was the matched frequency in MIDI number.

Source	DF	F ratio	P value
Reference	2	2.2166	0.1327
Frequency	**1**	**19.0718**	**0.0011**
Reference * frequency	**2**	**8.8727**	**0.0015**
Level	1	2.568	0.1374
Reference * level	2	2.0684	0.1503
Frequency * level	1	0.012	0.9149
Reference * frequency * level	2	1.2817	0.2975
Adjustable	**1**	**43.0739**	**<0.0001**
Reference * Adjustable	2	1.3289	0.2852
Frequency * Adjustable	1	2.2644	0.1605
Reference * frequency * adjustable	2	0.6179	0.5481
Level * adjustable	1	0.0099	0.9227
Reference * level * adjustable	2	0.5096	0.6076
Frequency * level * adjustable	1	0.3882	0.5459
Reference * frequency * level * adjustable	2	1.7635	0.1948

corresponding complex tone match (Fig. 6). The mean pitch match for the pure tone was 72.6 MIDI number compared to 66.3 MIDI number for the complex tone.

The interaction between reference frequency and complexity of the reference sound was also significant [F(2,22) = 8.87; p = 0.0015]. The interaction plot in Fig. 7 shows that regardless of the complexity of the reference sound, the high frequency references always yielded significantly higher pitch matches than the low frequency references, but the difference in matched pitch was much smaller for the piano notes (2.6 semitones) than for the tones (8.3 semitones) and single electrode (10.2 semitones). A Tukey-Kramer HSD *post hoc* analysis on the subset of data composed only of matched

stimuli to the low frequency reference, showed a significant difference between the fundamental frequency of the piano note reference and the pure tone (p = 0.0004) and the piano and the single electrode (p = 0.0013). No significant difference was found between the pure tone and the single electrode (p = 0.94). A similar analysis on the subset of data composed only with stimuli matched to the high frequency reference shows only a significant difference between the single electrode and the piano note (p = 0.030).

The main effect of level on the perception of pitch and its interaction with complexity reference or frequency was not significant. Therefore, no systematic change (decrease or increase) in the perception of pitch with increase in level was found.

B. Similarity rating results

ANOVA was performed on the similarity rating. In order to convert the percentage data to a normal distribution, the rating scores were transformed through a hyperbolic arctangent function. Table II shows statistically significant main effects for the complexity of reference, the level and spectral envelope of adjustable sound, and a significant interaction between the reference and the frequency with p < 0.05. The mean similarity ratings and standard errors, before transformation, for the three reference sounds were 78.5% ± 1.2%, 80.9% ± 1.0%, and 69.2% ± 1.5% for the single electrode, pure tone, and piano notes, respectively. For the pure tone and complex tone adjustable sounds, the mean similarity ratings and standard errors were 78.6% ± 1.0% and 73.9% ± 1.1%, respectively. The mean similarity ratings and standard errors for high and low level reference stimuli were 73.5% ± 1.1% and 78.9% ± 1.0%, respectively. Figure 8 shows the mean and standard deviation for the similarity ratings for the reference sounds and adjustable sounds, averaged over level and frequency.

A Tukey HSD *post hoc* analysis revealed that the rating of similarity for piano notes was significantly lower than the
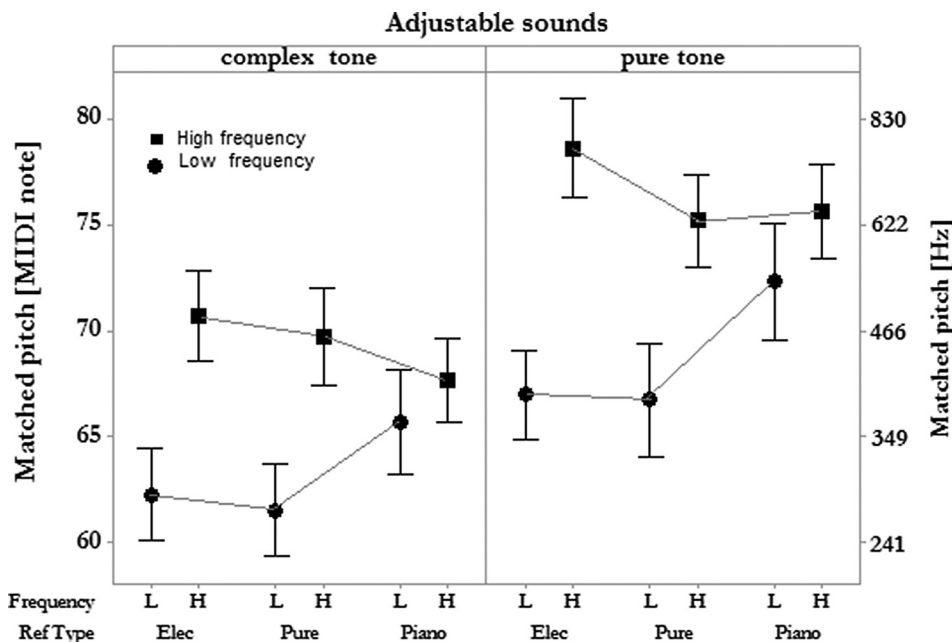

FIG. 7. Interaction plot of reference frequency (low and high) and reference complexity (single electrode, pure tone and piano note) showing 95% confidence intervals for each mean value.

2860   J. Acoust. Soc. Am. **142** (5), November 2017

Maarefvand *et al.*

TABLE II. ANOVA for the similarity ratings (see caption Fig. 2).

Source	DF	F ratio	P value
Reference	**2**	**8.0722**	**0.0023**
Frequency	1	0.4041	0.538
Reference ∗ Frequency	**2**	**4.072**	**0.0313**
Level	**1**	**6.0156**	**0.0321**
Reference ∗ level	2	0.7215	0.4972
Frequency ∗ level	1	0.3	0.5948
Reference ∗ frequency ∗ level	2	1.139	0.3383
Adjustable	**1**	**6.7491**	**0.0248**
Reference ∗ adjustable	2	0.0146	0.9855
Frequency ∗ adjustable	1	0.3177	0.5843
Reference ∗ frequency ∗ adjustable	2	1.5879	0.2269
Level ∗ adjustable	1	1.8837	0.1973
Reference ∗ level ∗ adjustable	2	0.0833	0.9204
Frequency ∗ level ∗ adjustable	1	2.4644	0.1447
Reference ∗ frequency ∗ level ∗ adjustable	2	0.0288	0.9716

pure tone (p = 0.0025) and the single electrode pulse train (p = 0.020). However, no significant difference was found between the pure tone and the pulse train (p = 0.65).

IV. DISCUSSION

A. The effect spectral envelope of adjustable

Adjustable pure and complex tones were used in an attempt to test the effect of timbre difference between electric and acoustic modalities on the pitch matching results. The specific acoustical properties of the adjustable complex tone were selected to minimise the timbre difference with a pulse train presented to the most apical electrode, based on the findings of Lazard et al. (2012). It was therefore expected that the match between the pulse train reference and the adjustable complex tone should have yielded the highest similarity ratings and the lowest F0 standard deviation. It is clear from the similarity ratings in Fig. 8 and the confidence intervals shown in Figs. 6 and 7 that this expectation was not correct. The inconsistency between the results of Lazard

et al. (2012) and this study may be due to the different participants and/or the different range of stimuli used in these two studies, although the stimulus generation methods and matching procedures were similar.

Figure 6 shows that regardless of the complexity, frequency, and level of the reference stimuli, the participants tended to decrease the F0s of the adjustable complex tones compared to the frequencies of the adjustable pure tones. This implies that an acoustic complex tone at a given F0 was perceived by the participants as having a higher pitch than a pure tone at the same frequency. It is worth noting that the average of the frequency difference between the stimuli matched to the pure and complex tones is about 6 semitones or a frequency ratio of 1.41. This ratio is roughly the same as the ratio of 1.64 between the spectral peak and the fundamental frequency of the complex tone (see Sec. II B 2). Therefore, the participants might have matched the acoustic stimuli based on timbre characteristics linked to the overall spectral envelope as was reported by Allen and Oxenham (2014).

B. The effect of frequency

As expected, frequency and place of electrode have a significant effect on the pitch matching results. However, as pointed out in Sec. III A, the matched frequencies of the adjustable tones were all higher than 250 Hz for the reference stimuli at the low frequency, and most were lower than 750 or 1000 Hz for the high frequency reference stimuli. The frequency-to-electrode map in the CI sound processor and the position of the electrode array in the cochlea may help to explain these results.

Although we had no information about the positions of individual participants' electrodes, it is possible to compare the average pitch matches in the right half of Fig. 6 with the predictions of Greenwood's (1961, 1990) function for a range of insertion depths and lengths of the basilar membrane. It is known that the length of the basilar membrane varies from person to person from about 30 to 35 mm
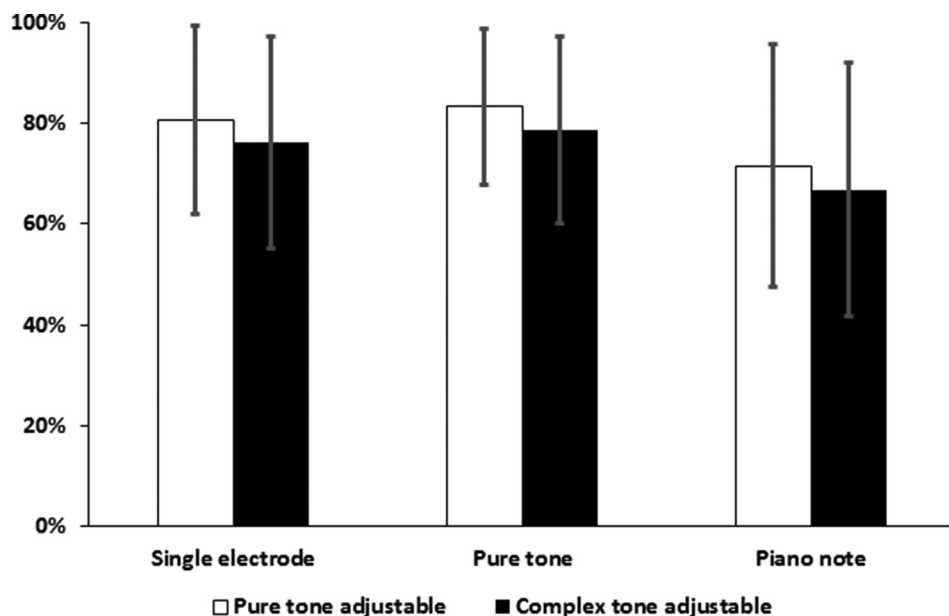

FIG. 8. Mean and standard deviation for the similarity ratings for three reference complexity (single electrode, pure tone, and piano note) and two adjustable spectral envelope (pure tone and complex tone).

J. Acoust. Soc. Am. **142** (5), November 2017
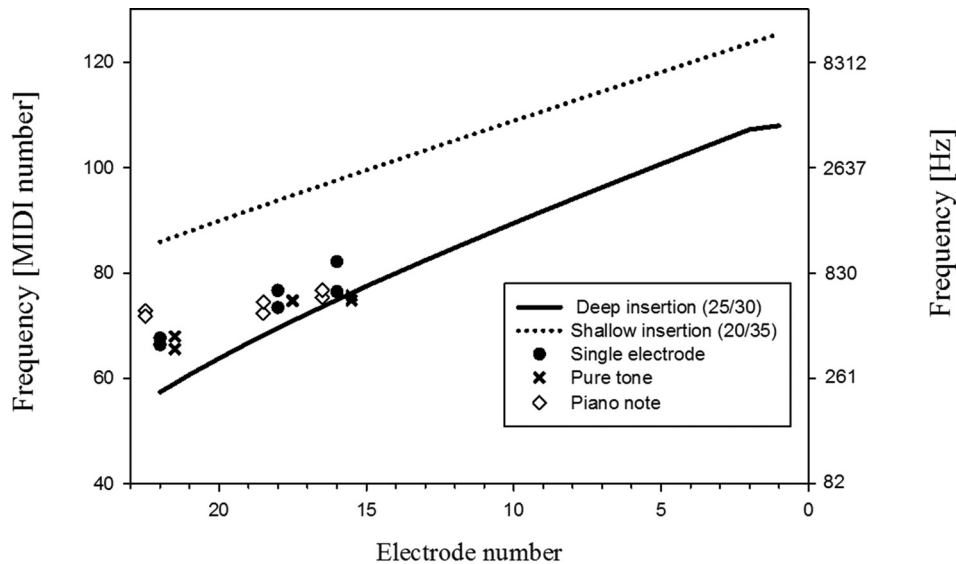
Maarefvand et al.   2861

FIG. 9. The mean of the matched frequency for loud and soft reference stimuli at high and low frequencies for pure tone adjustable sounds in relation to the range which is expected from Greenwood's function. The dashed line shows the highest estimation of the stimulated electrodes for a cochlea with 35 mm length and a short insertion depth of 20 mm. The solid line indicates the lowest pitch estimation based on Greenwood's function provided that there is deep insertion of the electrode array (25 mm) within a short cochlea (30 mm).

(Von Békésy, 1960; Stakhovskaya *et al.*, 2007). In addition to this variability, the insertion depth of the electrode array of a cochlear implant can influence the place of stimulation. Different combinations of the cochlea length and insertion depth are possible. For example, a short cochlea (30 mm) which has been implanted with a deep electrode array insertion (25 mm) or a long (35 mm) cochlea with short insertion of the electrode array (20 mm). Figures 9 and 10 show the range of possible characteristic frequencies corresponding to each electrode between these two extremes estimated using Greenwood's function, respectively, when the adjustable sounds were pure and complex tones.

The experimental data points lie close to the deep insertion/short cochlea line in Figs. 9 and 10. This suggests that electrode position and frequency-to-electrode map may account for some of the difference between the CI pitch matches and the values that would be expected for NH people performing binaural pitch matches. It is unlikely that all

of the participants in this study would have deep insertions and short cochleae, so it can be assumed that the pitch match values are lower than the predictions that would be calculated based on Greenwood's function for at least some of the individual participants' electrode positions. This observation is in line with studies which measured the actual positions of the electrodes and concluded that the perceived pitch was always lower than the prediction by about an octave (Blamey *et al.*, 1996; Dorman *et al.*, 2007; Baumann *et al.*, 2011). The reasons for the discrepancy between the pitch matches and the predictions of Greenwood's function include the difference in lengths of the electrode array, the basilar membrane, and Rosenthal's canal that contains the ganglion neurons. As the ganglion cells fill more turns of the cochlea than the electrode arrays, and the low frequency ganglion cells are clustered around a position away from the apex of the cochlea in the middle turn, the most apical electrode would be adjacent to the low frequency ganglion cells
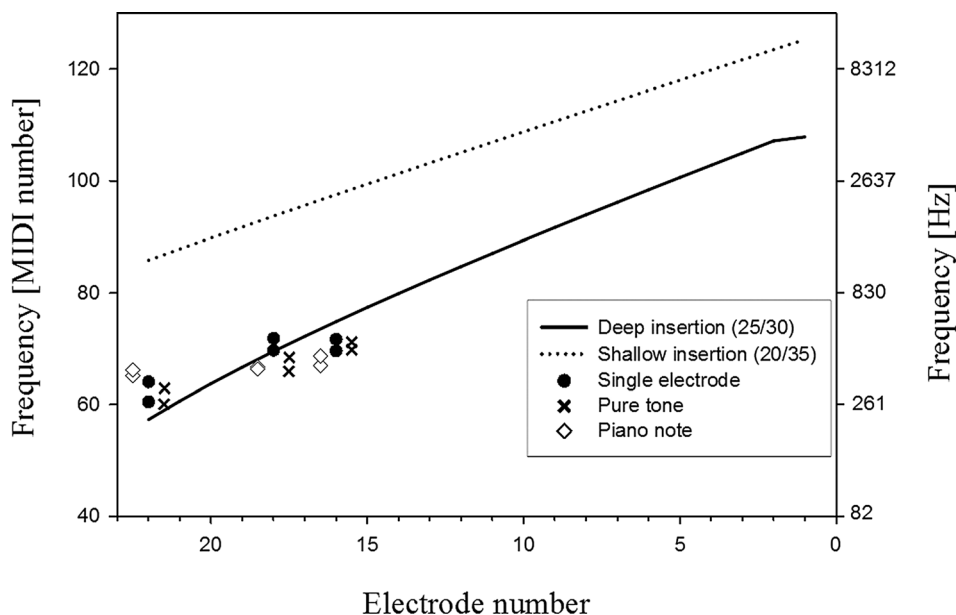

FIG. 10. The mean of the matched frequency for loud and soft reference stimuli at high and low frequencies for complex tone adjustable sounds in relation to the range which is expected from Greenwood's function.

and would therefore elicit lower pitches than expected (Stakhovskaya *et al.*, 2007). Blamey *et al.* (1996) speculated that four factors might account for the lower-than-Greenwood results, including the fact that the non-implanted ears had impaired hearing, that listeners may have adapted to the sounds over time, that the ganglion cells closest to apical electrodes would normally innervate hair cells at a more apical position and should therefore be interpreted by the brain as corresponding to a lower frequency stimulus, and that the low pulse rates used in Blamey's experiment might have lowered the perceived pitch. The last of these four suggestions does not apply to the present study that used high pulse rates, but the other three may still be relevant.

The frequency difference between the low and high frequency reference sounds in this study was either 19 or 24 semitones. Generally, NH listeners can perceive this difference. However, the difference between low and high perceived pitches was less than 12 semitones for CI users. This result supports the hypothesis that the effect of frequency on pitch perception is different for CI and NH listeners. The Greenwood's functions shown in Figs. 9 and 10 predict that the perceived pitch difference between stimuli presented to electrode 22 and electrode 16 or 18 should be less than 1 octave, regardless of the depth of insertion and length of the basilar membrane, consistent with the study results.

The interaction of the complexity and frequency of the reference sound was significant. The most likely explanation for this lies in the sound processing algorithms that do not preserve the fine temporal structure required to recognise the periodic nature of harmonic complexes like piano notes. Thus CI users must rely primarily on the place cues for pitch perception. However, as shown in Fig. 4, the electrical activation patterns of two broad-band sounds separated by a large F0 difference can still overlap significantly (Marozeau *et al.*, 2014) reducing the perception of pitch differences.

C. The effect of current level

As the intensity of the reference stimuli increased, the current levels applied to the CI electrodes also increased. It was expected that the resultant current spread and electrode activation spread would affect the pitch matching, but no significant effect of level on the pitch matches was observed. This result is in line with the results of Reiss *et al.* (2007). In contrast to the current study, some researchers have found small but significant effects of current level on pitch (Arnoldner *et al.*, 2008; Carlyon *et al.*, 2010a; Green *et al.*, 2012). Furthermore, it was expected that the piano stimuli would be more affected by an increase in level (see Fig. 4) due to the presence of the spread of neural and electrode activation around each harmonic. The results were not consistent with this expectation. They indicate that the spread of activation at neural level did not change the perception of pitch significantly on average. It is notable that level had a significant effect on the similarity ratings, indicating significantly higher ratings for the loud stimuli than the soft. It might indicate that the participants were more satisfied by their match for loud stimuli than soft.

D. The effect of the complexity of the reference sound

The analysis indicated lower rating scores for piano notes than pure tone and single electrode pulse trains. The lower rating for the piano note is not surprising, given that the spectral and temporal envelopes of the two adjustable sounds were very different from a piano.

E. Implications for bimodal and CI music perception

The pitch percept produced when an ear is stimulated with a cochlear implant is different from the pitch experienced when the opposite impaired ear is stimulated with the same sound via a HA. The results of this study show that the amount of pitch difference between the ears depends on the frequency and spectral envelope of sound that is presented to the two ears. The pitch difference between the ears also may depend on factors such as length of the cochlea and depth of insertion of the electrode array that can vary from one CI patient to another. Furthermore, the pitch changes between successive notes that are necessary for the appreciation of melody are different in an implanted ear from a non-implanted ear. These differences between electrical and acoustic pitch perception have major implications for the perception of music in bimodal listeners and in CI users.

Melody recognition for most music types is very difficult with a CI on its own. The results of this study show that the difference in perception of pitch for pure tones with different frequencies is reduced in CI listeners compared to NH listeners, and almost reduced to zero for piano notes which have similar spectral spread, regardless of the fundamental frequency of the note. Given that the spectral spread is more often associated with the perception of timbre than the perception of pitch, the continual variation in frequency and timbre that occurs in orchestral music and other polyphonic music must be quite confusing for CI listeners who are trying to pick out the melody.

On the other hand, an impaired ear can recognise melodies reasonably well from the acoustic signal. It has been reported that bimodal listeners have greater music appreciation ability than people who use a CI alone (El Fata *et al.*, 2009). The CI in the bimodal combination also makes a significant contribution, particularly in cases of vocal music that includes lyrics.

F. Potential improvements of CIs to enhance music

The results of this study suggest that potential improvements to music perception with CIs will require more than just a pitch shift. Although there was a shift toward low frequencies in the perceived pitch for the most apical electrode relative to Greenwood's function, this shift was not constant for all electrodes. For an increase of 2 octaves in the F0s of the acoustical inputs (24 semitones), the perceived pitch changed just 13.5 semitones for the adjustable pure tones and 10 semitones for the adjustable complex tones. This means that low frequency sounds were perceived as being higher, and high frequencies were perceived as being lower in the implanted ear than in the non-implanted ear. The amount of frequency compression was even greater for the

J. Acoust. Soc. Am. **142** (5), November 2017

Maarefvand *et al.*   2863

piano note reference sounds. This kind of compression has been reported in another study as well (Dorman *et al.*, 2007). In the study of Reiss *et al.* (2015) such frequency compression led to a mismatch between perceived pitches in the implanted and non-implanted ears in bimodal CI users which did not reduce over time. Frequency compression would affect the relative distance between the harmonics of a complex signal and change the timbre of sounds used in instrument recognition. One way to improve the representation of the harmonic structure of complex sounds maybe to use frequency expansion. For instance, if the distance between the harmonics were increased, it may be possible to resolve the harmonic structure of complex tones and subsequently pitch perception may be improved. Omran *et al.* (2011) used frequency expansion, and their results showed improved pitch difference limens in NH listeners who were tested with vocoder sounds. Kasturi and Loizou (2007) also showed improvements in melody recognition with such expansion in CI users. Swanson *et al.* (2009) reported better melody recognition when the notes of a melody were expanded up to 5 times although Luo *et al.* (2014) reported that CI users might not necessarily prefer more expanded intervals than NH listeners. Other ways to improve music perception might include an electrode array with a larger number of electrodes, more closely spaced and with narrower electric current distributions which are inserted deeper within the cochlea.

V. CONCLUSION

The hypotheses that CI patients' pitch perception would be affected by the frequency, and timbre of sounds, and that the sound processor and electrode neural interface would also affect pitch perception were supported by the data and analyses in this study. The systematic investigation of pitch matching for twelve different reference stimuli differing in frequency, level, and timbre helped to understand the factors affecting pitch perception in CI patients. Since a frequency compression was observed in the coding of sounds by the sound processor, a frequency expansion in the frequency-to-electrode map could be helpful for music perception. Using frequency expansion may also help to make timbre differences between musical notes more natural in CIs, and increase the similarity of the two sounds in opposite ears of bimodal listeners.

Allen, E., and Oxenham, A. (**2014**). "Symmetric interactions and interference between pitch and timbre," J. Acoust. Soc. Am. **135**(3), 1371–1379.

Arnoldner, C., Riss, D., Kaider, A., Mair, A., Wagenblast, J., Baumgartner, W.-D., Gstöttner, W., and Hamzavi, J.-S. (**2008**). "The intensity–pitch relation revisited: Monopolar versus bipolar cochlear stimulation," Laryngoscope **118**(9), 1630–1636.

Baumann, U., Rader, T., Helbig, S., and Bahmer, A. (**2011**). "Pitch matching psychometrics in electric acoustic stimulation," Ear Hear. **32**(5), 656–662.

Blamey, P. J., Dooley, G. J., Parisi, E. S., and Clark, G. M. (**1996**). "Pitch comparisons of acoustically and electrically evoked auditory sensations," Hear. Res. **99**(1-2), 139–150.

Blamey, P. J., and Martin, L. F. (**2009**). "Loudness and satisfaction ratings for hearing aid users," J. Am. Acad. Audiol. **20**(4), 272–282.

Blamey, P. J., Martin, L. F. A., and Clark, G. M. (**1985**). "A comparison of three speech coding strategies using an acoustic model of a cochlear implant," J. Acoust. Soc. Am. **77**(1), 209–217.

Carlyon, R. P., Lynch, C., and Deeks, J. M. (**2010a**). "Effect of stimulus level and place of stimulation on temporal pitch perception by cochlear implant users," J. Acoust. Soc. Am. **127**(5), 2997–3008.

Carlyon, R. P., Macherey, O., Frijns, J. H., Axon, P. R., Kalkman, R. K., Boyle, P., Baguley, D., Briggs, J., Deeks, J., Briaire, J., Barreau, X., and Dauman, R. (**2010b**). "Pitch comparisons between electrical stimulation of a cochlear implant and acoustic stimuli presented to a normal-hearing contralateral ear," J. Assoc. Res. Oto-laryngol. **11**(4), 625–640.

Dorman, M. F., Spahr, T., Gifford, R., Loiselle, L., McKarns, S., Holden, T., Skinner, M., and Finley, C. (**2007**). "An electric frequency-to-place map for a cochlear implant patient with hearing in the nonimplanted ear," J. Assoc. Res. Oto-laryngol. **8**(2), 234–240.

Eddington, D., Dobelle, W., Brackmann, D., Mladejovsky, M., and Parkin, J. (**1978**). "Auditory prostheses research with multiple channel intracochlear stimulation in man," Ann. Otol. Rhinol. Laryngol. **87**(6), 1–39.

El Fata, F., James, C. J., Laborde, M.-L., and Fraysse, B. (**2009**). "How much residual hearing is 'useful' for music perception with cochlear implants?," Audiol. Neuro-Otol. **14**(1), 14–21.

Fletcher, H. (**1935**). "Loudness and pitch," Bell Lab. Rec. **13**(5), 130–137.

Green, T., Faulkner, A., and Rosen, S. (**2012**). "Frequency selectivity of contralateral residual acoustic hearing in bimodal cochlear implant users, and limitations on the ability to match the pitch of electric and acoustic stimuli," Int. J. Audiol. **51**(5), 389–398.

Greenwood, D. (**1961**). "Critical bandwidth and the frequency coordinates of the basilar membrane," J. Acoust. Soc. Am. **33**(10), 1344–1356.

Greenwood, D. (**1990**). "A cochlear frequency-position function for several species—29 years later," J. Acoust. Soc. Am. **87**(6), 2592–2605.

Kasturi, K., and Loizou, P. C. (**2007**). "Effect of filter spacing on melody recognition: Acoustic and electric hearing," J. Acoust. Soc. Am. **122**(2), EL29–EL34.

Kong, Y.-Y., Cruz, R., Jones, J. A., and Zeng, F.-G. (**2004**). "Music perception with temporal cues in acoustic and electric hearing," Ear Hear. **25**(2), 173–185.

Lazard, D. S., Marozeau, J., and McDermott, H. J. (**2012**). "The sound sensation of apical electric stimulation in cochlear implant recipients with contralateral residual hearing," PLoS One. **7**(6), e38687.

Luo, X., Masterson, M., and Wu, C. (**2014**). "Melodic interval perception by normal-hearing listeners and cochlear implant users," J. Acoust. Soc. Am. **136**(4), 1831–1844.

Maarefvand, M., Marozeau, J., and Blamey, P. J. (**2013**). "A cochlear implant user with exceptional musical hearing ability," Int. J. Audiol. **52**(6), 424–432.

Marozeau, J., Simon, N., and Innes-Brown, H. (**2014**). "Cochlear implants can talk but cannot sing in tune," Acoust. Austr. **42**(2), 131–135.

Millar, I. B., Tong, Y. C., and Clark, G. M. (**1984**). "Speech processing for cochlear implant prostheses," J. Speech Lang. Hear. Res. **27**(2), 280–296.

Nimmons, G. L., Kang, R. S., Drennan, W. R., Longnion, J., Ruffin, C., Worman, T., Yueh, B., and Rubinstein, J. T. (**2008**). "Clinical assessment of music perception in cochlear implant listeners," Otol. Neurotol. **29**(2), 149–155.

Omran, S., Lai, W., Buchler, M., and Dillier, N. (**2011**). "Semitone frequency mapping to improve music representation for nucleus cochlear implants," Eurasip J. Audio Speech Music Process. **2**, 1–11.

Pijl, S. (**1997**). "Pulse rate matching by cochlear implant patients: Effects of loudness randomization and electrode position," Ear Hear. **18**(4), 316–325.

Reiss, L. A. J., Ito, R. A., Eggleston, J. L., Liao, S., Becker, J. J., Lakin, C. E., Warren, F. M., and McMenomey, S. O. (**2015**). "Pitch adaptation patterns in bimodal cochlear implant users: Over time and after experience," Ear Hear. **36**(2), e23–e34.

Reiss, L. A. J., Turner, C. W., Erenberg, S. R., and Gantz, B. J. (**2007**). "Changes in pitch with a cochlear implant over time," J. Assoc. Res. Oto-laryngol. **8**(2), 241–257.

Shannon, R. V. (**1983**). "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics," Hear. Res. **11**, 157–189.

Stakhovskaya, O., Sridhar, D., Bonham, B. H., and Leake, P. A. (**2007**). "Frequency map for the human cochlear spiral ganglion: Implications for cochlear implants," J. Assoc. Res. Oto-laryngol. **8**(2), 220–233.

Swanson, B., Dawson, P., and McDermott, H. (**2009**). "Investigating cochlear implant place-pitch perception with the Modified Melodies test," Cochlear Implants Int. **10**(S1), 100–104.

Townshend, B., Cotter, N., Compemolle, D. V., and White, R. L. (**1987**). "Pitch perception by cochlear implant subjects," J. Acoust. Soc. Am. **82**(1), 106–115.

Vermeire, K., and Van de Heyning, P. (**2009**). "Binaural hearing after cochlear implantation in subjects with unilateral sensorineural deafness and tinnitus," Audiol. Neuro-Otol. **14**(3), 163–171.

Verschuure, J., and Meeteren, A. A. v. (**1975**). "The effect of intensity on pitch," Acustica **32**(1), 33–44.

Von Békésy, G. (**1960**). *Experiments in Hearing* (McGraw-Hill, New York).

Zarate, J. M., Ritson, C. R., and Poeppel, D. (**2013**). "The effect of instrumental timbre on interval discrimination," PLoS One. **8**(9), e75410.