TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 PROBLEM STATEMENT

1.1.1 Context of the problem

Cochlear implant (CI) devices afford many profoundly deaf individuals worldwide partially restored hearing ability (Wilson & Dorman 2008). Although CI users achieve remarkable speech perception with contemporary multichannel CIs (Gifford et al. 2008), reports of generally unsatisfactory music perception (for example Gfeller et al. 1997; Leal et al. 2003; Looi et al. 2004; McDermott 2004; Gfeller et al. 2005; Looi et al. 2008) have prompted research efforts directed at improving CI users' music listening experience.

One of the main challenges facing scientists involved in CI research is that the input delivered to the central auditory system after peripheral electrical stimulation is deprived compared to that for normal hearing (NH). Owing to only 22 electrodes being available to handle much of the frequency range relevant to human auditory tasks, acoustic–electrical signal conversion results in severely limited spectral resolution, compounded by mismatched frequency-to-place mapping (Fu & Shannon 2002) and channel interaction (Hughes & Abbas 2006). However, CI-mediated perception of temporal cues, as measured during gap detection tasks, is comparable to that of NH listeners (Shannon 1989). In a functionally hierarchically organised perceptual system, such as the auditory system (Pandya 1995; Kaas et al. 1999; Wessinger et al. 2001; Zatorre et al. 2002; Griffiths & Warren 2002; Semple & Scott 2003; Stewart et al. 2008; Warren 2008; Woods & Alain

2009), the outcome of events at earlier levels serves as input to subsequent processing levels. It follows that impoverished input deriving from an early level will propagate through the entire processing system, possibly to the detriment of ultimate perceptual outcome. For successful CI-mediated hearing it is therefore necessary not only to encode stimulus attributes accurately during acoustic–electrical conversion, but also to know their perceptual outcomes.

This context is pertinent to understanding CI-mediated perception, and specifically music perception, in real-world listening conditions. Successful perception of specifically Western tonal music relies largely on accurate processing of spectral information for pitch perception, but with important contributions from temporal cues for the perception of rhythm and metre. Music perception has consequently been proposed to be facilitated by a modular, hierarchically organised processing system that initially consists of two independent but parallel streams (Peretz & Coltheart 2003), of which the outputs later combine to create a unified musical percept (Peretz 1990). In order to improve music perception experience for CI listeners, it is thus necessary to understand how impoverished information – both as isolated cues and in interaction – translate to perceptual outcome in real-world listening conditions (Middlebrooks et al. 2005; Wilson & Dorman 2008; Moore & Shannon 2009). Such understanding rests on three pillars: (i) understanding the underlying requirements for and mechanisms of music perception to allow relevant information to be extracted peripherally and presented for further auditory processing, (ii) linking perceptual experience of music-relevant stimuli in real-world listening conditions to the psychophysical ability afforded by the implant device and (iii) investigating and understanding CI-mediated music perception as the outcome of systems-based processing of an auditory input.

1.1.2 Research gap

Improved CI-mediated music perception ability requires that the underlying constraints hindering processing of music-relevant information need to be understood. This implies the need for systematic investigation of CI-mediated perception of simple music-relevant stimuli, as isolated elements as well as in combination. Several previous studies have addressed CI-mediated perception of music-relevant stimuli, but have focused on using either isolated cues (for example Gfeller & Lansing 1991; Gfeller & Lansing 1992; Pijl & Schwarz 1995b; Pijl 1997; Gfeller et al. 1997; Kong et al. 2004) or familiar melodies comprising complex tones and combined pitch, rhythm and timbre information (for example Gfeller et al. 2003; Looi et al. 2004; Gfeller et al. 2005; Vongpaisal et al. 2009). Although such studies provide insightful comment regarding CI users' limited music perception ability, results were not interpreted in the context of a complete sensory perceptual system as operates in real-world listening nor related to the psychophysical ability allowed by the implant technology. Since CI sound processors are generally developed to extract speech cues, progress towards improving specific processing deficits underlying poor music perception will remain elusive without understanding the constraints related to CI-mediated extraction of musical cues from auditory input, their presentation to the processing system and subsequent cognitive processing. By considering music as consisting of multiple components that are presented to the auditory periphery as a unified whole, an approach to understanding its perception by deconstructing the elements as well as combining them, may improve our understanding of how elementary perceptual ability afforded by the CI device may translate into real-world like melodic sequences.

Earlier studies regarding CI-mediated music perception, especially those aimed at investigating real-world music listening ability, have often used familiar melodies as stimuli during assessment of CI listeners music perception ability (see for example Gfeller et al. 1997; Leal et al. 2003; Kong et al. 2004; McDermott & Looi 2004; Gfeller et al. 2005; Looi et al. 2008). Listeners are usually asked to identify a melody from a list of well-known tunes following an adjustment to the sound processing algorithm or the listener's map settings. However, owing to only a set number of response choices being available a listener may have to draw on several cognitive processing strategies, including mnemonic analysis, to support the response decision. This is not an undesirable ability in real-world listening; however, when it forms the basis of a task to probe music listening success, the behavioural outcome may not be able to be linked unambiguously to perception of a specific component of the input. In this sense the task of music perception then is rather one of music recognition, meaning that music perception ability is assessed in relation to *familiarity with a music excerpt*, rather than the successful recognition of specific musical features. Again, results from such investigations contribute to assessing CI users' music listening behaviour, but do not facilitate a deeper understanding of the underlying processing that facilitates the observed behaviour.

Whether CI users experience a sequence of successive tones of specific frequency and duration and separated by fixed silences as melody-like has not been tested previously. Yet, by using familiar melodies as a probe to gauge CI listeners' melody perception ability, it is assumed that they do. It may thus be valuable to take one step back in the quest to understand CI-mediated music perception and first determine whether the output of signal processing at the periphery, which serves as input to cortical processing, contains sufficient auditory information to allow a listener to infer melodic quality of a tone sequence.

Implant users' desire for restored music listening ability may prompt successful music perception to become the next benchmark for CI success (Limb et al. 2010). To this end, two clinically relevant assessment tools, which both offer valuable entry points to the assessment of CI-mediated music perception, have recently been proposed (Cooper et al. 2008; Kang et al. 2009). However, given the multifaceted nature of music as auditory stimulus, objectively gauging the real-world perceptual effect of manipulations to processing algorithms, user settings or biophysical constraints will require an understanding not only of the perception of deconstructed musical features but also of their interaction within a complete perceptual system. This study thus aims to put forward a systematic investigation into CI-mediated music perception by investigating perception of specific cues, in isolation and combination, as they may contribute to music perception. Use of pure tone signals presented in sound-field listening conditions allows perceptual outcomes to be interpreted in consideration with signal processing strategies as well as systems-based processing of electrically delivered auditory information. The approach thus allows an intermediate step to assess perception of a subjective auditory quality in as objective a manner as possible in order to facilitate improved understanding of CI-mediated music perception (also see paragraph 1.5 and Figure 1.2).

1.2 BACKGROUND LITERATURE REVIEW

Figure 1.1 provides a schematic representation of the respective study fields that inform the understanding of music perception in electrical hearing. Broadly, these fields can be clustered into two streams, one relating to the understanding of music perception in NH and the other to the study of perceptual ability following electrical peripheral auditory stimulation. Each is informed by an understanding of the underlying neurobiology of hearing to in turn inform the understanding of CI-mediated music perception. It is important to note that the schema is not an exhaustive representation of the processing modules that contribute to a listener’s forming a final, unified musical percept, but rather a conceptual model that guided the thought process during the current study. Although an important part of music perception, timbre perception is not included in the schema as it was not probed as one of the empirical investigations during the course of the study. The schema is anchored in the model put forward by (Peretz & Coltheart 2003), which provides a more comprehensive overview of processing modules that form part of the music perception pathway.

The next sections will provide a brief literature review to sketch the context within which the relevance of specific components of the schema can be understood. Emphasis is specifically on components included in this schema; the literature review does therefore not attempt to provide a comprehensive discussion of extant literature regarding music perception in either normal or electrically mediated hearing. Authoritative and expansive reviews regarding these aspects already exist (Deutsch 1998; Loizou 1999; Krumhansl 2000; Peretz & Zatorre 2003; McDermott 2004; Rubinstein 2004; Limb 2006a).

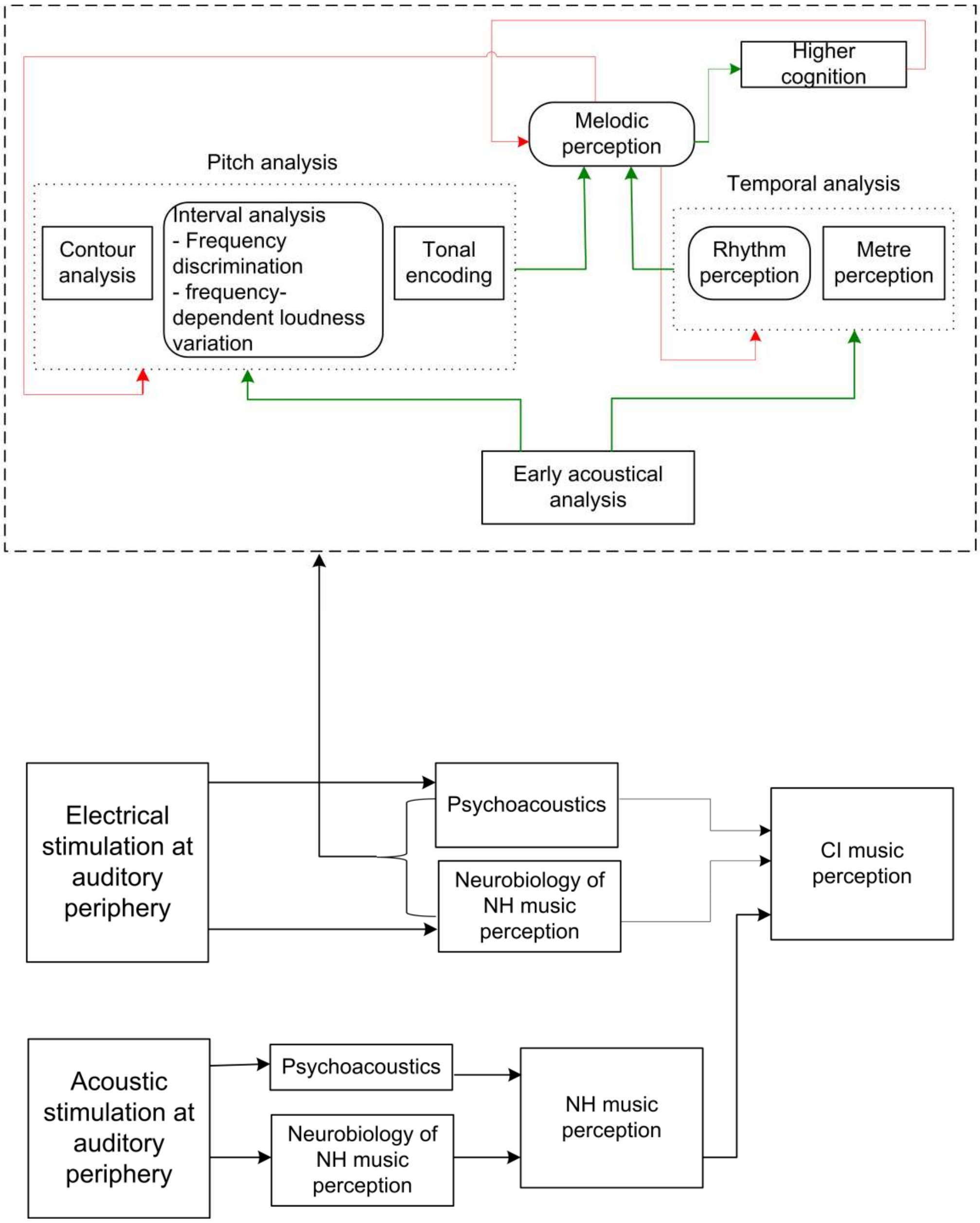

Figure 1.1: Schematic representation of the context of the study. The approach to improve understanding of CI music perception is informed by knowledge regarding psychoacoustics and neurobiology underlying music perception. The conceptual framework on which systematic investigation in this study is based is developed in the dashed frame. Processing modules that contribute to independent pitch and time-related analysis are enclosed in dotted frames. Resulting output feeds into a processing module at a higher hierarchical level, from where output is sent further again. Green arrows represent forward information flow; red arrows represent modulatory feedback. The rounded blocks indicate specific processing steps investigated during this study.

1.2.1 Music as auditory input

Owing to the substantial acoustic similarity between speech and music it is difficult to define why music is a unique auditory input. Both stimuli employ sequences of successive frequency changes that develop over time to communicate meaning according to a rule-based syntax (Limb 2006a), yet music is accompanied by an aesthetically pleasurable experience quite different from language. Perhaps the distinction resides in the perceptual meaning of music being conveyed through the relationship of individual tones to preceding and following ones within the context of the presented sequence (Peretz & Zatorre 2005). As such music is probably best defined as a collection of sound patterns built on the physical dimensions frequency and duration to convey perceptually meaningful pitch and time-related information organised according to the principles of pitch, rhythm and harmony (Krumhansl 2000; Gray et al. 2001; Limb 2006b).

Although pitch and rhythm are posited as the primary dimensions of music (Krumhansl 2000), successful music perception appears to rely foremost on analysis of frequency-related information (Zatorre 2001). Recent studies have shown that the inability to recognise music is the result of pitch perception deficits at various levels (Peretz 2002; Peretz & Hyde 2003; Foxton et al. 2004), which provides empirical support for the central position afforded to pitch in a theoretical framework of musical structure (Krumhansl 1990). According to this framework, different pitches, each associated with a specific frequency, are organised along musical scales. Within a scale each pitch is in turn associated with a tonal priority as dictated by the implied melodic scale (Krumhansl 1990), and so the relevance of a specific tonal pitch (and thus the associated frequency) within a tonal sequence depends on the relationship between frequencies rather than absolute frequencies per se (Krumhansl 1990; Trainor et al. 2002; Peretz & Zatorre 2005). The size of each melodic pitch interval is thus associated with a fixed frequency ratio (Hartmann 1998).

Similar to the tonal dimension, temporal cues in music (i.e. metre and rhythm) also find perceptual relevance only when duration of tones and neighbouring silences are heard in relation to those of the rest of the sequence (Peretz 1990; Peretz & Zatorre 2005).

However, unlike tonal information, temporal information does not convey the defining characteristics of music, but rather facilitates and supports music perception at an organisational level. This is in line with Jones's (1993) theory of joint accent structure, which posits that temporal accents coinciding with tonal accents increase perceptual saliency of melodic moments and so promote melody perception. The notion of temporal information fulfilling a supporting role during music perception has been underlined by studies which showed CI listeners' recognition of familiar melodies to improve in the presence of tonal as well as temporal cues, as opposed to recognition of melodies when only the pitch contour was preserved (Kong et al. 2004; Galvin et al. 2007).

1.2.2 Neurobiological foundations of music perception

Music perception is believed to be underlain by a modular cortical processing architecture, evident at various organisational levels (Peretz & Coltheart 2003; Koelsch & Siebel 2005). Firstly, processing of music and language information appears to be performed by largely independent processing networks despite some localised functional overlap between cortical areas. This view is based on collective data from studies with amusic subjects, who, following cortical lesions, presented with impaired music perception abilities while speech perception remained intact (Peretz et al. 1994; Nicholson et al. 2003).

The notion of a modular processing architecture underlying music processing is further supported by observations that cortical lesions may lead to impaired pitch-related perception without compromising the perception of temporal properties of music, and vice versa. As shown in Figure 1.1, the approach suggests that music processing modules are organised along two independent subsystems operating in parallel – one handling the pitch content of the acoustic input and the other dealing with the temporal content – that consist of several individual, hierarchically organised information processing steps (Peretz & Coltheart 2003; Koelsch & Siebel 2005; Warren 2008). The systems have been shown to be hierarchically organised with regard to both anatomy and function; elementary spectrotemporal properties appear to be processed in primary auditory cortex, followed by analyses of relationships between successive acoustic events, short-term retention of these patterns and other higher cognitive functions in a widely distributed network beyond primary auditory cortex (Zatorre et al. 1994; Platel et al. 1997; Patterson et al. 2002; Scott & Johnsrude 2003; Limb 2006b; Warren 2008).

The functional distinction between music- and language processing networks, as well as the distinction between pitch- and rhythm processing modules may be related to functional hemispheric specialisation. Observations of perceptual deficits with regard to tonal and temporal aspects of music in patients with unilateral brain damage (Peretz & Babaï 1992; Schuppert et al. 2000), together with results from neuroimaging studies (Zatorre et al. 1994; Liégeois-Chauvel et al. 1998; Schonwiesner et al. 2005), point to possible hemispheric specialisation for processing of tonal and temporal properties of complex sounds. On closer inspection, such lateralised specialisation appears to be linked to the differential contribution from temporal and spectral properties to speech and music and presents a possible framework according to which shared functionality between the two domains can be explained. Given that properties relevant to speech recognition are mostly contained in rapid temporal variations, while fine spectral structure conveys music-relevant information (Peretz & Hyde 2003; Zatorre 2003), a processing system that can handle both input types with similar precision would be required. Owing to the inverse relationship between frequency and time, improved spectral or temporal resolution will come at the expense of resolution in the complementary dimension if both were to be handled by a processing system wired only for one dimension. Zatorre and colleagues (2002) suggested that hemispheric specialisation developed to allow the auditory system to process both speech and music with the necessary precision. Such a hypothesis is supported by neuroimaging studies that show complementary responsiveness to fine temporal and spectral variations in the left and right auditory cortex respectively (Zatorre & Belin 2001). It thus appears that lateralisation is not directed at speech and music perception per se, but rather at functions dealing with temporal and spectral processing that, as argued earlier, form the underlying basis of the distinction between speech and music as auditory input.

1.2.3 Perceptual ability afforded by the CI device

In the case of sensorineural deafness, the cochlea, responsible for converting acoustic signals to a neurally encoded representation, is severely impaired (Loizou 1999). Auditory information is therefore not represented accurately and may be severely degraded, which in turn affects the fidelity with which the auditory nerve conveys information to the central auditory nervous system. Cochlear implants aim to circumvent this problem by decomposing a sound signal into its frequency components and subsequently stimulating the auditory nerve electrically.

The device consists of two units: the (speech) signal processor worn externally, and an implanted receiver–stimulator unit. The external components consist of a microphone and speech processor, usually worn behind the ear, and a transmitting coil. These are responsible for converting acoustic signals to electrical stimuli and subsequent transmission of the signals to the internal components. The internal unit is responsible for electrical stimulation of the auditory nerve and consists of a receiver coil, a decoder that generates pulsatile electrical stimuli (both implanted behind the ear), and an electrode array inserted in the scala tympani to stimulate auditory nerve fibres originating in the cochlea.

Place pitch has been shown to be an important pitch perception mechanism (Oxenham et al. 2004) and may be especially relevant for successful CI-mediated music listening (Townshend et al. 1987; Swanson et al. 2009), as purely temporal pitch is available only to about 300 Hz (Zeng 2002). Yet in CI-mediated hearing the resolution of the place pitch mechanism is governed by 22 discrete electrodes placed, in the case of the Nucleus device, at 0.75 mm intervals on the basilar membrane and reaching maximally only about two-thirds into the cochlea. Each electrode is associated with a frequency band, of which the width, in turn, is determined by the filter bank chosen for the specific signal processing strategy. For the Nucleus device the smallest filter width equals 125 Hz, and is applied only to the first eight or ten electrodes. Later electrodes (i.e. coding for higher frequencies) are associated with wider filters (and thus stimulatory frequency ranges). The place pitch mechanism in CI-mediated hearing thus allows only limited spectral resolution, which is probably reduced even further by factors such as filter overlap and unfocused current distribution (Hughes & Abbas 2006). The combined effect of acoustic–electrical conversion, reduced spectral resolution and mismatched frequency-to-place mapping (Fu & Shannon 2002) severely hinders CI users’ pitch perception ability.

Such coarse spectral resolution does not support CI-mediated music perception ability (Shannon 2005), especially when one considers that in almost 70% of all melodic intervals in Western tonal music, the pitch difference between two successive tones appears not to exceed two semitones (Vos & Troost 1989), translating to a frequency ratio of only 1:1.122. Since perception of the relative pitch changes between successive tones is so important to music perception, it follows that a perceptual system without sufficient pitch processing resolution will be unable to process an essential part of musical structure (Peretz & Hyde 2003). This is possibly one of the main reasons for poor music perception abilities of CI users.

Acoustic–electrical conversion, however, does yield sufficient information to support adequate speech perception. Speech appears to be fairly robust to spectral degradation (Shannon et al. 1995), probably since speech cues contained in spectral fine structure can be conveyed adequately through coarse spectral recoding. Further properties relevant to speech recognition are contained in rapid temporal variations. Taken together this implies that speech perception requires temporal cues to be delivered with high fidelity and less focus on fine spectral resolution (Shannon 2005).

Insights regarding hemispheric specialisation (as discussed earlier) and the associated processing benefits can advance the understanding of CI-mediated music perception. For example, based on the work of Mazziotta et al. (1982, as cited in Platel et al. 1997) and Phelps and Mazziota (1985, as cited in Platel et al. 1997), Platel et al. (1997) argue that difficult pitch perception tasks during music listening may induce analytical listening strategies subserved by left hemisphere activity. Considering then that speech perception (and per implication temporal processing) is subserved predominantly by left hemisphere structures, that several speech areas have been implicated in music perception (Koelsch et al. 2002; Levitin & Menon 2003) and that CI users receive regular activation of speech structures, it is possible that CI users' left hemispheric speech structures may be accustomed to receiving and dealing with impoverished information. This may in turn facilitate analytical processing to some extent during CI listeners' music perception when rhythmic cues (Fujita & Ito 1999; Kong et al. 2004) and lyrics (Fujita & Ito 1999; Gfeller et al. 2003) are available.

Recent imaging studies by Koelsch et al. (2004) and Limb et al. (2010) provide further support for the relevance of understanding processing constraints of the CI system in the context of the underlying neurobiology. Both studies showed that similar cortical structures are activated during CI-mediated music perception as for NH listeners, but that the pattern and intensity of activation differ. These findings underline that music perception is possible in electrical hearing, but that the quality of the input signal influences processing at cortical level, especially since the music processing system is posited to be subject to both bottom-up input and top-down feedback control (Purwins et al. 2008).

1.3 RESEARCH OBJECTIVES

1.3.1 Hypothesis

The hypothesis of this study is that investigation of CI-mediated music listening with simple yet musically relevant stimuli in sound-field conditions can provide useful insights that may allow perceptual outcome to be linked to the psychophysical abilities afforded by the CI device. Such insights may help to guide systematic efforts at improving signal processing strategies. The specific research questions put forward form only a limited part of a wider research effort regarding CI users' music perception abilities and experiences.

Figure 1.1 is a schematic representation of the thought process guiding the study. It is important to note that research endeavours regarding music perception in normal and electric hearing have not developed in parallel owing to each discipline addressing research questions specific to behavioural outcome experienced by the distinct end users. However, as shown in Figure 1.1, investigation into CI-mediated music perception can draw on a similar approach, using results from psychoacoustic experiments as well as insights regarding the underlying neurobiological mechanisms, to inform understanding of CI music perception in the context realistic listening conditions confronting a CI user.

1.3.2 Research questions

The context described in the earlier literature review and developed schematically in Figure 1.1 serves to guide formulation of research questions which address elements of both pitch and temporal analysis during CI-mediated music perception. Recombination of the two dimensions after initial separation at an earlier level in the processing hierarchy follows during melody perception. Output from this stage feeds forward to higher-order cognition.

The investigation aims to put forward an approach to provide a deeper understanding of factors that contribute to CI listeners' ability to perceive music-relevant auditory information. Systematic investigation of information processing as occurring in sound-field conditions allows CI-mediated perception of musical information to be interpreted with consideration to a hierarchically organised perceptual system, and so may provide a link, in future studies, for inferring perceptual outcome from adjustments that apply to processing strategy design.

Research questions relate to elements shown in rounded blocks within the dashed frame in Figure 1.1.

1. Considering that successful melody perception in the Western tonal tradition requires the pitch changes of successive tonal events to be interpreted in relation to those of flanking events (Krumhansl 2000; Peretz & Hyde 2003), what is the typical frequency discrimination ability of CI users in sound field-listening conditions, specifically measured over an extensive range of base frequencies?[1] Can the perceptual outcome be explained with regard to signal processor design, and if so, to which design features? How would considering CI-mediated perceptual

---

[1] At the time when this research question was drafted and the investigation initiated (2005), such frequency discrimination data had not been available.

outcome as the result of systems-based processing contribute to the understanding of perception of music-relevant information in electric hearing?

2. An earlier study regarding amusic listeners' rhythm perception abilities showed significantly poorer perception of pitch-varying rhythmic patterns than experienced by NH listeners (Foxton et al. 2006). Given that pitch and temporal information are posited to be handled by independent processing streams (Peretz & Coltheart 2003), a pitch processing deficit was not expected to influence rhythm perception ability. However, the results of the study by Foxton et al. (2006) showed that adding pitch complexity to tone sequences adversely affected amusic listeners' rhythm perception abilities. Since CI listeners can be regarded as functionally amusic owing to pitch processing constraints, it raises the question of whether their rhythm perception abilities are influenced to the same extent as that of amusic listeners owing to pitch processing constraints. Therefore, how do CI listeners typically perform on rhythm perception tasks when confronted with simple tone patterns in sound-field conditions when pitch and rhythm cues covary?

3. Music as a perceptual input is difficult to define. Investigating its perception is complicated by not knowing whether the listener experiences a tone sequence as a melody. Even more difficult is to assess perceptual success with regard to musical input if the accuracy of transfer of components that contribute to forming a musical percept is unknown. To improve the ability of the CI device to convey music-relevant information it is necessary to determine whether the CI device conveys sufficient information to facilitate perception of the melody-like character of a tone sequence. How can the transfer of melodic character during electrically mediated hearing be probed objectively? To what extent can CI listeners judge musical characteristics as defined by musical syntactic congruency in short Western tonal melodies compared to NH listeners? How can such insights contribute to the understanding of CI-mediated music perception?

4. With reference to the relationship between pitch and intensity perception (Stevens 1935) and its relevance in electric hearing (e.g. Arnoldner et al. 2006), how much

frequency-dependent loudness variation do CI listeners typically experience when presented with musically relevant stimuli in sound-field conditions? How does this compare with that experienced by NH listeners?

1.3.3 Approach

The specific research questions put forward were addressed through perceptual experiments (see dashed frame in Figure 1.1). To this end, task-specific stimuli and new test procedures, specifically for sound-field conditions, were developed. Since the ultimate goal of cochlear implantation is to create an artificial perceptual system that mimics the normal auditory system, results obtained from CI users in the respective tasks were compared to those from NH listeners to allow meaningful interpretation of systems-based CI-mediated music perception. Specific methodological designs as applied to address the respective research questions are discussed in more detail in relevant later chapters.

1.4 RESEARCH CONTRIBUTION

According to the context outlined in section 1.3, this study focussed on the systematic investigation of aspects that may further current understanding of CI-mediated music perception in sound-field listening conditions. Specific research contributions are listed below.

1. Sound-field frequency discrimination thresholds for CI users were determined across an extensive range of base frequencies, with specific consideration to the position of the base frequency relative along the frequency response curves associated with the filter bank. Results showed that, in listening conditions that resemble those experienced in daily life, the CI technology as is currently available may provide pitch resolution that is better than what would be expected with place coding of pitch, which is considered to be governed by the output of the filter bank to listeners. Findings were interpreted with regard to signal processor filter

properties. The results were encouraging considering the importance of fine pitch resolution for successful perception of Western tonal music and may help to guide future signal processing strategies aimed specifically at improving CI-mediated music perception. Typical thresholds were in line with results by Gfeller et al. (2007). The work was published in *Hearing Research* (see Pretorius & Hanekom 2008).

2. Rhythm perception abilities were investigated using music-relevant tone sequences under conditions of varying pitch and rhythmic complexity. This allowed the specific effect of a varying pitch pattern on rhythm perception ability to be investigated. Compared to results for amusic listeners (Foxton et al. 2006) in a similar study, the observations point to a possible influence of the position at which an auditory deficit originates on the extent of constrained perceptual outcome.

   The experimental design also allowed investigation of CI users' ability to perceive rhythmic patterns in the context of a tone sequence. The findings may help to further contemporary understanding of CI users' ability to perceive and understand music in realistic listening conditions. The work is currently under review for publication in *Cochlear Implants International* (see Pretorius et al. 2010).

3. CI listeners' ability to perceive musical character as conveyed during electrically mediated hearing was investigated using tasks that probed their ability to judge syntactic congruency of simple, single-voice melodies. Using novel melodies allowed CI listeners' perception of real-life melodies to be investigated without the confounding constraints associated with using familiar musical excerpts. The work presents a first approach to developing test modules to gauge whether sufficient music-relevant information is conveyed to support successful melody perception. Upon further refinement such modules may add and complement existing tests that assess CI-mediated music perception success. An article about this work has been submitted to peer review (see Pretorius & Hanekom 2011).

4. Investigation into the frequency-dependent loudness variation CI listeners experience when listening through their own processors in quiet but realistic listening conditions appeared to be similar to that experienced by NH listeners. The finding suggests that clinically assigned processor settings adequately mimic loudness perception abilities. The findings may help to guide further investigation into the influence of unbalanced loudness cues during music listening, for comment on the relevance of rigorous loudness balancing prior to commencing with music perception tests in sound-field listening conditions.

1.5 OVERVIEW OF THE STUDY

As described earlier (paragraph 1.2) the understanding of music perception in electric hearing is broadly informed by knowledge regarding music perception in NH and investigation into perceptual ability following electrical stimulation of the auditory periphery. The present study focused on the latter as shown in Figure 1.1.

As can be seen from Figure 1.1 the study is only a partial investigation of aspects that contribute to successful music perception. They are bound together by their functioning within the greater processing system underlying CI-mediated music perception and specifically since all investigations were performed in sound-field listening conditions. The approach presents an intermediate step in a hierarchy of possible approaches to reach a better understanding of music perception facilitated by electric hearing.

Figure 1.2 shows that existing findings from basic psychoacoustic experiments, conducted both at the electrode level and in sound-field listening conditions, were used as starting point to inform the present study. Investigation into basic psychoacoustic abilities of CI users when confronted with simple cues in sound field conditions, whether in isolation or in combined form in melody-like tone sequences, presents an intermediate level in the approach hierarchy.

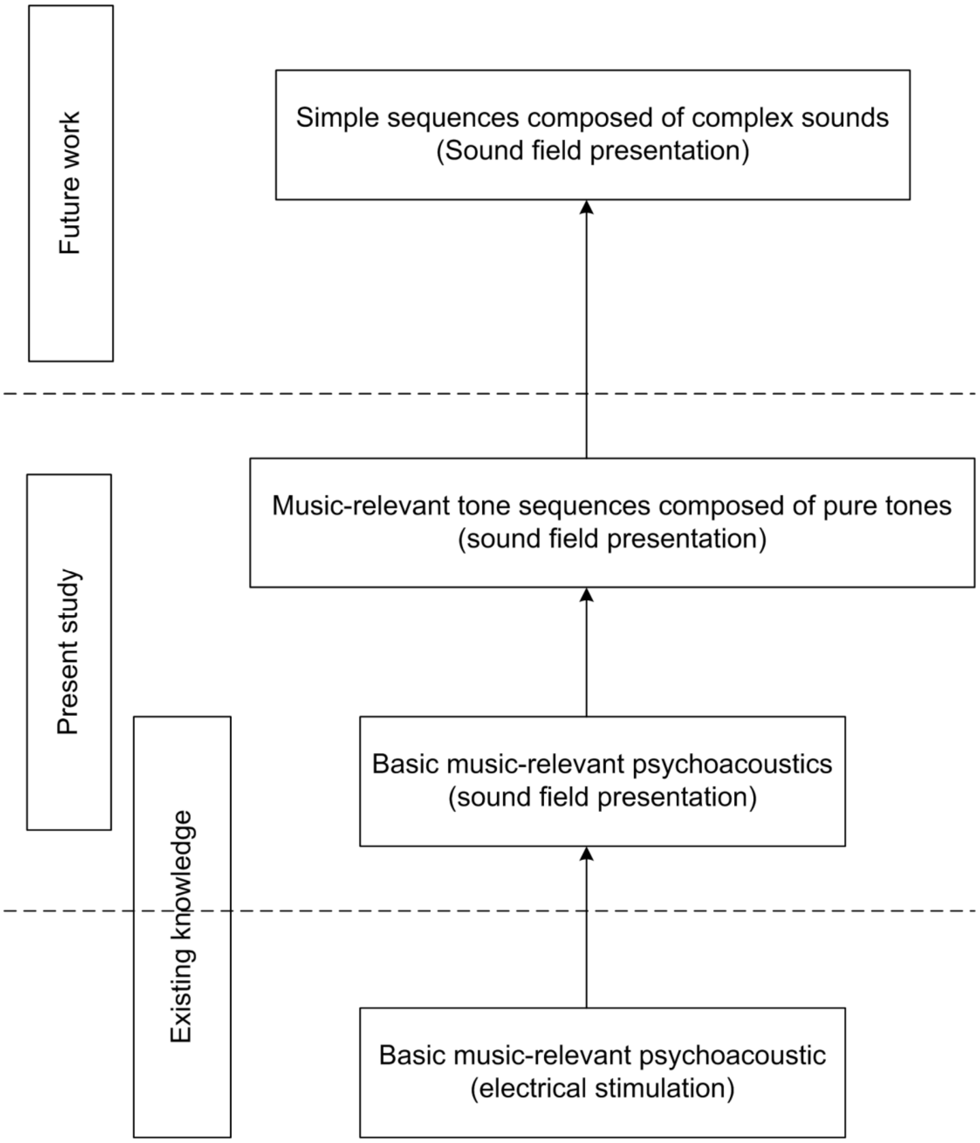

Figure 1.2: Hierarchical approach for studying CI-mediated music perception. The respective investigations reported in this study represent an intermediate step to link findings from electrode-level investigations to real-world perceptual outcomes.

In **Chapter 2** sound-field frequency discrimination was investigated and findings were interpreted with regard to speech processor design. Outcomes from investigation at this level can inform study at a subsequent level where basic cues are combined into simple, musically relevant stimuli presented in sound-field conditions. This was implemented during the investigation of CI users' rhythmic perception ability with regard to simple tones sequences where pitch and rhythm cues covaried (**Chapter 3**). The extent to which combined basic cues as provided by electric hearing facilitate melodic judgements was investigated in **Chapter 4**. The perceptual relationship between pitch and loudness and prompted a small investigation into frequency-dependent loudness variation as experienced in sound-field conditions (**Chapter 5**). Although loudness is not a direct cue for music listening, its relationship to pitch perception warrants investigation in a study of this kind.

Chapter 6 concludes this thesis with a brief general discussion of the study and concluding remarks, and offers recommendations for future work.

Chapter 2

FREQUENCY DISCRIMINATION ABILITIES OF COCHLEAR IMPLANT USERS IN SOUND FIELD

Pretorius, L.L. & Hanekom, J.J. (2008) Free field frequency discrimination abilities of cochlear implant users. *Hearing Research, 244,* 77-84

2.1 INTRODUCTION

CIs provide many profoundly deaf individuals worldwide with partially restored hearing ability. Although not quite comparable to NH, present multi-channel CI technology can transmit sufficient speech information to allow useful perception of speech (Wilson & Dorman 2007) and environmental sounds (Reed & Delhorne 2005). Music perception abilities, however, remain poor (Gfeller et al. 1997; Leal et al. 2003; Gfeller et al. 2005). Several music perception studies have shown that some cochlear implantees are able to perceive features such as tempo, rhythm (Gfeller & Lansing 1991; Kong et al. 2004) and lyrics (Looi et al. 2004; Gfeller et al. 2005) fairly well, but that perception of pitch-related features proves extremely challenging to most listeners (Gfeller & Lansing 1991; Gfeller et al. 1997; Kong et al. 2004; Sucher & McDermott 2007; Galvin et al. 2007). Even when pitch cues are presented in the absence of any confounding temporal cues, discrimination resolution of less than several semitones is not consistently observed in sound-field conditions (Pijl 1997).

The inability of otherwise NH listeners to perceive music successfully has been attributed to pitch perception deficits at various levels within the auditory processing system (Peretz 2002; Peretz & Hyde 2003; Foxton et al. 2004). Although both speech and music contain frequency information, fine spectral analysis appears to be more important for successful music perception than it does for speech perception. Adequate speech understanding can be achieved with as little as four spectral channels (Shannon et al. 1995), while at least 32 channels seem to be necessary for successful music perception (Smith et al. 2002; Kong et al. 2004). Shannon (2005) further argues that speech and music have different requirements for spectral resolution, which is in line with findings from studies regarding the neurocognition of speech and music in NH listeners (Zatorre et al. 2002).

Given the importance of pitch for successful music perception (in the Western tonal tradition), it follows that insufficient transmission of frequency information to the auditory system will adversely affect the quality of music perception. Despite pitch perception being mediated by both a place and a rate code (and possibly also a combination of the two) (Smith et al. 2002; Moore 2003) correct tonotopic representation appears to be especially important to successful pitch perception in NH listeners (Oxenham et al. 2004). CI-mediated transmission of frequency information also relies on the place pitch mechanism to a large extent (Swanson et al. 2009). After a signal has been passed through a set of bandpass filters, specific electrodes associated with these are activated, with the stimulus strength corresponding to the energy in the particular filter band. Spectral peak extracting algorithms (e.g. SPEAK and ACE) may then select the number of filters containing the strongest output and only apply stimuli to the associated electrodes. The electrodes in turn stimulate neural populations at (ideally) discrete positions along the length of the cochlea. However, because much of the frequency range relevant to human auditory tasks is handled by only a few electrodes during electrically mediated hearing, only limited spectral resolution is available to CI users. Correct transmission of frequency information is further hindered by channel interaction (Hughes & Abbas 2006) as well as mismatched frequency-to-place mapping (Fu & Shannon 2002), causing unintended neural populations to be stimulated.

The work reported in this chapter investigated the frequency discrimination abilities of cochlear implant users when presented with pure tone stimuli, across a wide frequency range, through their own processors in sound-field conditions. Because music is a multifaceted and acoustically complex auditory input, efforts to improve music perception abilities of CI users may benefit from studying music in terms of its comprising components, similar to the modular approach suggested for studying music-related deficits in neurologically impaired individuals (Peretz & Coltheart 2003).

With reference to the conceptual framework introduced in Figure 1.1 (Chapter 1), frequency discrimination represents an important module within the pitch processing stream during music perception. As shown in Figure 1.1, frequency discrimination contributes to the pitch analysis module at a processing level where the two primary dimensions of music information occur independently before combining at the next hierarchical level to effect “whole” melody perception. Assessing sound-field frequency discrimination abilities may thus present an early step in a modular approach towards music perception in CIs, since it would evaluate the system’s ability to transmit one of the fundamental components of music in a setting which a user would perceptually be confronted with. By using simple stimuli, results of behavioural experiments can potentially be interpreted in terms of sound processing and associated neural stimulation, possibly providing a link between results from pitch perception studies performed with stimuli delivered at electrode level (Pijl & Schwarz 1995a; Pijl 1997) and those conducted at a perceptual level where stimuli consist of true music examples presented through users’ speech processors (e.g. Gfeller et al. 2005).

The aim of the work presented in this chapter was twofold. Firstly, the typical frequency discrimination ability of CI users in sound field was investigated. At the time the investigation was initiated such data had not yet been measured systematically over an extensive frequency range and it was therefore deemed appropriate to measure. (Such data have in the meantime also been reported by other investigators (e.g. Gfeller et al. 2007) and the results of the investigation reported here are in line with those earlier ones.) Secondly, the investigation aimed to determine whether the observed frequency discrimination ability could be influenced by the stimulus frequency's position relative to filter response curves (see Figure 2.1). Several studies have shown that intermediate pitch percepts can be generated by dual electrode stimulation (McDermott & McKay 1994; McKay et al. 1996; Donaldson et al. 2005; Kwon & Van den Honert 2006). The reasoning behind the approach was that if the intermediate pitch percepts obtained from electrode level stimulation are linked to the differential distribution of current within a filter band, it should be possible to replicate the trend in sound field, since stimulus frequencies at different positions along the filter response curve would cause differential activation of more than one electrode quasi-simultaneously.

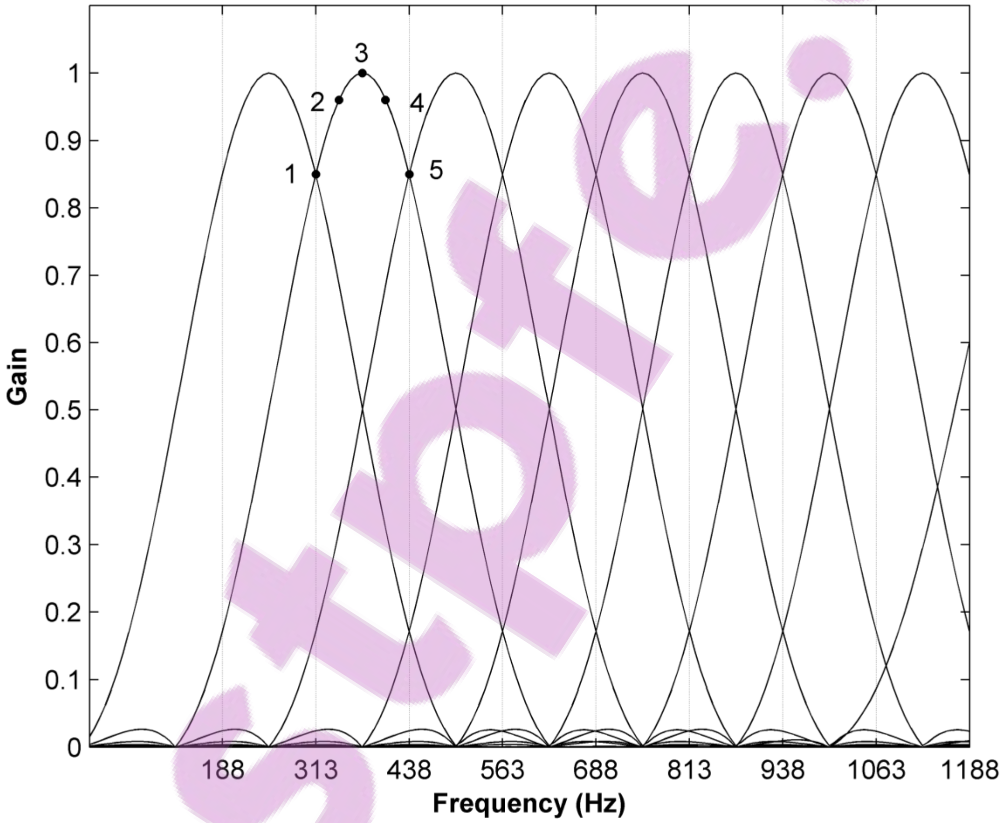

Figure 2.1: Frequency response associated with different filters as implemented in the ACE processing strategy. Numbers represent the different positions relative to which frequency discrimination thresholds were measured (indicated for one filter only).

2.2 METHOD

2.2.1 Subjects

Five post-lingually deafened adults implanted with the Nucleus 24 array participated in the investigation. All subjects had experienced profound hearing loss for more than 10 years and had more than six months' experience with the implant system. Subjects all used ACE as their normal speech processing strategy, with either the Esprit 3G or the Freedom speech processor, in a monopolar stimulation mode. Three participants were implanted bilaterally, but for these subjects only the implanted ear subjectively regarded by the subject as the clearer of the two was used during the course of the study. Electrodes 1 and 2 were inactive for S3, S10 and S14, while electrodes 1 and 22 were inactive for S13. All electrodes were active for S11. All subjects gave written informed consent, according to the guidelines from the relevant ethics committee, for their participation prior to commencement of the study and were compensated for their time at the end. Three subjects have participated in previous CI studies at our laboratory. The context in which the sample (which was one of convenience) was recruited is described in the Appendix and other relevant demographic details are provided in Tables A1 and A2 (Appendix A).

2.2.2 Stimuli

Since the study concerned assessment of CI users' frequency discrimination abilities in sound-field conditions, stimuli were chosen and presented to approach everyday listening conditions. Firstly, individualised reference frequency sets were compiled for each subject according to frequency responses of filters within the test range, as specified by the subject's clinical map. The test frequencies ranged between 250 and 1250 Hz, i.e. the frequency range with linearly spaced filters. This frequency range corresponds to fundamental frequencies of musical notes between C4 and D6.

All stimuli were 500 ms pure sine tones generated in Matlab 6.5[2] on a personal computer. Harmonics were suppressed by 45 dB or more across the frequency range used, as determined with a Larson-Davis 824 type 1 sound level meter. Amplitude ramps (30 ms) were included at the beginning and end of each tone to reduce onset clicks. Reference frequencies corresponded to frequencies at the peak of each filter, the cross-over between adjacent ones, and two points midway between the peak and cross-over positions on either side of the peak (see Figure 2.1). Thus, depending on the settings (filter width and active electrodes) specified by the subject’s clinical map, a set of between 23 and 27 reference frequencies could be compiled for each subject. A subject’s frequency discrimination ability was determined relative to each reference frequency within his/her individual set (also see Figure 2.2 for exact reference frequencies). Initial probe frequencies were always 100 Hz higher than the reference frequency, from where the difference between reference and probe frequencies was adapted according to listener response. Since the tested frequency range only spanned the region where filters are linearly spaced, using a fixed initial step size ensured that approximately equally spaced neural populations were targeted initially, regardless of reference frequency. More details regarding probe frequencies are provided in the section on experimental procedure.

Because stimuli were to be presented in sound field, it was important to present stimuli at the same perceived loudness level to all subjects, rather than at a fixed intensity level. A loudness estimation procedure was hence used to compile subject-specific sets of intensity levels corresponding to different loudness levels, by scaling 1 kHz tones ranging between the lowest and highest, yet comfortable, loudness levels. Stimuli were presented 20 times at each of 10 equidistant intensity levels within this range to derive an average estimated loudness percept at each level. All stimuli were subsequently presented at the intensity level corresponding to a subject’s 75% perceived loudness level, interpolated from the curve obtained with the 1 kHz tones. These were 80, 90, 80, 85 and 77 dB for subjects 3, 10, 11, 13 and 14 respectively.

Stimuli were not loudness balanced to one another. Although balancing is usually performed to ensure that loudness is not used as a cue (e.g. Collins & Throckmorton 2000) it was

---

[2] www.mathworks.com

reasoned that for the present experiments, electrode loudness balancing performed during the normal clinical mapping procedure would be sufficient (and preferred, as argued later) for sound-field testing conditions. After initial electrode-specific determination of threshold (T) and comfort (C) values, the clinical mapping procedure entails balancing the loudness of electrodes at C-level by adjusting C-values, and balancing the electrodes at 25% and 50% above T-level by adjusting T-values. Electrodes are subsequently compared in groups of five with an overlap of one electrode. The T- or C-level of a particular electrode can then be adjusted if it does not sound equally loud as its neighbours.

It was further reasoned that electrode loudness balancing would be adequate for the present experiments using pure tone stimuli, as these would primarily activate one electrode with weaker activation of neighbouring electrodes. Assumedly loudness-balanced electrodes should result in pure tone stimuli that are largely loudness balanced. To test the assumption, and whether frequency discrimination thresholds were influenced by using unbalanced stimuli, two subjects (S11 and S13) participated in an additional experiment to compare frequency discrimination abilities at a number of frequencies from the original range. Both loudness-balanced and unbalanced stimuli[3] were used. In the unbalanced condition electrodes were loudness balanced according to the usual clinical mapping procedure, but not the stimuli. Loudness balancing was performed by comparing sound-field loudness of pure tones at 12 evenly spaced frequencies between 200 and 1200 Hz, to that of a 1 kHz tone presented at the subjects' respective 75% loudness levels. The listener's task was to adjust probe stimuli to be equally loud, just louder and just softer than the reference stimulus. Three repetitions per frequency had to be completed per task, from which the average deviation from the 75% loudness level was calculated from all the data at a particular frequency.

The two subjects were always consistent in their adjustment of softer, louder and equally loud. Results showed that no more than a 4.7 dB (on average 1.8 dB) deviation from the reference 75% loudness level was necessary for equal sound-field loudness perceptions across the frequency range tested. Furthermore, as will be shown later, no significant differences in frequency discrimination thresholds were found for clinically loudness-

---

[3] This is defined as electrodes being loudness balanced according to the usual clinical mapping procedure, but not stimuli.

balanced maps (unbalanced condition) compared to those with loudness-balanced stimuli. Together, these results served as further motivation for not performing loudness balancing of stimuli in the main experiments. (The loudness balancing tasks performed here prompted a subsequent investigation regarding frequency-dependent loudness variation experienced by CI users in sound-field listening conditions, as described in Chapter 5.)

All stimuli were presented in a sound-proof booth. Standing waves produced in the sound booth may be a possible concern in sound-field listening conditions. Pure tones (10 s signal duration) were recorded across the frequency range and analysed. At all intensities used, the measured intensities were within 0.5 dB of the setpoint, with no measureable variation in the signal intensity, so that a possible influence of standing waves was ruled out. Total harmonic distortion was never larger than 3.5% (1.2% on average) across the entire range of intensities and frequencies used. The second harmonic was always the largest and was always suppressed by more than 35 dB. These measurements were taken with the same sound level meter as mentioned earlier.

The speaker (Yamaha MS101 II) was located approximately 1 m in front of the subject, but on the side of the implanted test ear. To ensure that the speech processor microphone did not produce significant harmonic components, harmonic distortion of the combination of loudspeaker and microphone was measured. Across the frequency range used, the largest harmonics were always suppressed by more than 35 dB.

2.2.3 Experimental procedure

Frequency discrimination thresholds in both the loudness-balanced and unbalanced conditions were determined using an adaptive two-alternative forced choice (2AFC) procedure. Each trial consisted of two 500 ms pure tones separated by a 1 s interstimulus gap, with one tone always corresponding to the reference frequency and the other being the probe frequency. Subjects were asked to indicate which of the two tones had the higher pitch by selecting either of two buttons indicated on the screen. No feedback was provided and subjects were not allowed to re-listen to a tone pair. However, a subsequent tone pair would not be presented before user response had been received, allowing subjects enough time to make a decision. Three practice trials, unknown to the subjects, were included at the start of each experiment session. These had to be completed successfully (i.e. correct response provided on three successive presentations) before the actual session commenced. Since a single session could be quite lengthy, an indicator showing estimated progress was included in the user interface.

An experiment session consisted of threshold determinations relative to three reference frequencies performed in triplicate. The three reference frequencies were preselected from the set of reference frequencies described earlier and were, as far as possible, of the same frequency class (i.e. peak, cross-over or midway frequencies relative to filter response curve). During threshold determinations reference frequencies were presented randomly to prevent the subject from tracking the change in probe frequency. For the same reason, reference and probe tones were also randomly assigned to either the first or the second presentation interval. The software controlling the experiment was, however, designed to keep track of subjects' previous responses at a specific reference frequency, in order to present stimuli correctly according to the adaptive technique. In order to reliably determine the average discrimination threshold relative to each reference frequency, two repetitions of each experiment session had to be completed, resulting in six threshold values at each reference frequency.

The first presentation of a probe frequency was always 100 Hz above the relevant reference frequency. Step size (difference between reference and probe frequency) was subsequently adjusted according to Levitt's (1971) transformed up-down staircase technique using a 2-down/1-up decision rule. A staircase technique allows the test stimulus to alternate between the values where it is just discernable from the reference stimulus and just not discernable, by tracking the listener's response behaviour. As soon as a non-discernable value is reached, the difference between probe and reference stimulus is increased until the probe can again confidently be distinguished from the reference. Each such response alternation is counted as a response reversal. A presentation run constitutes a succession of stimulus presentations between two response reversals. According to the decision rule used here, the difference between the probe and reference frequencies would only be reduced after two successive correct responses, while it would increase after a single incorrect response. This allowed the

71% correct point to be approached, which was accepted as the threshold. By adapting the step size according to a specific factor after successive responses, the discrimination threshold could be homed in on accurately. Step size changed by a factor of 2 (i.e. halved or doubled relative to the previous difference) during the first six response reversals, then by 1.7 for reversals 7 through 12 and by 1.3 for the last six reversals. An experiment session terminated once 18 response reversals had been recorded for each reference frequency. The threshold relative to each reference frequency was calculated as the average of the last five mid-run estimates.

Experimental sessions were scheduled according to subjects' availability, and hence data collection was completed over the course of several weeks. Subjects were briefed on the aim of the study and the nature of the experimental procedure prior to commencement, but were not aware of the algorithm controlling stimulus presentation.

2.3 RESULTS

Figure 2.2 shows the average normalised frequency discrimination thresholds obtained for the individual subjects relative to each reference frequency. Averages were calculated from six data points at each reference frequency. Because behavioural data obtained during psychophysical studies are prone to some variability, the Extreme Studentised Deviation (ESD) technique was used to test for outliers, which were then excluded from the average. No more than one outlying observation was found in cases where outliers were detected, leaving at least five data points from which to calculate the average.

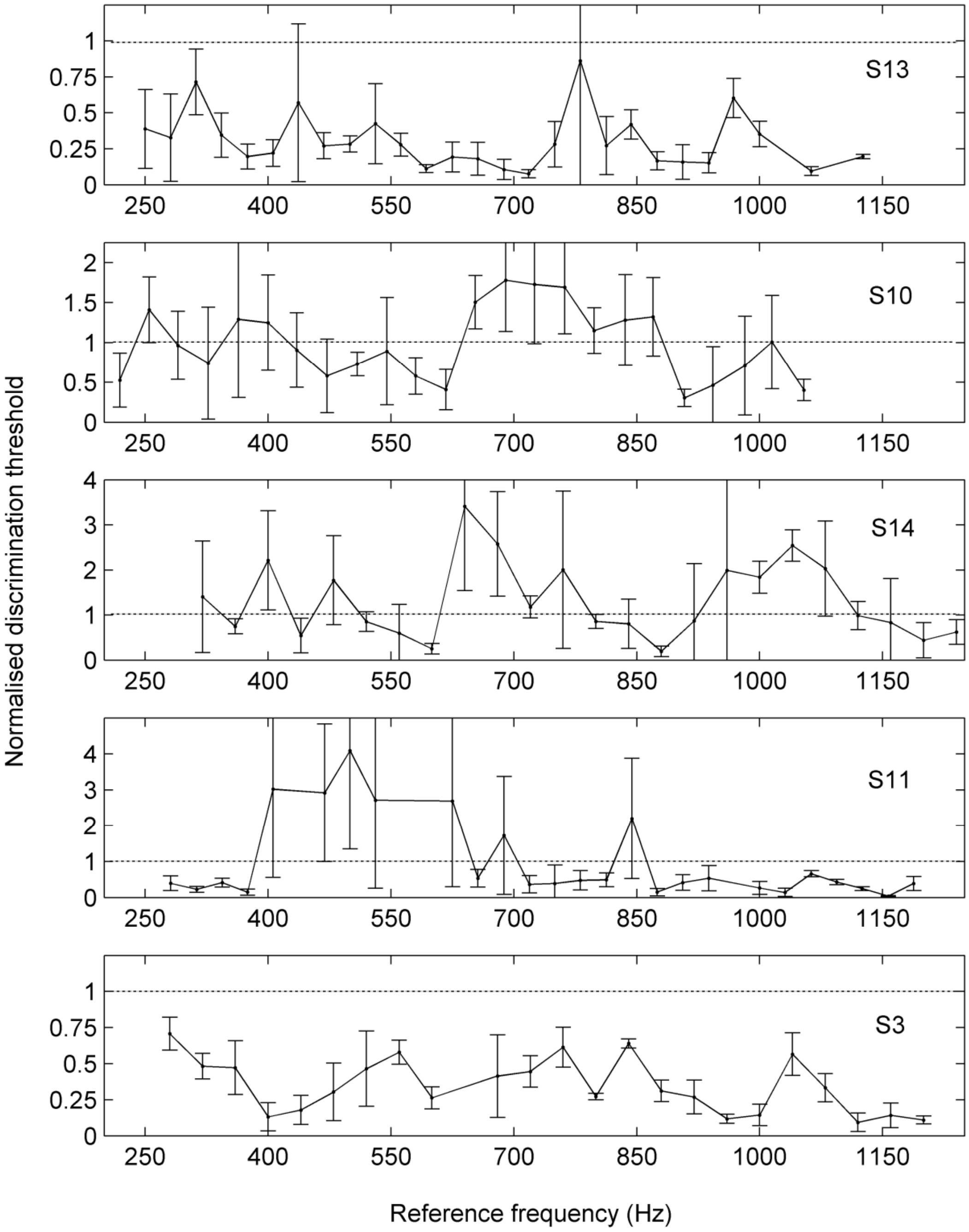

Figure 2.2: Average normalised frequency discrimination thresholds as a function of reference frequency. Thresholds are normalised to filter width as specified by the applicable frequency allocation table. The dashed line indicates one filter width. Individual subjects' results are shown in the respective panels.

Cases for which a frequency discrimination threshold could not be determined,[4] resulting in fewer than five replications for a specific reference frequency, were excluded from the analysis. Less than 3% of the total number of data points was excluded. Because of substantial intersubject variability, discrimination thresholds were normalised relative to filter width to allow all subjects' results to be viewed on the same scale. For the purposes of this study, a filter's width was defined as spanning the frequency range between the cross-over frequencies of two flanking filters. It should also be noted that because subjects used different frequency allocation tables, reference frequencies were not the same for all subjects. Figure 2.2 thus serves as a general view of the frequency discrimination trend across the frequency range tested. The dashed line indicates thresholds equal to one filter width.

It is evident from the data that not all subjects performed equally well during frequency discrimination tasks, nor was frequency discrimination behaviour similar across the frequency range tested. For two subjects, average discrimination thresholds were always less than one filter width, while results of the other three subjects showed discrimination thresholds of more than one filter width at several data points. For S11 large thresholds were consistently observed in the frequency region between 400 and 625 Hz, which may be attributed to subject-specific anatomical or physiological factors, since thresholds larger than those observed at other reference frequencies also presented in this region in two re-test sessions. However, all subjects seemed to exhibit a similar frequency discrimination trend in that some reference frequencies were associated with small discrimination thresholds, while others were associated with larger thresholds. Despite this observed variability, a large proportion of threshold values were smaller than one filter width, which a binomial analysis showed to be highly significant ($p < 0.001$). The analysis is essentially a hypothesis test whereby the probability associated with the mean not being equal to a forecasted value owing to a systematic influence of a variable is determined. The result indicates that these CI users were often able to discriminate pure tones less than one filter width apart in a sound-field situation.

---

[4] Frequency limits, above which the experiment would terminate, were set at initial experiment set-up. It sometimes happened that listeners failed to approach a frequency discrimination threshold, usually observed as very large step sizes. In such a case the specific run would terminate and it would be recorded in the results sheet that a frequency limit had been reached during the specific run. Frequency limits were always set at least 1000 Hz above the reference frequencies tested.

As mentioned earlier, minimal deviation from the 75% loudness level was necessary to achieve equally loud stimuli across the frequency range tested. Results from two subjects who repeated the frequency discrimination task with both balanced and unbalanced stimuli to test the influence of loudness imbalance are shown in Figure 2.3. Frequency discrimination thresholds smaller than one filter width were consistently observed with both balanced and unbalanced stimuli. For S11 the four (unbalanced) data points larger than one filter width presented in the same frequency region in which large frequency discrimination thresholds were observed during the initial discrimination tasks. Results of a Mann-Whitney test showed no statistically different frequency discrimination behaviour between balanced and unbalanced stimuli ($p > 0.05$).

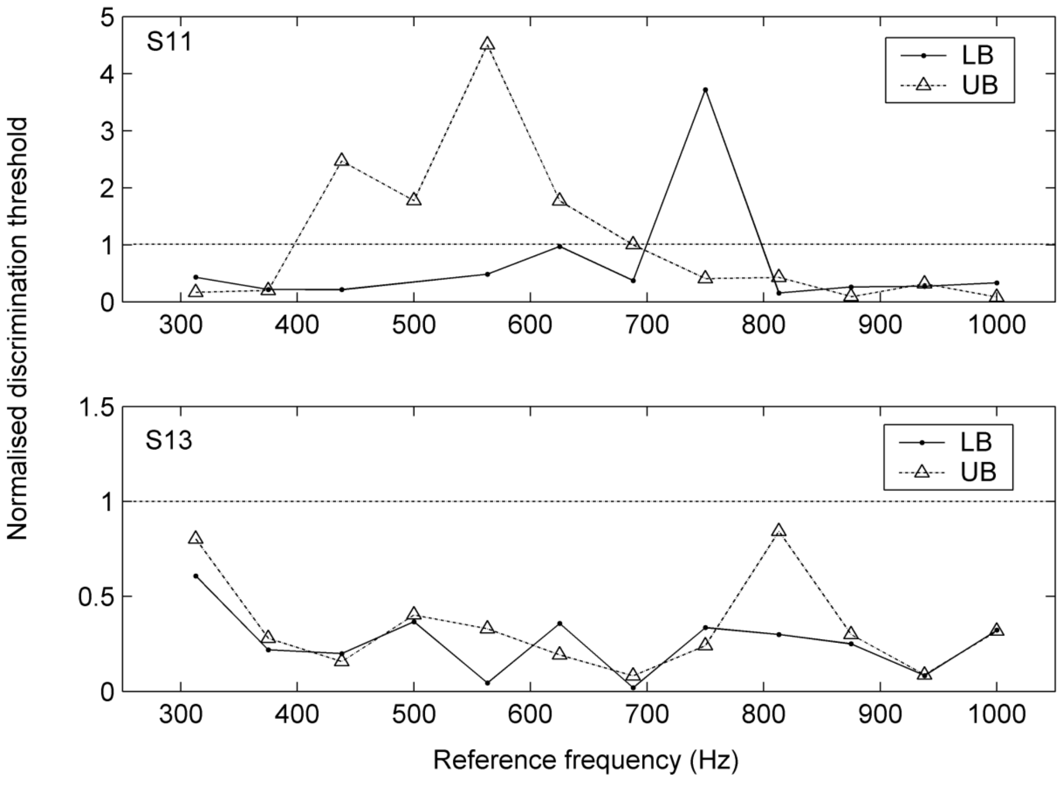

Figure 2.3: Average normalised frequency discrimination thresholds as obtained with loudness-balanced and unbalanced pure tone stimuli for two subjects.

In order to determine whether the frequency discrimination behaviour shown in Figure 2.2 could have been influenced by speech-processor-related parameters, the effect of filter location (a low, middle and high frequency region from the tested frequency range) as well as reference frequency position relative to the filter's frequency response curve were investigated. Figure 2.1 provides a schematic representation of the reference frequency positions and filter locations. Positions 1 and 5 correspond to the left and right cross-over frequencies respectively, with position 3 associated with the peak frequency. Positions 2 and 4 are frequencies halfway between the peak and left and right cross-over frequencies respectively. Owing to different absolute reference frequencies, as well as some missing value cases (as explained earlier) the data set had to be reduced somewhat to obtain a complete set for analysis purposes.

Figure 2.4 shows average normalised thresholds at five positions along the filter response curve for filters located at a low, middle and high frequency (relative to the tested range) for each subject. Results from a binomial test using the complete data set showed that a statistically significant proportion (73%, $p < 0.05$) of discrimination thresholds was smaller than one filter width. Furthermore, 64% of thresholds were smaller than half a filter width at position 3 ($p < 0.05$). This compares to significant proportions (78%, 71%, 67% and 67%, $p < 0.05$) of discrimination thresholds at positions 1, 2, 4 and 5 being smaller than one filter width, but not smaller than half a filter width. The result shows that these listeners were often able to reliably discriminate different pitches falling within the same filter band, especially when the reference frequency corresponded to that of the filter's peak (position 3).

A general linear model with a three-factor design further showed that both filter location and position of the reference frequency relative to a filter's frequency response curve, as well as their interaction, have significant effects on frequency discrimination behaviour ($p < 0.02$). However, further analyses indicated that differences with regard to frequency discrimination behaviour were not always the same for a specific position, but rather that the position effect appeared to be dependent on the particular filter.

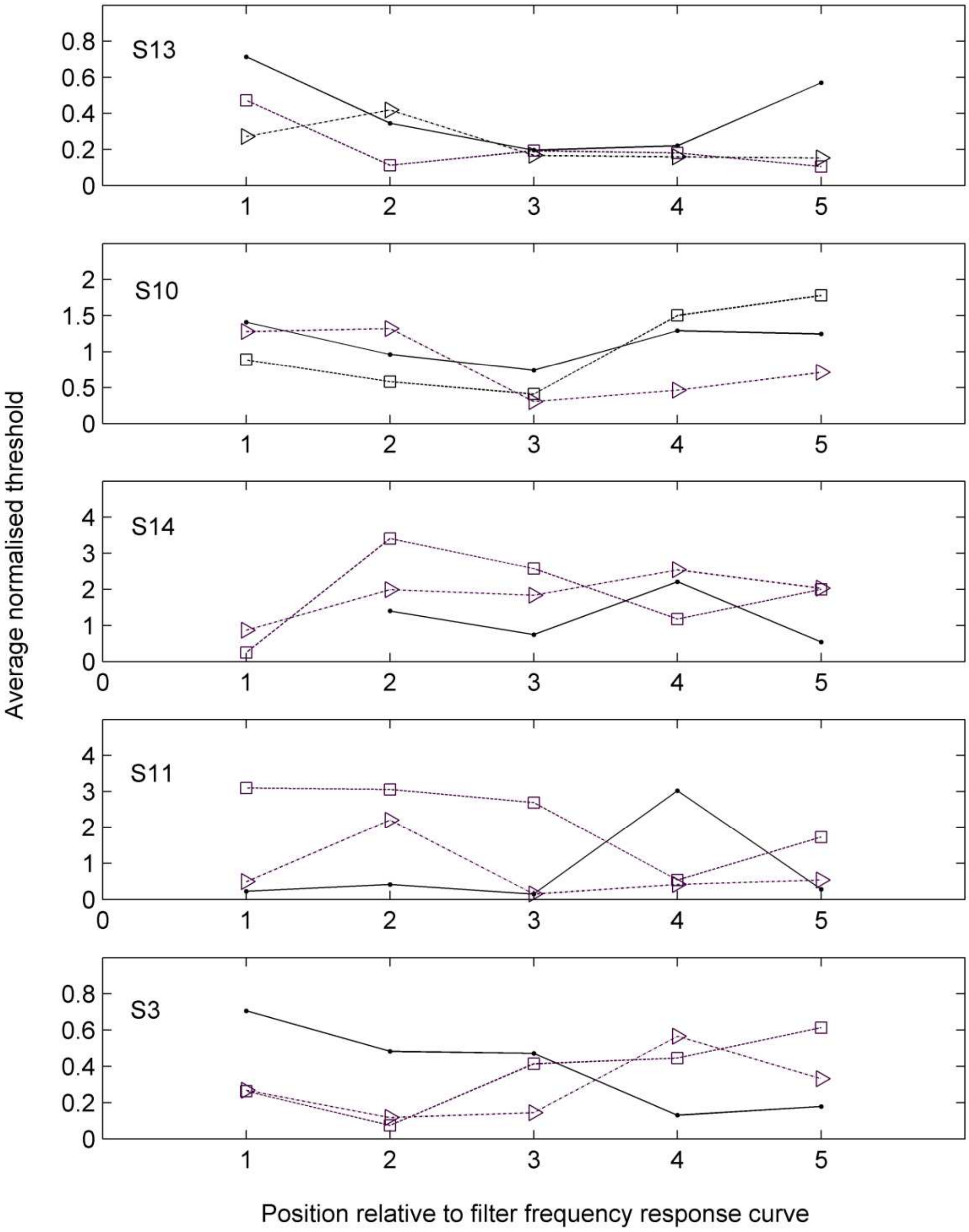

Figure 2.4: Average normalised frequency discrimination thresholds as a function of position along filter frequency response curve. Different lines represent frequency discrimination behaviour from filters representing low (dots), middle (open squares) and high (open triangles) frequencies in the range tested. Individual subjects' results are shown in the respective panels.

2.4 DISCUSSION

The results from this investigation show that (i) CI users were able to discriminate between different pure tone frequencies falling within one filter band and (ii) that discrimination behaviour was influenced by the position of reference frequency relative to the filter's frequency response curve. Taken together, the results indicate that finer frequency resolution than would be expected, could be available to CI users.

This is an unexpected result, considering the dominant place pitch effect in electrically mediated hearing (McDermott & McKay 1997) and, given fixed stimulation rate, the limited temporal pitch cues available in the present experimental conditions. However, earlier pitch perception studies conducted at the electrode level have shown that dual stimulation of adjacent electrodes, either simultaneously (Townshend et al. 1987; Donaldson et al. 2005; Firszt et al. 2007) or separated by short gaps (McDermott & McKay 1994; Kwon & Van den Honert 2006) may generate more pitch percepts than dictated by the number of electrodes. Firszt et al. (2007) showed that the effect could be associated with the proportion of current delivered to two simultaneously activated electrodes, so that different pitch percepts could effectively be created by deliberate manipulation of current on adjacent electrodes. This is in line with results from studies regarding vowel discrimination where Dorman et al. (1992; 1996) suggested that the frequency of vowel formants could be signalled by the energy balance between two adjacent filter bands.

The Nucleus device does not allow true simultaneous dual electrode activation and therefore improving frequency resolution through current manipulation may not be feasible. However, quasi-simultaneous electrode activation patterns can activate different neural populations within the same stimulation cycle. The neural activation strength is strongly affected by the stimulating frequency's position relative to the filter response curve. Frequency input at different positions along the filter response curve may thus naturally create a neural targeting mechanism by causing small shifts in the activated neural population. It is possible that the smaller than expected frequency discrimination thresholds observed here may be attributed to operation of such a mechanism.

Each position along the filter response curve investigated in the present investigation represented a frequency that falls within two or three adjacent filter bands. As shown in Figure 2.1, no more than 6 dB attenuation (gain of 0.5) will be experienced at any of these positions. This should be sufficient to contribute substantially towards activation of these filters' associated electrodes, as this would normally be within the dynamic range of an electrode. Such differential activation of neighbouring electrodes may result in slight shifts in the activated neural population upon slight adjustment of the pure tone input frequency. This would in turn cause an overlapping neural population to receive substantial activation from more than one electrode within the same stimulation cycle, thereby increasing the strength of activation in the specific neural population. Given that stronger activation increases action potential frequency, stronger activation in a specific neural population might be interpreted as a more salient frequency cue. A schematic explanation of the mechanism is offered in Figure 2.5, showing how filter output is associated with electrode activation and consequently drives neural activation. As can be seen in the schematic representation, a frequency that corresponds to the peak of the frequency response curve would activate three electrodes simultaneously (represented by the red, green and blue ranges), and consequently three overlapping neural populations. The neural population receiving strongest activation (i.e. most overlap) may thus generate a more salient frequency message that feeds into subsequent processing modules.

It is interesting to note that in all the aforementioned studies, improved frequency resolution was observed when stimuli were manipulated deliberately at the outset of the experiment so that different current proportions would be delivered to adjacent filter bands. Here, however, stimuli were presented in sound field and electrode activation was governed completely by speech processor output. This suggests that proportional current delivery as determined by the speech processor may translate effectively to differential neural activation. A differential neural activation mechanism as might be in play in these experimental conditions may therefore have a similar perceptual effect as deliberate current steering. It further appears that the effect is strong enough to translate to a perceptual outcome even when it functions within the complete, electrically stimulated peripheral auditory system. Although the number of additional pitch steps that can be made available through such a mechanism was not specifically determined in this study, the observation of a statistically significant proportion of discrimination thresholds being smaller than half a filter width at position 3 implies that at least double the present frequency resolution may be available at some positions. It thus appears that frequency resolution may be improved by exploiting the characteristics of existing speech processors. It should be noted, however, that physiological and anatomical characteristics, such as neural survival, position of electrode array relative to the basilar membrane and formation of scar tissue may influence individual performances and the benefit derived from a possible functionally available endogenous steering mechanism as described here.

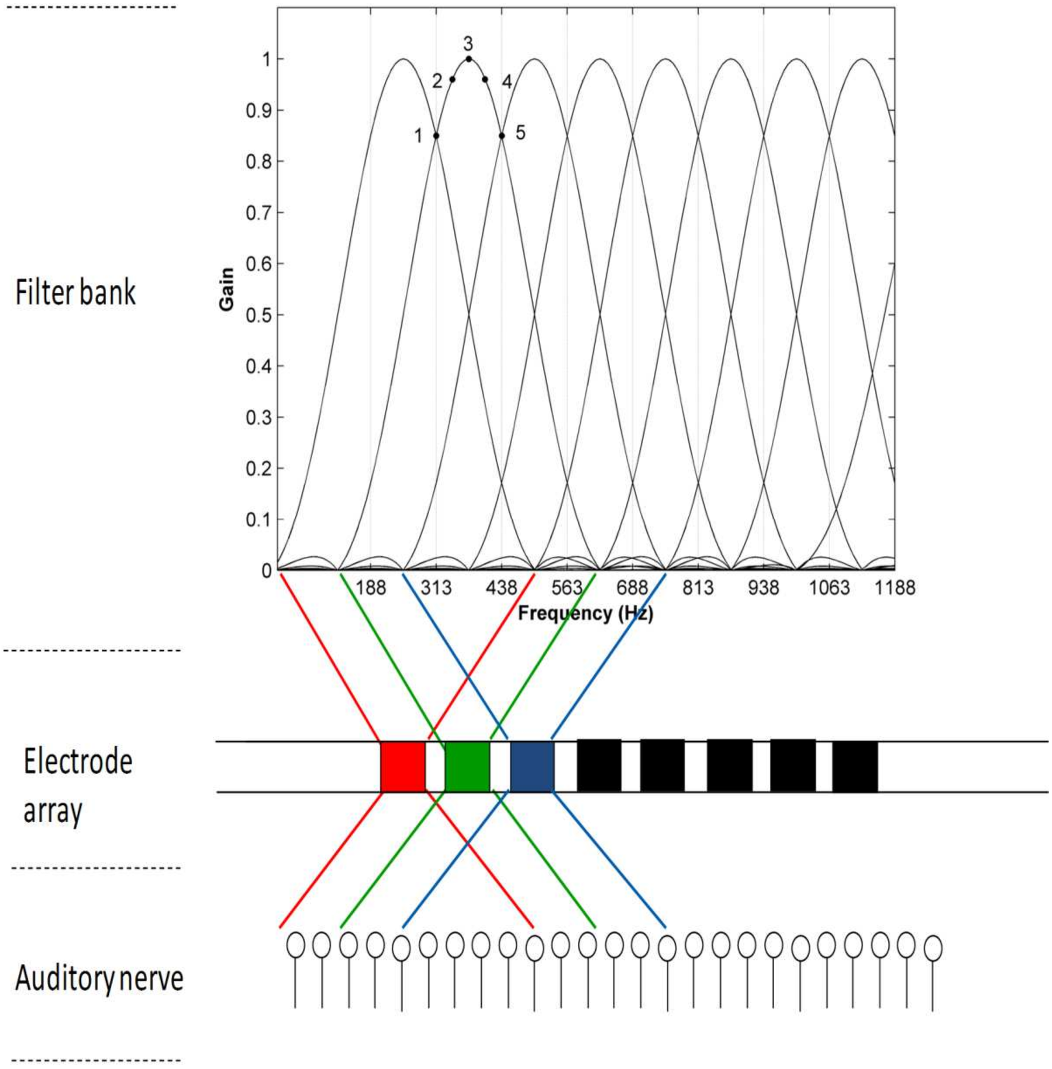

Figure 2.5: Schematic representation of possible mechanism of the neural steering mechanism as may be in operation during sound-field frequency discrimination

The frequency discrimination behaviour observed in this investigation also hints at CI users' possible music perception abilities. Although musical tones are rarely simple sinusoidal tones as used in this study, the results may be a relevant first indication of the level of pitch resolution that may be available to CI users when listening to music.

In Figure 2.6 frequency discrimination is expressed as the ratio of the actual threshold frequency to the reference frequency. This translates to the discrimination threshold in number of semitones, as shown on the right-hand ordinate.

Successful melody perception, regarded as central to music perception in the Western tonal style, relies strongly on the correct recognition of such relative rather than absolute pitch distance between two successive tones (Peretz 1990; Peretz & Zatorre 2005). The frequency of two tones an octave (12 semitones) apart will always be related in a 1:2 ratio, while two tones one semitone apart stand in a 1:1.06 ratio. Although the absolute frequency difference between two tones an octave apart increases exponentially from the lower to the upper end of the musical spectrum, the frequency ratio between them always remains constant.

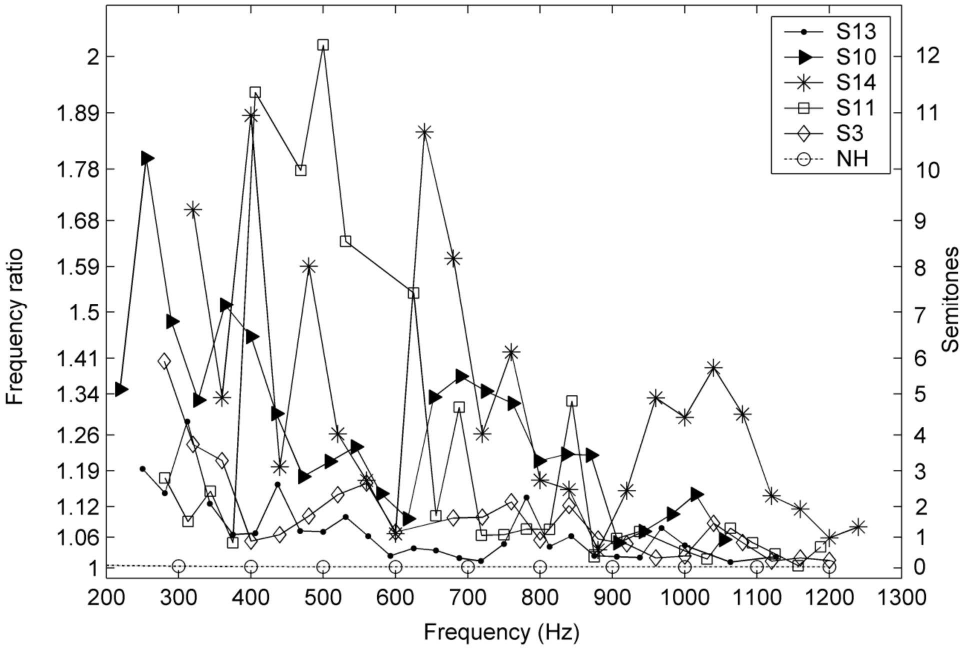

Figure 2.6: Ratio of absolute threshold frequency to reference frequency plotted as a function of reference frequency. The number of semitones associated with each ratio is shown on the right-hand ordinate. The dashed line represents frequency discrimination behaviour of NH listeners across the same range (Zwicker & Fastl 1999).

It follows then that to perceive such melodic contours successfully, pitch distances of at least one semitone should be recognised consistently across the entire musically relevant frequency range. This is indeed the case for NH listeners (Zwicker & Fastl 1999). However, as shown in Figure 2.6, CI listeners in this investigation displayed marked variation in their frequency discrimination ability across the range of base frequencies, with discrimination thresholds of the sample as a whole not consistently approaching one semitone across the frequency range tested. There is, however, a general trend of the ratio between frequency difference limen and reference frequency, and hence the number of semitones that can be discriminated, to decrease for CI users across the frequency range tested. Despite this being a direct consequence of filters being spaced several semitones apart at lower frequencies, while only a few semitones apart at higher frequencies, it does not change the implication that frequency differences between two tones may not be discriminated equally well in low and high registers. This observation is in line with findings that CI users achieve somewhat better sound-field melodic contour recognition of tone sequences toward the upper end of the musical spectrum (Galvin et al. 2007; Singh et al. 2009). Such inconsistent recognition of pitch distances may contribute to unsatisfactory music perception and subsequent appreciation.

Finally, it is necessary to consider the influence of loudness balancing on frequency discrimination thresholds. More important than loudness balancing between electrodes (as achieved in clinical mapping), would be loudness balancing between stimuli (which was not done in the main experiments of this investigation). A particular pure-tone stimulus activated not only a single electrode, but also adjacent electrodes as can be inferred from Figure 2.1. Therefore, some criticism may be expressed about not performing loudness balancing of stimuli in the present investigation.

Results suggest, however, that loudness balancing was not necessary for the particular task. Not only did stimulus loudness have to be adjusted only minimally from the unbalanced condition to achieve loudness balance, but results also showed no statistically significant difference in frequency discrimination behaviour for balanced and unbalanced stimuli. This suggests that sufficient loudness balancing was achieved in the clinical mapping procedure so as not to influence frequency discrimination behaviour significantly in sound-field conditions. This is an important result, considering that the objective was specifically to assess implant users’ behaviour in conditions similar to those they would be confronted with during everyday listening using their standard maps. It was argued that, if a sound-field frequency discrimination test were to become part of a music perception test for cochlear implantees, it would be inappropriate to perform loudness balancing before commencement of such a test, as one would wish to measure music perception ability with a particular speech processor and a particular clinical map.

Furthermore, since frequency discrimination thresholds were often smaller than one filter width, the reference and probe stimuli would in these cases have activated the same electrodes, with loudness effects unlikely to influence frequency discrimination behaviour. This is borne out by the data of S13 shown in Figure 2.3. However, frequency discrimination thresholds larger than one filter width obtained for S11 between 400 and 625 Hz suggest that frequency discrimination data may have been influenced by loudness effects to some extent. Since these larger discrimination thresholds were repeatedly observed only in an isolated region during presentation of unbalanced stimuli, it is possible that an existing physiological or anatomical constraint, together with loudness effects, may have resulted in confusion. Possible loudness effects in these regions appear to be localised and subject-specific, and therefore do not influence the general conclusions of this investigation.

2.5 CONCLUSION

Sound-field frequency discrimination abilities of CI implant users appear better than expected. Although discrimination resolution is not equally good at all reference frequencies across the range investigated, it seems that two pure tones falling within the same filter band can often be discriminated successfully. The findings may be linked to the distribution of energy across adjacent filter bands, resulting in differential electrode and subsequently differential neural activation patterns. Attempts to improve CI users’ music perception abilities may benefit from further investigation into exploiting existing sound processing features to improve spectral resolution, which in turn may contribute to better music perception for CI users.

Chapter 3

RHYTHM PERCEPTION BY COCHLEAR IMPLANTEES IN CONDITIONS OF VARYING PITCH

Pretorius, L.L, Hanekom, J.J. & Venter, P.J. (2010) Rhythm perception by cochlear implantees in conditions of varying pitch (in review, *Cochlear Implants International*)

3.1 INTRODUCTION

A melody can be described as a succession of tonal events that unfold over time, with individual events related to one another according to a tonal and temporal structure. This implies that pitch and rhythm information both contribute to constituting a perceptually meaningful melody. Moreover, pitch and rhythm patterns seem to function in an interdependent (Kidd et al. 1984; Boltz 1999; Lebrun-Guillaud & Tillman 2007) and often complementary (Jones 1993) way to create musically coherent wholes for perception. This is supported by the modular approach to music processing, which, based on results from neurocognition and lesion studies, suggests that separate yet parallel cortical processing pathways exist for time- and pitch-related information (Peretz & Coltheart 2003). The pathways are believed to be organised in a serial, hierarchical manner, with regard to both structure and function, which allows processing output from the two domains during earlier stages to combine at a higher hierarchical level to generate a coherent musical percept (Peretz & Kolinsky 1993; Schuppert et al. 2000; Peretz & Coltheart 2003; Griffiths 2003; Semple & Scott 2003).

Successful melody perception therefore requires adequate perception of pitch as well as rhythm information. It follows that constrained information transfer of either (or both) of these dimensions will hamper melody perception. Temporal resolution afforded by a CI, as inferred from gap detection thresholds, appears to be similar to that of NH listeners (Moore & Glasberg 1988; Shannon 1989; Grose & Buss 2007). These authors concluded that temporal resolution is probably determined at a post-cochlear stage, whereas pitch resolution is determined already at the cochlear level. It is thus reasonable to expect that CI users should perform better on melodic rhythm than on pitch perception tasks, and possibly even at a level comparable to NH listeners, since temporal information is probably transmitted to the auditory cortex relatively unconstrained.

However, studies regarding implant users' music perception abilities (e.g. Gfeller & Lansing 1991; Gfeller et al. 1997; Kong et al. 2004; Galvin et al. 2007; Cooper et al. 2008; Looi et al. 2008) have shown that although significantly better than pitch perception, rhythm perception is not unambiguous. Using pairs of simple monotonic rhythmic patterns, which subjects had to judge as the same or different, Gfeller et al. (1997) showed that CI users performed similarly to NH listeners. In the same study, however, when a discrimination judgement had to be made for more complex rhythmic patterns, implant users performed significantly poorer than NH listeners. Similarly, Kong et al. (2004) found, albeit in a small sample, that CI users performed 5–25% worse than an NH control group on a rhythm identification task.

Observations that simple familiar melodies are better recognised when pitch as well as rhythmic cues are available (Gfeller et al. 2002; Galvin et al. 2007) highlight the important contribution of rhythmic cues to CI-mediated melody perception. It is possible that when confronted with limited pitch information during melody perception, the auditory system relies on rhythm information to supplement insufficient information to achieve the best possible melody perception in constrained conditions. In this regard recent findings regarding the rhythm perception abilities of amusic listeners are important. Amusic listeners often find it difficult to follow pitch changes in a melody, which possibly stems from deficits in processing pitch information (Foxton et al. 2004). Within the framework of early independent processing of pitch and rhythm information (Peretz 1990; Peretz & Coltheart 2003) one would expect that rhythm perception should be unaffected by pitch processing deficits. However, Foxton et al. (2006) showed that amusic listeners' perception of pitch-varying rhythmic patterns was significantly poorer than that of NH listeners. The authors suggested that when pitch and rhythm information covary in the same sound sequence, the information is encoded as a unified representation rather than activating two separate processing systems. Insufficient pitch processing hence appears to interfere with rhythm perception.

CI users also experience pitch processing difficulties, although not derived from central processing impairments (see for example Koelsch et al. 2004) but rather owing to insufficient delivery of pitch-related information to the central auditory system (Rubinstein 2004; Sucher & McDermott 2007), and as such can be regarded as functionally tone-deaf. This raises the question whether implant users' rhythm perception ability is subject to the same constraints as those experienced by amusic listeners. If so, it could mean that already challenging melody perception, due to insufficient transmission of pitch information, may be complicated even further. In this regard it is important to keep the schema outlined in Figure 1.1 (Chapter 1) in mind. Rhythm perception represents processing of musical information features at an early hierarchical level from where output is sent forward to a subsequent hierarchical level to contribute to "whole" melody perception. However, based on results regarding amusic listeners' rhythm perception abilities, it is important to consider the possible influence of pitch processing deficits on rhythm perception, despite the two dimensions' processing being posited as largely independent at this hierarchical level. Thus according to the hierarchical processing framework as proposed in Figure 1.1, investigating rhythm perception at a perceptually early level while considering a possible cross-influence from the pitch dimension, could contribute to perceptual outcome at later levels being better interpreted.

In previous studies specifically aimed at investigating rhythm perception of CI users (e.g. Gfeller et al. 1997; Kong et al. 2004; Looi et al. 2008), pitch was not varied. However, covarying pitch and rhythm cues in a controlled manner may provide improved understanding of CI users’ rhythm perception ability with regard to music-like tone sequences. The work presented in this chapter investigated CI users’ rhythm perception for sound-field tone sequences of varying rhythmic complexity, in either monotonic (pitch-constant) or polytonic (pitch-varying) conditions. The objectives were (i) to determine what the typical rhythm perception ability of CI users is for simple sound-field tone sequences, and (ii) to assess whether pitch processing deficits influence rhythm perception to the same extent as for amusic listeners. Interpreting findings in the context of the neurocognition of music may lead to a better understanding of some of the underlying information processing that contributes to CI-mediated melody perception.

3.2 METHODS

3.2.1 Subjects

Seven post-lingually deafened adult users of the Nucleus device (24-electrode array) participated in the study. All subjects had experienced profound hearing loss for more than 10 years and had more than a year’s experience with the CI system. Except for one Sprint and one Esprit 3G user, all subjects used a Freedom processor. All but one of the CI listeners used the ACE strategy. Four subjects were implanted bilaterally, but used only the ear subjectively regarded as allowing clearer perception during the investigation in order to create a homogenous monaural listening test group.[5] None of the subjects had advanced musical training before onset of deafness, although S3 was a member of a church choir as teenager and S11 often attends folk music recitals. Other relevant detail is provided in Tables A1 and A2 (Appendix A).

---

[5] S11 used the right ear during this investigation (compare to Chapter 2). This was because she had been fitted with new maps on both implants shortly before this investigation commenced and found the right ear to give a more “natural” sound.

Seven age-matched NH control listeners also participated in the study. Normal hearing was defined as achieving audiometric thresholds of 30 dB HL or better at six octave frequencies from 250 to 8000 Hz. Only one NH listener had formal music training, but all other participants indicated that they enjoyed listening to Western tonal music and could informally participate in singing. All but one of the control listeners had previously participated in research in our laboratory.

As required by the relevant ethics committee, all subjects gave written informed consent for their participation prior to commencement of the study. Subjects were compensated for their time on completion of the tasks.

3.2.2 Stimuli

Stimuli consisted of pairs of simple tone sequences generated in Matlab 6.5 on a personal computer. Each sequence comprised five 100 ms pure sine tones separated by four brief intertone intervals (ITIs). Sequences were strung together prior to each presentation round to prevent the duration of the ITIs from being influenced by processing. Amplitude ramps (30 ms) were included at the beginning and end of each tone to reduce onset clicks. During each presentation round one sequence (probe) contained a longer ITI presented at a specific position, while the other (reference) did not. The longer ITI was presented either between the second and third, or fourth and fifth tones. The probe and reference sequences were separated by a 1500 ms gap.

A combination of rhythmic and pitch pattern complexity (two levels each) created four conditions of varying perceptual demand. Rhythmic complexity was determined by the ITI, which, together with the tone duration, determined the inter-onset interval (IOI). Gauging responses to IOI changes is a useful indicator of rhythm perception, since the rhythm of a sequence is primarily determined by the time between tones, rather than the duration of the tones themselves (Krumhansl 2000). Since tone duration was always 100 ms, the ITI was altered to create different IOI patterns. An isochronous sequence with a constant ITI of 300 ms represented the simple rhythmic condition, while an anisochronous pattern with ITIs alternating between 300 ms and 600 ms represented the complex rhythmic pattern. Base ITI in the isochronous condition was 300 ms, while 600 ms in the anisochronous condition. In the monotonic pitch condition all tones were presented at 500 Hz. The polytonic pitch condition involved varying pitch patterns, with the first tone always presented at 500 Hz and subsequent tones varying between 200 Hz and 2000 Hz at 100 Hz increments to ensure that pitch differences would be clearly audible to CI listeners. The frequency of successive tones could change in any direction (up, down or not at all), but never differed by more than 200 Hz from the preceding tone. The stimuli design strongly resembles that used by Foxton et al. (2006). Figure 3.1 provides a schematic representation of stimuli specifics.

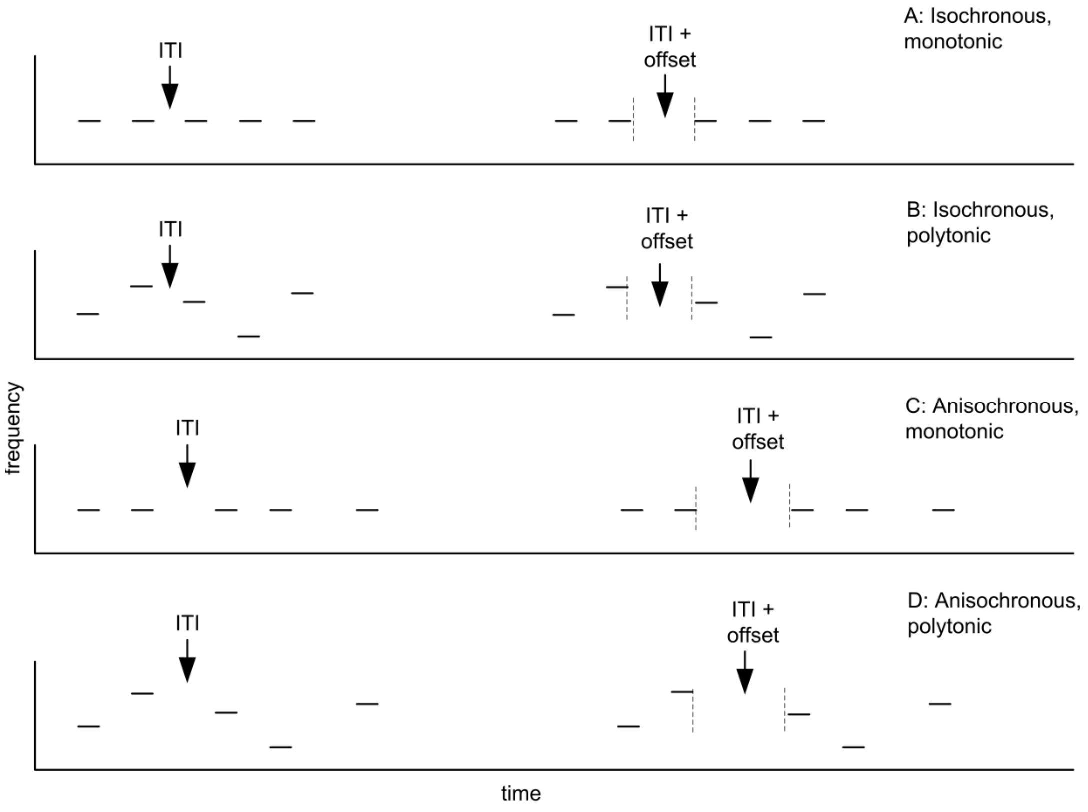

Figure 3.1: Schematic representation of test conditions. IOI = Inter-onset interval; ITI = intertone interval. Because tone duration always remained constant, ITI was adjusted to create different IOI patterns. Reference ITI duration was 300 ms in the isochronous conditions (A and B) and 600 ms in the anisochronous conditions (C and D).

All stimuli were presented in sound field. Subjects used their clinically assigned processors and settings. Stimuli were presented through a Yamaha MS101 II loudspeaker, which was placed approximately 1 m away, in front of the listener on the side of the test ear. Because of the sound-field set-up it was important to present stimuli at the same subjective loudness to each participant, rather than at a fixed intensity level. A loudness estimation task was performed prior to the rhythm perception tasks. Each listener’s 75% perceived loudness level was determined for a 1 kHz pure sine tone, to which intensities of all other tone frequencies were subsequently loudness balanced.

Loudness balancing was performed to rule out any confounding loudness cues, especially considering the pitch-varying nature of the polytonic conditions. Frequencies selected at 100 Hz increments across the range to be used during the rhythm perception tasks were loudness balanced relative to a 1 kHz tone at the intensity earlier determined to represent the 75% perceived loudness level for each listener. The balancing task required listeners to perform three comparisons: adjusting the probe stimulus to sound equally loud, just louder and just softer compared to the reference stimulus. Each frequency was randomly presented three times during the course of the respective loudness adjustment tasks. The average deviation from the reference intensity was calculated from these nine samples and presentation intensity for frequencies selected for rhythm perception stimuli were stored in a look-up table for later use.

3.2.3 Experimental procedure

Rhythm perception ability was assessed in terms of ITI duration discrimination, using an adaptive 2AFC procedure. The probe sequence, randomly assigned to either of the two presentation intervals, was identical to the reference sequence, barring target ITI duration. The target ITI was also randomly presented between the second and third, or fourth and fifth tones, requiring subjects to listen to the whole sequence rather than focus on a specific ITI position. Together these measures reduced the possibility of listeners tracking the pattern of duration offsets, which could have introduced bias into the results. However, an equal number of response reversals had to be completed for each of the two target ITI positions, and results were analysed separately for the two positions.

The target ITI duration offset (the difference between the reference and probe ITI) was adapted according to subject response. In both isochronous and anisochronous conditions the initial ITI offset was 600 ms to allow subjects to hear the difference in ITI duration clearly. Thus for the isochronous condition, total target ITI amounted to 900 ms; for the anisochronous condition, the total intial target ITI duration was 1200 ms (600 ms base + 600 ms offset). A transformed up-down staircase procedure (2-down/1-up decision rule) was used to adjust ITI duration offset, converging on 71% correct (Levitt 1971). A test run concluded after 15 response reversals. For the first five reversals, the ITI duration offset was adjusted by a factor of 2; for the second five and last five reversals, the ITI duration offset was adjusted by a factor of 1.7 and 1.3 respectively. The average of the last five reversals was calculated as the average ITI duration threshold for each test run.

Each combination of rhythmic and pitch complexity represented a test condition and was tested six times (three sets of two test runs for each combination) with CI listeners. Since NH listeners' results showed less variation, they completed only three repetitions of each test condition. Depending on subjects' available time, more than one test condition could be tested during a single test session. Subjects were told that the target (lengthened) ITI could be between either the second and third or the fourth and fifth tones, and that the probe sequence could randomly be presented as either the first or the second presentation interval. Rhythmic patterns were also tapped out to subjects prior to commencement of the study. Subjects neither received feedback regarding their responses nor were allowed to re-listen to stimuli. Test trials were self-paced and a new stimulus was not presented until subjects had responded to the previous stimulus. Unbeknown to subjects, three practice trials were included at the beginning of each run to familiarise listeners with the stimuli and test procedure. Subjects were required to complete the practice trials successfully before actual testing was begun. Since an experiment run could last up to 20 minutes, a progress indicator was shown on the computer screen to keep listeners motivated.

3.3 RESULTS

Figure 3.2 shows mean individual ITI duration discrimination thresholds, as well as results averaged over all participants in a group, for the four experimental conditions. The top and bottom panels depict results for the target interval placed between the second and third, and fourth and fifth tones respectively. Individual subjects’ results represent the average of six repetitions (or three, in the case of NH listeners) per stimulus condition.

Since subjective response behaviour as measured in this study can be prone to between-session variability, each subject’s respective data sets were inspected for outliers using the ESD technique. Only three outliers were found among the 336 CI listener responses (six repetitions x four conditions x two target positions x seven subjects), in three different data sets. The reported average thresholds for those sets were thus calculated from the remaining repetitions. No outliers were found among the 168 NH listener responses (three repetitions x four conditions x two target positions x seven subjects). Rhythmic complexity (isochronicity or anisochronicity), pitch complexity (monotonic or polytonic) and position of target interval (between second and third, or fourth and fifth tones) were defined as within-subject variables, while hearing ability (CI or NH) was defined as the between-subject variable in a three-way repeated measures analysis of variance

CI and NH listeners performed similarly in all conditions ($F(1,12) = 1.62$, $p = 0.23$), indicating that the CI device allows uncompromised perception of the rhythmic information presented in this investigation. The finding is supported by the observation that none of the main effects produced a statistically significant interaction with hearing ability.

The influence of additional rhythmic and pitch complexity are noteworthy. Manipulating rhythmic complexity had a pronounced effect on rhythm perception ability for both CI and NH listeners ($F(1,12) = 89.77$, $p < 0.001$), as seen from discrimination thresholds that were three to four times higher for anisochronous than isochronous conditions. Manipulating pitch complexity, however, exerted less influence on rhythm perception ability.

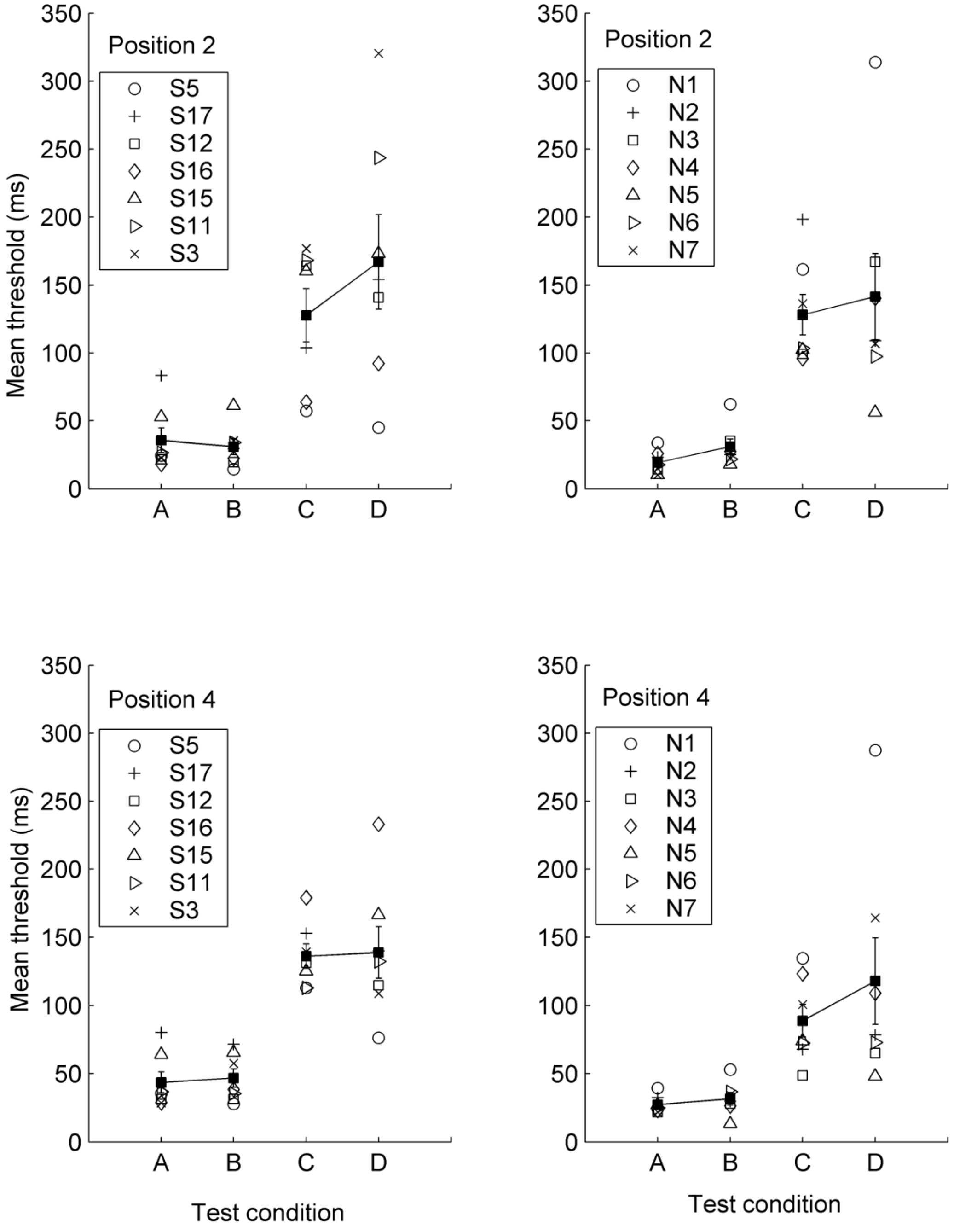

Figure 3.2: Absolute mean ITI duration discrimination thresholds in the four experimental conditions (see Figure 3.1). The top panel shows thresholds for the target interval between the second and third tones; the bottom panel shows thresholds for the target ITI between the fourth and fifth tones. Filled squares show group mean thresholds. Error bars represent standard errors.

Although anisochronous polytonic sequences appeared to present the most difficult rhythm perception condition for all listeners, discrimination thresholds were in general similar for polytonic and monotonic sequences of a specific rhythmic condition ($F(1,12) = 2.43$, $p = 0.15$). This finding indicates that rhythmically complex sequences represented more demanding listening conditions to subjects than sequences with a complex pitch pattern.

Performance trends appeared to be insensitive to the target interval’s position within the tone sequence. Subjects performed similarly when the target interval was placed in position 2 or position 4, irrespective of rhythmic or pitch complexity or hearing ability ($F(1,12) = 0.52$, $p = 0.49$). This indicates that listeners were able to identify the target interval in the context of the sequence and not merely with regard to its position.

3.4 DISCUSSION

Three main findings emerged from this investigation: (i) CI and NH listeners showed similar rhythm perception ability for tone sequences as used here, (ii) pitch complexity did not adversely influence CI users’ rhythm perception ability, and (iii) rhythm perception ability appeared to be insensitive to the relative position of the target interval within the tone sequence. Although the observation that CI listeners experience similar rhythm perception ability as NH listeners supports the general trend regarding CI listeners’ rhythm perception abilities (Gfeller & Lansing 1991; Zeng 2004; Cooper et al. 2008; Looi et al. 2008) the investigation extends contemporary understanding by commenting on the influence of covarying pitch patterns on rhythm perception ability and CI users’ ability to identify rhythmic irregularity in the context of a tone sequence.

CI users’ rhythm perception ability as observed here may be associated with generally acceptable speech perception abilities afforded by modern implant devices (Wilson & Dorman 2007). Language processing has previously been shown to be subserved by a predominantly left hemisphere-based cortical network (Zatorre & Belin 2001; Zatorre et al. 2002) and in a recent neuroimaging study, where CI users performed similarly to NH listeners on a rhythm task, speech perception was also associated with predominantly left hemisphere cortical activation patterns (Limb et al. 2010). Considering that left-hemisphere-dominant activation, specifically in speech/language processing areas, has previously been observed during perception of rhythm patterns (Platel et al. 1997) the authors suggested that the intact rhythm perception ability of CI users may be related to speech perception abilities. It is thus plausible that regular activation brought on by speech stimuli may help to familiarise left-hemisphere-based cortical areas involved in rhythm perception with the nature of input derived from electrical stimulation at the auditory periphery.

The finding that rhythmic rather than pitch complexity adversely influenced the rhythm perception ability of CI users was somewhat surprising, especially considering a recent finding that amusic listeners' compromised rhythm perception abilities for pitch-varying tone sequences may be attributed to earlier pitch processing deficits (Foxton et al. 2006). Although pitch and rhythm patterns are thought to be processed separately during early analysis, the information probably combines at a later processing stage to form a unified percept (Peretz & Kolinsky 1993; Peretz & Coltheart 2003). Conceivably, the position of the deficit in the auditory pathway may influence the behavioural outcome. Degraded auditory input originates early during auditory processing for cochlear implantees, whereas for amusic listeners, the information processing difficulty appears to stem from a cortical deficit (Peretz et al. 2009). It is thus plausible that the nature of the input received by higher-order processing stages determines the neural activation patterns. Limb et al. (2010) indeed recently found more intense neural activation patterns across a more distributed cortical network in implant listeners compared to NH listeners, despite similar behavioural outcome on rhythm tasks. Heightened neural activity, especially in frontal areas, possibly reflects effortful processing and associated neural plasticity to allow successful behavioural response to degraded information. Although the findings by Limb et al. (2010) do not comment on neural activation patterns in amusic listeners, they do support the notion that degraded auditory information can elicit differential neural activity among listener groups, depending on the specific auditory deficit, which may explain the differential effect of additional pitch complexity on rhythm perception ability in amusic and CI listeners.

Lastly, the finding that rhythm perception ability was insensitive to the position of the target interval in the sequence suggests that implant listeners were able to extract rhythm cues within the context of a tone sequence, rather than focusing only on a specific intertone interval. Since rhythm can be regarded as patterns of temporal information (Samson et al. 2001), it follows that successful rhythm perception requires temporal relationships between successive tonal events to be established (Krumhansl 2000). In terms of pattern perception, rhythm perception possibly represents an intermediate processing stage – later than mere feature extraction but before semantic processing (Griffiths 2003; Koelsch & Siebel 2005). It thus appears that good early temporal resolution afforded by the implant device, as reflected by small within-channel gap detection thresholds (Moore & Glasberg 1988; Shannon 1989; Grose & Buss 2007), may propagate sufficiently to a subsequent processing level to support rhythm perception in the context of a tone sequence.

3.5 CONCLUSION

Several earlier studies have shown that CI listeners exhibit rhythm perception abilities comparable to those of NH listeners when confronted with isolated rhythmic components. However, to extend the understanding of cochlear implantees' music perception ability, rhythm perception needs to be investigated in a context that more closely represents real-world listening conditions. The finding reported here showed that CI users were indeed able to perceive rhythmic patterns equally well as the NH control group in such listening conditions. The results thus confirm that rhythm perception is not the main contributor to unsatisfactory perception in CI-mediated music listening.

Chapter 4

PERCEPTION OF SOME MELODIC CHARACTERISTICS BY COCHLEAR IMPLANT USERS

Pretorius, L.L. & Hanekom, J.J. (2011) Perception of some melodic characteristics by cochlear implant users: an exploratory study (submitted for review to *Journal of Speech, Hearing and Language Research)*

4.1 INTRODUCTION

Multichannel cochlear implants (CIs) afford many profoundly deaf listeners partially restored hearing ability. Despite deriving good speech understanding (Gifford et al. 2008; Wilson & Dorman 2008), CI users generally experience only limited music perception ability and enjoyment (Gfeller et al. 2000; Leal et al. 2003; Gfeller et al. 2003; Mirza et al. 2003; McDermott 2004; Zeng 2004; Gfeller et al. 2005). Unsatisfactory music perception may be attributed to the limited ability of the implant device to convey enough high-fidelity information relating to both the pitch and temporal domains to allow a unified and coherent musical percept to be formed (Limb et al. 2010).

Despite several studies assessing CI listeners' music perception behaviour (e.g. Fujita & Ito 1999; Gfeller et al. 2002; Leal et al. 2003; Gfeller et al. 2003; Mirza et al. 2003; Gfeller et al. 2005; Looi et al. 2008; Vongpaisal et al. 2009; Singh et al. 2009), little is known about the underlying constraints imposed by cognitive and perceptual factors, and especially their interaction, in a music listening setting. The difficulty with objectively assessing music listening success stems from the nature of musical sounds. In its simplest form music can be defined as a succession of tonal events bound together over time to form a coherent perceptual entity (Patel 2003; Limb 2006b). This implies that successful music listening involves not only analysing pitch-related and temporal information separately, but also processing their interactions.

In Chapters 2 and 3 pitch- and rhythm perception abilities of CI users were investigated at a processing level where the two dimensions are regarded to be analysed by independent processing modules within the bigger processing scheme proposed for music perception (also see Figure 1.1, Chapter 1). Since coherent music perception is the outcome of a hierarchical processing system (Peretz & Coltheart 2003), low-level deficits will undoubtedly propagate to later processing stages where processing is, in turn, linked to and influenced by other processing modules (Peretz & Coltheart 2003; Koelsch & Siebel 2005). A number of previous psychophysical assessments of CI-mediated music perception ability have investigated the perception of separate music-relevant input parameters. These studies showed that perception of pitch-related cues is severely hampered in CI users when listening in sound-field conditions (Gfeller & Lansing 1991; Gfeller et al. 1997; Kong et al. 2004; Gfeller et al. 2007; Galvin et al. 2007), while perception of temporal information (rhythm, metre and tempo) appears, if not always comparable to normal-hearing (NH) listeners' ability, at least superior to CI users' pitch perception ability (Gfeller & Lansing 1991; Gfeller et al. 1997; Kong et al. 2004; Looi et al. 2008). Given the relative importance of pitch-related information for music perception (Zatorre 2001; Zatorre et al. 2002) in especially the Western tonal tradition, it is understandable that attempts at improving the music perception ability of CI users would focus on ways to convey more intact pitch information to the electrically stimulated auditory system (Laneau et al. 2006; Milczynski et al. 2009). However, reports of slightly improved music perception when rhythm cues are available (Gfeller et al. 2002; Kong et al. 2004; Galvin et al. 2007) suggested that investigating CI perception of Western tonal music with stimuli that incorporate information from both pitch and temporal dimensions may contribute to the understanding of constraints involved in contemporary CI music perception. Studies using real-world music tokens were subsequently performed to assess the music perception ability of CI users when subjected to covarying pitch and rhythm cues (Gfeller et al. 2005; Gfeller et al. 2007).

Although such studies provided valuable empirical support for CI users experiencing only limited success when confronted with real-world music listening situations, there are some weaknesses associated with approaches that are based on recognition of familiar melodies from a specified set to evaluate perception of music-relevant information. It is not certain that the musical character of a tone sequence is retained after acoustic–electrical conversion of auditory input. Reports of reasonable individual success in studies using familiar melodies to gauge music perception ability (e.g. Gfeller et al. 2000; Kong et al. 2004; Galvin et al. 2007; Singh et al. 2009) do indeed provide useful assessment of a CI listener’s ability to match a melody to one of a limited set of stored representations supported by the available cues, but do not comment on the listener’s ability to use the available cues to drive perception of musical character and then identify a melody as one from any number of stored representations as a result. Such tasks thus do not provide a true handle for probing perception of musical character independent of memory representations, and as such offer limited possibility for tracking improvements in processing strategies or personal improvement afforded by alternative map parameters or implementation of user-specific device settings.

With many CI users expressing the wish for better music perception ability, standardised and clinically practical music perception tools are needed to allow for standardised assessment and subsequent comparison across different CI user groups, processing strategies and rehabilitation facilities. Two approaches deserve mention. Kang et al. (2009) recently put forward the University of Washington Clinical Assessment of Music Perception test as a possible clinically relevant music perception test. The test comprises pitch direction discrimination as well as recognition of commonly heard familiar melodies and identification of musical instruments from closed sets. Although this test is regarded as the most useful clinically applicable yardstick of CI music perception at present, owing to it being selfadministered and rapidly completed, it does not allow the quality of music-relevant information as conveyed by the implant to be gauged.

Cooper et al. (2008), in turn, employed the Montreal Battery for Evaluation of Amusia (MBEA) (Peretz et al. 2003) to probe CI-mediated music perception ability. This measure, based on neurocognitive principles underlying music perception, uses six tests to assess different aspects of music perception along the melodic (pitch interval, scale and contour) and temporal dimensions (rhythm and metre). It also includes a melodic memory test to assess music perception ability when confronted with co-varying pitch and rhythm information. It requires short-term memory for melodies presented in the context of the test battery, rather than the listener having to match a present melody to a representation of a familiar melody stored before becoming deaf.

It is proposed here that a music perception assessment tool that takes into account the neurocognitive principles underlying music perception, such as the MBEA test battery, is well suited to provide an overview of CI users’ general music perception ability, independent of memory representations. Moreover, to assess the quality of transfer of musical character after acoustic–electrical signal conversion such tests can be expanded by including tasks that require listeners to judge syntactic congruency of short melodies. In Western tonal music, perception of musical syntactic congruency is based on the recognition of tonal relationships between successive notes. Such ability has been shown to exist in musically untrained listeners and children (Koelsch et al. 2000; Koelsch et al. 2005) and is believed to be based on inherent musical knowledge, which can be expressed without training (Trainor & Trehub 1994; Tillman et al. 2000). Tonal relationships reinforce the implied tonality and associated tonal hierarchy (Krumhansl 1979; 1990), which in turn facilitates expectancy for specific tonal events (Marmel et al. 2008).

Preceding context of a melodic line can generate expectancies for a specific musical event both at the end of a phrase and within the tone sequence itself. Both are based on the implicit knowledge of melodic key derived from the relationships between preceding notes. Out-of-key notes within a melodic line violate tonal expectancy (Brattico et al. 2006) and it follows that if such out-of-key notes can be identified correctly, the preceding notes must have built up a musical context that allows these notes to be identified as odd. Similarly, expectancy for melodic endings is generated by both the tonal and the temporal structure of the preceding context (Boltz 1989; 1993). Tonal expectancy stems from priority assigned to specific notes within a tonal hierarchy, which in turn determines their stability in a sequence (Krumhansl 1979; 1990). It has been shown that listeners expect unstable notes to resolve to stable ones to create a feeling of completion at the end of a phrase or melodic line and conversely that sequences ending on unstable (and hence unexpected) notes are judged to have a low degree of completion (Meyer 1956; Boltz 1989; Bigand & Pineau 1997). Temporal relationships between successive notes have been shown to generate similar expectancies within the temporal domain (Boltz 1993; Nittono et al. 2000) and contribute to judgements of melodic completion possibly through creating accents that direct a listener’s attention to a specific event (Boltz 1989; Jones 1993; Boltz 1993). Melodies ending on prolonged final notes, especially when these are stable notes, have been shown to be judged more complete than those without a lengthened final note (for a description see Boltz 1989).

Melodic expectancies contribute to the build-up of melodic context for music listening, which can facilitate processing of future bottom-up information (Schulkind 2004). It thus follows that without proper perception of the underlying musical information a musical context will not be established, which in turn may lead to impaired higher order auditory processing of music. Since CI listeners experience only limited melodic perception, probing their perception of melodic expectancy as a marker of melodic context synthesis may be a valuable approach.

According to Koelsch et al. (2004) syntactic violations elicit similar brain responses in CI listeners and NH control subjects, indicating that the neural mechanism underlying detection of music syntactic congruency is present in CI listeners. The finding suggests that tasks which require perceptual judgements to be made based on musical expectancy may (i) provide a useful indication of CI listeners’ ability to extract musical context from a sequence of preceding events when presented with unfamiliar real-world-like melodies and (ii) contribute to the development of new or further refining of existing clinically practical test batteries to compare music perception performance across CI user groups and gauge perceptual outcomes of future processing strategies.

The investigation described in this chapter explores CI users' ability to perceive simple, coherent melodies in which pitch and rhythm information contribute simultaneously to form a unified musical token. It is based on the notion of the music processing system being a modular, hierarchically organised perceptual system as a departure point. According to this view, forming a melodic percept is the outcome of a processing step that occurs beyond separate pitch and rhythm analysis but needs direct input from these processing modules. Given the multifaceted nature of music as well as the human auditory processing system's ability to adapt to less-than-optimal input, it is worthwhile to explore CI users' ability to judge musical character of simple melodies as a window into the system-dependent processing of electrically mediated auditory input. In view of this explorative approach, CI listeners' judgement of musical syntactic congruency either at the end of melodic lines (Experiment 1) or within the tone sequence (Experiment 2) is assessed. The two perceptual tasks presented in the respective experiments described here are offered as alternative approaches to assessing the extent to which useful musical context is built up in CI-mediated sound-field hearing.

4.2 EXPERIMENT 1: PERCEPTION OF MELODIC COMPLETION

In Western tonal music, relationships between seven diatonic tones are used perceptually to establish the key of an unfolding melody. Within the framework of the established key each tone is perceptually assigned a priority, which translates to a certain level of melodic stability within the key (Krumhansl 1979; 1990). The relative stability or instability of a note within a key, coupled with rhythmic patterns, in turn generates expectancy in the listener for a specific note to follow, based on the preceding context.

Expectancy judgements have previously been used to investigate the information processing and neurocognitive principles underlying music perception (e.g. Boltz 1989; Boltz 1993; Bigand & Pineau 1997; Koelsch et al. 2000; Koelsch et al. 2005; Brattico et al. 2006; Marmel et al. 2008). Brattico et al. (2006) showed that judgements regarding congruency violations at the end of melodic phrases are possibly facilitated by the output of earlier feature extraction, which suggests that perception of harmonic belonging may depend on prior successful perception of musical scale belonging. In the context of CI music perception research, markers of processing stages can be usefully applied to investigate to what extent the information conveyed by the implant device supports music perception. Experiment 1 investigated CI users' perception of melodic completion to determine whether the information conveyed by the implant device allows successful processing at this relatively advanced processing stage.

Since CI users find melody recognition easier when both pitch and rhythmic cues are available (Gfeller et al. 2002; Kong et al. 2004; Galvin et al. 2007), it was deemed valuable to investigate whether rhythmic information can influence perception of melodic completion to a similar extent as for NH listeners (Boltz 1989; 1993). Perception of melodic completion was therefore investigated with and without temporal cues to signal melodic completion.

4.2.1 Methods

Subjects

The CI test group consisted of seven post-lingually deafened adult users (mean age = 49.3 yrs) of the Nucleus 24 or Nucleus 22 electrode array device. All participants had more than two years' experience with their devices and had experienced profound hearing loss for more than 10 years. CI listeners all used the Freedom processor; five fitted with the ACE processing strategy and two with the SPEAK strategy. Three of the participants were implanted bilaterally, but only the ear subjectively regarded as the better of the two was used during testing. Only one of the CI participants had formal music training prior to the onset of deafness (S10), while two others had been choir members during their youth (S3 and S8). Two more (S21 and S18) indicated that they regularly try to listen to music or the radio, despite perception often being unsatisfactory. All but one (S21) of the participants had participated in earlier CI research at our laboratory. Other relevant demographic details are provided in the Appendix (Table A1 and explanatory description). An overview of participants' general listening success with the CI, as rated by the treating audiologist, is given in Table A2. The questionnaire was adapted from the Abbreviated Profile of Hearing Aid Benefit questionnaire (Cox & Alexander 1995).

Seven age-matched NH control group listeners (mean age = 48.6 yrs) also participated in this investigation. Normal hearing was defined as achieving audiometric thresholds of 30 dB HL or better at five octave frequencies from 250 to 4000 Hz. Subject N3 achieved only a 35 dB HL threshold at 4000 Hz , but since stimulus frequencies never exceeded 1600 Hz (see section on stimuli), he was included in the control group. Only one NH listener had formal music training, but all other participants indicated that they enjoyed listening to Western tonal music and could informally participate in singing. Two of the control group listeners had previously participated in research at our laboratory.

Owing to the exploratory nature of this investigation, the same group of listeners participated in all tasks described here. All participants (CI and NH) gave written informed consent prior to commencement of the investigation according to the requirements of the relevant ethics committee. Participants were compensated for their time at the conclusion of the investigation.

Stimuli: general considerations

The tasks described here are suggested as a first step towards probing a subjective perceptual quality in as objective a manner as possible. The explorative nature of the investigation creates the backdrop for the selection of test stimuli. Firstly, aural training material (Horacek & Lefkoff 1970; University of South Africa 1970; Van Zuilenberg 1996) was considered to provide a suitable pool of short, simple, single-voice melodies with easy-to-follow pitch contours and rhythmic structure, especially since all stimuli were chosen from the primary grades’ material. It was deemed more feasible to use existing melodies that were all of the same difficulty level than composing melodies anew and run the risk of them not being of similar perceptual difficulty. Secondly, use of aural training material avoids the risk of the stimuli being familiar to listeners, but simultaneously provides the assurance that the tokens are true melodies created according to established principles of melody composition. All stimuli were reviewed for their melodic quality and ‘listening ease’ by a colleague at our university's music department during design of the investigation and comments and recommendations were duly addressed. Thirdly, only melodies in simple time, starting on the tonic and of which the shortest note value was an eighth note were included in the stimuli set. This was done to allow listeners a reasonable amount of melodic information on which to base their decisions, without burdening them with too rapid or intricate rhythmical patterns. The resulting set of 20 melodies was regarded to provide a suitably varied stimuli set without dragging on too long and so risk listeners losing concentration, based on personal experience working with the listener profile available at our implant centre over several years. All stimuli and tasks were piloted with a NH listener before commencement of the study.

Stimuli: specific considerations

A core set of 20 melodies were adapted from material used for graded aural training in music education (as described above) to ensure novelty to listeners. The melody set is shown in Figure 4.1(a) and (b).

Although it is customary to perform experiments of melodic completion perception with chord sequences, unaccompanied tonal melodies consisting of pure sine tones generated in Matlab 6.5 were used in this experiment to prevent unnecessary spectral information from introducing additional processing difficulty. Melodies were between 15 and 23 notes long (average length = 17.4 notes) and contained only half, quarter and eighth notes, separated by 90 ms gaps (Cowan 1984). Amplitude ramps (30 ms) were included at the beginning and end of each tone to reduce onset clicks. Melodies were presented in simple duple, triple or quadruple time at a tempo of 150 beats per minute.
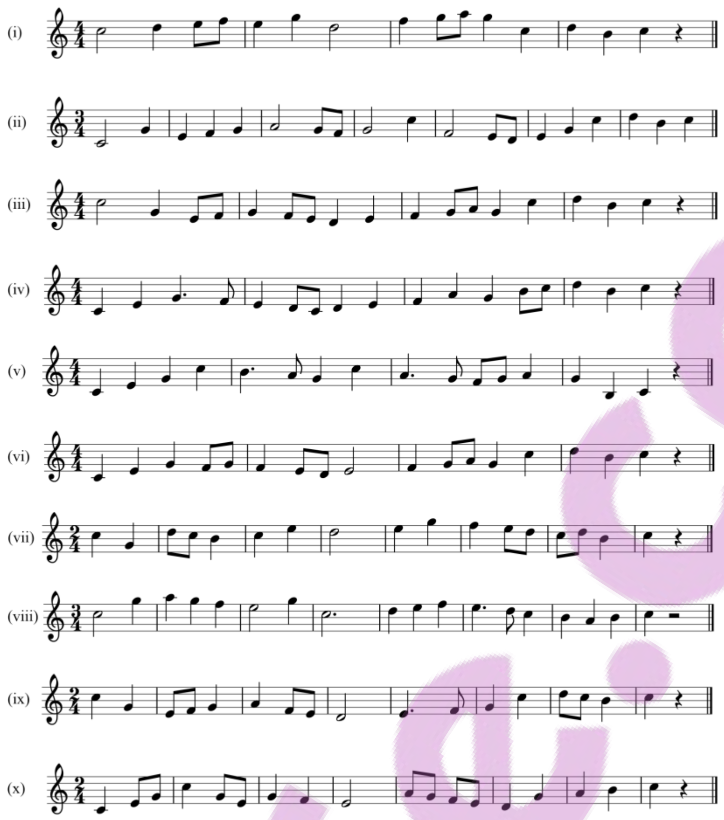
Figure 4.1a: Ten melodies with leading tone–tonic final progressions, from the set of 20 melodies used in the melodic completion task.

Melodies always started on the tonic, but never exceeded a 1.5-octave span. Since each melody was presented in C, F and G major, the frequency span ranged from 220 Hz (A3) to 1318.51 Hz (E6). This range falls roughly within the cochlear region where frequency filters of the speech processor are linearly spaced. An arpeggio, consisting of the tonic, mediant, dominant and end-octave tonic tone, was presented before each melody to establish its implied key.
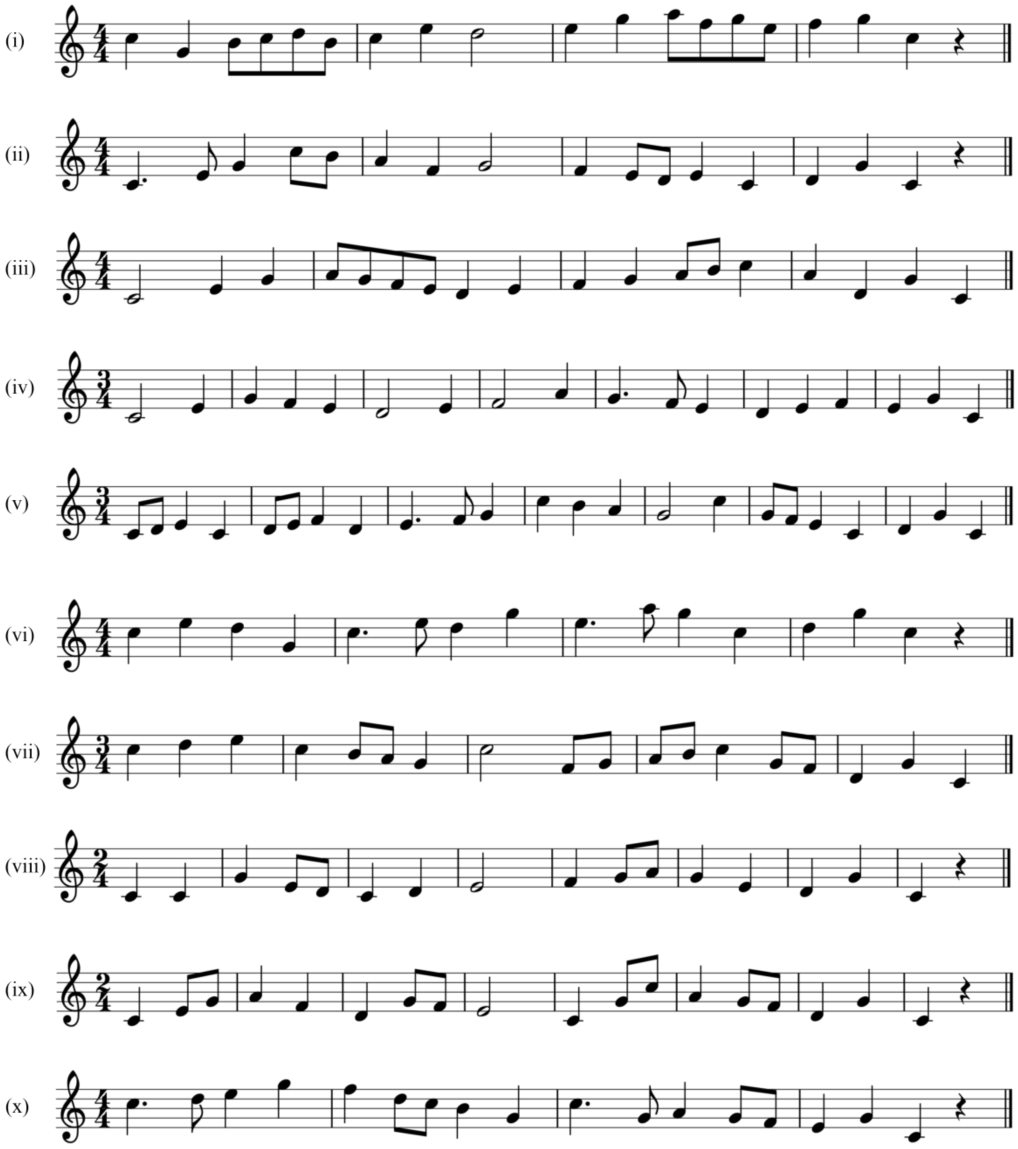
Figure 4.1b: Ten melodies with dominant–tonic final progressions, from the set of 20 used in the melodic completion task.

Each of the 20 melodies was presented with a complete and incomplete ending during the course of the experiment. However, each of these (40) melodies was presented in three keys and the stimuli sets were therefore divided into subsets of ten melodies each so as to prevent an experiment run from becoming too tedious. Complete endings were created by using either a leading tone–tonic or a dominant–tonic progression, while the order of the last two notes was reversed to create incomplete endings. To prevent progression direction from unduly being used as a completion cue, owing to melodies with complete and incomplete endings differing only with regard to the order of their last two notes, equal numbers of samples of both progression types (leading tone–tonic and dominant–tonic) were presented in each subset of complete melodies and then paired with a complementary subset of incomplete melodies. Examples are shown in Figure 4.2(a) to (d).
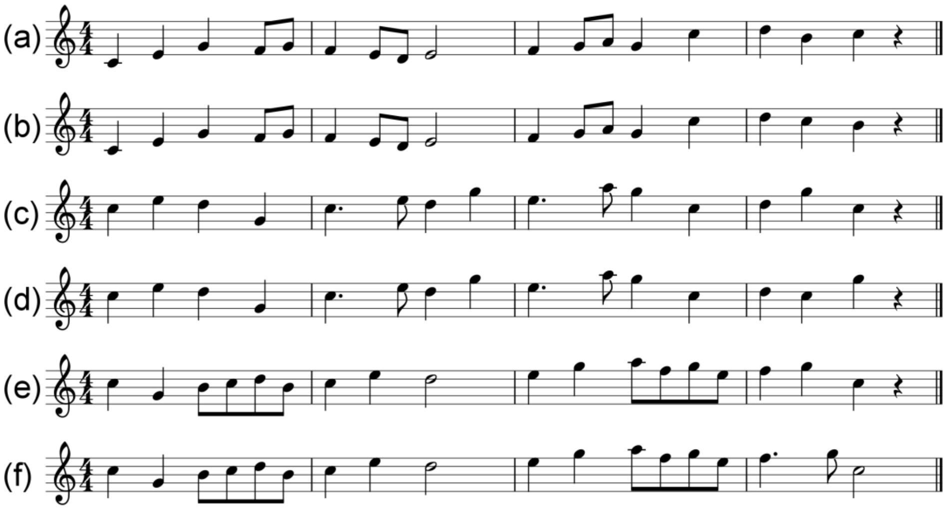
Figure 4.2: Examples of stimuli used in the melodic completion task. Melodies with leading tone–tonic and dominant–tonic final progressions are shown in staves (a) and (c) respectively, while their complements with incomplete final progressions are shown in (b) and (d). Staves (e) and (f) show an example of a melody without and with a lengthened final note.

Perception of melodic completion was investigated both with and without temporal cues in the final progression to determine whether rhythmic information would influence completion judgements to the same extent as for NH listeners. However, separate stimuli sets (consisting of the core set of 20 melodies) were compiled for testing with and without rhythmic cues and were not mixed during an experiment run. For stimuli without rhythmic cues, the notes of the final progression were equally long (quarter notes), while for stimuli with rhythmic cues the length of the last note (half note) was four times that of the penultimate one (eighth note). Examples are shown in Figure 4.2(e) and (f).

Since all experiments were to be conducted in sound field, with subjects using their clinically assigned speech processors and settings, it was important to present stimuli at comfortably audible loudness levels, rather than at a fixed intensity. Participants therefore had to perform a loudness estimation task during which the subjective loudness of a 1 kHz tone presented at 10 intensities between minimum and maximum sound-field loudness had to be estimated. Each intensity level was presented 20 times in random order and participants had to assign a value of between 1 and 100 to describe the tone’s loudness. A value of 1 corresponded to a sound that was just audible, while 100 corresponded to one that would be regarded as being of maximum comfortable sound-field loudness. A comfortably audible loudness level was defined as between 50% and 70% as determined from a resulting loudness estimation curve.

Loudness balancing was performed in order to rule out any confounding loudness cues that could influence melody perception performance. Frequencies selected at 100 Hz increments from the range relevant to the melody perception tasks were loudness balanced to the 1 kHz tone at the level earlier determined to represent a comfortably audible loudness. Participants were required to adjust each of the probe tones to sound equally loud to, just louder and just softer than the 1 kHz reference tone. Probe frequencies were presented in random order in triplicate during each adjustment task and the average deviation from the reference intensity was subsequently calculated from the nine loudness judgments. Presentation intensities for each frequency were stored in a look-up table for use during the melody perception tasks.

Experimental procedure

To complete a full set of data responses for a specific experiment condition, participants were required to complete two experiment runs of three task repetitions each. During a single experiment run each task consisted of ten melodies with complete endings, paired to a complementary set of ten melodies with incomplete endings and presented in three keys each, for a total of 60 trials per task. The remaining ten melodies with complete endings from the core set of 20 were presented similarly during the second experiment run and again paired to a set of ten complementary melodies with incomplete endings. A full data set thus consisted of responses to 180 melodies with complete endings (20 melodies x 3 keys x 3 repetitions) and an equal number of melodies with incomplete endings.

All experiments were performed in sound field. Stimuli were presented through a Yamaha MS101 II loudspeaker, which was placed approximately 1 m in front of the listener on the side of the test ear. Subjects were instructed that they would hear a short melody and had to judge whether its ending sounded complete or incomplete. An incomplete ending was described as one that would end abruptly or unexpectedly, leave the listener “hanging”, or induce a feeling in the listener of wanting to add additional notes to round off the melody. Conversely, a complete ending was described as one that would leave the listener satisfied that nothing more was needed or expected to round off the melody to a meaningful unit.

The experiments were of a yes/no design. Participants were required to indicate whether the ending of the melody sounded complete by selecting either the “yes” or “no” button on a graphical user interface (GUI). All stimuli were presented in random order. Three practice trials were presented prior to commencement of each task. All practice trials had to be completed successfully before data recording would start. Participants were aware of the practice trials, but did not know how many would be presented before each task. No feedback was provided, but participants could monitor the task progress via a progress bar included on the screen. All tasks were self-paced and a new melody would not be presented before listener response had been received to the previously presented stimulus. Each task repetition lasted approximately 20 minutes and a single experiment run thus lasted almost an hour.

Although these were quite long sessions, participants were urged to take short breaks between each of the tasks.

4.2.2 Results

Figure 4.3 shows average completion perception ability of the respective listener groups (NH and CI listeners) for each combination of melodic ending and progression type (experiment condition A–D). Data responses were organised according to experiment condition and each listener’s performance was thus calculated as the percentage correct responses over 90 trials in each of the four conditions (10 melodies x 3 keys x 3 repetitions x 4 conditions).[6] The top and bottom panels represent results for melodies with and without rhythmic cues respectively. Melodic ending (complete or incomplete), progression type (leading tone–tonic or dominant–tonic pairs) and final note duration (lengthened or not lengthened) were defined as within-subject variables, while hearing ability (CI or NH) was defined as the between-subjects variable in a three-way repeated measures analysis of variance.

NH listeners achieved (average) scores of 83–96% correct, irrespective of final note duration, while CI listeners performed close to or below chance level for all but dominant–tonic progressions in both note length conditions. The difference in performance level between listener groups was highly significant ($F(1,12) = 104.70$, $p < 0.001$).

Although no significant main within-subjects effects were found in either of the listener groups, several interactions were found to be significant at the 0.05 decision level. Significant interaction between hearing ability and progression type ($F(1,12) = 6.57$, $p < 0.05$) indicates that performance was differently influenced by the progression type for the two listener groups.

---

[6] S8 completed only 45 trials (five melodies x three keys x three repetitions) in each of the four conditions owing to his limited time availability.

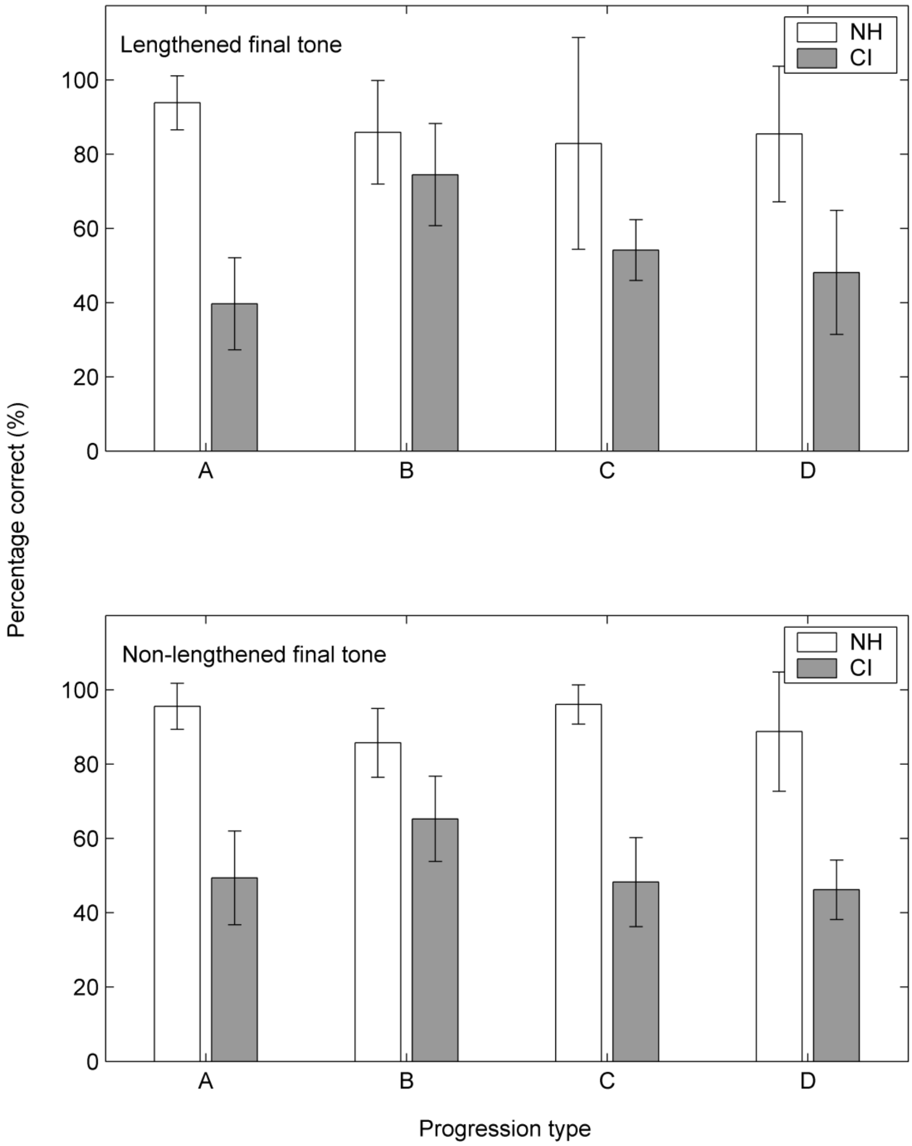

Figure 4.3: Average performance across listener groups in melodic completion task, with final note lengthened and not lengthened in top and bottom panels respectively. A = leading tone–tonic; B = dominant–tonic; C = tonic–leading tone; D = tonic–dominant. Error bars indicate one standard deviation from the mean.

CI listeners were better able to use dominant–tonic progressions for judgements regarding melodic completion, while NH listeners used leading tone–tonic progressions better to this end.

The significant interaction between hearing ability, melodic ending and progression type (F(1,12) = 10.94, $p < 0.01$) shows that, as for the interaction between hearing ability and progression type, progression type influenced performance in the two listener groups differently. Moreover, progression type exerted its most pronounced influence in melodies with complete endings. Taken together, these results indicate that CI listeners were better able to recognise complete endings involving dominant–tonic progressions, while NH listeners were better able to recognise complete endings that involved leading tone–tonic progressions. A similar effect was not seen for melodies with incomplete endings.

Significant interaction was also found for final note duration and progression type (F(1,12) = 7.19, $p < 0.05$). Both listener groups found judging melodic ending slightly easier when a lengthened final note was used together with a dominant–tonic progression than with a leading tone–tonic progression. However, a similar facilitating effect of progression type was not seen for melodies without a lengthened final note.

4.2.3 Discussion

Results of Experiment 1 show that information required for successful and reliable judgement of congruent melodic completion was not available to CI listeners to the same extent as to NH listeners. The effect of hearing ability appears to have masked any within-subject variables from significantly influencing perceptual outcome during these tasks, which may point to either insufficient or inappropriate information reaching higher-order cortical processing after electrical stimulation of the auditory system. Judgement of the congruency of a melodic ending depends largely on the assignment of tonal priorities according to an implicit tonal hierarchy, which in turn is based on successful perception of pitch relationships. Considering the limited pitch resolution afforded by the CI device and ensuing broad neural activation patterns, such insufficient or inappropriate information reaching an advanced music processing stage may result from impaired feature extraction already during early auditory processing stages, especially within the framework of a hierarchical, modular music processing system (Zatorre et al. 2002; Peretz & Coltheart 2003; Koelsch & Siebel 2005).

The observation of CI users' better performance with regard to melodies with dominant–tonic final progressions compared with those with leading tone–tonic progressions may be attributed to the wider frequency difference between the two consecutive final tones. A dominant–tonic progression is seven semitones apart, while a leading tone–tonic progression spans only two semitones. Depending on the key (which would dictate the frequency of the tonic) two tones seven semitones apart would amount to a 130–195 Hz frequency difference – more than one filter width of the speech processor – whereas a leading tone–tonic progression would not exceed 47 Hz. Previous studies regarding musical pitch perception have found that, on average, CI users struggle to reliably discriminate tonal intervals of less than five semitones (Gfeller et al. 2007; Galvin et al. 2007; Looi et al. 2008; Pretorius & Hanekom 2008), which has been suggested to be related to electrode-associated filter width and resulting overlapping neural activation patterns (Pretorius & Hanekom, 2008). However, the wider frequency separation resulted in a significant effect only in melodies with complete endings. If preceding melodic context were built up accurately, the same effect would have been expected for melodies with incomplete endings and it is hence reasonable to infer that the effect of wider frequency separation seen with complete endings may have resulted more out of bias towards positive responses than proper interpretation of preceding melodic context. This underlines the uncertainty experienced by CI users when having to make melodic judgements during music listening.

4.3 EXPERIMENT 2: PERCEPTION OF MUSICAL KEY VIOLATION

Experiment 1 showed that sufficient information for successful judgement of melodic completion is not available to CI users. Use of contextual melodic information was therefore investigated also at a different music processing stage. Experiment 2 sought to determine whether CI users are able to establish pitch relations between tones of a melody such that violation of the underlying musical key can be inferred. The perception of tones belonging to a specific musical key is governed by implicit rules of tonality of Western tonal music (Krumhansl 1990) and has been shown to be processed pre-attentively by listeners without music training (Tillman et al. 2000; Brattico et al. 2006).

Perception of musical key belonging represents an intermediate stage of music processing, occurring later than feature extraction, but earlier than syntactic judgement (Koelsch & Siebel 2005). Since pitch information thus needs to be integrated according to the pattern established by preceding tones, (global) contour analysis provides a reference point for (local) interval perception (Peretz 1990; Stewart et al. 2008). Considering the hierarchical framework proposed for music processing (Peretz & Coltheart 2003; Koelsch & Siebel 2005; Brattico et al. 2006), it follows that melodic character would not be perceived without successful earlier (primary) feature extraction. When applied to CI-mediated music perception research, an approach that considers stepwise build-up of musical context may provide useful beacons in search of establishing improved understanding of the underlying mechanisms that govern CI-mediated music perception.

Perception of musical key belonging, as a measure of melodic character perception, was investigated under two experimental conditions. Since accents serve as perceptual markers within a melody (Jones 1993), key-deviant notes were placed in either an accented or an unaccented position. Perceptual accents are generated relative to the preceding melodic context, both at pitch level (contour and interval patterns) and temporal level. Considering that coinciding pitch and temporal accents increase the saliency of melodic markers (Jones 1993), CI listeners’ perception of musical key violation was investigated with and without melodic accents to assess the extent to which melodic context can assist music perception by this listener group.

4.3.1 Methods

Subjects

The same group of CI and NH listeners who participated in Experiment 1 also participated in Experiment 2. Informed consent granted at the beginning of the investigation covered participation in both experiments. Subjects were again compensated for their participation.

Stimuli

The same core set of 20 melodies used for Experiment 1 was used during Experiment 2. Again all stimuli consisted of pure sine tones, generated in Matlab 6.5, at frequencies associated with tones between A3 and E6 and at intensity levels as determined for Experiment 1. All melodies were presented with final progressions that satisfied the requirements for harmonically complete endings and without any lengthening of the final note. Key deviations were created by tuning a target tone one semitone up or down, to violate belonging to the implied diatonic scale, but not that of the more generally implied chromatic scale. An arpeggio, consisting of the tonic, mediant, dominant and end-octave tonic tone, was presented before each melody to establish its implied key.

Target tones were placed well into the melody but not too close to the final progression, to ensure that sufficient melodic context would be available to direct listeners' decisions. Tonic and dominant tones (and if possible all three members of the tonic triad) featured melodically earlier than the target tone to reinforce the tonal context established by the broken chord.

Target tones at unaccented positions qualified as being neither pitch nor temporally conspicuous according to Jones's (1993) Joint Accent Structure, while accented positions satisfied the requirements for both pitch and temporal accent. The melody sets are shown in Figure 4.4(a) and (b).
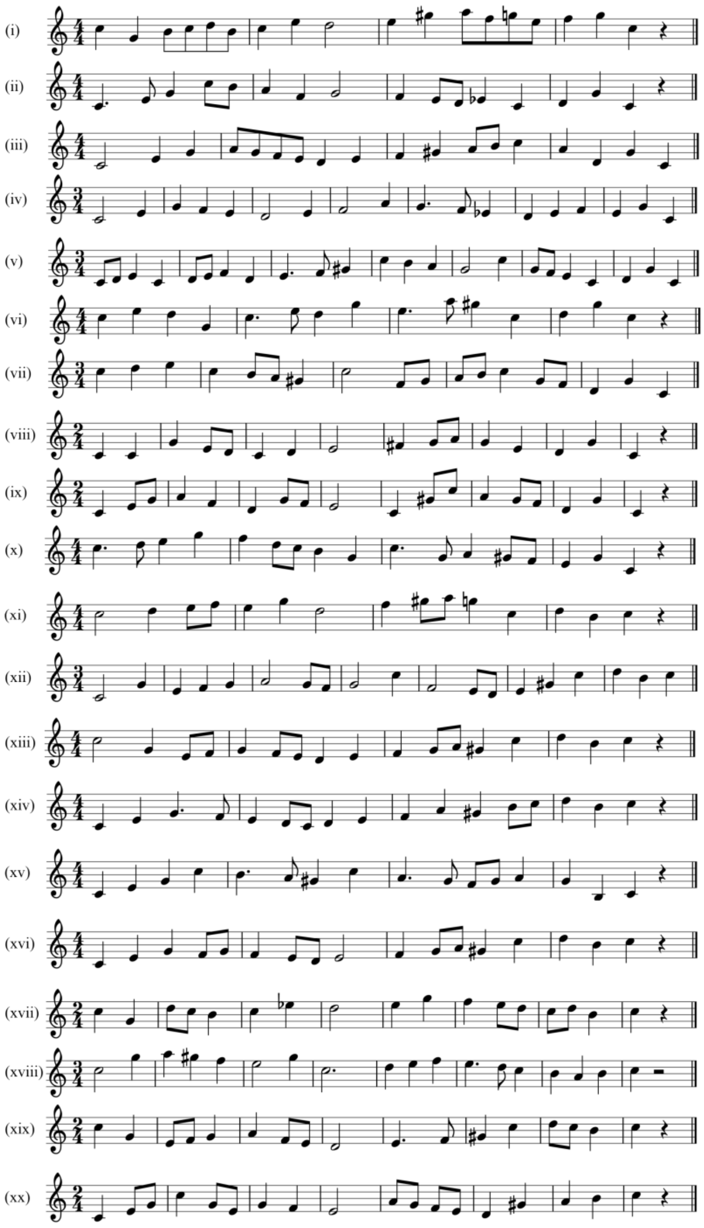
Figure 4.4a: Set of 20 token melodies with the target tone in an unaccented position.
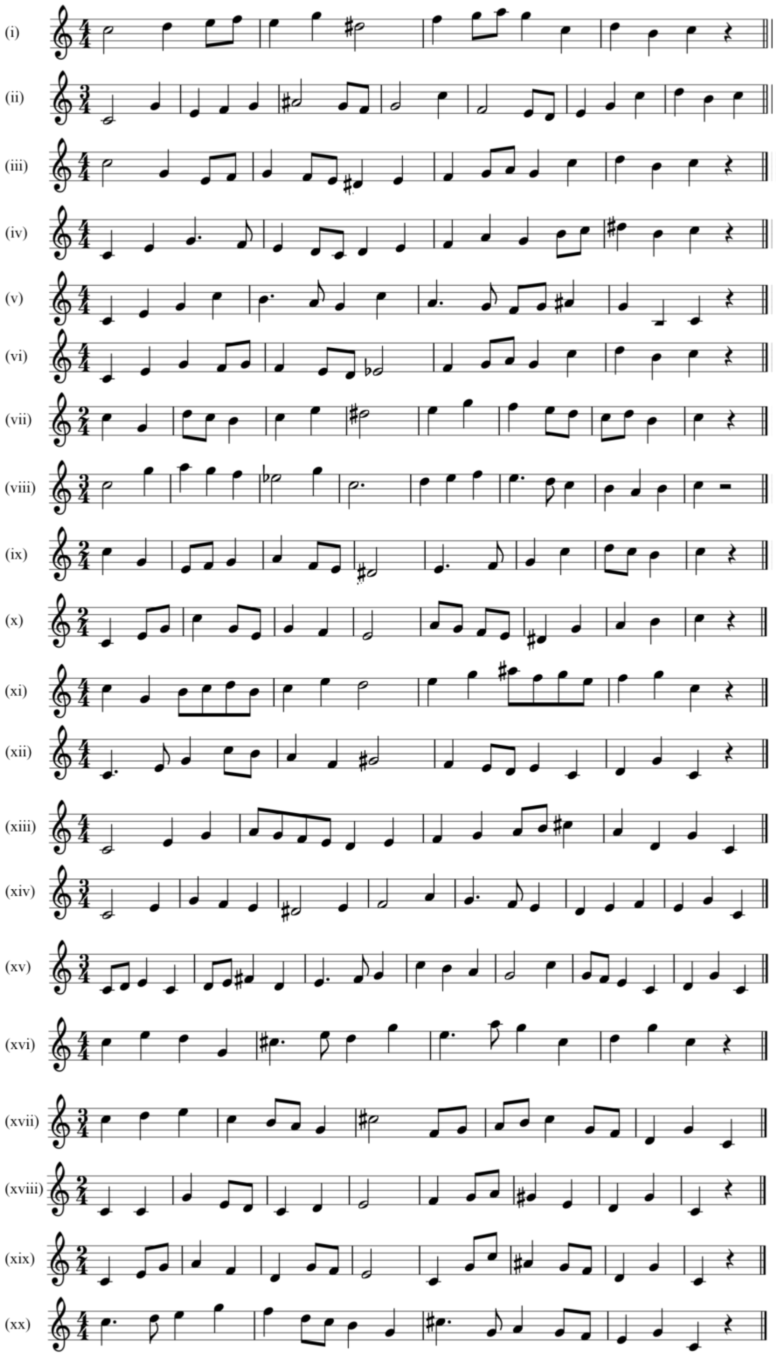
Figure 4.4b: Set of 20 token melodies with the target tone in an accented position.

Similar to the set-up used in Experiment 1, each of the 20 melodies was again presented in C, F and G major, resulting in a total stimulus set of 60 melodies. As during Experiment 1, the set was divided into two subsets, each containing an equal number of melodies ending with leading tone–tonic (upward) and dominant–tonic (downward) progressions. A melody subset containing key-deviant target tones was always paired with a complementary subset of melodies which did not contain any key deviant tones of the same accent condition (either accented or unaccented target tones), so that a single melody would not be presented with and without a key violation during the same experiment run.

Experimental procedure

The procedure for Experiment 2 was similar to that of Experiment 1. Within an experimental condition (accented or unaccented) subjects had to complete two experiment runs of three task repetitions each to complete a full set of data responses. Each task comprised 60 melodies, of which 30 (10 melodies presented in three keys each) contained key-deviant target tones. A full data set per condition thus comprised responses to 180 melodies containing a key-deviant target tone (20 melodies x 3 keys x 3 repetitions) and 180 melodies without a key-deviant target tone.

The experiment was performed in sound field in a quiet room without background noise. A Yamaha MS101 II loudspeaker was placed approximately 1 m in front of the listener on the side of the test ear. Subjects were instructed that they would hear a melody from the same set used during Experiment 1 and had to judge whether a specific tone violated the tonality rules of the underlying key. If the tone sounded out-of-key, subjects had to select the “yes” button on the GUI, while an in-key tone had to be associated with the “no” button. The position of the target tone was indicated on screen by a visual marker displayed for the duration of each melody, while a second marker indicated the progress of the melody. When the two markers aligned they turned red to alert subjects to the target tone being presented. All stimuli were presented in random order.

Practice trials, feedback and progress monitoring options were implemented as for Experiment 1. Experiment sessions again lasted approximately one hour and subjects were urged to take short attention breaks as necessary.

4.3.2 Results

Figure 4.5 shows average key violation perception scores for CI and NH listener groups. Results for melodies with and without key-deviant target tones are shown separately in both accented (top panel) and unaccented (bottom panel) conditions. Individual scores were calculated as the average percentage correct after 180 stimulus presentations (20 melodies x 3 keys x 3 task repetitions) in each experiment condition. Key violation status (in-key or out-of-key) and accent condition (accented or unaccented target tone) were defined as within-subject variables, while hearing ability (CI or NH) was defined as the between-subject variable in a two-way repeated measures analysis of variance.

Similar to their performance in Experiment 1, CI listeners performed significantly worse during the key violation perception experiment than NH listeners ($F(1,12) = 227.46$, $p < 0.001$). It is interesting to note, however, that performance trends were similar across the two listener groups. Both groups performed 12–15% worse for melodies with a key-deviant target tone present than for those without key violation. Key violation status was subsequently found to have a statistically highly significant effect on perception ability ($F(1,12) = 14.994$, $p < 0.005$). Perception ability was not significantly different for accented or unaccented target tones ($F(1,12) = 0.314$, $p > 0.05$). No interactions between variables were significant at the $p < 0.05$ decision level.

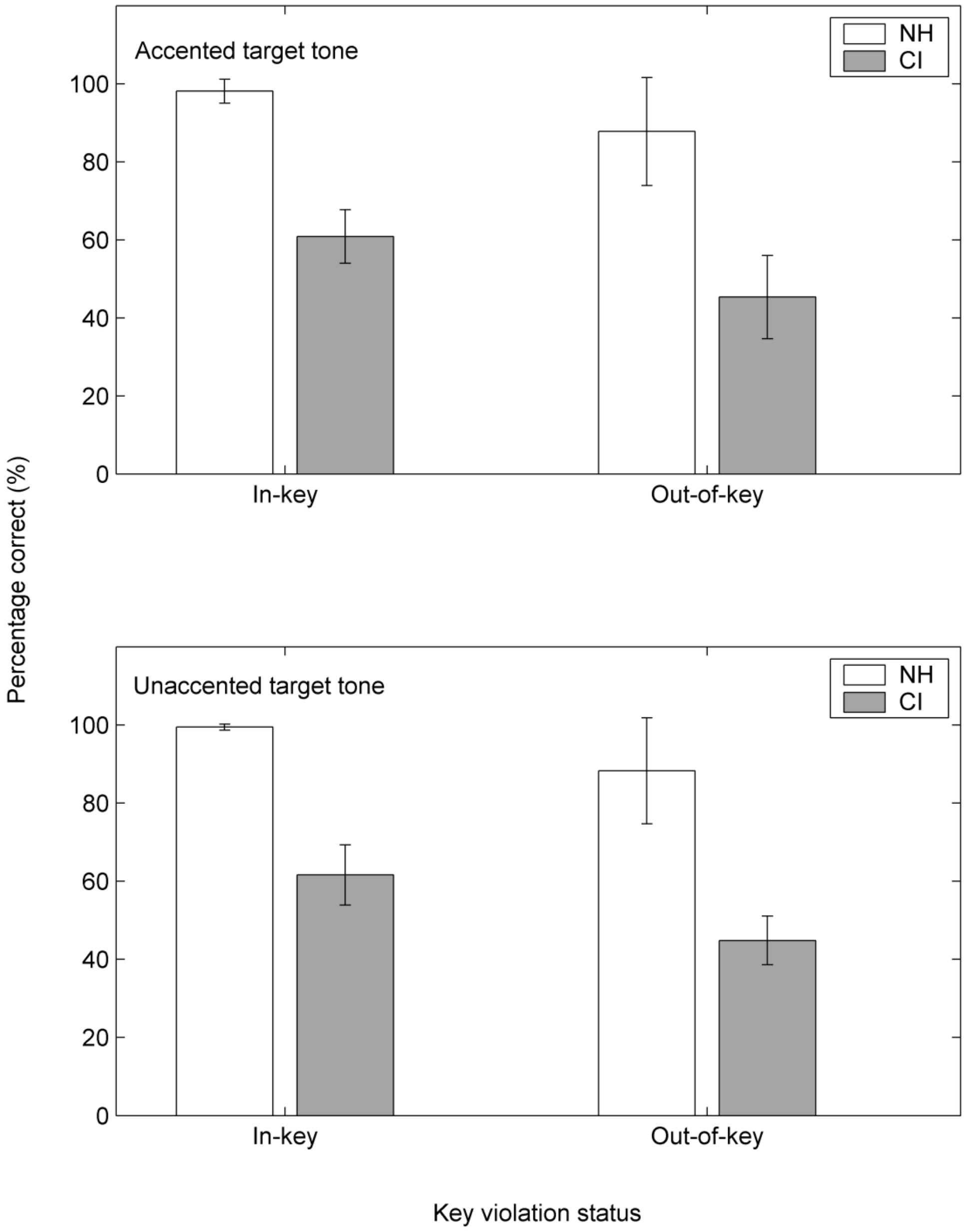

Figure 4.5: Average performance across listener groups in key violation task for accented and unaccented target tones in top and bottom panels respectively. Error bars indicate one standard deviation from the mean.

In view of these results, a brief experiment was performed using musical scales and arpeggios as stimuli. This was done to determine whether CI users are better able to use melodic context for recognition of a key-deviant tone when listening to a highly regular and expected tone sequence. Stimuli consisted of ascending and descending C, F and G major scales and arpeggios, each randomly presented with and without a key-deviant target tone. As for the previous experiment, a key violation was generated by tuning the target tone a semitone up or down from its normal pitch. Since accent condition did not significantly affect perception ability in the melody task, no such distinction was used for the present experiment. A complete stimulus set contained 24 tone sequences, of which subjects had to complete three repetitions. A full experiment run lasted approximately 20 minutes. Experiment design and set-up, task instructions, tone characteristics and frequency ranges were otherwise as for the melody task.

Results of the scale/arpeggio task are shown in Figure 4.6. Individual scores were calculated as the average of 36 data responses (12 stimuli x 3 repetitions) per condition. Each bar represents the group average per condition. Results were compared to those obtained during the melody task (accent condition), using violation status (in-key or out-of-key) and stimulus type (scale or melody) as within-subject variables and hearing ability (NH or CI) as between-subject variable during a two-way repeated measures analysis of variance. Since the target tone in the scale task was a tonic triad member (mediant or dominant), which qualifies as a pitch accent owing to tonal priority (Meyer 1956; Krumhansl 1979), results from the accented melody task were used for comparison during analysis.

Similar to the effect seen in the melody task, stimuli with a key violation resulted in significantly worse perception ability than those without ($F(1,12) = 8.72$, $p < 0.05$). The trend was similar across both listener groups, as confirmed by the non-significant interaction between hearing ability and violation status. Neither stimulus type nor its interaction with violation status was found significant at the 0.05 decision level. As in both earlier experiments, hearing ability exerted a highly significant influence on perception ability ($F(1,12) = 146.33$, $p < 0.001$).

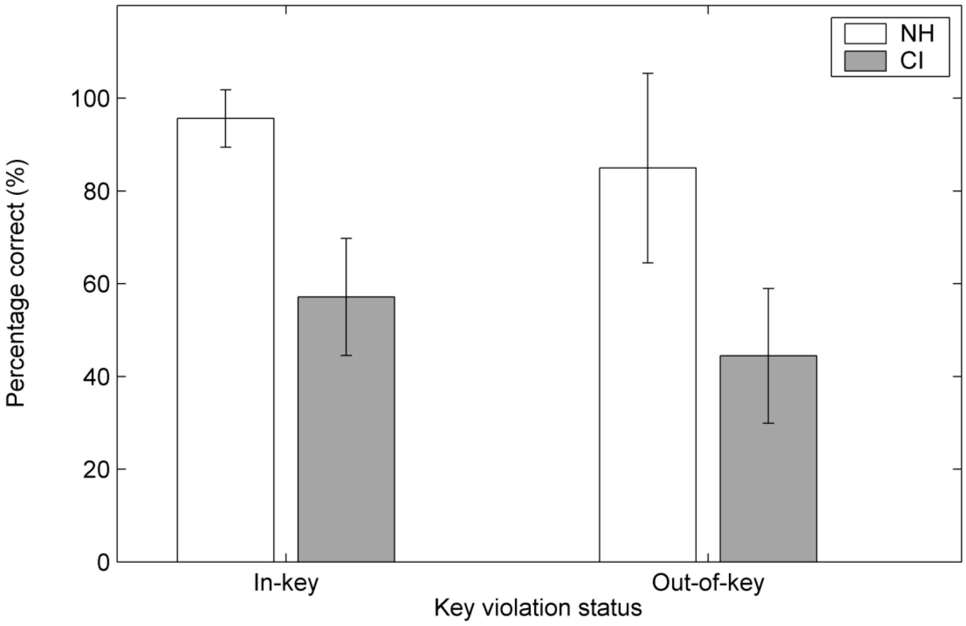

Figure 4.6: Average performance across listener groups in scale/arpeggio task. Error bars indicate one standard deviation from the mean.

4.3.3 Discussion

Taken together, results from Experiment 2 show that CI listeners could not access information needed for judgement of key violation, whether in the context of a melody or a more regular scale/arpeggio, to the same extent as NH listeners. This indicates that CI listeners were not able to build up sufficient melodic context from an unfolding tone sequence to drive melodic perception as experienced by NH listeners. Since a pitch difference of only one semitone, which corresponds to a frequency ratio of 1:1.06, had to be perceived during this experiment, it is likely that the pitch information reaching the cortical music processing system was inadequate to support perception of unfamiliar, real-world melodies. Compromised pitch quality may in turn be attributed to the limited spectral resolution available following conversion of acoustic to electrical signals already during early auditory processing (Shannon et al. 1995; Smith et al. 2002; Shannon 2005).

The finding that accent condition did not afford significantly different perception ability supports the powerful contribution of pitch information to successful music perception (Zatorre 2001; Zatorre et al. 2002). According to Jones's (1993) theory of joint accent structure, coinciding pitch and rhythmic accents serve as perceptual markers during melody perception, thereby promoting contour perception, which in turn facilitates local interval perception (Peretz 1990). However, it appears that pitch information may have masked facilitating effects of the accent structure of the tone sequence, with contrasting end results. For NH listeners, the available pitch information seem to have adequately signalled key violation status, without support from additional melodic markers needed for judgement; for CI listeners, however, the effect of compromised pitch information reaching this processing stage may have proved too strong for cues from the accent structure to render any helpful effect. This underlines the need for improved pitch resolution of the implant device.

The marked performance difference associated with violation status seen in both listener groups was somewhat surprising. If the results were due to pitch resolution alone, performance would be expected to have been similar for tone sequences either with or without a key violation. As such, NH listeners would have been expected to perform equally well – and CI listeners equally poorly – in both conditions, owing to the respective groups' earlier demonstrated frequency discrimination thresholds (for NH listeners see Wier et al. 1977; for CI listeners see e.g. Gfeller et al. 2007; Pretorius & Hanekom 2008). The finding that both listener groups performed worse on tone sequences with key-deviant tones (irrespective of stimulus type) suggests that in the presented context key-deviant target tones may not always have been unambiguously salient (Brattico et al. 2006). The result may thus also reflect response bias introduced by salience ambiguity, rather than mere (in)ability to detect a key violation. Close inspection of stimuli also showed that in two of the melodies used in this experiment, the target tone represented the first occurrence of that specific tone in the melody. If only the preceding context were used to guide the judgement of key violation, the tone could have implied the minor instead of major version of the key, and so may not have been regarded as key-deviant. This may have added to listeners' response uncertainty. In theory, presenting the task stimuli with chord accompaniment may have made more acoustic information available upon which listeners could base their perceptual decisions. However, since the neural activation patterns and associated behavioural outcome resulting from simultaneously presenting multiple tones of different frequencies have, to our knowledge, not yet been determined, chord accompaniment may have introduced an additional variable in the current study design. Despite these methodological concerns, it should be noted, though, that NH listeners still achieved between 85% and 88% correct (for scales/arpeggios and melodies respectively) in key-deviant conditions, whereas CI listeners performed close to chance. This indicates that although both groups' performance was influenced by response uncertainty, NH listeners were affected to a lesser extent than CI listeners.

4.4 GENERAL DISCUSSION

Results show that CI listeners were not able to complete either of the melodic perception tasks presented in this investigation successfully, which suggests that musical context information is not available to CI listeners to the same extent as to NH listeners. The results may be attributed to several causes, notably related to poor pitch resolution afforded by current CI technology. Frequency-to-place mismatch may result in pitch interval being nonlinearly distorted, more limited spectral resolution compared to that available to NH listeners reduces the extent to which different frequencies can be resolved and channel interaction may cause widespread neural activation that would limit extraction of a specific frequency based on the place code for pitch. Together these constraints may present an overwhelming challenge to the auditory processing system with regard to pitch perception of input as required for successful melody perception.

Although the results confirm poor melody perception ability of CI users as evidenced by previous studies, the approach used here may be valuable in disambiguating reasons for such observed behaviour. CI listeners' inability to use the preceding context and syntax markers of an unfolding tone sequence to drive melody perception can possibly be attributed especially to degraded and compromised pitch information being passed on from early feature extraction stages. Since a hierarchical information processing strategy, as has been proposed for music perception (Peretz & Coltheart 2003; Koelsch & Siebel 2005; Brattico et al. 2006), relies on integration of bottom-up information and top-down modulatory influences of especially pitch cues (Stewart et al. 2008; Balaguer-Ballester et al. 2009) In Western tonal music, it is clear that compromised low-level pitch information propagating through the music processing system will reduce the chance of music perception similar to that experienced by NH listeners.

Investigating CI-mediated music perception ability according to an approach that takes the neurocognitive principles of music perception into account provides insights that are not available from studies that consider perception of music-relevant components, such as pitch and rhythm, in isolation. Moreover, the present design goes beyond the conventional investigation of music perception in that it employs unfamiliar melodies to determine whether information that supports generic music listening is conveyed by the implant device.

Results from the studies described in the earlier chapters of this thesis are pertinent to the discussion of the observed outcome. Given that (i) an endogenous mechanism may exist to facilitate better pitch perception than would be expected from the operation of predominantly place pitch, as posited in Chapter 2, and (ii) given that CI users' contextual rhythm perception ability is comparable to that experienced by NH listeners, regardless of additional pitch complexity that may demand extra cognitive resources to be allocated to the stimulus, why can CI users not extract the melodic characteristics of a tone sequence? Or, on the contrary, if tasks that probed perception of melodic characteristics are too difficult for CI users, why is it that some CI listeners do score favourably on closed-set melody recognition tasks? I believe that the answer lies in the use of a closed set. If rhythm cues are the most salient cue reaching the melody processing stage, and it is sufficiently salient to be matched to a memory representation of one of the response options given by the closed, it could well facilitate higher-order cognitive functions to "fill in the gaps", and in a way "reconstruct" the tune sufficiently in the listener's mind to allow a perception of successful recognition. In that sense, music is in the ear of the beholder. However, when the aim is to determine how CI sound processing algorithms can be changed to convey sufficient auditory information to the central auditory processing system so that an unfamiliar melody, which conforms to the general compositional rules of Western tonal music, can be interpreted, the reasons for familiar melody recognition – whether successful or unsuccessful – need to be known. The results of the investigation described in this chapter therefore reflect descriptive music listening ability as a result of using novel melodies in an open-set design, rather than prescriptive music listening as dictated by using a closed-set, familiar melody recognition task.

A measure to evaluate CI listeners' ability to use contextual information during music listening may thus contribute to available test batteries (e.g. Cooper et al. 2008; Kang et al. 2009) by allowing objective and quantitative comparison across different processing strategies, CI user groups and rehabilitation facilities, and tracking of individual user improvement. Given the exploratory nature of the present investigation it should be emphasised that the tasks proposed here are not presented as being an absolute, 'end-of-line' assessment tool, but rather as modules that may, upon further refining and tailoring, contribute to existing test batteries and so provide a more comprehensive assessment of a CI listener's perceptual ability when confronted with music stimuli in a setting akin to everyday listening conditions.

It should also be noted that the task demands may have been too steep for the CI participants and that it may have been difficult for them to have judged, specifically, completion of a melody. However, since the CI participants were all post-lingually deafened listeners, who had all indicated during introductory interviews before commencement of the investigation that they had enjoyed listening to music before having lost their hearing, it was deemed a suitable exploratory approach. The task instruction that a complete melody would likely evoke a sensation of having ended satisfactorily, without the need for additional tones to "let it finish", while an incomplete melody was expected to create a sensation of melodic tension wanting to resolve to a more stable scale tone, was deemed appropriate owing to earlier reports regarding perception of musical expectancy and its violations (also see paragraph 4.2).

Since both tasks used during this investigation assessed listeners' ability to use preceding melodic context during music listening, incorporating a specific one into an existing music perception test battery would depend on the auditory processing level at which perception ability needs to be evaluated. Tasks based on melodic completion may be better suited to gauging listeners' use of global melodic cues, while tests evaluating perception of key violation could provide insight into context-based pitch perception at a local melodic level. It should be noted, however, that the design of the key violation test may have to be improved to accommodate response bias associated with the yes/no design used in this investigation. Furthermore, in future application of the approach explored here, specific attention should be given to (re)designing tasks that are brief enough to be relevant in a clinical setting; in their present form the tasks are too tedious to be incorporated into a music-specific test battery.

4.5 CONCLUSION

This investigation assessed whether CI listeners can use melodic context to aid perception of simple melodies in a setting similar to what would be experienced in real life. Results showed that helpful contextual information is not available to CI users to the same extent as to NH listeners, which points to the pronounced influence of compromised low-level information being propagated through a hierarchical processing system. The approach used during this investigation may contribute to existing measures to assess and improve CI-mediated music perception ability.

Chapter 5

FREQUENCY-DEPENDENT LOUDNESS VARIATION IN SOUND-FIELD LISTENING CONDITIONS

5.1 INTRODUCTION

In electric hearing cochlear conversion of acoustic input to nerve impulses is taken over by direct stimulation of the auditory nerve. The input signal is thus only an approximation of the naturally received stimulus and as such physical stimulus characteristics need to be coded appropriately to achieve approximately natural perceptual experience of the associated physical dimensions. In the case of loudness it is important to apply electrical signals to the electrode array such that the acoustic loudness will be conveyed accurately (Dorman et al. 1993).

The power relationship between loudness and stimulus intensity (Stevens 1955), combined with the absence of the natural compression mechanism of the basilar membrane and refractory control of neurotransmitter release, mean that a small change in applied current can result in a large change in perceived loudness (Moore 2003). The built-in loudness growth function of the speech processor compensates for unnatural loudness growth to some extent, but allows only a much reduced perceptual acoustic dynamic range – between 35 and 45 dB

(McDermott & Sucher 2007; McDermott & Varsavsky 2009) compared to a 60 dB range for speech and music in NH (Zwicker & Fastl 1999).[7]

Stimulus level has been shown to aid cochlear implantees’ speech recognition ability (Fu & Shannon 1998; Franck et al. 2002) and although not a critical factor in quiet listening conditions (Fu & Shannon 1998; Fu & Shannon 2000), the benefit of accurately encoded loudness cues may become evident in everyday listening conditions where limited spectral selectivity is available (Shannon et al. 2004, p. 339). Considering further also the perceptual relationship between pitch and loudness (Stevens 1935; Arnoldner et al. 2006), it follows that loudness effects may influence the outcome of psychophysical studies involving frequency-dependent acoustic input such as music.

This context is especially relevant for sound-field studies where a listener’s performance in everyday listening conditions, using a clinically assigned processor and settings, is tested. Such testing conditions assume that clinical amplitude mapping (as performed by an audiologist) generates balanced loudness percepts over the frequency range encountered in daily listening. Perceived loudness during CI-mediated hearing is usually controlled by establishing maximum (or comfort (C)) and minimum (or threshold (T)) stimulation levels for each electrode and then applying a single loudness growth function to determine the actual level of stimulation. This ensures that all electrodes are stimulated at a current level that produces an acceptable loudness percept between the T- and C-levels, but does not necessarily mean that each electrode produces an equally loud percept and thus smooth variation of loudness across different electrodes at intermediate stimulation levels may not be guaranteed (Blamey et al. 2000).

Since CI-mediated hearing strives to restore natural sound sensation (Hoth 2007), it is important to quantify the extent to which the perceptual experience of specific stimulus parameters in CI-mediated hearing differs from NH. Furthermore, within the hierarchical framework that underlies music perception and which guided the approach to this study (see

---

[7] These values refer to perceptual acoustic dynamic range and should not be confused with the input acoustic dynamic range, which is set as a processing algorithm parameter.

Figure 1.1, Chapter 1), it is important to be aware of factors that may influence processing at early hierarchical levels and the extent to which their influence propagates to subsequent processing levels. The perceptual association between pitch and loudness and specifically the encoding of loudness cues in electrical hearing, serve as motivation for the work reported on in this chapter. The aim of this investigation was twofold: (i) to investigate loudness estimation behaviour of CI users in sound-field conditions for pure tone signals using their clinically assigned processors and settings and (ii) to compare loudness balancing over a wide frequency range in sound-field conditions with trends observed for NH listeners. The two objectives were addressed in the loudness estimation and loudness balancing tasks respectively.

5.2 METHODS

5.2.1 Subjects

The CI group consisted of eight post-lingually deafened adult users (mean age = 45.6 years) of the Nucleus implant system. All subjects had experienced profound deafness for more than 10 years and had at least two years' experience with their implant device. Seven subjects were implanted with the CI24 array, while one was implanted with the CI22 array. Other than two implantees who used the 3G processor, all subjects used the Freedom processor, fitted with either the ACE or SPEAK speech processing strategy. For three subjects who were bilaterally implanted, only the ear subjectively regarded as the better of the two was used during the experiments. All subjects had participated in earlier research in our laboratory. Other relevant demographic details are provided in Tables A1 and A2 (Appendix A).

Eight age-matched NH listeners formed the control group (mean age = 48.9 years). All subjects achieved audiometric thresholds of 30 dB HL or better at five octave frequencies (250 Hz to 4000 Hz). Four subjects of the control group have also participated in other psychoacoustic research in our laboratory.

All subjects (NH listeners and cochlear implantees) gave written informed consent in line with the requirements of the relevant ethics committee prior to commencement of the investigation. Subjects were compensated for their time at the end of the investigation.

5.2.2 Experimental procedure: Loudness estimation task

All experiments were performed in sound field in a quiet listening environment. CI listeners used their own speech processors and clinically assigned maps as used during everyday listening. Stimuli were presented through a Yamaha MS 101 II loudspeaker placed approximately 1 m in front of the listener on the side of the test ear.

Each listener’s threshold and maximum comfort levels were determined acoustically before commencement of the experiment using a 1 kHz pure tone reference stimulus. The resulting acoustic dynamic range was then divided into 10 equal steps for presentation of a probe stimulus of the same frequency. Twenty probe tokens were presented at each of the ten stimulus levels in random order. Listeners were asked to assign a value between 1 and 100 to describe the loudness of the probe stimulus, with 1 being just noticeable (but certain) and 100 corresponding to the loudest sound that could comfortably be tolerated during the course of the experiment. Responses were submitted via a GUI.

To prevent decision drift over the course of the experiment, listeners were asked to judge the loudness of the probe stimulus relative to that of a reference stimulus. Each presentation round therefore consisted of a reference and a probe stimulus. The initial presentation round presented a 1 kHz reference tone at the acoustical maximum comfort level, which was by default assigned a loudness rating of 100, after which a probe stimulus followed. Listeners used a GUI-based slider bar to assign a rating to the probe stimulus. This probe stimulus, with its associated loudness rating shown on the screen, was then presented at the beginning of the next stimulation round to serve as reference for the next probe stimulus. Two different colours were used to denote presentation of the reference and probe stimuli on the screen.

Reference and probe stimuli were both 1000 ms long and separated by a 1000 ms inter-stimulus gap. The task was self-paced and a new stimuli pair was not presented before user response had been received. It took approximately 12 minutes to complete the task.

5.2.3 Experimental procedure: Loudness balancing task

The loudness balancing task required listeners to compare the loudness of probe tones at frequencies between 200 and 6100 Hz to that of a 1 kHz reference stimulus. The frequency range for the probe tone was divided into three sets (200–1250 Hz; 1300–2900 Hz; and 3300–6100 Hz), which were presented blockwise. For the three sets, probe tones were selected at 50 Hz, 200 Hz and 400 Hz increments respectively. Listeners had to perform three separate comparison tasks over the course of the experiment, respectively setting the probe equally loud to, just louder and just softer than the reference stimulus. Probe tones were always presented randomly and three repetitions at each frequency had to be completed in each of the comparison tasks. The average comparative loudness at each frequency was subsequently calculated from nine responses (three repetitions x three comparison tasks).

Both tones of a specific reference–probe combination were initially presented at the intensity associated with a listener’s pre-determined comfortable loudness level (70–75% of acoustic dynamic range). The intensity of the probe stimulus had to be adjusted to satisfy the specific task requirement using GUI-based controls. Listeners were allowed to relisten to the stimuli pair after each adjustment and submitted their final judgement only once they were satisfied that the loudness of the probe stimulus compared correctly to that of the reference stimulus. To prevent undue response bias, the probe tone was randomly assigned to either the first or the second presentation slot in a presentation round. This presentation order held for the entire adjustment cycle of the specific reference–probe combination. The reference and probe stimuli were associated with green and red on-screen signals respectively to ensure that listeners correctly aimed their adjustments at the probe stimulus of the pair. Both the reference and the probe stimulus were presented for 350 ms, separated by a 1000 ms interstimulus interval. Listeners completed each block of comparison tasks in approximately 15 minutes.

5.3 RESULTS AND DISCUSSION

5.3.1 Loudness estimation task

Loudness estimation results are shown for CI and NH listeners in Figure 5.1 and Figure 5.2 respectively, with the logarithm of the numerical loudness rating shown on the y-axis and stimulus intensity level shown on the x-axis. NH listeners’ behaviour generally followed a linear trend, while CI listeners’ loudness growth estimation was generally linear only for part of the dynamic range (except for S4) before levelling off for input levels above 60–70% of the dynamic range.

Linear regression was performed to compare the rate of loudness growth for the two groups. Since CI listeners experience loudness growth only for part of the dynamic range, least squares curve fitting was applied only to the linear section of CI listeners’ loudness estimation curves. The transition point where loudness growth entered the plateau phase was defined as the point after which the ratio between at least two consecutive loudness ratings differed by less than 5%. Fitted lines and the associated regression coefficients are shown on each graph. As shown by the fitted lines, CI listeners experienced steeper loudness growth than NH listeners (also see Table 1), which a t-test showed to differ significantly from NH listeners’ response curves ($t(14) = 2.71$; $p < 0.02$).

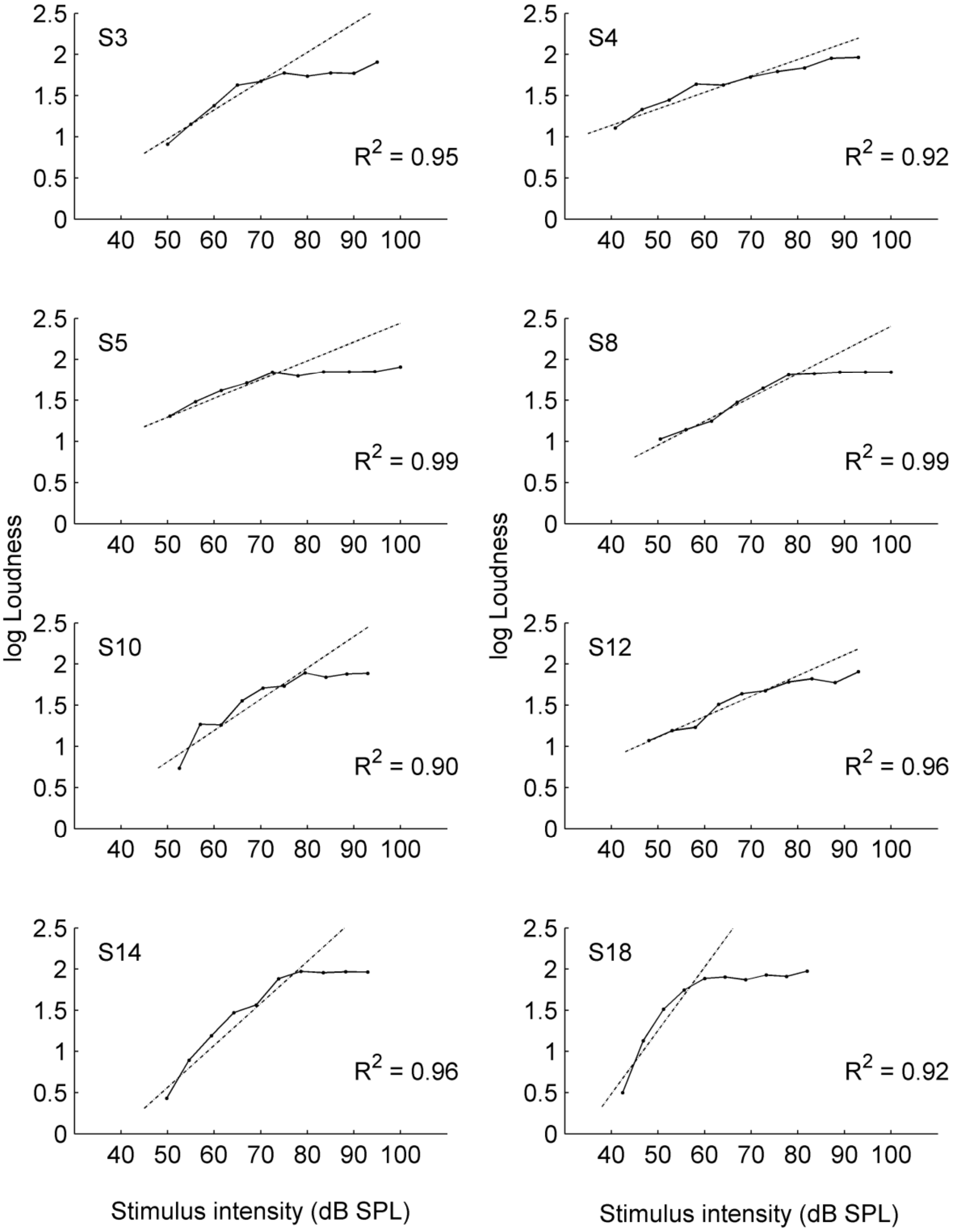

Figure 5.1: Loudness estimation results of individual CI listeners. Regression coefficients for a linear model fit only through the data points corresponding to loudness growth are shown.

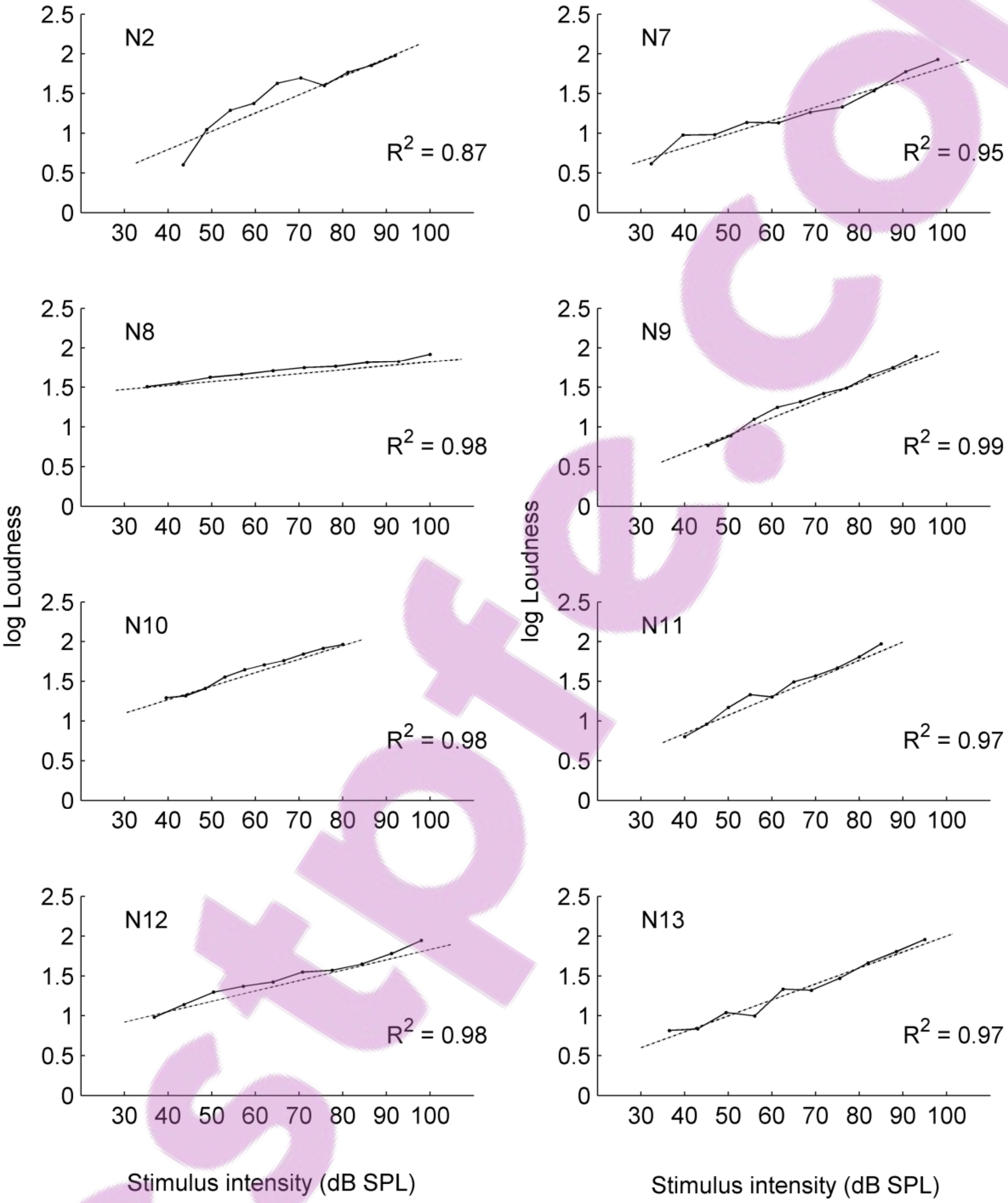

Figure 5.2: Loudness estimation results of individual NH listeners. Regression coefficients for a linear model fit through the data points corresponding to loudness growth are shown.

Table 1: Slopes of linear sections of loudness growth curves of CI and NH listeners and associated regression coefficients. CI users are denoted by S; NH subjects are denoted by N.

Subject	**Slope**	**Regression coefficient**
S3	0.035	0.95
S4	0.019	0.92
S5	0.024	0.99
S8	0.029	0.99
S10	0.038	0.90
S12	0.025	0.96
S14	0.052	0.96
S18	0.077	0.92
N2	0.024	0.99
N7	0.017	0.95
N8	0.006	0.98
N9	0.023	0.87
N10	0.018	0.98
N11	0.024	0.97
N12	0.014	0.98
N13	0.020	0.97

The results are borne out by the average behaviour of the two listener groups, as shown in Figure 5.3.

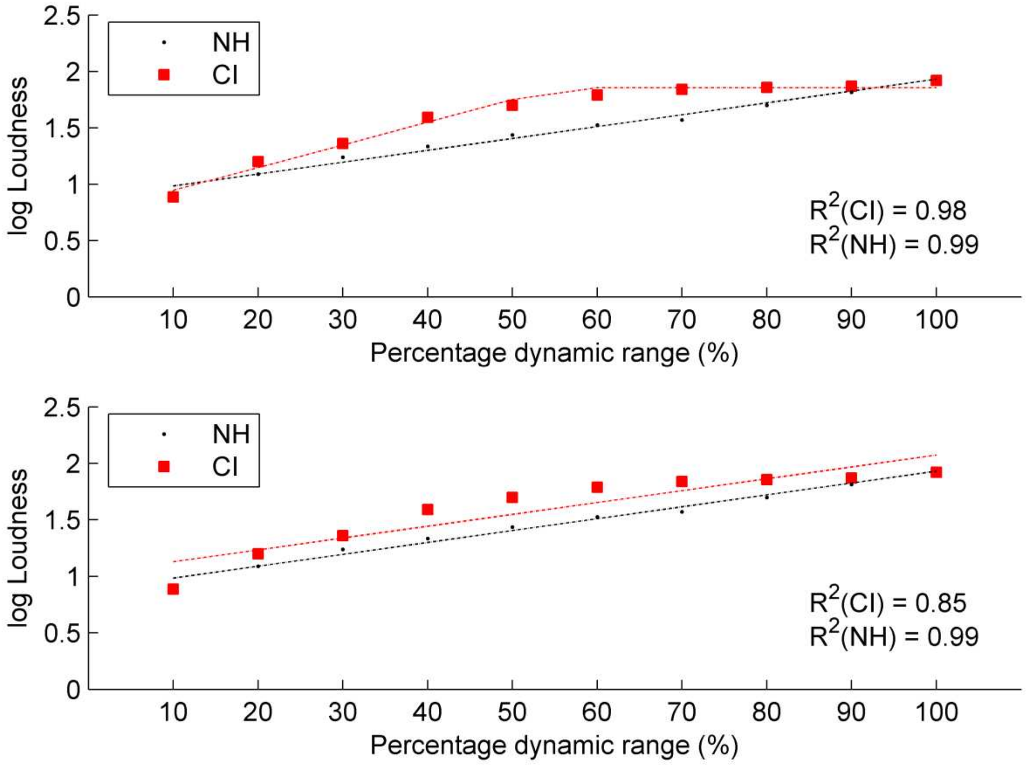

Figure 5.3: Average loudness estimation results for the two listener groups. Since the absolute values of perceptual minima and maxima are not the same for all subjects, data are plotted against percentage dynamic range. The top panel shows CI data fitted according to a two-stage linear model (dashed line), whereas the same data are fitted according to a monotonically increasing linear curve in the bottom panel. For comparison, the NH data set is fitted according to a monotonically increasing linear model in both panels. Regression coefficients for the respective fits are shown in each panel.

Considering the constraints on dynamic range, least squares fitting was performed on average CI data according to

$$f(x) = \begin{cases} ax + c, & x < k \\ c', & x > k \end{cases}$$

with a describing the slope, c the y-intercept of the line with positive slope, k the transition point and c' the y-intercept for the plateau section. Optimised parameter values (as found by using Microsoft Excel Solver function) showed $a = 0.02$ for CI data. Average data from NH listeners were fitted according to $f(x) = ax + c$, with $a = 0.01$ (top panel). The bottom panel shows CI and NH data both fitted to $(x) = ax + c$, with $a = 0.01$. Although the slope of true loudness growth for CI listeners was thus twice that for NH listeners, loudness growth for CI listeners, when considered across the entire available dynamic range, is comparable to that experienced by NH listeners. This may explain the similar loudness balancing behaviour observed for the two groups (see later).

These psychophysical results obtained in sound-field listening conditions are in line with those of McDermott and Sucher (2007), as well as the trends predicted by a computational model of expected acoustic loudness proposed by McDermott and Varsavsky (2009).

5.3.2 Loudness balancing task

The results of the loudness balancing task, averaged over the two groups, are shown in Figure 5.4. Since the absolute reference level was not the same for all subjects (see Methods), results are expressed as the deviation (in dB SPL) from the loudness associated with the reference tone. Relevant musical note names are shown on the secondary x-axis.

The similarity of the loudness balancing behaviour of the two listener groups was evaluated using the Pearson correlation coefficient, similar to the approach used by Throckmorton and Collins (2001). The patterns follow fairly similar trends across the frequency range ($r = 0.65$, $p < 0.0001$), although slightly more variable responses are seen at the higher end of the frequency range. Deviation from the reference level was seldom more than 3 dB SPL, except for the CI listeners at 6100 Hz and NH listeners between 2300 Hz and 3300 Hz. The observation may be attributed to listeners finding the reference balancing method more difficult (Collins & Throckmorton 2000), especially for loudness comparisons between tones at large frequency separations (Lim et al. 1977). The specific pattern of both curves displaying a general peak between 400–500 Hz and a trough between 1100–1200 Hz, suggests an effect of room acoustics.

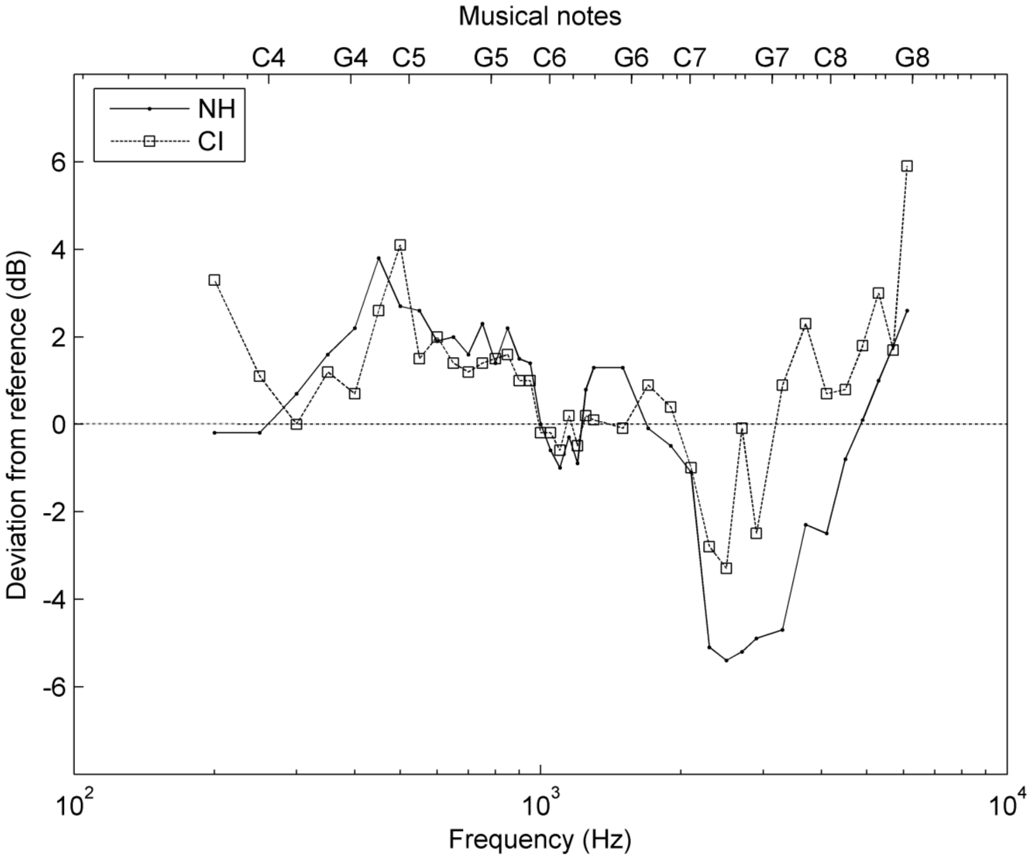

Figure 5.4: Average loudness balancing results for the two listener groups, expressed as deviation from the reference loudness (dashed horizontal line). The frequency range is expressed as musical notes on the secondary x-axis.

The practical relevance of the results, however, needs to be considered in the context of the dynamic range. In line with typical results reported for perceptually relevant listening in everyday conditions (e.g. Fu & Shannon 1998; McDermott & Sucher 2007), NH and CI listener groups in this investigation displayed average acoustic dynamic ranges of 60 dB SPL and 33 dB SPL respectively. (These acoustic dynamic ranges were determined before commencement of the loudness estimation task, see paragraph 5.2.2.) For both listener groups deviations from the reference level thus translate to 10% or less of the acoustic dynamic range, with the exception of the last data point for the CI listeners. Furthermore, for at least 85% of all the successive frequency values used in this study, the deviation differed by less than 2 dB SPL. Vos and Troost (1989) showed that in almost 70% of melodic intervals regularly found in Western tonal music the pitch difference does not exceed two semitones. When the frequency range is thus expressed in musically relevant units, the perceptual effect of loudness deviation is expected to be minimal across a melodic line (consisting of pure tones) presented in sound-field conditions.

5.4 CONCLUSION

The study reported here investigated a group of CI listeners' loudness perception abilities with regard to pure tone signals in sound-field conditions using their clinically assigned processors and compared the results to those of a group of NH listeners. Taken together, the results suggest that CI listeners would experience similar loudness deviation effects as NH listeners, despite reduced acoustic dynamic range and consequently steeper loudness growth curves. It furthermore appears that the loudness deviations experienced by CI users would not adversely affect their listening experience in a frequency-dependent sound-field listening task using pure tone stimuli. The findings do not, however, negate the importance of loudness balancing to rule out loudness as a cue during specific psychophysical tasks (Throckmorton & Collins 2001), but merely suggest that loudness mapping as performed during processor fitting may provide sufficient frequency-dependent loudness balancing for general listening experience (such as music listening) in everyday conditions. It would be interesting to see whether a task using complex harmonic sounds yields similar results.

Chapter 6

GENERAL DISCUSSION AND CONCLUSION

6.1 RESEARCH OVERVIEW

Despite allowing remarkable speech recognition success (Gifford et al. 2008), contemporary multichannel CIs generally afford post-lingually deafened users only limited music perception ability. To date research efforts have investigated CI-mediated music perception ability at two seemingly disconnected perceptual levels, considering perceptual outcomes either after finely controlled feature manipulation at electrode level or after presentation of complete music tokens that include several concurrently varying features. However, a systems approach, which considers electrically mediated peripheral hearing and subsequent central auditory processing together as the link between auditory input and perceptual outcome, may help to shape development of music processing strategies that are applicable to CI listeners' real-world listening conditions.

This study aimed to put forward a systematic approach for investigating cochlear implantees' music perception abilities. By investigating their perception of simple but realistic music-like stimuli in sound-field listening conditions, the study allowed insight into how central processing shapes primary auditory input derived from artificial peripheral stimulation into final behavioural outcome. Although the study still focussed on separate musical features, they were investigated in a musically relevant listening context and interpreted within the framework of the implant's processing capabilities, and so the investigation represents a bridge between understanding the perception of finely controlled elementary musical features and "whole" music listening. Findings from this and similar experimental investigations may contribute to the development of music-specific test batteries that can ultimately help to develop improved music processing strategies.

This study does not cover the entire music listening experience of a CI user, but rather presents an initial exploration into the perceptual outcome brought about by electrical hearing in real-world music listening conditions. Rather than adopting a progressive approach where the findings of a prior analysis serve as starting point for a subsequent investigation, the study focused on four separate aspects that may contribute to understanding CI-mediated music perception. Owing to the importance of successful pitch perception during music listening (Zatorre 2001; Peretz 2002; Peretz & Hyde 2003; Foxton et al. 2004), investigation of CI users' sound-field frequency discrimination abilities (Chapter 2) was a fitting starting point for the study. This was followed by an investigation into implantees' rhythm perception abilities when confronted with short tonal sequences of different rhythmic and pitch complexity (Chapter 3), based on an earlier report (Foxton et al. 2006) about rhythm perception difficulties experienced by tone-deaf listeners. Frequency-related loudness deviation as may be applicable to electrical hearing was also investigated (Chapter 5). Although loudness is not a primary contributor to the musical percept, unbalanced loudness cues associated with different pitch sensations as heard during an unfolding melody may influence a CI user's overall music listening experience.

The work described in Chapters 2, 3 and 5 examined CI users' perception of music-relevant stimuli at a feature analysis level (Griffiths 2003). It was, however, also relevant to explore how, in view of a modular music processing architecture (Peretz & Coltheart 2003), information from early analysis stages affect later perceptual stages. To this end, CI users' ability to infer musical meaning from the preceding context in a novel, single-voice melody with covarying pitch and rhythm cues was investigated (Chapter 4).

6.2 SUMMARY OF RESULTS AND RESEARCH CONTRIBUTIONS

6.2.1 Sound-field frequency discrimination abilities of CI users

In Chapter 2 the frequency discrimination abilities of CI users confronted with pure tone signals in sound-field listening conditions was tested. The investigation aimed, firstly, to put forward typical frequency discrimination threshold values for sound-field listening conditions (since such values had not been available at the time of the study) and secondly, to explain pitch perception outcomes in the context of the filter frequency response curves of the specific speech processing strategy (Research objective 1). Results generally showed that finer frequency resolution than would be expected in sound-field listening conditions may be available to CI listeners; the participants in this investigation were regularly able to discriminate between pure tone frequencies within a single filter band. The frequency discrimination behaviour furthermore appeared to have been influenced by the position of the reference frequency relative to the applicable filter frequency response curve.

The observed frequency discrimination resolution was somewhat unexpected, especially since pitch perception under the specific experimental conditions would predominantly be governed by the place pitch mechanism. Taken together, the findings are thought to point to a possible differential neural activation mechanism, which, based on the current distribution pattern applied to the recruited neural population, may contribute to conveying the frequency information. It was further particularly satisfying to find that such a mechanism may be activated by processor output in sound-field listening conditions without any deliberate attempt to manipulate electrode stimulation, as it suggests that the effect of such a mechanism may translate to perceptual outcome in the context of the complete electrically stimulated auditory system. Such a finding underlines the relevance of a systems approach whereby perceptual abilities of CI users are assessed in listening conditions relevant to their daily auditory environment.

6.2.2 Rhythm perception by cochlear implantees in conditions of varying pitch

Chapter 3 focused on the rhythm perception abilities of CI users when confronted with short tonal sequences of varying rhythmic and pitch pattern complexity (Research objective 2). The investigation was prompted by earlier findings, which showed that listeners who suffer from congenital amusia also experience rhythm perception difficulties (Foxton et al. 2006). These results were surprising in view of earlier observations of selective sparing of rhythm perception abilities when an impairment of the pitch processing stream was experienced (Hyde & Peretz 2004) and independent processing modules being proposed for pitch- and time-based musical relations (Zatorre 2001; Zatorre & Belin 2001; Peretz & Coltheart 2003; Peretz & Zatorre 2005).

Since CI listeners can be regarded as functionally tone-deaf (amusic) owing to limited, and possibly confounded, pitch processing abilities, it was possible that they might experience similar rhythm perception deficits as amusic listeners when confronted with tone sequences of covarying pitch and rhythm complexity (as would be heard during real-world music listening). Some ambiguity regarding cochlear implantees' rhythm perception abilities does indeed exist (see discussion in Chapter 3) and it was therefore deemed worthwhile, especially within the systems-based approach strived for in this study, to investigate CI-mediated rhythm perception using a similar approach as the one adopted by Foxton et al. (2006).

The results of the investigation showed that CI and NH listeners performed similarly on all the rhythm perception tasks, and specifically that additional pitch complexity did not adversely affect CI listeners' rhythm perception ability. In addition, CI users exhibited favourable rhythm perception ability regardless of the position of temporal irregularity within anisochronous tone sequences. Taken together, the results suggest that at least for temporal information, input derived from electrical stimulation at the auditory periphery supports adequate central processing to facilitate behavioural outcome comparable to that of NH listeners. Considering the results as the outcome of a hierarchical processing system, and subsequently drawing on findings regarding cortical activation patterns elicited by time-based auditory stimuli during interpretation, contemporary understanding of CI-mediated rhythm perception is extended. Such insights may help to shape efforts to develop music-specific signal processing strategies that are relevant to real-world listening conditions.

6.2.3 Perception of some melodic characteristics by CI users

Listeners' judgement of syntactic congruency was evaluated in Chapter 4 to investigate whether contextual information yielded by an unfolding melody is available to CI users during music listening. The investigation was developed on the strength of the notion that people have an innate ability to judge melodic endings as complete or incomplete and identify violation of the melodic key, based on the musical context provided by the preceding notes. This approach circumvents the need for closed-set recognition tasks and so frees the perceptual outcome from being influenced by memory representation. Results from such perceptual tasks can thus yield insight into the extent to which the information delivered by electrical stimulation can facilitate general music listening success in real-world listening conditions (Research objective 3).

The investigation comprised two experiments as alternative windows into investigating contextual perception of melodic lines as facilitated by electrically delivered auditory information. Perception of melodic completion is thought to be the outcome of processing effected at an advanced hierarchical level (Brattico et al. 2006) and if CI listeners were able to distinguish between complete and incomplete melodic endings correctly, it could indicate that the higher auditory system is able to assimilate sufficient contextual information for successful music listening, despite unnatural electrical stimulation at the periphery.

Implantees performed close to chance or below for all but the dominant–tonic final progressions, regardless of additional contextual cues such as final note duration. NH listeners achieved between 83% and 96% success on all tasks. The result may point to inadequate pitch analysis experienced by CI listeners at feature extraction level (as described in Chapter 2) creating a substantial perceptual impairment during higher processing stages. However, the finding that the nature of tonic–dominant final progressions was poorly identified while dominant–tonic final progressions appeared to be easier to interpret, indicates that perceptual uncertainty may compound perceptual challenges already presented by insufficient pitch perception ability. The finding that temporal markers provided insufficient contextual support to facilitate perception of melodic completion suggests that within the context of a processing system of which early input is constrained, as is operating in the case of a CI listener, the constraints imposed by insufficient resolution in one dimension overwhelms adequate perception (akin to NH ability) of a complementary stimulus dimension during perception of Western tonal melodies. Understanding CI-mediated hearing as the outcome of an intricate processing system in which several processing modules function in parallel and so is subject to forward feeding as well as modulatory feedback, rather than just as the function of peripheral processing of which the output feeds into individual processing modules, furthers the understanding of the influence early-level processing constraints may impose on eventual perceptual outcome. Although CI users' rhythm perception of tone sequences was similar to that of NH listeners' ability when the task was specifically aimed at perceiving temporal parameters, as shown in Chapter 3, such perception success was not powerful enough to support perception of melodic character of tone sequences where pitch and temporal cues were presented simultaneously as in real-world melodies.

Experiment 2 sought to establish whether CI listeners were able to establish adequate pitch relations between tones of an unfolding tone sequence for supporting identification of key violations. Identification of key violation represents processing at an earlier stage of contextual perception than higher-order syntactic congruency judgement (such as during a melodic completion task) (Koelsch & Siebel 2005). Results showed that sufficient melodic context could not be built up when listening to any of the presented tone sequences, regardless of attention-guiding accent tones. The findings pointed to a sustained effect of inadequate early-level pitch processing during electrical hearing.

Taken together, the results from these melodic listening tasks confirmed that cochlear implantees were not able to use contextual musical information to the same extent as NH listeners. Such inability may stem for compromised pitch perception that originates already at early feature extraction level. When the auditory system is considered as being hierarchically organised, the effect of early processing deficits at integratory processing stages becomes evident. Considering CI-mediated hearing with respect to underlying neurocognitive mechanisms may thus contribute to an improved understanding of the mechanisms affecting CI-mediated music perception.

6.2.4 Frequency-dependent loudness variation in sound field listening conditions

In electrical hearing, stimulus loudness is encoded as the amount of current applied to an electrode (Moore 2003). Owing to the exponential relationship between applied current and perceived loudness (Chatterjee 1999), a slight increase in amplitude may lead to a substantial change in perceived loudness. Further, given the known relationship between pitch and loudness perception (Stevens 1935; Arnoldner et al. 2006) and CI listeners' already limited pitch perception ability, uneven loudness percepts experienced for different frequencies may further worsen their music listening experience (Research objective 4). Loudness in CIs is usually controlled by establishing threshold and comfort levels associated with each electrode and then applying a single loudness growth function to determine the stimulation level on each electrode (Blamey et al. 2000). However, it is possible that controlling loudness in this way may not produce equal loudness across frequencies associated with the different electrodes. In order to put forward a realistic assessment of factors that may contribute to CI listeners' unsatisfactory music listening experience, it is thus pertinent to establish the variation experienced for frequency-dependent loudness perception in sound-field listening conditions.

Results showed that although CI listeners experienced steeper loudness growth than NH listeners, the two groups experienced similar loudness deviation across a frequency range spanning almost five octaves. In the context of the frequency variations in typical Western tonal melodies (for a discussion, see Vos & Troost 1989), such loudness deviation should have a minimal effect during music listening. The findings thus suggest that loudness mapping as performed during processor fitting provides sufficient frequency-dependent loudness balancing for satisfactory general listening.

6.3 GENERAL DISCUSSION

This study put forward a systematic approach to guide investigation into CI-mediated music perception. As outlined in Figure 6.1, the approach builds on the modular information processing scheme underlying the neurocognition of music perception as proposed by Peretz and Coltheart (2003) and findings from various psychoacoustic studies.

Figure 6.1 shows that auditory input is separated along two analysis streams, which deal with pitch and temporal information respectively, already at an early stage of central auditory processing. Within each stream distinct modules for handling separate characteristics of the auditory input exist. However, despite their apparent independence, these modules contribute collectively to processing of pitch and temporal information, and the resulting output combines to drive melody perception at a subsequent processing level. From there output, in turn, feeds into yet higher cognitive association processing modules to eventually create the aesthetically pleasing entity we perceive as music.

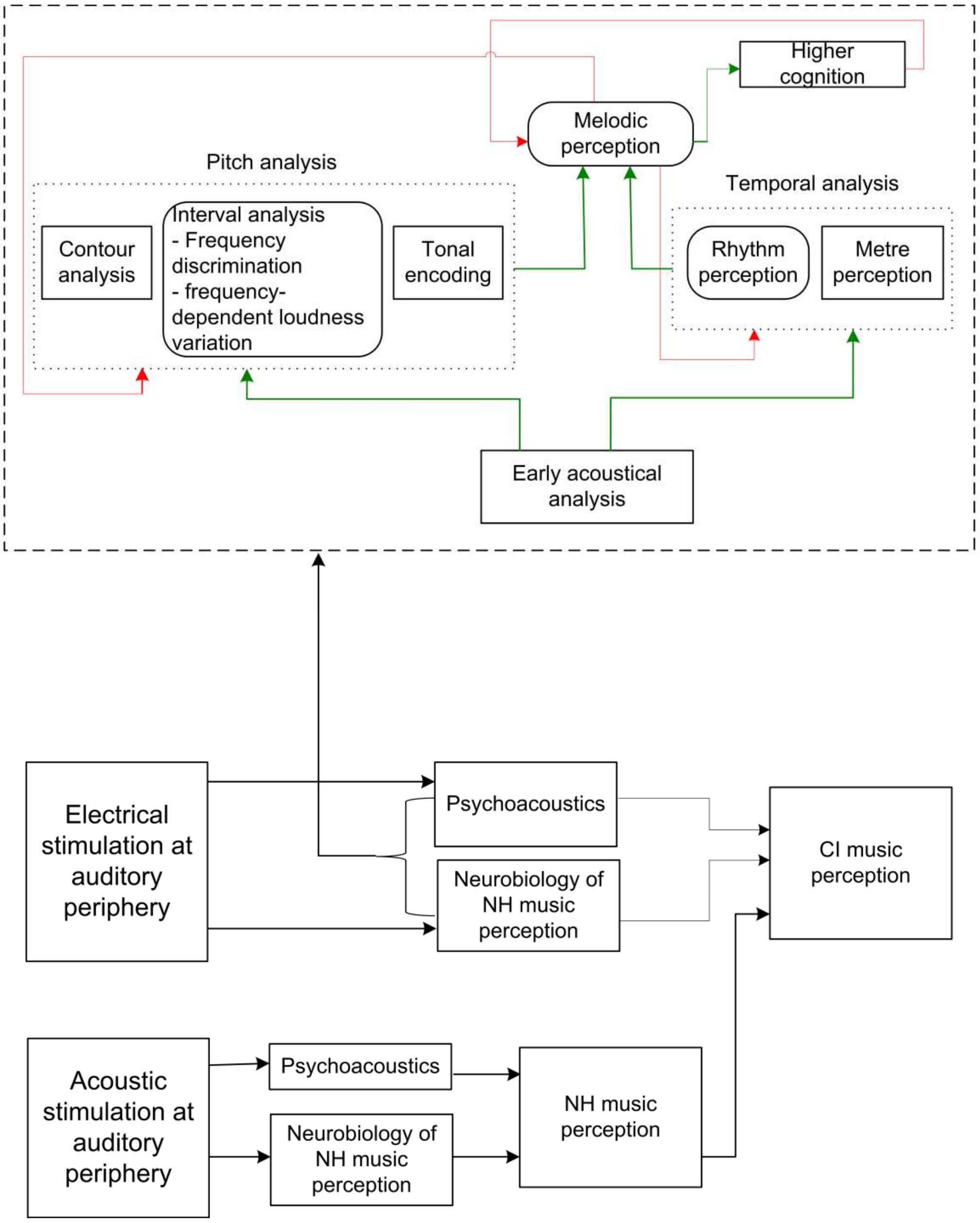

Figure 6.1: Schematic representation of the context of the study (repeat of Figure 1.1). The approach to improve understanding of CI music perception is informed by knowledge regarding psychoacoustics and neurobiology underlying music perception. The neurobiological framework on which psychoacoustic investigation in this study is based is developed in the dashed frame. Processing modules that contribute to independent pitch and time-related analysis are enclosed in dotted frames. Resulting output feeds into a processing module at a higher hierarchical level, from where output is sent further again. Green arrows represent forward information flow; red arrows represent modulatory feedback. The rounded blocks indicate specific processing steps investigated during this study.

To date CI-mediated music perception has not been investigated within the framework of a hierarchically organised perceptual system, nor has the outcome of sound-field listening been considered in this scheme. However, using a systematic, hierarchically organised approach to guide psychoacoustical experiments as proposed here may produce valuable insights to understanding CI-mediated music perception by allowing peripheral signal processing to be linked to eventual perceptual outcome.

This study investigated CI users' perception of musically relevant stimuli presented as isolated cues at early analysis levels as well as in combined form at perceptually higher levels in sound-field listening conditions. Although sound-field stimulus presentation is often regarded as allowing the experimenter too little control for interpreting perceptual outcome, implementing sound-field presentation according to an approach informed by the underlying systematics of music perception in NH in this study allowed insights that may not have been found otherwise. For example, investigation into sound-field frequency discrimination (Chapter 2) showed that CI users may achieve finer frequency resolution than a position-governed pitch-coding mechanism would be expected to make available Interpretation of results with consideration to the filter properties associated with the electrodes suggested that stimulation of the auditory nerve as occurs during sound-field listening may activate a differential neural activation mechanism to help signal frequency cues despite pitch-coding constraints imposed by tonotopic mismatch (Fu & Shannon 2002), unfocused neural targeting (Hughes & Abbas 2006) and limited spectral resolution (Shannon 2005).

Similarly, studying CI users' rhythm perception in the context of short tone sequences presented in sound field (Chapter 3) reaches beyond inferring CI users' perception of time-based cues during music listening from their performance on isolated rhythm perception tasks. Results from this investigation showed that CI listeners are able to perceive rhythmic patterns and irregularities equally well as NH listeners, irrespective of tonal information presented simultaneously. The finding suggests that CI-mediated transfer of temporal information perception is adequate to support rhythm perception as necessary during simple music listening. Furthermore, by considering the results against the background of rhythm perception difficulties experienced by amusic listeners whose amusia stems from pitch processing difficulties, it was concluded that the processing stage at which the pitch processing deficit occurs may influence rhythm perception. Within the framework proposed in Figure 6.1 it is thus possible that a central pitch processing impairment may silence the contribution of earlier extracted temporal information during music perception, while CI-mediated temporal information may proceed to later analysis stages unhindered owing to pitch and temporal cues already being separated at the auditory periphery.

Yet, despite finer pitch resolution being possible during electrically mediated hearing and transfer of temporal information allowing satisfactory rhythm perception, the investigation into CI users' ability at a higher perceptual level (where they had to perceive the melodic character of a simple, single-voice melody) shows that current implant technology supports neither satisfactory melody perception nor useful judgement of musical character (Chapter 4). When CI listeners were presented with stimuli that included covarying pitch and rhythm cues in tone sequences that obeyed the theoretical principles of Western tonal music, they were unable to perceive the melodic characteristics musically untrained NH listeners were able to interpret with ease. Findings from neurological studies (Koelsch et al. 2004; Limb et al. 2010), however, show that CI listeners' cortical processing structures are similarly susceptible to processing musical information as their NH counterparts'. When results from the frequency discrimination and rhythmic perception investigations (Chapters 2 and 3 respectively) are thus considered together with findings from the melodic perception study, these point towards inadequate and inaccurate early pitch processing exerting its effect well into higher cortical processing stages. Thus although this investigation confirms that limited pitch processing ability is central to CI users' poor music perception, it presents a context in which a systems-effected perceptual outcome allows signal processing strategies to be adjusted more efficiently. This may allow for either more objective assessment of user performance over time or better comparative studies regarding different signal processing strategies.

In Chapter 5 frequency-dependent loudness variation was investigated. Although loudness is not a primary contributor to music perception, stimulation of cochlear electrodes controlled only by signal processor output in sound-field listening conditions may result in uneven loudness cues being associated across the tonotopically arranged electrodes. The investigation showed that frequency-dependent loudness variation resulting from electrical stimulation is similar to that experienced by NH listeners under the same listening conditions. The finding suggests that loudness settings as set during conventional mapping procedures allow adequately even loudness percepts across a wide frequency range and therefore loudness cues as delivered during sound-field listening are unlikely to significantly confound pitch perception in everyday listening conditions. The outcome may be useful for design of music perception test strategies in that overly rigorous and time-consuming loudness balancing may not be necessary.

6.4 CONCLUSION

6.4.1 Concluding remarks

This study put forward an exploratory investigation into CI-mediated music perception in listening conditions relevant to those an implant user would generally encounter. It drew on knowledge regarding the neurocognition of music perception to guide experimental investigation into how electrical stimulation at the auditory periphery shapes eventual perceptual outcome. Using a systems-based approach, it was possible to explore the extent to which compromised transfer of music-relevant information influences processing at different stages, leading to a complete musical percept. Results showed that low-level pitch processing deficits exert a sustained effect on subsequent higher-order processing and thereby severely complicates successful music perception.

Although the respective investigations described in Chapters 2 to 5 addressed aspects of separate processing modules as would be operating during CI-mediated music perception, they are bound together by their being investigated within a system-relevant context. The overarching rationale behind the approach was that speech processor output at the periphery feeds into the central auditory processing system where it functions within a complete perceptual system where output of several processing modules is integrated to effect a whole music percept. As such CI-mediated music perception was regarded as a function not only of peripheral sound processing, but rather as the *result of speech processor output as handled within an integrated, modular and hierarchically organised central processing system*. It was hence deemed applicable to investigate aspects that may contribute to music perception while considering the central processing mechanism, but also to try to relate them to design parameters as implemented in current sound processing strategies since that is the processing level at which outcome of research regarding CI-mediated music perception can be implemented.

In Chapter 2 frequency discrimination abilities of CI users was investigated in listening conditions that resembled those they are confronted with in daily life. Results were in line with those of previous studies (e.g. Gfeller et al. 2007) in that CI users' frequency discrimination abilities were markedly poorer than had been shown for NH listeners previously. However, the investigation delved deeper than only determining typical frequency discrimination thresholds over a range of frequencies by relating the observed behaviour to a characteristic of the frequency response curves as implemented in the filter bank of the sound processing algorithm. When expressed relative to filter width, a significant number of frequency discrimination thresholds were found to be smaller than one filter width. Considering that each filter spans a range of input frequencies but is associated with one electrode (also see Figure 2.5), discrimination between frequencies falling within the same filter band was unexpected, especially given that the place pitch mechanism was deemed to dominate pitch perception under such conditions. By considering the nature of overlap between filter frequency response curves and how it would influence current delivery to electrodes – and hence neural activation patterns that feed into central processing steps – it was satisfying to find that a mechanism that could effect improved pitch perception in CI-mediated hearing may be available in current-generation sound processing algorithms. Had CI-mediated pitch perception thus been considered only as the outcome of peripheral, device-governed processing, without considering its influence on further processing steps, such hidden pitch perception possibility would not have been found. The findings thus may contribute to future efforts to improve pitch perception abilities of CI users despite the constraints of reduced spectral resolution owing to only a limited number of electrodes being available to handle the entire spectrum of cues delivered to the speech processor in everyday listening conditions.

Rhythm perception ability of CI users was investigated in Chapter 3 with specific interest in determining whether good rhythm perception ability, as has been shown in earlier studies, holds when the task is context based. The investigation was prompted by findings of a study by Foxton et al. (2006), which showed that amusic listeners displayed rhythm perception difficulties which likely stemmed from a pitch processing deficit. This was an unexpected finding owing to the early independence of the pitch and temporal processing streams facilitating music perception, as suggested by Peretz and Coltheart (2003). Seeing that CI users can also be regarded as functionally amusic owing to the limited pitch perception ability afforded by the device, it was deemed suitable to determine whether pitch cues presented together with rhythm cues gave rise to similar rhythm perception difficulties in CI users as amusic listeners. The approach was deemed valuable especially considering that several earlier studies specifically aimed at investigating CI users' rhythm perception abilities were conducted without covarying pitch changes within a tone sequence. Results of the present investigation showed that pitch processing deficits did not influence rhythm perception ability of CI users to the same extent as experienced by amusic listeners in the study by Foxton et al. (2006); rather, rhythm perception behaviour under all test conditions appeared similar to that of NH listeners. Interpreting the results with consideration to the (good) temporal resolution afforded by the implant device suggested that the level at which the pitch processing deficit occurs along the auditory processing pathway may influence the extent to which processing of one dimension influences processing of the other, possibly because the good temporal resolution is able to propagate and sustain its effect throughout subsequent processing steps. Had rhythm perception abilities been considered only as a function of peripheral processor output, without consideration to how it functions within later central auditory processing steps, such understanding of systems-level rhythm perception in CI-mediated hearing may not have precipitated.

Chapter 4 described CI users' abilities to perceive some melodic characteristics of novel, single-voice melodies. The rationale for the investigation was that melody perception, specifically in the Western tonal tradition, requires both pitch and rhythm cues to be perceived successfully to allow perception of a unified musical percept. However, owing to earlier studies of CI users' music perception ability often employing familiar melodies and some favourable outcomes being observed, it was deemed appropriate to disambiguate the reason for such recognition success. The rationale was that having to select a response option from a closed set of familiar melodies may activate higher cognitive processing modules such as memory and/or associative emotional affects to contribute to successful recognition. Although drawing on additional information to help one make a perceptual decision is not undesirable in real-world listening, it does not allow the perceptual outcome to be linked to the nature of information delivered to the central auditory processing system by electrical handling of input at the periphery. It was therefore deemed valuable to determine first whether the information reaching the processing stage at which melody perception is effected is sufficiently salient to signal that a tone sequence is melody-like, for if perception of the melodic character of a tone sequence is not possible, successful melody recognition of familiar melodies as observed in earlier studies is likely to be due to use of other supporting perceptual strategies.

Although the perceptual tasks investigated in this thesis were not meant to follow in a progressive sequence, perception of melodic character was deemed an appropriate perceptual task after investigation of perception of separate elementary stimuli. Furthermore, since sound-field frequency discrimination ability was found to be better than expected (Chapter 2) and rhythm perception behaviour within the context of a tone sequence was similar to that of NH listeners (Chapter 3) it was not an unreasonable question to ask whether the output of those respective processing steps could combine to support perception of the melodic character of a tone sequence. Results showed, however, that despite pitch perception being better than would have been expected when considering the dominant place pitch mechanism at play, constraints such as limited spectral resolution, mismatched frequency-to-place mapping and channel interaction imposed on pitch perception ability may prove too overwhelming during the melody perception stage to allow NH-hearing-like rhythm perception ability to make a substantial contribution to perception at this level. The value of the approach of this investigation lies therein that the ability of CI-mediated auditory input to convey melodic character could be investigated in an objective manner, without being confounded by contribution of additional information derived from higher-order processing modules that form part of the music perception system.

The approach implemented in the study described in Chapter 4 was exploratory and as such the same CI listener group did not participate in each investigation (also see Appendix A and Table A1). Reflection after the investigation had been completed showed that a different study design may have been useful. It is possible that the tasks were too difficult for CI listeners, especially given their pitch perception difficulties and some may argue that such an outcome could have been foreseen. However, since the rationale underlying the entire study was that eventual perceptual outcome is not only the “linear” sum of the outcome of separate processing steps, the investigation was deemed to be justified in the overarching quest of improving current understanding of factors that contribute to (or distract from) successful CI-mediated music perception.

The choice of stimuli may also have contributed to the difficulty of the tasks and it may be argued that using stimuli which included harmonic accompaniment could have provided listeners with more contextual information upon which to base their decisions. However, to deduce processing of peripheral input based only on eventual perceptual outcome means that the researcher sacrifices some experimental control over variables. As applied to CI-mediated perception it implies that auditory input which delivered several frequency components simultaneously may have complicated interpretation of results, since it would not have been known to what extent overlapping neural activation, frequency mismatch or nonlinear frequency distortion may have influenced eventual perceptual outcome. In view of such a design constraint it was thus deemed relevant to attempt the investigation in a rather simplistic way, which, admittedly, may not be realistic for true music listening situations, but which could provide some elementary insight upon which further investigations can build.

In hindsight, designing investigations described in the respective chapters such that the outcomes could have been correlated statistically may have helped to bind the findings together more strongly and so have contributed to delineating the relative contributions of specific processing modules’ outcome in supporting CI-mediated music perception (also see section 6.4.2). However, despite this limitation, the current study design did allow an alternative view into CI-mediated music perception, which may serve as a departure point for future investigations, specifically with the aim of tracking individual listener improvement or functional perceptual benefit afforded by future-generation processing algorithms.

6.4.2 Future directions

As shown through the earlier discussion, this study was not an exhaustive investigation of all the underlying aspects that contribute to CI-mediated music perception, but considered some that may be pertinent to a better understanding of music perception in electric hearing. As such, further investigation into cochlear implantees' perception of musical characteristics such as timbre, metre and melodic contour, performed in listening conditions that approach those an implant user would generally encounter, may further shape understanding of how to translate electric auditory input to useful, behaviourally relevant perceptual output. Especially relevant may be combining investigation of sound-field perceptual ability of CI users with imaging studies in order to form a more comprehensive picture of how the central auditory processing system handles electrically mediated input originating at the peripheral processing stage.

At a more immediate level, further research may involve the following:

- Similar experiments as performed during this study, but with harmonic complexes or true instrument sounds. Results may show how filter output following extraction of harmonic components activates several neural populations, and how such combined neural activity in turn shapes perceptual outcome.

- Delineating the relative contribution of specific music-relevant input cues. Findings may be useful in developing a quantitative, clinically useful measure of general music perception ability of CI users. Such a quantitative measure may help to compare objectively different processing strategies, track changes in user performance over the course of rehabilitation or evaluate the perceptual effect of specific adjustments to processing algorithms or stimulation procedures. Current test batteries for assessing CI users' music perception abilities (e.g. Cooper et al. 2008; Kang et al. 2009) evaluate perception of musical cues in an isolated manner and connecting the outcomes of the different test modules may be useful in establishing a general view of a CI users' music perception ability.

REFERENCES

Arnoldner, C., Kaider, A. & Hamzavi, J. (2006). The role of intensity upon pitch perception in cochlear implant recipients. *Laryngoscope*, *116*, 1760-1765.

Balaguer-Ballester, E., Clark, N.R., Coath, M., et al. (2009). Understanding pitch perception as a hierarchical process with top-down modulation. *PLoS Computational Biology*, *5*, e1000301-

Bigand, E. & Pineau, M. (1997). Global context effects on musical expectancy. *Perception and Psychophysics*, *59*, 1098-1107.

Blamey, P.J., Dooley, G.J., James, C.J., et al. (2000). Monaural and binaural loudness measures in cochlear implant users with contralateral residual hearing. *Ear and Hearing*, *21*, 6-17.

Boltz, M. (1989). Rhythm and "good endings": Effects of temporal structure on tonality judgments. *Perception and Psychophysics*, *46*, 9-17.

Boltz, M. (1993). The generation of temporal and melodic expectancies during musical listening. *Perception and Psychophysics*, *53*, 585-600.

Boltz, M.G. (1999). The processing of melodic and temporal information: independent or unified dimensions? *Journal of New Music Research*, *28*, 67-79.

Brattico, E., Tervaniemi, M., Näätänen, R., et al. (2006). Musical scale properties are automatically processed in the human auditory cortex. *Brain Research*, *1117*, 162-174.

Chatterjee, M. (1999). Effects of stimulation mode on threshold and loudness growth in multielectrode cochlear implants. *Journal of the Acoustical Society of America*, *105,* 850-860.

Collins, L.M. & Throckmorton, C.S. (2000). Investigating perceptual features of electrode stimulation via a multidimensional scaling paradigm. *Journal of the Acoustical Society of America*, *108*, 2353-2365.

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 and Hearing*, *29*, 618-626.

Cowan, N. (1984). On long and short auditory stores. *Psychological Bulletin*, *96*, 341-370.

Cox, R.M. & Alexander, G.C. (1995). The abbreviated profile of hearing aid benefit. *Ear and Hearing*, *16*, 176-186.

Deutsch, D. (1998). *The psychology of music*. San Diego: Academic Press.

Donaldson, G.S., Kreft, H.A. & Litvak, L. (2005). Place-pitch discrimination of single- versus dual-electrode stimuli by cochlear implant users (L). *Journal of the Acoustical Society of America*, *118*, 623-626.

Dorman, M.F., Smith, L. & Parkin, J.L. (1993). Loudness balance between acoustic and electric stimulation by a patient with a multichannel cochlear implant. *Ear and Hearing, 14*, 290-292.

Dorman, M.F., Smith, L., Smith, M., et al. (1992). The coding of vowel identity by patients who use the Ineraid cochlear implant. *Journal of the Acoustical Society of America*, *92*, 3428-3431.

Dorman, M.F., Smith, L.M., Smith, M., et al. (1996). Frequency discrimination and speech recognition by patients who use the Ineraid and continuous interleaved sampling cochlear-implant signal processors. *Journal of the Acoustical Society of America*, *99*, 1174-1184.

Firszt, J.B., Koch, D.B., Downing, M., et al. (2007). Current steering creates additional pitch percepts in adult cochlear implant recipients. *Otology and Neurotology*, *28*, 629-636.

Foxton, J.M., Dean, J.L., Gee, R., et al. (2004). Characterization of deficits in pitch perception underlying "tone deafness". *Brain*, *127*, 801-810.

Foxton, J.M., Nandy, R.K. & Griffiths, T.D. (2006). Rhythm deficits in "tone deafness". *Brain and Cognition*, *62*, 24-29.

Franck, K.H., Xu, L. & Pfingst, B.E. (2002). Effects of stimulus level on speech perception with cochlear prostheses. *Journal of the Association for Research in Otolaryngology*, *4*, 49-59.

Fu, Q.J. & Shannon, R.V. (2000). Effects of dynamic range and amplitude mapping on phoneme recognition in Nucleus-22 cochlear implant users. *Ear and Hearing*, *21*, 227-235.

Fu, Q.J. & Shannon, R.V. (2002). Frequency mapping in cochlear implants. *Ear and Hearing*, *23*, 339-348.

Fu, Q.J. & Shannon, R.V. (1998). Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. *Journal of the Acoustical Society of America*, *104*, 2570-2577.

Fujita, S. & Ito, J. (1999). Ability of Nucleus cochlear implantees to recognize music. *Annals of Otology, Rhinology and Laryngology*, *108*, 634-640.

Galvin, J.J., Fu, Q.J. & Nogaki, G. (2007). Melodic contour identification by cochlear implant listeners. *Ear and Hearing*, *28*, 302-319.

Gfeller, K., Christ, A., Knutson, J., et al. (2003). The effects of familiarity and complexity on appraisal of complex songs by cochlear implant recipients and normal hearing adults. *Journal of Music Therapy*, *40*, 78-112.

Gfeller, K., Christ, A., Knutson, J.F., et al. (2000). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of the American Academy of Audiology*, *11*, 390-406.

Gfeller, K. & Lansing, C. (1992). Musical perception of cochlear implant users as measured by the *Primary Measures of Music Audiation*: An item analysis. *Journal of Music Therapy*, *29*, 18-39.

Gfeller, K. & Lansing, C.R. (1991). Melodic, rhythmic, and timbral perception of adult cochlear implant users. *Journal of Speech and Hearing Research*, *34*, 916-920.

Gfeller, K., Olszewski, C., Rychener, M., et al. (2005). Recognition of "real-world" musical excerpts by cochlear implant recipients and normal-hearing adults. *Ear and Hearing*, *26*, 237-250.

Gfeller, K., Turner, C., Mehr, M., et al. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants International*, *3*, 29-53.

Gfeller, K., Turner, C., Oleson, J., et al. (2007). Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear and Hearing*, *28*, 412-423.

Gfeller, K., Woodworth, G., Robin, D.A., et al. (1997). Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear and Hearing*, *18*, 252-260.

Gifford, R.H., Shallop, J.K. & Peterson, A.M. (2008). Speech recognition materials and ceiling effects: Considerations for cochlear implant programs. *Audiology and Neurotology*, *13*, 193-205.

Gray, P.M., Krause, B., Atema, J., et al. (2001). The music of nature and the nature of music. *Science*, *291*, 52-54.

Griffiths, T.D. (2003). The neural processing of complex sounds. In I. Peretz & R. J. Zatorre (Eds). *The cognitive neuroscience of music* (pp. 168-177). Oxford: Oxford University Press.

Griffiths, T.D. & Warren, J.D. (2002). The planum temporale as a computational hub. *Trends in Neurosciences*, *25*, 348-353.

Grose, J.H. & Buss, E. (2007). Within- and across-channel gap detection in cochlear implant listeners. *Journal of the Acoustical Society of America*, *122*, 3651-3658.

Hartmann, W.M. (1998). *Signals, sound, and sensation*. New York: Springer Science+Business Media.

Horacek, L. & Lefkoff, G. (1970). *Programed ear training*. New York: Harcourt Brace Jovanovich.

Hoth, S. (2007). Indication for the need of flexible and frequency specific mapping functions in cochlear implant speech processors. *European Archives of Oto-Rhino-Laryngology*, *264*, 129-138.

Hughes, M.L. & Abbas, P.J. (2006). Electrophysiologic channel interaction, electrode pitch ranking, and behavioral threshold in straight versus perimodiolar cochlear implant electrode arrays. *Journal of the Acoustical Society of America*, *119*, 1538-1547.

Hyde, K.L. & Peretz, I. (2004). Brains that are out of tune but in time. *Psychological Science*, *15*, 356-360.

Jones, M.R. (1993). Dynamics of musical patterns: How do melody and rhythm fit together? In T. A. Tighe & W. J. Dowling (Eds). *Psychology and music: The understanding of melody and rhythm* (pp. 67-92). Hillsdale: Lawrence Erlbaum Associates.

Kaas, J.H., Hackett, T.A. & Tramo, M.J. (1999). Auditory processing in primate cerebral cortex. *Current Opinion in Neurobiology*, *9*, 164-170.

Kang, R., Nimmons, G.L., Drennan, W.R., et al. (2009). Development and validation of the University of Washington Clinical Assessment of Music Perception test. *Ear and Hearing*, *30*, 411-418.

Kidd, G., Boltz, M. & Jones, M.R. (1984). Some effects of rhythmic context on melody recognition. *American Journal of Psychology*, *97*, 153-173.

Koelsch, S., Fritz, T., Schulze, K., et al. (2005). Adults and children processing music: an fMRI study. *NeuroImage*, *25*, 1068-1076.

Koelsch, S., Gunter, T., Friederici, A.D., et al. (2000). Brain indices of music processing: "Nonmusicians" are musical. *Journal of Cognitive Neuroscience*, *12*, 520-541.

Koelsch, S., Gunter, T.C., Van Cramon, Y., et al. (2002). Bach speaks: a cortical "language-network" serves the processing of music. *NeuroImage*, *17*, 956-966.

Koelsch, S. & Siebel, W.A. (2005). Towards a neural basis of music perception. *Trends in Cognitive Sciences*, *9*, 578-584.

Koelsch, S., Wittfoth, M., Wolf, A., et al. (2004). Music perception in cochlear implant users: An event-related potential study. *Clinical Neurophysiology*, *115*, 966-972.

Kong, Y.Y., Cruz, R., Jones, J.A., et al. (2004). Music perception with temporal cues in acoustic and electric hearing. *Ear and Hearing*, *25*, 173-185.

Krumhansl, C.L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

Krumhansl, C.L. (1979). The psychological representation of musical pitch in tonal context. *Cognitive Psychology*, *11*, 346-374.

Krumhansl, C.L. (2000). Rhythm and pitch in music cognition. *Psychological Bulletin*, *126*, 159-179.

Kwon, B.J. & Van den Honert, C. (2006). Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users. *Journal of the Acoustical Society of America*, *120*, EL1-EL6.

Laneau, J., Wouters, J. & Moonen, M. (2006). Improved music perception with explicit pitch coding in cochlear implants. *Audiology and Neurotology*, *11*, 38-52.

Leal, M.C., Shin, Y.J., Laborde, M.L., et al. (2003). Music perception in adult cochlear implant recipients. *Acta Oto-Laryngologica*, *123*, 826-835.

Lebrun-Guillaud, G. & Tillman, B. (2007). Influence of a tone's tonal function on temporal change detection. *Perception and Psychophysics*, *69*, 1450-1459.

Levitin, D.J. & Menon, V. (2003). Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *NeuroImage*, *20*, 2142-2152.

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*, 467-477.

Liégeois-Chauvel, C., Peretz, I., Babaï, M., et al. (1998). Contribution of different cortical areas in the temporal lobes to music processing. *Brain*, *121*, 1853-1867.

Lim, J.S., Rabinowitz, W.M., Braida, L.D., et al. (1977). Intensity perception VIII. Loudness comparisons between different types of stimuli. *Journal of the Acoustical Society of America*, *62*, 1256-1267.

Limb, C.J. (2006b). Structural and functional neural correlates of music perception. *Anatomical Record - Part A Discoveries in Molecular, Cellular, and Evolutionary Biology*, *288*, 435-446.

Limb, C.J. (2006a). Cochlear implant-mediated perception of music. *Current Opinion in Otolaryngology Head and Neck Surgery*, *14*, 337-340.

Limb, C.J., Molloy, A.T., Jiradejvong, P., et al. (2010). Auditory cortical activity during cochlear implant-mediated perception of spoken language, melody, and rhythm. *Journal of the Association for Research in Otolaryngology*, *11*, 133-143.

Loizou, P.C. (1999). Introduction to cochlear implants. *IEEE Engineering in Medicine and Biology Magazine*, *18*, 32-42.

Looi, V., McDermott, H.J., McKay, C., et al. (2008). Music perception of cochlear implant users compared with that of hearing aid users. *Ear and Hearing*, *29*, 421-434.

Looi, V., McDermott, H.J., McKay, C., et al. (2004). Pitch discrimination and melody recognition by cochlear implant users. *Proceedings of the VIII International Cochlear Implant Conference*, 197-200.

Marmel, F., Tillman, B. & Dowling, W.J. (2008). Tonal expectations influence pitch perception. *Perception and Psychophysics*, *70*, 841-852.

McDermott, H. & Varsavsky, A. (2009). Better fitting of cochlear implants: modeling loudness for acoustic and electric stimuli. *Journal of Neural Engineering*, *6*, 1-8.

McDermott, H.J. (2004). Music perception with cochlear implants: A review. *Trends in Amplification*, *8*, 49-82.

McDermott, H.J., Looi, V. (2004). Perception of complex signals, including musical sounds, with cochlear implants. *International Congress Series*, *1273*, 201-204.

McDermott, H.J. & McKay, C.M. (1994). Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea. *Journal of the Acoustical Society of America*, *96*, 155-162.

McDermott, H.J. & McKay, C.M. (1997). Musical pitch perception with electrical stimulation of the cochlea. *Journal of the Acoustical Society of America*, *101*, 1622-1631.

McDermott, H.J., Sucher, C.M. (2007). Speech processing for cochlear implants and bimodal stimulation. *8th EFAS Congress/10th Congress of the German Society of Audiology*, 1-4.

McKay, C.M., McDermott, H.J. & Clark, G.M. (1996). The perceptual dimensions of single-electrode and nonsimultaneous dual-electrode stimuli in cochlear implantees. *Journal of the Acoustical Society of America*, *99*, 1079-1090.

Meyer, L.B. (1956). *Emotion and meaning in music*. Chicago: University of Chicago Press.

Middlebrooks, J.C., Bierer, J.A. & Snyder, R.L. (2005). Cochlear implants: the view from the brain. *Current Opinion in Neurobiology*, *15*, 488-493.

Milczynski, M., Wouters, J. & Van Wieringen, A. (2009). Improved fundamental frequency coding in cochlear implant signal processing. *Journal of the Acoustical Society of America*, *125*, 2260-2271.

Mirza, S., Douglas, S.A., Lindsey, P., et al. (2003). Appreciation of music in adult patients with cochlear implants: A patient questionnaire. *Cochlear Implants International*, *4*, 85-95.

Moore, B.C.J. (2003). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology and Neurotology*, *24*, 243-254.

Moore, B.C.J. & Glasberg, B.R. (1988). Gap detection with sinusoids and noise in normal, impaired and electrically stimulated hearing. *Journal of the Acoustical Society of America*, *83*, 1093-1101.

Moore, D.R. & Shannon, R.V. (2009). Beyond cochlear implants: awakening the deafened brain. *Nature Neuroscience*, *12*, 686-691.

Nicholson, K.G., Baum, S., Kilgour, A., et al. (2003). Impaired processing of prosodic and musical patterns after right hemisphere damage. *Brain and Cognition*, *52*, 382-389.

Nittono, H., Bito, T., Hayashi, M., et al. (2000). Event-related potentials elicited by wrong terminal notes: effetcs of temporal disruption. *Biological Psychology*, *52*, 1-16.

Oxenham, A.J., Bernstein, J.G.W. & Penagos, H. (2004). Correct tonotopic representation is necessary for complex pitch perception. *Proceedings of the National Academy of Sciences of the United States of America*, *101*, 1421-1425.

Pandya, D.N. (1995). Anatomy of the auditory cortex. *Revue Neurologique, 151*, 486-494.

Patel, A.D. (2003). Language, music, syntax and the brain. *Nature Neuroscience*, *6*, 674-681.

Patterson, R.D., Uppenkamp, S., Johnsrude, I.S., et al. (2002). The processing of temporal pitch and melody information in auditory cortex. *Neuron*, *36*, 767-776.

Peretz, I. (2002). Brain specialization for music. *Neuroscientist*, *8,* 374-382.

Peretz, I. (1990). Processing of local and global musical information by unilateral brain-damaged patients. *Brain*, *113*, 1185-1205.

Peretz, I. & Babaï, M. (1992). The role of contour and intervals in the recognition of melody parts: Evidence from cerebral asymmetries in musicians. *Neuropsychologia*, *30*, 277-292.

Peretz, I., Brattico, E., Järvenpää, M., et al. (2009). The amusic brain: in tune, out of key and unaware. *Brain*, *132*, 1277-1286.

Peretz, I., Champod, A.S. & Hyde, K. (2003). Varieties of musical disorders: The Montreal Battery of Evaluation of Amusia. *Annals of the New York Academy of Sciences*, *999*, 58-75.

Peretz, I. & Coltheart, M. (2003). Modularity of music processing. *Nature Neuroscience*, *6*, 688-691.

Peretz, I. & Hyde, K.L. (2003). What is specific to music processing? Insights from congenital amusia. *Trends in Cognitive Sciences*, *7*, 362-367.

Peretz, I. & Kolinsky, R. (1993). Boundaries of separability between melody and rhythm in music discrimination: a neuropsychological perspective. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *46*, 301-325.

Peretz, I., Kolinsky, R., Tramo, M., et al. (1994). Functional dissociations following bilateral lesions of auditory cortex. *Brain*, *117*, 1283-1301.

Peretz, I. & Zatorre, R.J. (2005). Brain organization for music processing. *Annual Review of Psychology*, *56*, 89-114.

Peretz, I. & Zatorre, R.J. (2003). *The cognitive neuroscience of music*. Oxford: Oxford University Press.

Pijl, S. (1997). Labeling of musical interval size by cochlear implant patients and normally hearing subjects. *Ear and Hearing*, *18*, 364-372.

Pijl, S. & Schwarz, D.W.F. (1995a). Intonation of musical intervals by musical intervals by deaf subjects stimulated with single bipolar cochlear implant electrodes. *Hearing Research*, *89*, 203-211.

Pijl, S. & Schwarz, D.W.F. (1995b). Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes. *Journal of the Acoustical Society of America*, *98*, 886-895.

Platel, H., Price, C., Baron, J.-C., et al. (1997). The structural components of music perception: A functional anatomical study. *Brain*, *120*, 229-243.

Pretorius, L.L. & Hanekom, J.J. (2011). Perception of some melodic characteristics by cochlear implant users: an exploratory study. *(In review)*

Pretorius, L.L. & Hanekom, J.J. (2008). Free field frequency discrimination abilities of cochlear implant users. *Hearing Research*, *244*, 77-84.

Pretorius, L.L., Hanekom, J.J. & Venter, P.J. (2010). Rhythm perception by cochlear implantees in conditions of varying pitch. *(In review)*

Purwins, H., Herrera, P., Grachten, M., et al. (2008). Computational models of music perception and cognition I: The perceptual and cognitive processing chain. *Physics of Life Reviews*, *5*, 151-168.

Reed, C.M. & Delhorne, L.A. (2005). Reception of environmental sounds through cochlear implants. *Ear and Hearing*, *26*, 48-61.

Rubinstein, J.T. (2004). How cochlear implants encode speech. *Current Opinion in Otolaryngology Head and Neck Surgery*, *12*, 444-448.

Samson, S., Ehrlé, N. & Baulac, M. (2001). Cerebral substrates for musical temporal process. *Annals of the New York Academy of Sciences*, *930*, 166-178.

Schonwiesner, M., Rübsamen, R. & Von Cramon, D.Y. (2005). Hemispheric asymmetry for spectral and temporal processing in the human antero-lateral auditory belt cortex. *European Journal of Neuroscience*, *22*, 1521-1528.

Schulkind, M.D. (2004). Conceptual and perceptual information both influence melody identification. *Memory and Cognition*, *32*, 841-851.

Schuppert, M., Münte, T.F., Wieringa, B.M., et al. (2000). Receptive amusia: Evidence for cross-hemispheric neural networks underlying music processing strategies. *Brain*, *123*, 546-559.

Scott, S.K. & Johnsrude, I.S. (2003). The neuroanatomical and functional organization of speech perception. *Trends in Neurosciences*, *26*, 100-107.

Semple, M.N. & Scott, B.H. (2003). Cortical mechanisms in hearing. *Current Opinion in Neurobiology*, *13*, 167-173.

Shannon, R.V. (1989). Detection of gaps in sinusoids and pulse trains by patients with cochlear implants. *Journal of the Acoustical Society of America*, *85*, 2587-2592.

Shannon, R.V. (2005). Speech and music have different requirements for spectral resolution. *International Review of Neurobiology*, *70*, 121-134.

Shannon, R.V., Fu, Q.J., Galvin, J., et al. (2004). Speech perception with cochlear implants. In F.-G. Zeng, A. N. Popper, & R. R. Fay (Eds). *Cochlear implants: auditory prostheses and electric hearing* (pp. 334-376). New York: Springer.

Shannon, R.V., Zeng, F.-G., Kamath, V., et al. (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 and Hearing*, *30*, 160-168.

Smith, Z.M., Delgutte, B. & Oxenham, A.J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*, 87-90.

Stevens, S.S. (1955). The measurement of loudness. *Journal of the Acoustical Society of America*, *27*, 815-829.

Stevens, S.S. (1935). The relation of pitch to intensity. *Journal of the Acoustical Society of America*, *6*, 150-154.

Stewart, L., Overath, T., Warren, J.D., et al. (2008). fMRI evidence for a cortical hierarchy of pitch pattern processing. *PLoS ONE*, *3*, e1470-e1475.

Sucher, C.M. & McDermott, H.J. (2007). Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. *Hearing Research*, *230*, 80-87.

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

Throckmorton, C.S. & Collins, L.M. (2001). A comparison of two loudness balancing tasks in cochlear implant subjects using bipolar stimulation. *Ear and Hearing*, *22*, 439-448.

Tillman, B., Bharucha, J.J. & Bigand, E. (2000). Implicit learning of tonality: a self-organizing approach. *Psychological Review*, *107*, 885-913.

Townshend, B., Cotter, N., Van Compernolle, D., et al. (1987). Pitch perception by cochlear implant subjects. *Journal of the Acoustical Society of America*, *82*, 106-115.

Trainor, L.J., McDonald, K.L. & Alain, C. (2002). Automatic and controlled processing of melodic contour and interval information measured by electrical brain activity. *Journal of Cognitive Neuroscience*, *14*, 430-442.

Trainor, L.J. & Trehub, S.E. (1994). Key membership and implied harmony in Western tonal music: developmental perspectives. *Perception and Psychophysics*, *56*, 125-132.

University of South Africa (1970). *Aural tests: For the guidances of teachers and candidates preparing for the examinations of the University of South Africa* (Primarynd ed.). Pretoria: University of South Africa.

Van Zuilenberg, P.L. (1996). *Practical musicianship: Pre-grade 1- Grade 3*. Pretoria: University of South Africa.

Vongpaisal, T., Trehub, S.E. & Schellenberg, E.G. (2009). Identification of TV tunes by children with cochlear implants. *Music Perception*, *27*, 17-24.

Vos, P.G. & Troost, J.M. (1989). Ascending and descending melodic intervals: statistical findings and their perceptual relevance. *Music Perception*, *6*, 383-396.

Warren, J. (2008). How does the brain process music? *Clinical Medicine, 8,* 32-36.

Wessinger, C.M., Vanmeter, J., Tian, B., et al. (2001). Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. *Journal of Cognitive Neuroscience*, *13*, 1-7.

Wier, C.C., Jesteadt, W. & Green, D.M. (1977). Frequency discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, *61*, 178-184.

Wilson, B.S. & Dorman, M.F. (2007). The surprising performance of present-day cochlear implants. *IEEE Transactions on Biomedical Engineering*, *54*, 969-972.

Wilson, B.S. & Dorman, M.F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research, 242*, 3-21.

Woods, D.L. & Alain, C. (2009). Functional imaging of human auditory cortex. *Current Opinion in Otolaryngology Head and Neck Surgery*, *17*, 407-411.

Zatorre, R. (2003). Neural specializations for tonal processing. In I. Peretz & R. Zatorre (Eds). *The cognitive neuroscience of music* (pp. 231-246). Oxford: Oxford University Press.

Zatorre, R.J. (2001). Neural specializations for tonal processing. *Annals of the New York Academy of Sciences*, *930*, 193-210.

Zatorre, R.J. & Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cerebral Cortex*, *11*, 946-953.

Zatorre, R.J., Belin, P. & Penhune, V.B. (2002). Structure and function of auditory cortex: Music and speech. *Trends in Cognitive Sciences*, *6*, 37-46.

Zatorre, R.J., Evans, A.C. & Meyer, E. (1994). Neural mechanisms underlying melodic perception and memory for pitch. *Journal of Neuroscience*, *14*, 1908-1919.

Zeng, F.-G. (2004). Trends in cochlear implants. *Trends in Amplification*, *8*, 1-34.

Zeng, F.-G. (2002). Temporal pitch in electric hearing. *Hearing Research*, *174*, 101-106.

Zwicker, E. & Fastl, H. (1999). *Psychoacoustics: Facts and models* (2nd ed.). Berlin: Springer.

Appendix A

COCHLEAR IMPLANT SAMPLE SELECTION

A.1 CONTEXT OF RECRUITMENT

The context in which the sample of CI participants was recruited and selected for participation deserves mention. The Pretoria Cochlear Implant Programme is one of only five academic CI centres in South Africa. Of these, the Tygerberg Cochlear Implant Unit and the Pretoria Cochlear Implant Programme are the oldest, with the other three having been established only after 2009. At the time of data collection for the current study (2005–2009) only a limited group of CI recipients were thus available for recruitment to participate in research studies, a constraint exacerbated by the decision to include only post-lingually deafened adult CI recipients in this study. Furthermore, most of the CI research conducted at these academic implant centres pertain to clinical application and rehabilitation; the research group at which the study reported in this thesis was conducted is the only in the country that focuses on the technology that supports CI hearing.

The benefit cochlear implantation affords profoundly deaf individuals paradoxically also complicates recruitment of participants for research that involves substantial time commitment from them. The improved quality of life afforded by cochlear implantation allows many implantees to stand in demanding jobs and lead busy social and family lives.

This greatly impacts on their time availability and although the majority have expressed willingness to contribute to CI research, the time commitment required from them in a study of the nature reported in this thesis, is impractical.

The nature of sound field studies further impacts participant availability, since experimental conditions as controlled as possible need to be created. Participants thus have to travel to the research facility, whereas electrode-level studies are less prone to such logistic constraints. Furthermore, even when participants are committed to participating in research and set time aside for experiments after or before work hours and over weekends (as several have done over the course of this study), the investigator should strive for study designs that are practical for both parties and which do not become too drawn out so as to minimise the risk of non-completion. Aligning all parties’ schedules becomes challenging when one considers these logistical constraints.

New CI centres established recently (as mentioned earlier) offers the opportunity for access to a larger potential local research population. Also, the insights and experience regarding relevant experimental design that have emerged from this study, creates an opportunity for improved intercentre collaboration, which would not only facilitate more authoritative studies to be conducted but also establish a stronger CI research community in South Africa.

A.2 PARTICIPANTS' DETAILS

Relevant demographic details of CI participants are provided in Table A1. Please note that, owing to listeners' time availability and the study having spanned several years, the same group of listeners did not participate in all investigations reported in the respective chapters. Each listener's participation in specific investigations is specified in Table A1.

Functional hearing ability of all CI participants, as rated by the treating audiologist at the local implant centre, is shown in Table A2. The rating instrument is based on the Abbreviated Profile of Hearing Aid Benefit questionnaire (Cox & Alexander 1995), but has been adapted slightly to allow for assessment by the audiologist. Assessment is done according to a 7-point Likert-type scale whereby the frequency of a listener's hearing success in various conditions is rated. The scale ranges from 1 (always) to 7 (never). The assessment regarding general perceptual success experienced with the CI device is a composite value based on the ratings for the preceding categories.

Table A1: Demographic particulars of participants. Some participants received new processor or map settings between investigation and hence have double entries. Asterisks indicate bilaterally implanted subjects.
F = female; M = male; YoB = year of birth; FAT = frequency allocation table (only applicable to investigation reported in Chapter 2); YoI = year of implantation

Subject (Gender)	YoB	Chapter	Implant	Processor	Strategy	FAT	YoI	Test ear
S3 (F)	1949	2	24R (CA)	Esprit 3G	ACE (500 Hz)	7	2004	Right*
		3, 4, 5	24R (CA)	Freedom	ACE (500 Hz)	NA	2004	Right*
S4 (M)	1970	5	CI24M	Freedom	ACE (900 Hz)	NA	2000	Left
S5 (F)	1967	3, 5	24M	Esprit 3G	Speak (250 Hz)	NA	1999	Right
		4	24M	Freedom	Speak (250 Hz)	NA	1999	Right
S8 (M)	1950	4,5	Nucleus 22	Freedom	Speak (250 Hz)	NA	1995	Right*
S10 (F)	1953	2	24M	Esprit 3G	ACE (900 Hz)	6	2000	Right
		4, 5	24M	Freedom	ACE (900 Hz)	NA	2000	Right
S11 (F)	1944	2	24 Freedom (CA)	Freedom	ACE (720 Hz)	22	2005	Left*
		3	24M	Freedom	ACE (900 Hz)	NA	1999	Right*
S12 (M)	1984	3, 4 & 5	Freedom	Freedom	ACE (1200 Hz)	NA	2006	Right
S13 (F)	1950	2	24R (CA)	Freedom	ACE (900 Hz)	20	2004	Left*
S14 (M)	1984	2, 5	24R (CS)	Esprit 3G	ACE (900 Hz)	7	2004	Right
S15 (F)	1988	3	24R (CA)	Freedom	ACE (900 Hz)	NA	1992	Left*
S16 (F)	1988	3	Not available	Freedom	ACE (Not available)	NA	1996	Right*
S17 (F)	1949	3	24R (CA)	Sprint	ACE (900 Hz)	NA	2005	Right

S18 (M)	1943	4, 5	Freedom (CA)	Freedom	ACE (1200 Hz	NA	2003	Right*
S21 (F)	1970	4	Freedom (CA)	Freedom	ACE (1200 Hz)	NA	2007	Left

Table A2: Rating of functional hearing ability of CI participants

	Participant						
Description of perceptual ability[a]	**S3**	**S4**	**S5**	**S8**	**S10**	**S11**	**S12**
Able to follow a conversation with familiar speaker, in quiet listening environment	2	1	3	2	4	2	2
Able to follow a conversation in quiet listening environment, even when speaker is not familiar	3	2	4	2	4	2	2
Able to follow a conversation/speech in room with substantial echoes	4	3	4	3	5	3	2
Able to follow a whisper conversation/soft speech	3	3	3	2	4	2	2
Able to follow a conversation/speech without visual cues (e.g. radio talk show)	3	3	3	2	5	3	3
Able to use a telephone successfully	5	5	5	3	6	3	3
Able to follow a conversation amidst other speech noise (e.g. at a party, in a restaurant)	4	5	5	3	6	3	3
Able to follow a conversation in an environment with substantial background noise (which is not speech)	4	4	5	3	6	3	3
Find environmental sounds (e.g. rain, wind, a passing aeroplane, thunder, construction noise, washing machine, etc.) disturbing	2	3	5	5	3	4	5
General CI-mediated perceptual success[b]	3	2	3	1	4	2	2

[a] The scale used for rating perceptual ability items is as follows:
1 = always; 2 = almost always; 3 = usually; 4 = half the time; 5 = occasionally; 6 = seldom; 7 = never.

[b] The scale for rating general perceptual success with the CI device is:
1 = excellent; 2 = very good; 3 = good; 4 = average; 5 = below average; 6 = poor; 7 = very poor

Table A2 (continued): Rating of functional hearing ability of CI participants

	Participant						
Description of perceptual ability[a]	**S13**	**S14**	**S15**	**S16**	**S17**	**S18**	**S21**
Able to follow a conversation with familiar speaker, in quiet listening environment	1	1	2	4	1	2	2
Able to follow a conversation in quiet listening environment, even when speaker is not familiar	2	2	2	4	2	2	2
Able to follow a conversation/speech in room with substantial echoes	3	3	2	5	2	3	2
Able to follow a whisper conversation/soft speech	2	2	2	4	2	2	2
Able to follow a conversation/speech without visual cues (e.g. radio talk show)	2	2	3	5	3	3	2
Able to use a telephone successfully	2	2	3	6	5	3	2
Able to follow a conversation amidst other speech noise (e.g. at a party, in a restaurant)	3	3	2	5	5	3	3
Able to follow a conversation in an environment with substantial background noise (which is not speech)	3	3	2	5	5	3	3
Find environmental sounds (e.g. rain, wind, a passing aeroplane, thunder, construction noise, washing machine, etc.) disturbing	4	5	5	4	3	5	5
General CI-mediated perceptual success[b]	1	1	1	4	4	2	1

[a] The scale used for rating perceptual ability items is as follows:
1 = always; 2 = almost always; 3 = usually; 4 = half the time; 5 = occasionally; 6 = seldom; 7 = never.

[b] The scale for rating general perceptual success with the CI device is:
1 = excellent; 2 = very good; 3 = good; 4 = average; 5 = below average; 6 = poor; 7 = very poor