Exploring Place Pitch & Temporal Pitch Perception with Cochlear Implants

Vijay Mohan Raaja Marimuthu

Audiology Section, Department of Linguistics
Faculty of Human Sciences
Macquarie University
Sydney, Australia

This thesis is submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD)

September 2012

Dedicated to my parents

&

my wife Anagha

Table of Contents

Declaration

I certify that this thesis entitled "Exploring Place Pitch and Temporal Pitch Perception with Cochlear Implants" is an original work and has not been submitted to any other universities and organizations for any degrees other than Macquarie University.

I also certify that the thesis has been written by me. Any help and assistance that I have received in my research work have been appropriately acknowledged.

The research presented in this thesis was approved by the Human Research Ethics Committee (HREC) - Macquarie University (Reference No: HE31JUL2009-D00040). Additional approval was obtained from the Sydney South West Area Health Service (SSWAHS) (Protocol No: X10-0281 & HREC/10/RPAH/497).

Signed:

Vijay Mohan Raaja Marimuthu

(Student No: 41567536)

Date:

Abstract

Cochlear Implant (CI) recipients are on par with their normal hearing counterparts in speech perception tasks in quiet, but their performance is at chance level for pitch perception tasks. Pitch in CI can be conveyed by either place cues or temporal cues. Although researchers have studied place and temporal pitch, there is scant evidence that reports the individual contribution of place and temporal pitch as a function of musical pitch. This thesis aims to investigate the role of place pitch and temporal pitch perception in CI recipients. It comprises three major experiments using four experimental procedures. The four experimental procedures were: 4AFC Discrimination, 2AFC Ranking, and 2AFC Modified Melodies Test – Backward, and 2AFC Modified Melodies Test – Warp modification.

Experiment I investigated temporal pitch sensitivity as a function of different stimulation patterns, base pulse rate, and electrode location. The four stimulation patterns were: single electrode stimulation (apical (Electrode E22) or middle (E12)), dual-electrode stimulation (E22 & E12), and multiple electrode stimulation (E22 to E12). Six post-lingually deafened CI subjects were tested using three of the four experimental procedures (i.e., excluding the discrimination task). The stimuli were presented as pulses at a base rate of 131 pulses per second (pps) (musical note C3 range) and 262 pps (C4 range). The temporal pitch sensitivity was not influenced by different stimulation patterns and the results showed no significant difference among the stimulation patterns across the three procedures. The results suggest that the CI recipients are unable to combine cues from different places in the cochlea to give a “stronger” cue.

Experiment II aimed to investigate the individual contribution of place pitch and temporal pitch in various pitch perception tasks. The performance of CI subjects was

tested using four experimental procedures and three stimulus types. The three stimulus types were: (1) Pure tones with base frequency of C5 (523 Hz), providing place cues only; (2) Harmonic tones with base frequency of C3 (131 Hz), providing temporal cues only; (3) Harmonic tones with base frequency of C4 (262 Hz), providing both place and temporal cues. The stimuli were presented via loudspeaker at a comfortable loudness level. The recipients used their own speech processor. The overall scores for discrimination and ranking were high for all three stimulus types, however, three subjects showed pitch reversals in ranking the C4 harmonic tones. In the Modified Melodies test, scores were similar for C5 pure tones and C3 harmonic tones, while scores using C4 harmonic tones were worse and mostly near chance. These results suggest that CI place pitch may convey melodic pitch information, but the contribution of brightness cannot be completely ruled out.

Experiment III sought to investigate the role of brightness in various pitch perception tasks in normal hearing individuals. The goal of this study was to investigate whether CI subjects in the previous experiment perceived place pitch as pitch rather than as a pattern of brightness changes. Eighteen normal-hearing adults participated in four experimental procedures using three stimulus types: brightness sequences (harmonic tones varying in brightness, with constant pitch), and noise sequences (Low-pass noise bands, varying in cut-off frequency) were compared with pitch sequences (harmonic tones varying in pitch, with constant brightness). Results showed that the subjects were able to discriminate and rank brightness, and were able to detect brightness contour changes, but were unable to make judgements of musical intervals for brightness. These results suggest that the cochlear implant recipients in the previous experiment may have perceived place cues as brightness rather than pitch.

In conclusion, a similar performance trend was seen between CI place pitch and NH brightness indicating that CI place pitch is akin to brightness and not to pitch. Conversely, CI performance was similar between temporal pitch and place pitch suggesting that CI place pitch can convey melodic pitch. However, the overall performance reveals that CI place pitch is more akin to brightness aspect of timbre than to pitch and additional studies need to affirm these findings. Unfortunately, the CI pitch performance was still significantly below NH performance.

Acknowledgements

I would like to acknowledge many people who have helped me during the course of this thesis. First and foremost, I would like to sincerely thank my primary supervisor, Dr. Robert Mannell, Department of Linguistics, Macquarie University for his constant support and motivation. His advice and support has helped me sail through the rough phases of my PhD. I would like to thank my adjunct supervisor, Dr. Brett Swanson for providing the Modified Melodies Test and Nucleus MATLAB Toolbox. His constant support, critical feedbacks, and assistance in programming have helped me extensively throughout my candidature. I also would like to thank my associate supervisor, Prof. William Forde Thompson for his valuable feedback and discussion.

I am grateful to Dr. Mridula Sharma, A/Prof. Catherine McMahon, Leanne Skinner, and Colleen Psarros for their help in accessing the Cochlear Implant Recipients. I owe a special note of gratitude to Andrew Vandali, John Galvin, Dr. Jeremy Marozeau, A/Prof. Bob Cowan, Dr. Vijay Narne, and Saji Maruthurkkara for the assistance, generosity, and advice they provided during my study. Special thanks to John Newell for helping me setup the lab for testing Cochlear Implant recipients (Chapter 6). I would sincerely thank from the bottom of my heart the Cochlear Implant recipients who have been enthusiastic participants throughout this study.

I would like to thank the administrative staffs at Macquarie University especially Sofia for her assistance during the course of my study. I would also like to thank HEARing CRC Manager at Macquarie University, Mo for her help, support, and proofreading my manuscripts. I acknowledge the financial support of the HEARing CRC. The Cooperative Research Centres (CRC) Program is an Australian Government Initiative.

I would like to thanks my colleagues Varghese, Yatin, Ankit, Balaji, Isabelle, Pia, Imran, Shannon, and Sara Cotterell for supporting me during the past three and half years of my study. I would also like to thank my beloved Mom and Dad, who are my everything and who were with me through success & failure. Without their encouragement & support, it was impossible for me to finish my studies. Special thanks to my brother for understanding me during the rough phase of my life. Last but not the least, my beloved wife Anagha, who stood beside me during the entire course of my study. Without her constant motivation and encouragement this thesis would not have been possible.

List of acronyms and abbreviations

2-AFC	Two-Alternative Forced Choice
4-AFC	Four-Alternative Forced Choice
ACE	Advanced Combination Encoder
AM	Amplitude Modulation
ANOVA	Analysis of Variance
AOS	Advance Off-Stylet Insertion
BP	Bipolar
CAMP	Clinical Assessment of Music Perception
CI	Cochlear Implant
CIS	Continuous Interleaved Sampling
CL	Current Levels
CSSS	Channel-Specific Sampling Sequences
DL	Difference Limen
DR	Dynamic Range
EAS	Electro-Acoustic Stimulation
ECAP	Evoked Compound Action Potentials
EEG	Electroencephalography
F0	Fundamental frequency
FDA	U.S Food and Drug Administration
FDL	Frequency Difference Limen
fMRI	Functional Magnetic Resonance Imaging
FSP	Fine Structure Processing
GLMM	Generalized Linear Mixed Model
IPI	Inter-Pulse Interval
JND	Just Noticeable Difference
LGF	Loudness Growth Function
MBEA	Montreal Battery of Evaluation of Amusia
MCI	Melodic Contour Identification
MDS	Multi-Dimensional Scaling
MEG	Magnetoencephalography
MP	Monopolar
MPP	Multiple Pulse Per Period

MPPH	Multiple Pulse Per Period- Half-Wave
MPPS	Multiple Pulse Per Period - Synchronized
MPPU	Multiple Pulse Per Period – Uniform-Sampling
NH	Normal Hearing
NMT	Nucleus MATLAB Toolbox
PA	Phased Array
PPS	Pulse Per Second
PSA	Pseudomonophasic pulses
pTP	Partial Tripolar
SAM	Sinusoidally Amplitude Modulation
SOE	Spread of Excitation
SPEAK	Spectral Peak
SPP	Single Pulse Per Period
STAR	Spike-Based Temporal Auditory Representation
TMTF	Temporal Modulation Transfer Function
TP	Tripolar
WHO	World Health Organisation

List of Figures

Conference Presentations

Marimuthu, V., Swanson, B., & Mannell, R. (2012). Investigating Cochlear Implant Place Pitch and Temporal Pitch Perception using different Pitch Perception Tasks. Paper presented at ARO, San Diego, California, USA.

Marimuthu, V., Swanson, B., & Mannell, R. (2011). Role of Brightness in Pitch Perception Tasks: Implication for Cochlear Implant Place Pitch. Paper presented at CIAP, Asilomar, California, USA.

Marimuthu, V., Swanson, B., Mannell, R., & Thompson, W .F. (2010). Can Brightness Changes Support Melody Perception in Normal Hearing Individuals. Paper presented at HEARing CRC Symposium, University of Melbourne, Melbourne.

Chapter 1: Introduction

Hearing Impairment is one of the major communication disorders seen across the entire age range (from young infants to older adults). According to the World Health Organisation (WHO) in 2005, about 278 million people had moderate to profound hearing loss and 80% of them live in low and middle income countries. In Australia, it is estimated that one in six people have a hearing loss and this is projected to increase to one in four people by 2050 (Listen Hear, 2006). Depending on the severity of hearing loss, people are aided either with hearing aids or cochlear implants (CI). Commercially, cochlear implants were first approved by the FDA (U.S Food and Drug Administration) in 1984. From then, there has been a remarkable increase in the number of people being implanted with this device and as of December 2010, approximately 219,000 people worldwide have received a CI.

A cochlear implant is an implantable auditory prosthesis which restores hearing in people with moderately severe to profound hearing loss or people who do not benefit from their hearing aids. Earlier CI devices did not convey adequate speech information to the CI users and as a result, they had to rely on other cues such as visual cues (facial expressions and gestures) to have an effective conversation. With the technological advancement in the past two decades, the CI devices have improved remarkably in the area of hardware design and acoustic signal processing. Current CI technology can convey adequate speech information to the recipients (Shannon et al., 1995; Dorman et al., 1997). Most of the CI recipients obtain good scores in speech perception tests in quiet, but their performance reduces greatly in speech in noise tasks or music perception tasks (Zeng, 2004;

McDermott, 2004). Unfortunately, not many CI recipients can have a successful conversation over a telephone. This accomplishment has led researchers to explore an intricate area of CI pitch perception which has gained a great deal of attention over the past few years.

The three major cochlear implant manufacturers are Cochlear Limited - Australia, Med-El - Austria, and Advanced Bionics Corporation - United States of America. A team approach is very vital for a successful cochlear implantation program. The team mainly comprises of Surgeons, Audiologists, Engineers, Therapists, Psychologists, Teachers of the deaf, and Nurses.

1.1. **Objectives of this thesis**

This thesis aims to investigate:

a) The role of cochlear implant place pitch and temporal pitch using various pitch perception tasks.
b) The role of temporal pitch as a function of electrode stimulation patterns, base pulse rate patterns, and electrode locations in CI recipients.
c) The role of pitch, brightness, and noise sequences in normal hearing individuals and its implication to CI place pitch.

1.2. **Outline of this thesis**

The thesis is divided into eight chapters and the objectives of this thesis are addressed systematically in various chapters. The brief outline of each chapter is listed below.

Chapter 1 introduces the main objectives of the thesis.

Chapter 2 describes sound processing in cochlear implant (CI) and normal hearing (NH) individuals. It also provides a brief history of the emergence of auditory prostheses and outlines the components of a CI device.

Chapter 3 provides a review of literature on pitch processing in cochlear implants. Place pitch and temporal pitch are reviewed in detail. An overview of loudness perception, timbre perception, and music perception in CI is also provided.

Chapter 4 reports the performance of normal hearing (NH) individuals on various pitch perception tasks using pitch sequences (C4 (262 Hz) harmonic tones). Four experimental procedures (Discrimination, Ranking, Modified Melodies Test (Backward Modification), and Modified Melodies Test (Warp Modification)) were used to explore place and temporal pitch perception in both normal and CI subjects. A detailed description of the four procedures is presented in this chapter. The NH performance on pitch sequences provided a baseline condition and their performance was compared with the CI subjects in the later chapters.

Chapter 5 addresses temporal pitch sensitivity of CI recipients as a function of different electrode stimulation patterns, base pulse rate, and electrode locations. The stimulation patterns were single electrode stimulation (apical electrode (E22) and middle electrode (E12)), dual electrode stimulation (E22 & E12), and multiple electrode stimulation (E22 to E12) and were explored using the base rate of C3 - 131 pulses per second (pps). Additionally, the single electrode stimulation was explored using the base rate of C4 – 262 pps. Overall there were 6 conditions which were investigated using the three experimental procedures (Ranking, Modified Melodies Test (Backward Modification), and Modified Melodies Test (Warp Modification)).

Chapter 6 explores the role of place pitch and temporal pitch perception in CI recipients using the four experimental procedures (Discrimination, Ranking, Modified Melodies Test (Backward Modification), and Modified Melodies Test (Warp Modification)). The three stimuli used in this study were C3 (131 Hz) harmonics tones, C4 (262 Hz) harmonic tones, and C5 (523 Hz) pure tones. The stimuli were presented through loudspeakers and the CI performance was reported as a function of three stimulus types.

Chapter 7 explores the performance of normal hearing subjects on various experimental procedures (Discrimination, Ranking, Modified Melodies Test (Backward Modification), and Modified Melodies Test (Warp Modification)) using brightness and noise sequences. The result obtained from brightness and noise sequences provides some additional information on the processing of place pitch in CI recipients.

Chapter 8 summarises the results of all experimental procedures across stimulus types in normal and CI subjects. This chapter concludes with some of the key findings of this thesis and proposes some ideas for future research. Finally, the pros and cons of the experimental designs are summarised.

Chapter 2: Introduction to Cochlear Implant Sound Processing

2.1.Introduction

In order to understand sound processing in CI users, it is necessary to understand sound processing in normal hearing individuals. This chapter provides an overview of sound processing mechanisms in normal hearing individuals, followed by a detailed description of the sound processing mechanisms in CI recipients.

2.2.Sound Coding in a Normal Auditory System

The sensation of hearing occurs when an acoustic signal enters the ear canal and impinges the tympanic membrane. The human auditory system is mainly divided into a peripheral system (External, Middle, and Inner Ear) and a central system (Auditory nerve, Cochlear Nucleus, Superior Olivary Complex, Lateral Leminiscus, Inferior Colliculus, Superior Colliculus, Medial Geniculate Body, and Auditory cortex).

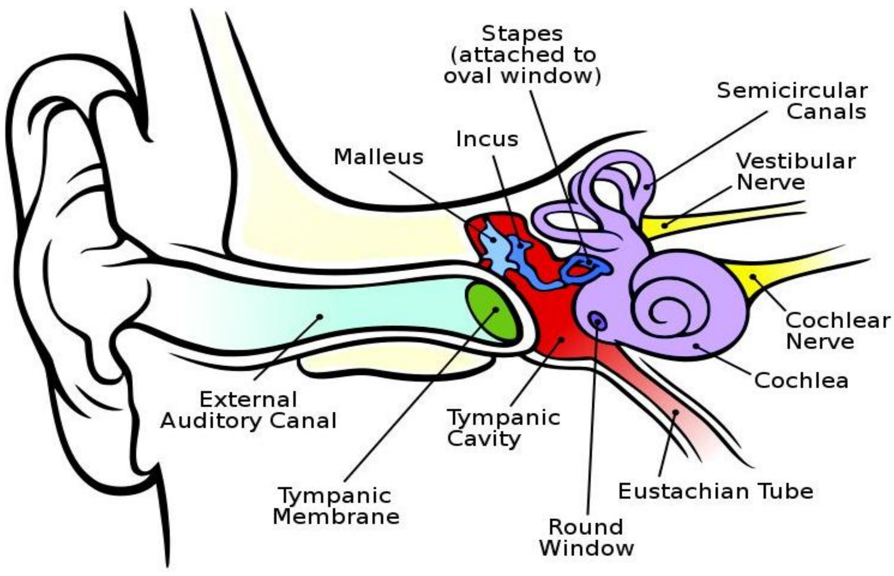


Figure 2.1: The Human Ear (Chittka & Brockmann, 2005).

The external ear consists of three main parts: the cartilaginous pinna, the resonant cavity called the concha and the external auditory meatus or ear canal leading to the tympanic membrane. The external ear plays two significant roles in the transduction of acoustic signal into the middle ear. Firstly, due to the resonance character of the meatus, there is an increase in the sound pressure at the level of tympanic membrane. Secondly, the directionality cues for identifying the direction of sound source is achieved with the help of pinna. The acoustic signal from the external ear travels through the external auditory meatus and causes the tympanic membrane to vibrate. The low-impedance vibrations from the tympanic membrane are transformed to the high-impedance oval window of the inner ear (cochlea) through the middle ear transformer function.

The middle ear uses two principles to transfer the sound to the inner ear (Figure 2.2). According to the hydraulic principle, the area of the tympanic membrane is larger than the area of the stapes footplate in the cochlea. Therefore, the force applied by the tympanic membrane is concentrated on a smaller area, thus increasing the pressure at the oval window. This process helps the sound to transmit through the fluid-filled cochlea and reduces the amount of sound being reflected from the oval window. The second principle is the lever action function of the middle ear bones. The arm of the incus is shorter than that of the malleus, and this produces a lever action that increases the force and decreases the velocity at the stapes (Pickles, 2008). Therefore, the middle ear acts as an impedance transformer and a mechanical lever.

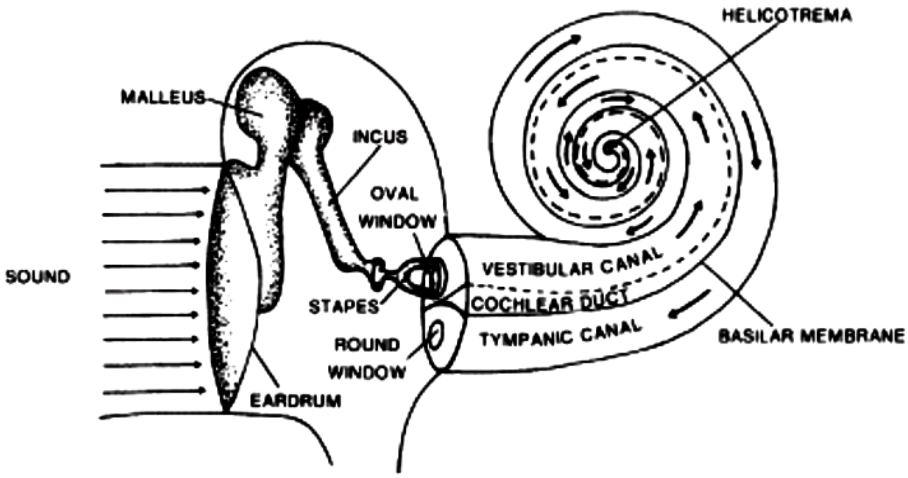


Figure 2.2: The Human Auditory System. The figure shows the sound transmission from the tympanic membrane to the inner ear (Warren, 2008).

The inner ear is divided into two systems based on their functions. The first system houses the organ of hearing (cochlea) and the second system houses the organ of equilibrium or balance (utricle, saccule, and the three semicircular canals). The entire inner ear is housed in the temporal bone, which is one of the hardest bones in the entire human body (Moller, 2006). The organ of hearing, the cochlea is a spirally shaped, fluid filled structure that has two and three quarter turns from the basal part of the cochlea. It is divided longitudinally into three scalae (scala vestibule, scala tympani, and scala media). The osseous spiral lamina divides the scala vestibuli from the scala tympani on the side near the modulus. The reissner's membrane separates scala media from scala vestibuli and the basilar membrane separates scala media from scala tympani (Figure 2.3 & Figure 2.4). The scala media narrows at the apex portion of the cochlea ending short of the bony labyrinth and creates an opening called the helicoterma (Figure 2.2). This opening allows communication between scala vestibuli and scala tympani. The inward and outward motion of the stapes into the oval window (opening of the scala vestibuli) displaces the fluid-filled cochlea and results in a corresponding displacement of the round window (opening of the scala tympani).

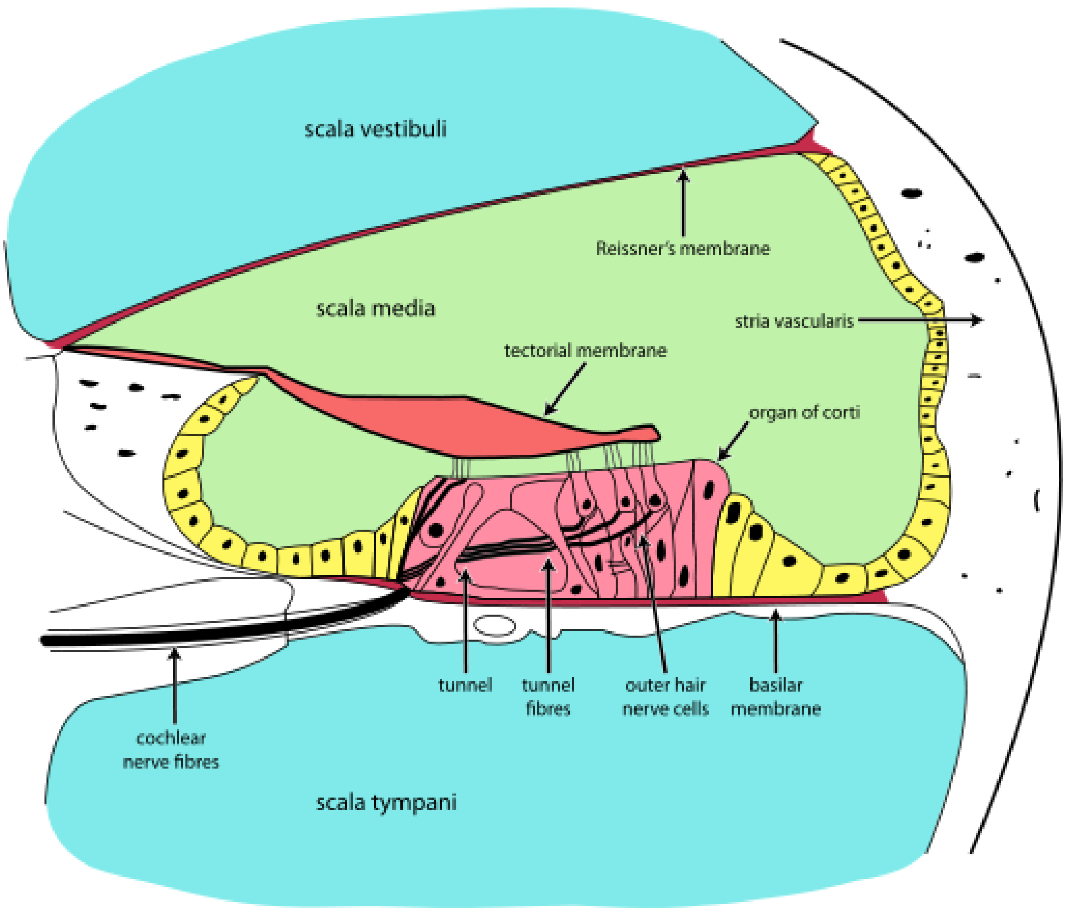


Figure 2.3: Cross section of cochlea (Ropshkow, 2010).

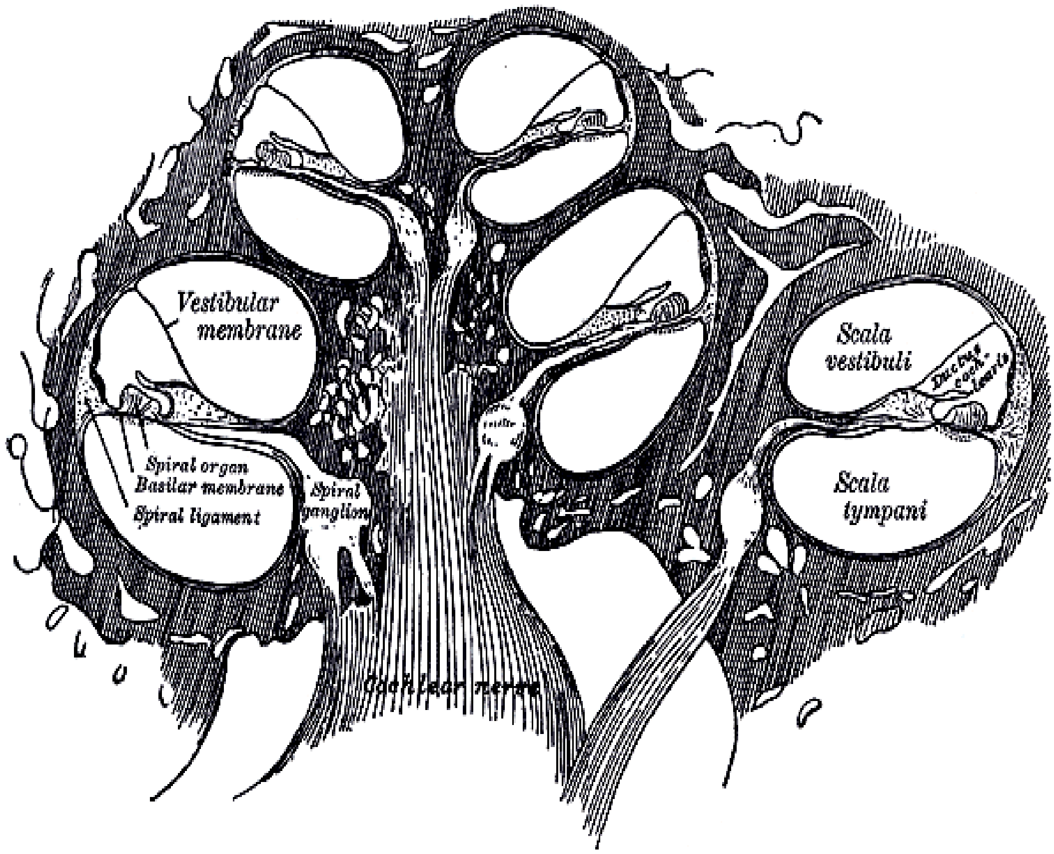


Figure 2.4: Longitudinal section of cochlea (Gray, 1918).

When the oval window is displaced, the pressure difference inside the fluid-filled cavity sets the basilar membrane in motion in the form of a travelling wave (Moore, 2003). Figure 2.5 schematizes four patterns of basilar membrane motion for sinusoids at successive instants of time. The different regions in the basilar membrane are excited by different frequency ranges. The basilar membrane resonates as a “travelling wave” that gradually grows in amplitude as it moves along the cochlear duct from the stapes (base) toward the helicotrema (apex). This mechanical property of the cochlea varies considerably from base to apex of the cochlea (Figure 2.5). Anatomically, the basilar membrane is narrow and stiff at the base, indicating that this region of the cochlea corresponds to high frequencies, while the apex part of basilar membrane is wide and less stiff, indicating that this region corresponds to low frequencies. The entire range of human hearing ranges from

20 Hz to 20,000 Hz. However, the mechanical tuning is not strong at the extremes of the range and each location along the cochlea is tuned to specific frequency and maximum excitation can be seen at this particular place and this frequency is termed as characteristic frequency.

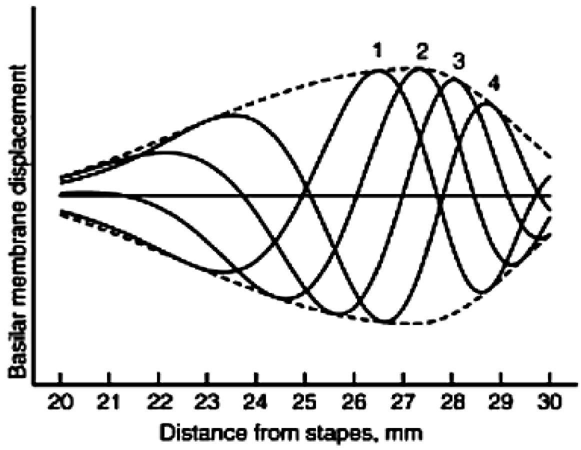


Figure 2.5: Travelling wave was first demonstrated by von Békésy. The instantaneous displacement of basilar membrane at four successive time intervals (solid lines). The x axis represents the distance from the stapes. The y axis is the basilar membrane displacement. The dash line represents the overall amplitude created by the four waveforms (Moore, 2003).

The cochlea behaves like a frequency-analysing system and cochlear mechanics can be explained in two ways: passive process and active process. The passive cochlear mechanics is explained based on the physical characteristics of the tuned component of the travelling wave. Pickles (2008) explained this theory with respect to waves on the surface of water. Once the energy is introduced on the surface of the liquid or water, it is carried passively along the wave by the inertia of fluid motion in the horizontal plane. The gravitational force acts perpendicularly in the vertical direction. This phenomenon is similar to passive cochlear wave, except that the restoring force comes from the stiffness of the cochlear partition (i.e., from the stiffness of the basilar membrane, organ of corti, and the tectorial membrane). The inertial forces are comprised of the mass of cochlear partition and the mass of the fluids. The passive travelling wave always propagates from the basal to the apical part of the cochlea and the amplitude of the wave grows as it passes down the cochlea. Once the maximum is attained, the amplitude drops sharply. The maximum excitation occurs for the characteristic frequency of the wave at a particular location inside the cochlea. This pattern of excitation from base to apex depends on the stiffness property of the cochlear partition. The variations in stiffness and mass affect the resonance properties of the basilar membrane. When the stiffness is high (low mass) it leads to high resonance frequencies at the basal part of cochlea and if the stiffness is low (high mass) it leads to low resonance frequencies at the apex part of cochlea. Therefore, acoustic signals of a certain frequency vibrate a specific location more than other locations of the basilar membrane. A study by Emadi et al. (2004) has confirmed that the cochlear partition is relatively more stiffness based near the basal region of the cochlea and more compliance based near the apical part of the cochlea. This property affects the way the sound travels along the basilar membrane. Near the base, where the stiffness is high, the vibrations are known as stiffness-limited. In contrast, at apex, the stiffness is relatively low and mass and inertia limits the vibration. This vibration is known as mass-limited. When a force is applied, the stiffness-limited system acts first followed by mass-limited system. This means that the stiffness dominated basal region responds first, followed by the mass dominated apical region of the cochlea. Thus, this results in the directionality feature of the cochlear where the wave always travels from the basal to the apical region of the cochlea and which depends on the compliance characteristics of the cochlea and not on the sound pressure.

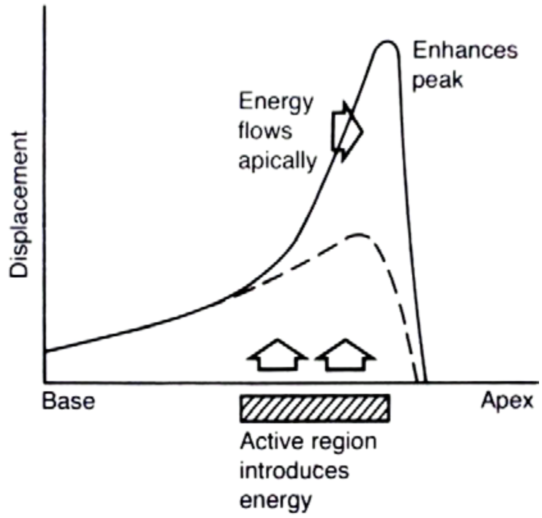


Figure 2.6: Excitation pattern possibly produced by active mechanism of the cochlea. The dotted line represents the excitation pattern of basilar membrane by passive mechanism (Moore, 2003).

The active process is comprised of two processes (Pickles, 2008). Firstly, the OHC (Outer Hair Cells) respond to basilar membrane motion by altering the micromechanical process. The lengthwise contraction and expansion of the OHC leads to intracellular depolarization and hyperpolarization. This displacement alters the travelling wave and the organ of corti, which in turn leads to feedback energy from the compound displacement resulting in an amplified travelling wave (Dallos, 1992). Secondly, the conformation of the mechanotransducer apparatus is altered by the Ca^{2+} ions which enter through the mechanotransducer channels. This leads to mechanical energy output which can feed back into the mechanical system (Kennedy et al., 2006).

Gain, tuning, and nonlinearity are the properties of active mechanism of the cochlea. In a normal ear, each location along the cochlea is sharply tuned, highly sensitive to a limited range of frequencies, and requires higher sound intensities to produce a response as the signal is moved outside the range (Moore, 2003). The basilar membrane acts as a nonlinear compressor, i.e., for very low input sound levels, below 20-30 dB SPL, the active mechanism amplifies the response (gain) by up to 50 dB or more on the basilar membrane (Figure 2.6). The gain drops proportionally for higher levels and the basilar membrane becomes a linear system for very high sound levels (above 90 dB SPL). Nonlinear compression is more important at the basal end of the cochlea and compression occurs when the stimulating frequency is close to characteristic frequency. At the apex, the compression nonlinearity is relatively different from the basal region. At the low frequency range, the compression is relatively uniform and surprisingly, the compression is not consistent at the characteristic frequency. The apical nonlinearity characteristics are not well explored, because of the difficulty in accessing this portion of the cochlea.

When a sound enters an ear, the external ear transfers the sound to the inner ear through middle ear transformer function. The tonotopic organisation begins from the inner ear, i.e., the high frequencies are represented at the basal part of the cochlea and the low frequencies are represented at the apical part of the cochlea. This tonotopic organisation is maintained beyond the cochlea and extends up to the primary auditory cortex. Each part in the central auditory pathway (Auditory nerve, Cochlear Nucleus, Superior Olivary Complex, Lateral Leminiscus, Inferior Colliculus, Superior Colliculus, Medial Geniculate Body, and Auditory cortex) is tonotopically organised. The review of the central auditory system is not in the scope of this thesis (for review see: Pickles, 2008). The coding of sound through cochlear implants will be explored in detail in the next section.

2.3. Sound Coding in Cochlear Implants

The sound coding in electrical hearing deviates from the normal hearing mechanisms. There are approximately 35,000 – 45,000 auditory nerve fibres in normal auditory nerve. These fibres are responsible for effective coding of sound signal from the peripheral auditory system and transfer the coded signal to the central auditory system. The cochlear implant bypasses the external ear and the middle ear and directly stimulates the auditory nerve. Therefore, the physiological salience of normal external and middle ear functions does not apply in CI recipients.

2.3.1. Brief History of Cochlear Implants

The concept of electrically stimulating the auditory nerve for hearing impaired people came into existence 200 years ago. Alessandro Volta, an Italian scientist in 1790 connected each end of two 50-volt batteries with a wire leading to a conducting rod and then he placed the rod in his ear canal. He received a jolt in his head and experienced an unpleasant sensation and described as “Boiling of thick soup” sensation. He immediately terminated the experiment and did not continue further. This was the very first documented event to demonstrate that electrical stimulation can evoke a crude form of auditory sensation. Stevens (1937) tested 2 subjects using an alternating electric current (AC) to directly stimulate the auditory nerve for investigating the underlying electrophonic hearing mechanisms. The subjects described the electrical stimulation as, “short single noise in quick rhythm”. The results showed steeper loudness growth for electrical stimulation and evoked non-auditory sensations (tickling, burning, and pricking sensations).

A French Physician, Djourno and his colleagues (Djourno et al., 1957a,b; Djourno & Eyries, 1957) in Paris reported the first successful restoration of hearing using electrical stimulation in a totally deafened subject. The implant device consisted of an active lead placed on the auditory nerve or adjacent to brainstem and the induction coil and indifferent electrode permanently placed beneath the temporalis muscle. The subject reported awareness of environmental sounds and was able to communicate using lip reading. The subject was unable to discriminate among speakers, but was able to discriminate speech sounds in a closed set paradigm. This device failed eventually and even the second implant did not last for a very long time. This was the first report ever published on electrical stimulation providing hearing sensation in deafened subjects.

The work by Djourno and his colleagues spurred researchers to work on electrically implantable devices for restoring hearing. In USA, William House was inspired by Djourno's work and began working with James Doyle (Neurosurgeon) and Jack Urban (Engineer) to develop an implantable device for restoring hearing in deaf individuals. In 1961, House was the first person in USA to perform the single channel cochlear implant surgery. This device consisted of a single gold electrode insulated with silicone rubber and placed in the scala tympani via the ear canal and round window. The sound was amplitude modulated and delivered to the electrode by a low frequency carrier wave (Square wave 40 – 200 Hz). Three profoundly deaf subjects were implanted only for an approximate duration of three weeks and reported a useful hearing sensation using electrical stimulation. Attempts were made by several researchers (Simmons et al., 1965; Michelson, 1971; Eddington et al., 1978; Hochmair et al., 1981), but no significant difference in terms of outcome was found among these studies.

Bilger et al. (1977) evaluated 13 cochlear implant recipients (11 CI subjects by House and 2 CI subjects by Mechelson) on their hearing abilities and found that the subjects' received useful hearing information from the single electrode cochlear implant. The subjects were able to identify environmental sounds and along with the help of lip reading these recipients established successful communication. Unfortunately, the subjects were unable to perform an open-set speech recognition task using the single electrode devices. Earlier attempts did provide substantial evidence that single electrode stimulation may evoke hearing sensation, but only awareness and discrimination of sounds was partially restored and recipients had to capitalise on these cues to have an effective communication. Thus indeed it was concluded that single electrode stimulation cannot allow the CI recipients to have an effective verbal communication.

In 1978, a multichannel CI developed at the University of Melbourne was successfully implanted in a postlingually deafened adult. The result showed increased abilities in speech reading and also in open set speech recognition task. The multichannel cochlear implants gained maximum attention scientifically, but commercially, it was House – 3M single electrode CI which obtained US FDA approval first in 1984. In the subsequent year, Cochlear Ltd. obtained clearance from FDA to implant deaf people from the age of 18 and above.

There were several other multichannel CI devices developed during this era. The Ineraid or Symbion device developed at the University of Utah (Eddington et al., 1978; Eddington, 1980) consisted of a six electrode implant array with a percutaneous plug connecting the external part of the implant array. This device did not obtain FDA clearance due to the safety issue over the percutaneous plug. The Laura device was developed at the University of Louvaine, Antwerp, Belgium. This device consisted of 8 bipolar channel or 15 monopolar channels (Peeters et al., 1987). These above devices are currently not available in the global market. The French company MXM developed the multichannel Digisonics devices (15 monopolar channels) marketed by Neurelec which are commercially available in some parts of Europe, Russia, Middle East, India, and Brazil. Recently, several companies are trying to develop low cost multichannel cochlear implants [Nurotron Biotechnology Inc. (Irvine, CA & Hangzhou, China), Advanced Cochlear Systems (Seattle, WA), and Neurobiosys Corporation (Seoul, Korea)]. Currently, there are three major CI companies around the world. Cochlear Ltd. (Australia) holds 70 – 80 % (Approx) of the global market and the rest are being shared between Advanced Bionics Corporation (USA) and Med-EL Corporation (Austria).

2.3.2. Components of Cochlear Implants

Cochlear implant devices are mainly divided into external parts (Microphones, Speech Processor, Battery, and Transmitting Coil) and internal parts (Receiver Coil, stimulator and Implant array) [Figure 2.7 & Figure 2.8]. The acoustic signal is picked up by the microphone and the Nucleus CI processors usually have two microphones. Dual microphones may assist the CI users to listen in adverse listening conditions (speech in background noise). The signal from the microphones are attenuated or amplified by the AGC (Automatic Gain Control) depending on the incoming acoustic signal. This signal is then converted into electrical signal by the sound/speech processor. The sound processor plays a significant role in extracting the key features of the sound and converts them into a biphasic electrical pulse train. The detailed information (pulse amplitude, pulse duration, pulse gap, active electrode, and return electrode) is encoded and transferred through the external transmitting coil on a radio frequency between 2 MHz – 10 MHz. The integrated circuit in the receiver stimulator decodes the transmitted signal and stimulates different intracochlear electrodes with appropriate current levels (Figure 2.7). The general working principle of cochlear implants is similar across the three CI manufacturers.

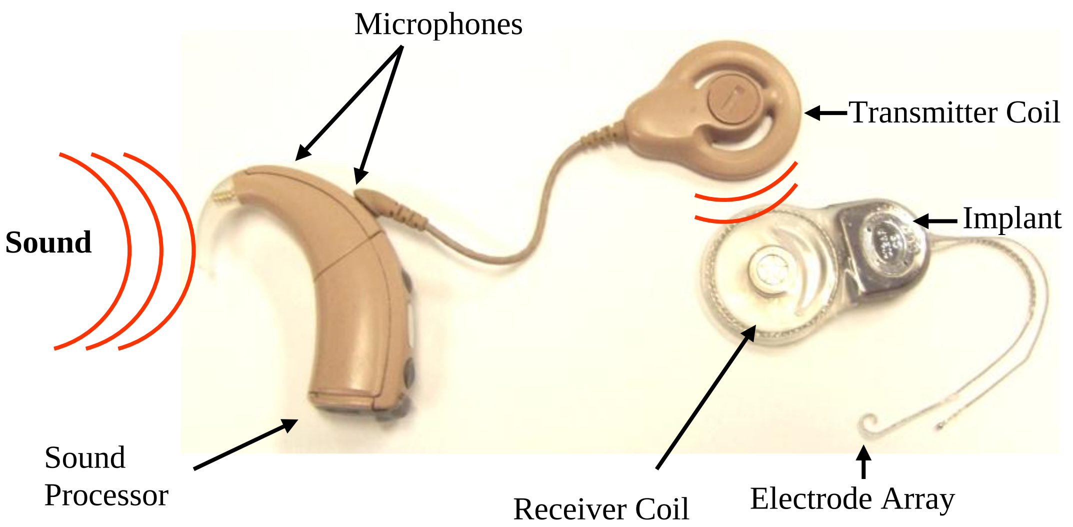


Figure 2.7: The Cochlear Implant System is usually divided into External part (Microphones, Sound Processor, and Transmitting coil) and Internal part (Receiver Coil and Electrode Array) (Courtesy: Cochlear Ltd).

2.3.2.1. Microphone

Microphones are usually housed within the behind-the-ear (BTE) component of the CI device. There are usually two types of microphones used in the commercial devices. Firstly, the omnidirectional microphone picks up the signal from all directions with equal sensitivity. The frequency response is usually flat across the wide frequency range. Secondly, the unidirectional microphones are more sensitive to sounds in front of the microphone compared to sounds behind the microphone. The two types of microphones assist the recipients in adverse listening conditions like speech in background noise. In the Nucleus freedom processors, there are two ports for directional microphones and one port for an omnidirectional microphone (Figure 2.8). This dual microphone setup permits a "beamforming" strategy that is more sensitive to front end sounds and reduces background noise substantially by a "nullification" method (Figure 2.9). In the beamforming method, a "null" position is chosen, based on the direction of an intrusive noise and the null position follows the noise source permitting maximum attenuation if noise is present.

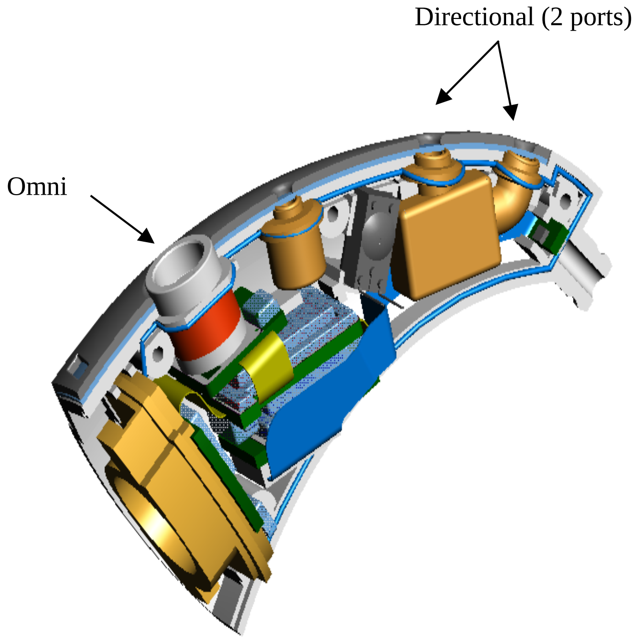


Figure 2.8: Nucleus Freedom (Dual Microphone Beamforming) processor showing two types of microphones (Courtesy: Cochlear Ltd).

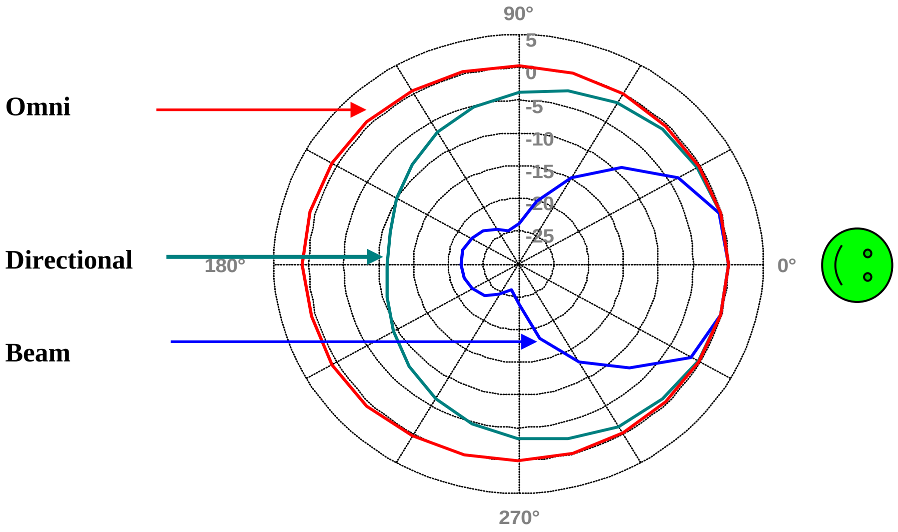


Figure 2.9: Polar plot demonstrating the sensitivity of three types of microphones (Courtesy: Cochlear Ltd).

2.3.2.2. *Sound Processor*

The sound processor acts as a mediator between the microphone and the electrode array. The main operation of the sound processor is to convert the acoustic signal into an electrical pulse train and directs the receiver stimulator to stimulate the intracochlear electrodes with appropriate current levels. The cochlear implant system consists of microphones, speech processor (front end, filter bank, sampling and selection, and amplitude mapping), RF encoder, Receiver stimulator, and intracochlear electrodes (Figure 2.10). The working mechanism of each component is explained below.

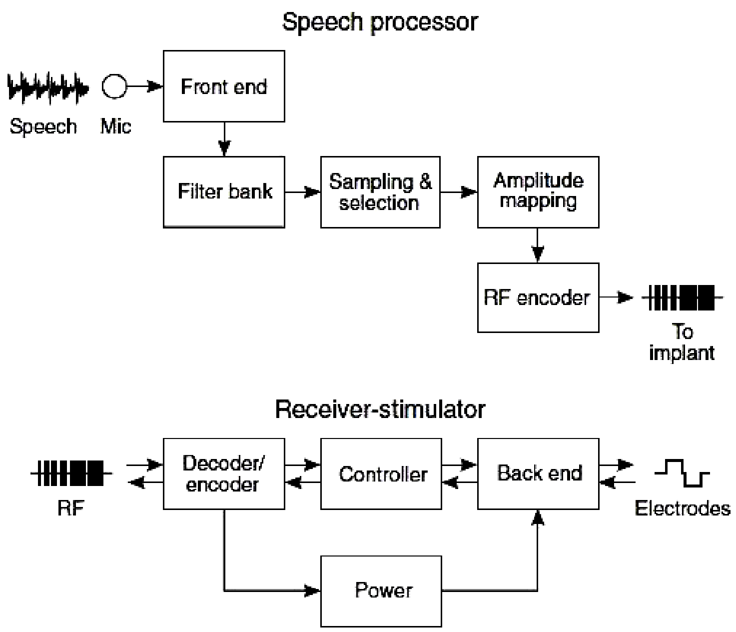


Figure 2.10: Block diagram showing the system architecture of a speech processor (Clark , 2003).

In the microphone and front end section, the sound vibrations are picked up by the microphones. The sound signals are either amplified or attenuated by the AGC system and transferred to the filter banks section. In the filterbank, the signal is filtered and analysed by numerous band pass filters mimicking the frequency analysis of the human ear. Each filter band corresponds to different characteristic frequency locations inside the cochlea. In the Nucleus freedom processor there is a 22 channel filter bank. The filters are linearly placed below 1000 Hz and logarithmically placed above 1000 Hz. The Nucleus Sprint and Freedom uses the FFT filterbank implementation because it is a computationally efficient algorithm. The FFT filterbank creates linear spaced filters spaced at 125 Hz for frequencies below 1000 Hz. For higher frequencies (above 1000 Hz), the filter spacing is doubled or more. During sampling and selection, the timing and stimulation patterns are determined by the filter envelopes and the implants are stimulated sequentially. It is in this stage, where the sound processing schemes (CIS, SPEAK, and ACE) vary depending on the stimulation patterns. The sampling rates vary across different speech processing schemes. Finally, the amplitude mapping determines the appropriate current levels for stimulating the electrodes. The dynamic range for electrical hearing is approximately 8 - 10 dB (difference between C (Comfort level) & T (Threshold level)) as compared to 100 dB in normal hearing individuals. Therefore, Nucleus devices use Loudness Growth Function (LGF) to compress the entire filterbank envelope signal into individuals' electrical dynamic range. The compressed amplitude provides the suitable C & T levels expressed in clinical units. The resultant biphasic pulse sequence is encoded and transmitted by RF transmission link into the internal components. The internal receiver stimulator decodes the signal and stimulates appropriate electrodes with accurate current levels. The Induction coupling method is used to power-up the internal receiver stimulator.

2.3.2.2.1. ***Overview of Speech Processing Strategies***

Earlier sound processing systems provided crude forms of speech information to the recipients. The advancement in microelectronics has improved the sound processing to a great extent. Current speech processing schemes are able to represent the acoustic signal in such a way that the auditory nerve can process it effectively. In fact, the sound processor is called the brain of CI (Zeng, 2004). Earlier processors relied on analog processing strategies, but the current speech strategies are based on pulsatile stimulation. The commonly used strategies are Continuous interleaved sampling (CIS), Spectral peak (SPEAK), and Advanced combination encoder (ACE). The detailed description of speech processing strategies can be found in Zeng, 2004. But in this thesis, the most commonly used strategies and the strategies which improve pitch (also music) perception are addressed below. These strategies are broadly divided into envelope based strategies and fine structure based strategies (Figure 2.11).

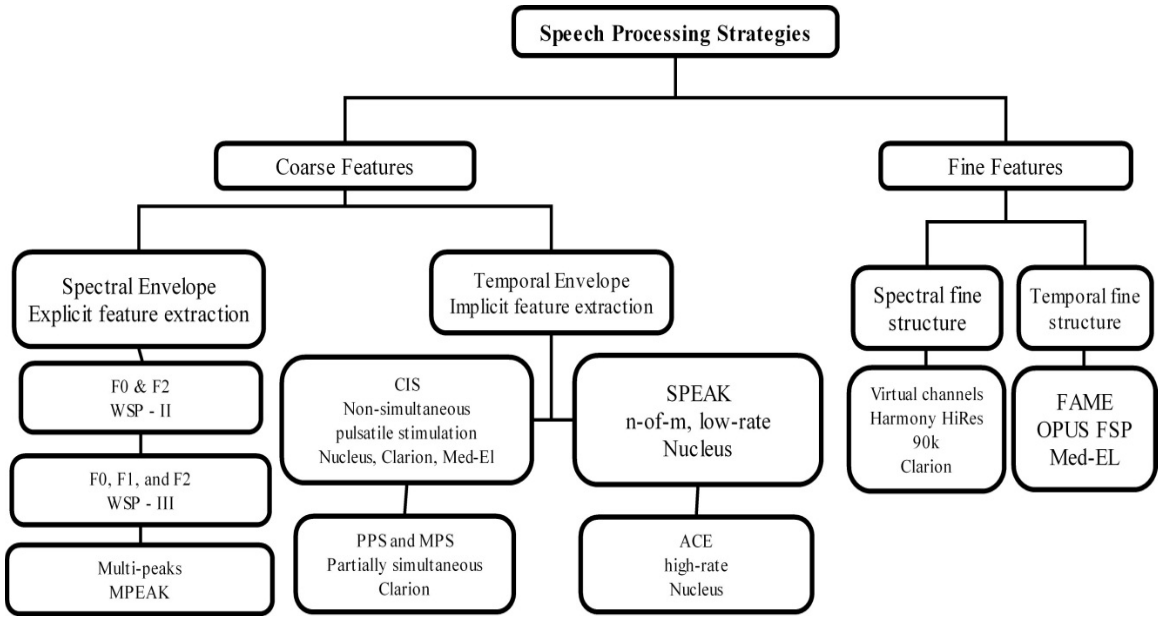


Figure 2.11: Classification of Speech Processing Strategies in Cochlear Implant (Zeng et al., 2008).

<u>*Continuous interleaved sampling (CIS)*</u>

CIS was developed by Wilson et al. (1991) and acts as a foundation for several other speech processing strategies. CIS relies on the principle of non-simultaneous stimulation of specific channels (8 or 12 frequency bands) at a high rate to code the rapid temporal fluctuations (Figure 2.12). The input signal is picked up by the microphone and AGC and the signal is attenuated below 1.2 kHz at 6 dB per octave. The significance of the preemphasis filer is to enhance the consonant components in the speech signals (Wilson, 2004). The output from the preemphasis filter is sent through the bandpass filter bank. The input of each channel is analysed separately by bandpass filtering, envelope detection, and compression. The bandpass filter usually corresponds to the physical electrodes inside the cochlea and the frequency band ranges from 125 Hz to 8 kHz. The envelope obtained from different channels is logarithmically compressed with respect to electrical dynamic range. Unlike normal hearing, the loudness grows rapidly and exponentially in electrical hearing (Zeng & Shannon, 1994). The compression acts similar to LGF (Loudness Growth Function) as in Nucleus devices for generating a normal loudness growth function in electrical hearing. The output thus generated is modulated as a sequence of biphasic symmetrical pulse trains stimulating apical electrodes if the bandpass channel contains envelope with low center frequency or stimulating basal electrodes if the bandpass channel contains envelope with high center frequency. CIS is a within channel high rate temporal coding strategy (Figure 2.13). In order to code the high rate temporal components of a signal, each channel is stimulated at a high pulse rate of 1000 pulses per second. Most importantly, the pulse rate must be twice the cutoff frequency to avoid aliasing and also to avoid rate related pitch cues. Typically, the low pass cut-off frequency is around 200 Hz. In the recent past, Hilbert transforms were used as a substitute for rectification and low pass filtering for extracting the envelope signal from different channels and obtained better results in COMBI 40 and COMBI 40 + devices (Helms et al., 2001). The CIS strategy is commonly used in Advanced Bionics and Med-El devices, but can also be implemented in Cochlear devices.

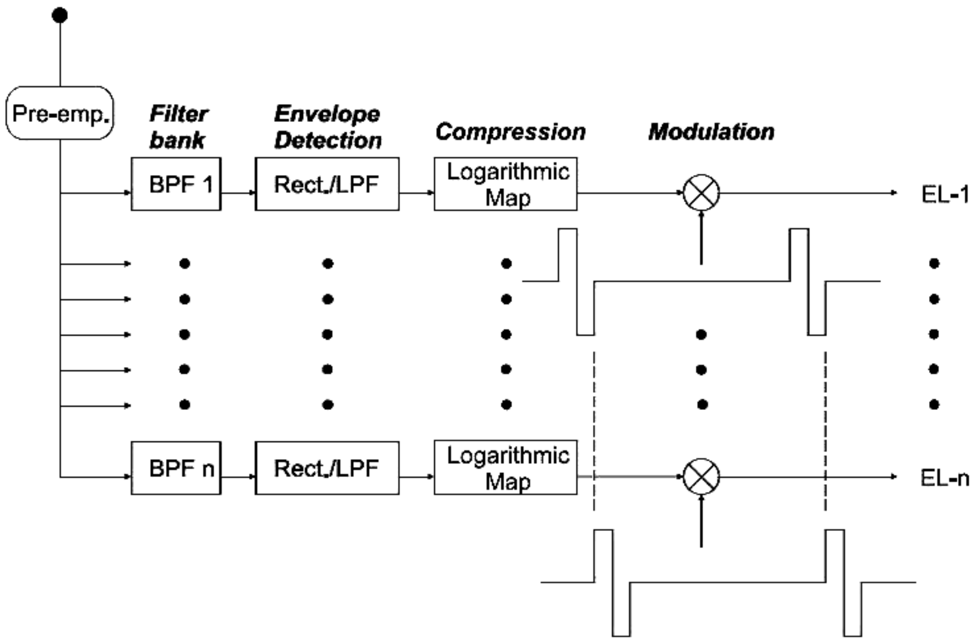


Figure 2.12: Block diagram of CIS strategy (Wilson, 2004).

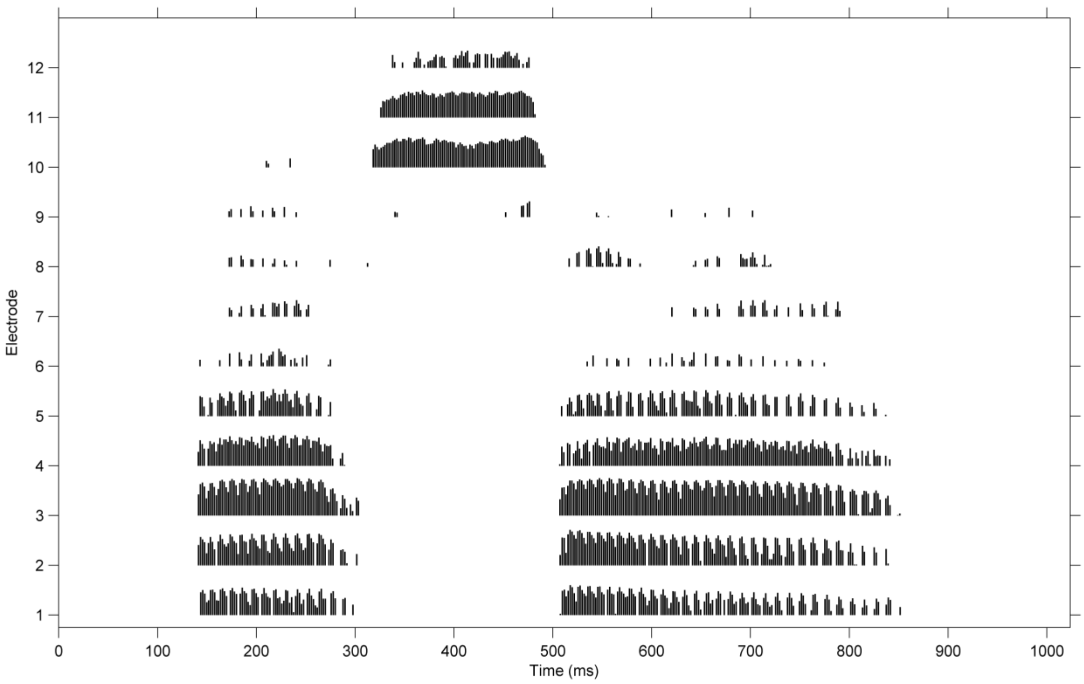


Figure 2.13: Stimulation pattern for the word "asa" using CIS strategy. The electrodes 1 to 12 are numbered from apical to basal portion of the cochlea.

<u>Advanced Combination Encoder (ACE)</u>

ACE is a default strategy for the current generation of Nucleus devices. The earlier Nucleus devices used the Spectral Peak Strategy (SPEAK) (Seligman & McDermott, 1995). SPEAK had a good spectral resolution (20 Channels), but had a slow analysis rate of 250 Hz. This stimulation rate is too slow to capture the rapidly changing temporal fluctuations in the signal. In order to improve the temporal resolution, the ACE strategy was implemented using the similar maxima selection algorithm of SPEAK, but with a high stimulation rate (Figure 2.14).

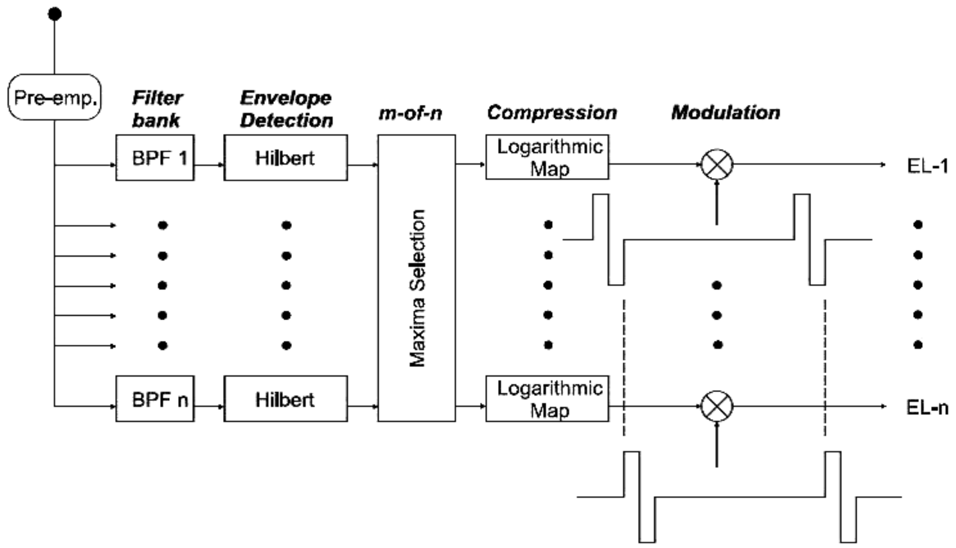


Figure 2.14: Block diagram of ACE strategy (Laneau, 2005).

The architecture of the ACE strategy (Arnd et al., 1999; Vandali et al., 2000) is very similar to the CIS strategy, but the former uses an additional paradigm called maximum selection and the latter does not use the maximum selection but instead stimulates all the channels sequentially. In every analysis period, the signal is processed by “M” (20 or 22) band pass filters and “N” (6 to 12) filter envelopes having the largest maxima are selected and the biphasic pulses are sequentially presented to corresponding “N” channels (Figure 2.15). This n-of-m approach is used to reduce the overall density of electrode stimulation and to increase the stimulation rate across different electrodes

(Wilson, 2004). In Nucleus 24 implants, the ACE strategy can stimulate an electrode as high as 2400 pps compared to 250 pps in SPEAK strategy and the maximum overall stimulation rate is 14400 pps which are sequentially presented to the corresponding "*N*" electrodes out of "M". When N = M the ACE strategy acts as a CIS strategy. The maximum overall stimulation rate for Nucleus Freedom implants is 31500 pps.

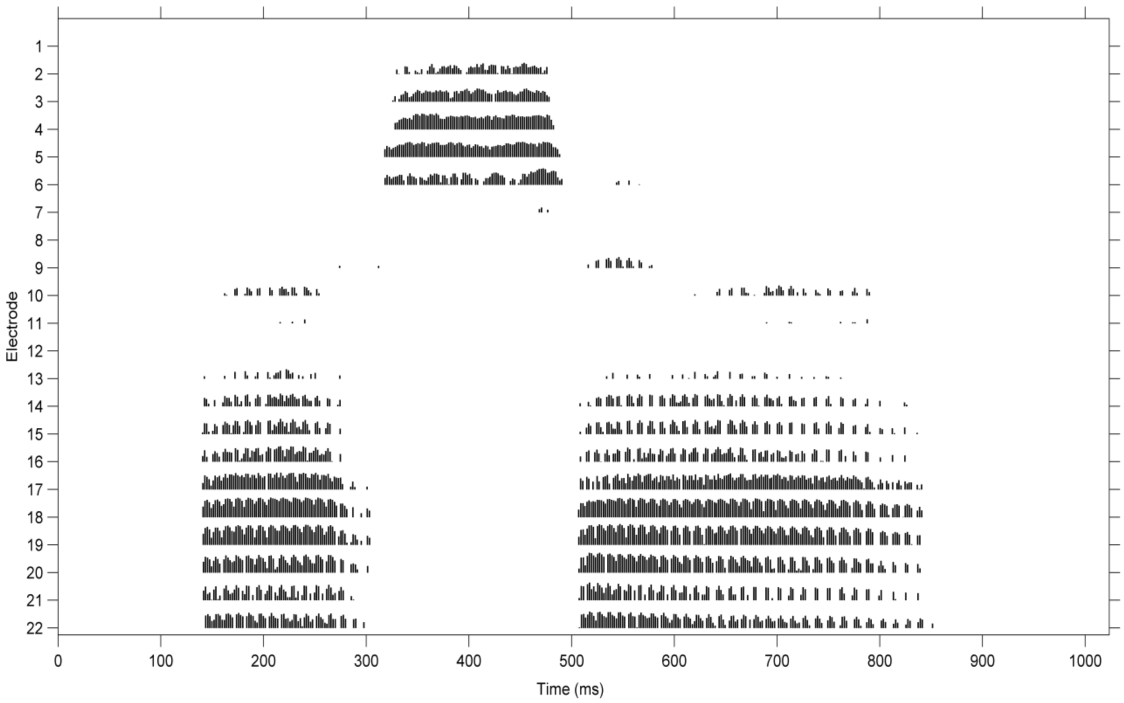


Figure 2.15: Stimulation pattern for the word 'asa' using ACE strategy. The electrodes 1 to 22 are numbered from basal to apical portion of the cochlea.

<u>Strategies to Improve Pitch or Music Perception</u>

Recently attempts have been made by several researchers to improve pitch perception in CI users. The above mentioned sound processing strategies are based on slow-varying envelope detection, but the fast-varying fine structure cues are pivotal for adequate pitch and music perception, speech in competing background noise, and tonal language perception (Smith et al., 2002). Therefore, several strategies have been developed to address these issues. However, only the commercially available strategies are discussed below.

HiRes 120 Strategy

The HiRes 120 strategy is a default strategy for Advanced Bionics devices (Frijns et al., 2004; Koch et al., 2004). This strategy was aimed at improving the spectral resolution of the signal by current steering techniques (Figure 2.16). The underlying principle of the current steering technique is that, when two adjacent electrodes are stimulated simultaneously with specified current ratio, then the place of stimulation is steered between the two electrodes.

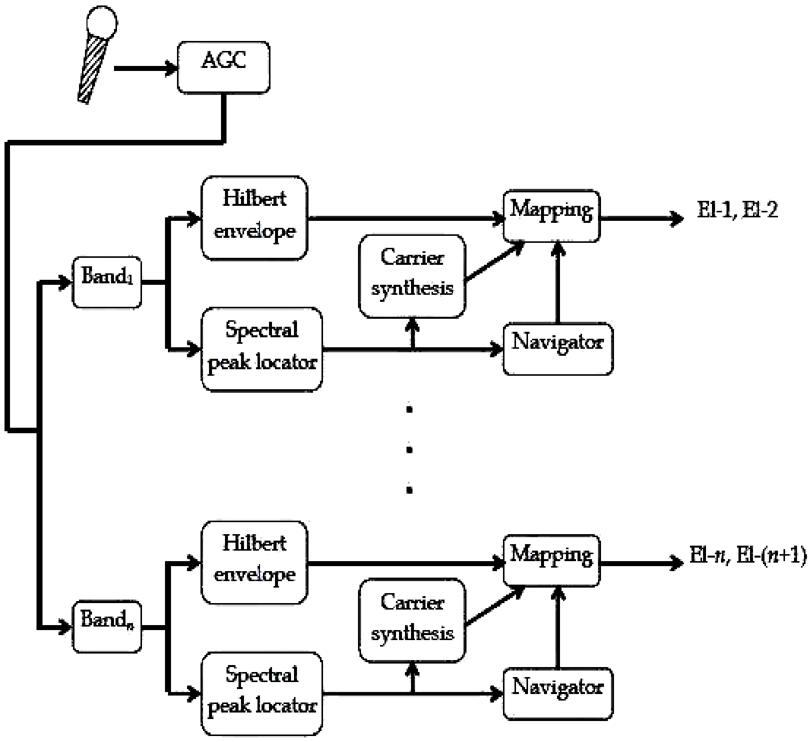


Figure 2.16: Block diagram of HiRes 120 Strategy (Choi & Lee, 2012).

The HiRes 120 strategy is an extension of CIS strategy and is implemented in CII and HiRes 90K implant devices. There are 16 intracochlear electrodes comprised of 15 electrode pairs which are used to steer the current to different intracochlear locations. Unlike the CIS strategy, the HiRes 120 strategy has spectral bands allocated at locations in between the two physical intracochlear electrodes. The current steering factor α (the proportion of current on the more basal electrode) can be calculated between the two electrodes. In HiRes 120 strategy α can be quantised ranging from one to eight spectral bands. Therefore, when 15 frequency bands are steered with a factor of $\alpha = 8$ ($15 * 8 = 120$), then 120 “sites” can be stimulated. The HiRes 120 strategy has substantially improved the spectral resolution by increasing the stimulation sites and the fidelity of temporal representations is ameliorated by increasing the stimulation rate from 2800 to 5600 pps per electrode. Overall they can stimulate at high rates of up to 90,000 pps across electrodes. These advancements did improve the speech perception skills in CI recipients, but the performance on music perception tasks (Nimmons et al., 2008) showed no advantage for HiRes 120 strategy in one of their participants (L7). In fact, the subject scored at chance for the closed set melody identification task.

Fine Structure processing (FSP)

The FSP strategy is implemented along with the CIS strategy for better coding of music and pitch information in CI users (Zierhofer, 2002). The FSP strategy is implemented in the OPUS speech processor (Med-El) and is based on the principle of using the timing of stimulation to code the temporal structure of the signal (Figure 2.17). The FSP uses Channel-Specific Sampling Sequences (CSSS) (Zierhofer, 2002) especially in low to mid frequency channels to improve temporal coding mediated by improved phase locking. CSSS are implemented in the apical channels (two or three) depending on the allocation of the bandpass filters. CSSS ranges from 70 – 350 Hz, so the low frequency temporal cues are emphasized and the higher harmonics are usually coded as place cues using sequential virtual channels. High rates (approximately four times the upper limit of the FSP frequency band) are required to implement FSP processing and the lower cutoff range is 70 Hz compared to 250 Hz in CIS. CSSS works on the principle of zero-crossing, i.e., a series of instantaneous stimulation pulses are triggered when a signal crosses from positive to negative phase of the signal in a specific bandpass filter. Therefore, the instantaneous repetition rate of these sequences is equal to the instantaneous fine structure frequency of the signal in the specific bandpass filter. Zero-crossing is directly proportional to the fine structure frequency, i.e., when the fine structure frequency is high, zero crossing occurs more frequently in time and sequences will be generated instantaneously.

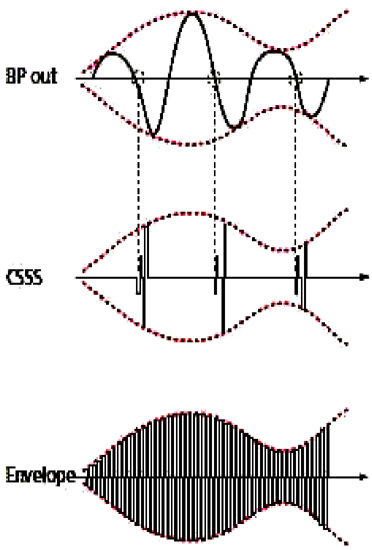


Figure 2.17: Fine Structure Processing (FSP) is achieved by using Channel-Specific Sample Sequences (CSSS) (Source: Med-El Ltd).

Arnoldner et al. (2007) investigated the performance of FSP strategy and found good scores in rhythmic tasks, which is not surprising because CI recipients are on par with normal hearing individuals for the perception of rhythm. But their performance was at chance for closed-set melody and timbre identification tasks. The result from this study suggests that there was no superiority of FSP strategy on music perception tasks.

The sound processing strategies can be broadly divided into coarsely grained coding strategy (CIS, ACE, and SPEAK) and finely grained coding strategy (HiRes and FSP). The coarsely grained strategies use slow-varying envelope detection to code the acoustic signal. This has several caveats. The rapidly varying fine structures are lost during the envelope filtering and do not convey subtle information about pitch, melody, prosody, tonal language, and speech in competing background noise. Surprisingly, the HiRes and FSP which convey fine structure information did not provide convincing results in music and pitch perception tasks.

2.3.2.3. Transmission Link and Receiver stimulator

The signal processed by the sound processor has to be transferred precisely and accurately into the internal receiver system (Figure 2.10). The signal can be transferred in two ways. Firstly, the percutaneous link is achieved by directly connecting the external components to the internal components. Due to safety issues, percutaneous connection is no longer available in the current CI devices. Secondly, the transcutaneous connection is attained by connecting the external coil to the internal receiver stimulator with the help induction coupling method. The function of transmission link is to transmit the vital parameters to the internal system, power-up the internal receiver stimulator, and acts as a bidirectional transmission of data. In other words, data can either be transferred from the external component to the internal component or vice versa (in case of intracochlear evoked potential recordings). The signal from the sound processor is converted into a series of digital data streams which are represented in terms of bits (1s and 0s). Bit coding is used to represent the signal with high fidelity and precision. The RF carrier frequency varies among companies, Clarion HiRes 90K uses 49 MHz, Med-El Sonata uses 12 MHz, and Nucleus 24 uses 5 MHz (Figure 2.18).

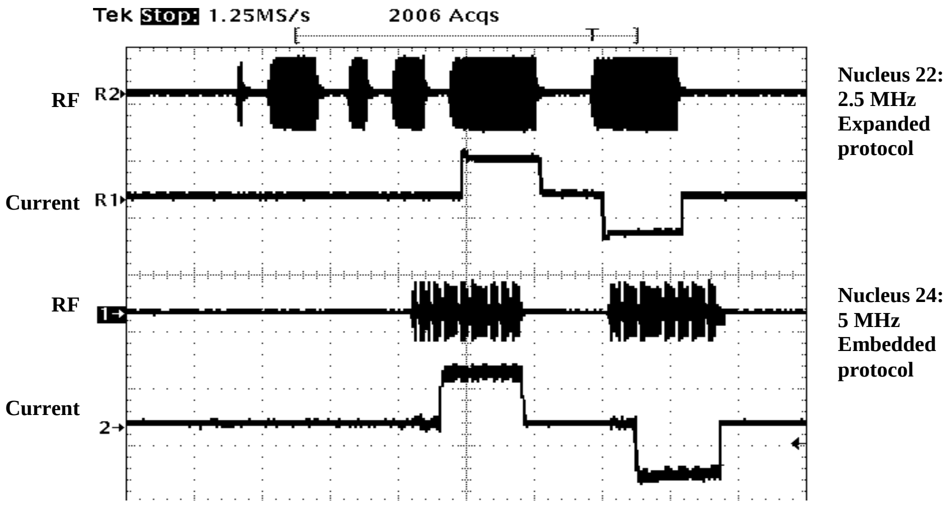


Figure 2.18: Radio Frequency Transmission for Nucleus Devices (Courtesy: Cochlear Ltd).

The Nucleus devices use the frame coding scheme to transmit stimulus parameters to the internal stimulator. These parameters are active electrode, return electrode, mode of stimulation, pulse amplitude, pulse phase, and inter phase interval. The two types of framing coding schemes are expanded and embedded coding protocol (Figure 2.18).

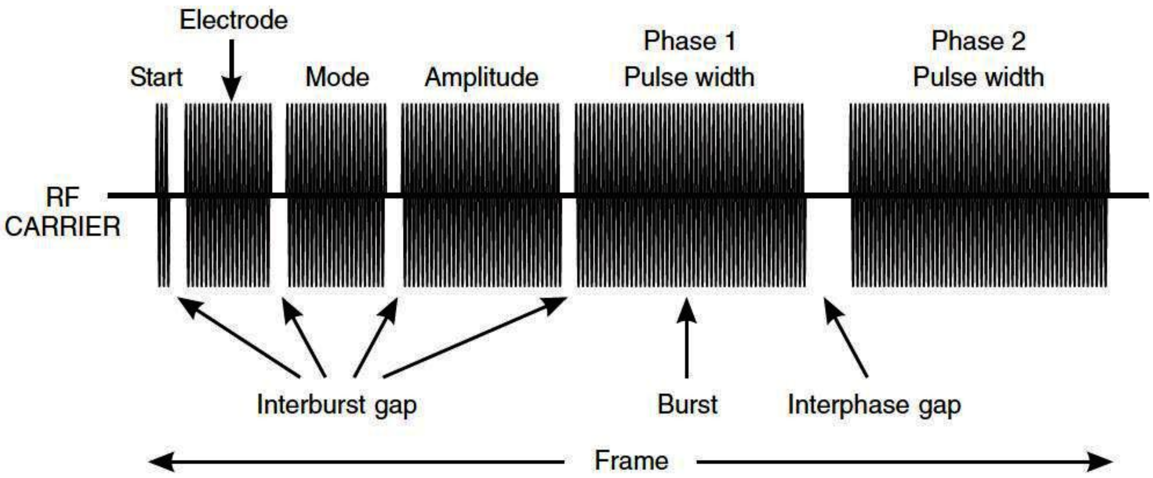


Figure 2.19: The expanded mode frame coding scheme in Nucleus device (Clark, 2003).

The expanded frame coding scheme consists of five crucial parameters for representing the biphasic pulse (Figure 2.19). The initial burst is a SYNC burst which lasts not more than seven RF clock cycles (Zeng et al., 2008). The next burst carries the information of the active electrode to be stimulated. The following burst contains the information regarding the electrode configuration (Monopolar or Bipolar). The amplitude burst contains the information regarding the pulse amplitude, i.e., the current level units (ranging from 0 – 255 current units). The pulse duration of phase 1 is present in phase 1 burst. The interphase gap is the duration of the gap between phase 1 and phase 2, i.e., from negative to positive phase of the pulse. The phase 2 duration is similar to phase 1. There is an inter frame gap of approximately 1.2 µs to 250 ms. This method of frame coding is relatively slow. Therefore, an embedded frame coding scheme was introduced in the latest Nucleus 24 devices, running at 5.0 MHz compared to 2.5 MHz in Nucleus 22 devices (Figure 2.18). Embedded protocol operates by providing specific parameters (electrode, mode, amplitude) to the succeeding frame (N + 1), while the present stimulus is being delivered to the frame (N). The embedded protocol validates the stimulus parameters effectively and reduces the errors. The current Nucleus CI24R and CI24RE implants are capable of using both expanded and embedded mode of frame coding.

2.3.2.4. Electrodes

Commercially available CI devices generally digitize the input acoustic signal into an electrical pulse train varying in amplitude, pulse width, rate, timing, and site of stimulation. Earlier intracochlear electrodes were made of copper and gold wires, but the current intracochlear electrodes are made of platinum or platinum-iridium alloy. Modern electrodes are durable and differ significantly in both geometric parameters and the stimulation mode (Zeng, 2004).

Figure 2.20: Intracochlear Electrode Representation for Cochlear Implants. The white rings on the black carrier represent the electrode contacts, which in turn stimulate the nearby auditory neuron in the modiolus. The electrode array is inserted via the scala tympani and folded into two complete turns (Zeng, 2004).

Current CI devices contain both extracochlear and intracochlear electrodes. The Nucleus implants have 22 intracochlear electrodes and 2 extracochlear electrodes. The intracochlear electrodes are numbered from E1 at the basal end to E22 at the apical end of the cochlea. Additionally, the two extracochlear electrodes act as ground electrodes. The first (ECE1) is a ball electrode connected to the simulator coil by a lead wire and placed beneath the temporalis muscle at the time of surgery. The second (ECE2) is a platinum plate mounted on the titanium package (Receiver-Stimulator) of the implant. The optimum intracochlear insertion depth ranges between 25mm to 31mm from the round window. The advancement in insertion techniques (AOS (Advance Off-Stylet Insertion)) has enabled atraumatic insertion for contour devices with reliable perimodiolar placement. The electrode arrays can be modified into short electrode array in case of combined acoustic and electrical stimulation (Electro-acoustic stimulation (EAS)) and double electrode array inserted separately into first and second turns of the cochlea in case of cochlear ossification.

In order for the current to flow in the implant, it is a prerequisite to have at least two electrodes (active electrode and a reference electrode). The complete path in which the current flows depends on the location of the reference electrode. Electrode configuration can vary in different stimulation modes. Multiple electrodes can be configured to deliver current to the auditory neuron in several ways. In a monopolar mode (MP), the current flows between an active intracochlear electrode and a reference extracochlear electrode (The reference electrode is usually the combination of ball electrode (ECE1) and the receiver stimulator (ECE2). The commonly used monopolar mode is MP1+2). Therefore, the spread of excitation is relatively broader in MP mode. This mode is available in the Nucleus 22, Nucleus 24, and Nucleus Freedom devices. In a bipolar mode, the current flows between an active and a reference intracochlear electrodes which are placed very close to each other. So they stimulate a spatially distinct set of neurons leading to perceptually discriminable auditory sensation. The spread of spatial excitation can be controlled by increasing the distance between the active electrode and the reference electrode. In Nucleus 22 device, BP represents bipolar mode stimulating two adjacent electrodes (E.g., E4 - E5). Similarly, BP + 1 indicates an additional spacing of one electrode (E.g., E4 - E6). Nucleus devices have a single current source and use electronic switches to route the current to the required electrodes. So, only one channel is stimulated at any given time; this is known as sequential or interleaved stimulation. The Nucleus devices are capable of producing monopolar and bipolar mode. The HiRes 90 K devices consist of 16 current sources corresponding to 16 physical electrodes and Sonata devices have 12 current sources. These two devices are capable of producing different electrode configuration other than monopolar and bipolar modes (Table 1). The Tripolar mode (TP) comprises of one current source with magnitude -i and two return electrodes surrounding the first with a magnitude of +i/2. The spread of excitation is substantially reduced in TP stimulation leading to focussed electrode stimulation and improved channel selectivity.

One of the main issues with the focused stimulation is that they are unable to maintain appropriate loudness across the channels, so higher current levels are required to accomplish the comfortable loudness levels (Litvak et al., 2007; Bonham & Litvak, 2008). To overcome this problem, an extra-cochlear electrode can be used as an additional ground electrode to form partial tripolar stimulation (pTP). When the ratio between intra cochlear and extracochlear electrode (σ) is 0, then the stimulation pattern is monopolar. When σ = 1 the stimulation pattern is completely TP. For accomplishing pTP, σ should be more than 0.5. The pTP comprises one current source with magnitude -i and three return electrodes (two intra-cochlear ground electrodes [+ (σ/2)i] and one extra-cochlear ground electrode [+(1-σ)i]) (Table 1). Unfortunately, studies implementing pTP in the sound processor did not facilitate the performances of CI recipients (Mens & Berenstein, 2005; Berenstein et al., 2008).

Table 1: Four pattern of electrode configurations. The amplitude of the first phase of the biphasic pulses is represented. The x-axis represents the electrode position (Landsberger & Srinivasan, 2009).

0	0	- i	0	+*i*	*Monopolar (MP)*
0	0	-i	+*i*	0	*Bipolar (BP)*
0	+*i/2*	-i	+*i/2*	0	*Tripolar (TP)*
0	+*(σ/2)i*	-i	+*(σ/2)i*	+(1 - σ)i	*Partial Tripolar (pTP)*
1	**2**	**3**	**4**	EC [Extra-Cochlear Electrode]	
	Electrodes				

Focused stimulation can be achieved by using the phased array (PA) stimulation (van den Honert & Kelsall, 2007). In the PA stimulation all the electrodes are stimulated simultaneously. The PA works on a principle that the electrical potential of all the electrodes except the centre contact electrode be zero instead of imposing an intracochlear current density pattern (Figure 2.21) and PA stimulation relies on the impedance measurement of the actual electrode (Frijns et al., 2011). The resultant electric field is indeed highly restricted and focused. Preliminary results using PA stimulation in percutaneous CI users revealed substantially reduced channel interactions. Additionally, the PA stimulation is capable of stimulating multiple sites simultaneously rather than sequentially (Frijns et al., 2011).

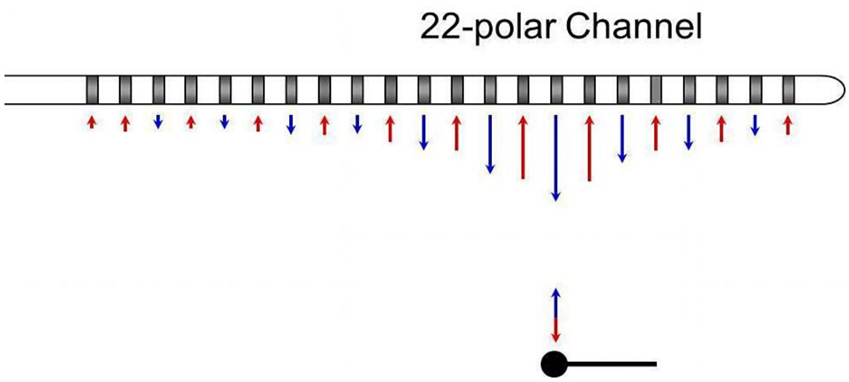


Figure 2.21: Phased Array Stimulation (Courtesy: Cochlear Ltd)

The above descriptions of the cochlear implant components provide a detailed understanding of the holistic working pattern of the cochlear implants. The next section addresses the candidacy issues in cochlear implants.

2.4. Candidacy for Cochlear Implants

The candidacy criteria for CI has been revised frequently with the advancement in device technologies, surgical techniques, and the sound processors (Zwolan, 2009). The selection criteria for adults differ from the criteria for children. Firstly, cochlear implantation is usually recommended for a person whose unaided threshold ranges from severe (> 70 dB) to profound (> 90 dB) hearing loss. Secondly, the person should have moderate hearing loss specifically at the low frequency region. Thirdly, there should be little or no benefit with hearing aids. Fourthly, the performance should be ≤ 50 % in open set speech perception task. Finally, there should not be any medical or radiological contraindication. Indeed, the test battery approach is crucial for finding the appropriate candidates for cochlear implantation program (Details on CI candidacy, see: Cooper & Craddock, 2006; Waltzman & Roland, 2006; Katz et al., 2010. Details on surgical issues, see: Niparko, 2009).

This chapter provided a detailed description of the sound coding in CI users and a brief summary on the audiological criteria for selecting the candidates for cochlear implantation program. In the next chapter, how pitch is processed in the CI users will be discussed.

Chapter 3: Pitch Processing in Cochlear Implant Recipients

3.1 Introduction

Cochlear Implant technology has surpassed a number of very challenging obstacles. The remarkable progress from hardly understanding speech to securing 100 % scores in speech perception tasks (in quiet) have demonstrated that the CI recipients are capable of having a successful conversation in quiet, like normal hearing counterparts. This extraordinary achievement guides us to the next level of improving the quality of life in CI users, which focusses on musical listening. Pitch is one of the fundamental elements in speech and music. In English language, pitch salience is vital for processing suprasegmental features like prosody, stress, and intonation. Conversely, in tonal languages, pitch is used as an integral part of the speech, i.e., pitch inflections can change the entire meaning of the word. For example, *ma* in Mandarin Chinese can be produced in four different ways [*mā, má, mǎ,* and *mà*] all varying in pitch patterns and all having an entirely different set of meanings. There are more than one billion people speaking tonal languages in 88 countries around the world (East Asia, Africa, part of Europe, Mexico, and part of South America).

The extensive research in the area of pitch and music perception in CI users provides a unanimous conclusion that their performance is typically much worse than normally-hearing subjects. Attempts have been made by several researchers to ameliorate the pitch coding for CI users. In the recent past, attempts have been made by commercial companies to improve the design of intracochlear electrodes for better pitch processing (CI24RE in Cochlear Ltd, HiRes 90K in Advanced Bionics, and Pulsar in Med-El) and others have worked on improving the sound coding schemes in the speech processor: Fine

Structure Processing (FSP) (Zierhofer, 2002), Spike-Based Temporal Auditory Representation (STAR) (Grayden et al., 2004), HiRes 120 (Koch et al., 2004), Modulation Depth Enhancement (MEM) (Vandali et al., 2005), Multi-Channel Envelope Modulation (Vandali et al., 2005), F0mod (Laneau et al., 2006), Half Wave Gating (Swanson, 2008), eTone (Vandali & van Hoesel, 2011). Unfortunately, these advanced electrode designs and newer sound coding schemes did not significantly improve pitch perception in CI users. This chapter discusses the pitch processing mechanisms in normal hearing individuals and CI recipients. Additionally, a detailed review on place pitch and temporal pitch mechanisms are addressed in this chapter.

3.2 Definition

There are two ways of defining pitch. Firstly, the latest definition, *"pitch [is] that attribute of auditory sensation in terms of which sounds may be ordered on a scale extending from low to high. Pitch depends primarily on the frequency content of the sound stimulus, but it also depends on the sound pressure and the waveform of the stimulus"* (ANSI, 1994). This definition provides a broad understanding of pitch, but it is apparent that even other dimensions such as loudness and timbre (brightness) can be ordered on a similar scale.

Secondly, the traditional and most widely accepted definition of pitch is "that attribute of auditory sensation in terms of which sounds may be ordered on a musical scale" (ASA, 1960). According to Plack & Oxenham (2005a), pitch is defined as " *that attribute of sensation whose variation is associated with musical melodies"*. In this thesis, an operation definition of pitch is that the variation in pitch can convey a melody (Moore & Carlyon, 2005). The reason for defining pitch from a music perspective will be discussed in-depth in the following chapters.

3.3 Overview of pitch processing in normal hearing individuals

An acoustic signal consists of three perceptual dimensions. They are pitch, loudness, and timbre and their acoustic correlates are frequency, intensity, and spectral centroid respectively. In the normal auditory system, pure tones are represented either by place or temporal mechanism. A pure tone consists of a single frequency component (no harmonic complexes) which produces maximum excitation in the specific region of the basilar membrane and this is attributed to place mechanism. The temporal mechanism usually occurs at the level of auditory nerve and is based on the principle of synchronous firing of the neurons in the auditory nerve. Researchers have reached a consensus that, in the normal auditory system, pure tones are phased locked up to 5 kHz but phase locking declines substantially from 2 KHz (Moore, 2003; Plack & Oxenham, 2005b) and pure tones above 5 kHz are represented using the place mechanism. The transition from the pure temporal to the place mechanism is poorly understood (Plack & Oxenham, 2005b). Interestingly, the upper limit for musical pitch is around 5 kHz, which is similar to the upper limit for phase locking in pure tones. It is also observed that the place mechanism comes into play at a rather lower frequency (below 5 KHz). Plack & Oxenham (2005b) argued that the place mechanism alone is not capable of decoding musical pitch and suggested the possibility of a combined mechanism of place and temporal pitch for coding low frequencies. In the “rate-place” mechanism (Plack & Oxenham, 2005b), the pure tones are represented based on the rate of neuronal firing corresponding to the excitation pattern across different regions of the basilar membrane.

In a harmonic tone complex, the pitch is determined by the low-numbered harmonics which are resolved by the normal auditory system (Plomp, 1967; Moore et al., 1985). The low-numbered harmonics usually ranges from 5 – 10 harmonics (Houtsma & Smurzynski, 1990; Bernstein & Oxenham, 2003). The high-numbered harmonics do not stimulate distinct places along the basilar membrane and therefore cannot be resolved by the auditory filter banks and are referred to as the unresolved harmonics. The auditory perception of resolved harmonics is significantly better than that of the unresolved harmonics (Houtsma & Smurzynski, 1990; Kaernbach & Bering, 2001). One of the problems with current CI devices is that they are not capable of resolving the harmonic contents of a signal.

In the real world, pure tones are uncommon and more realistic stimuli are harmonic tone complexes. A typical harmonic tone comprises of fundamental frequency (the first harmonic (F0) in the harmonic complex, which is the lowest frequency of a signal) and other additional higher harmonics. The pitch percept corresponds to the fundamental frequency (F0) of a harmonic tone. Pitch increases relative to increase in F0 until up to 5 kHz. Pitch is unaffected even when the amplitude of the harmonics and the phase is modified. The underlying mechanisms of pitch are still unclear and debated by several researchers till date. For example, pitch across the two musical instruments may remain constant, but the timbre is modified. The pitch extraction mechanisms can be explained by the autocorrelation (or All-Order Inter Spike Interval) model (Licklider, 1951; Meddis & Hewitt, 1991; Cariani & Delgutte, 1996; Meddis & O'Mard, 1997). According to this model, nerve spikes are generated at every time delay which corresponds to an acoustic signal. This time delay is attributed to the fundamental frequency (F0) of the signal. In the complex harmonics, the F0 is determined by taking the inverse of the time delay of the first largest peak in the autocorrelation function (ACF) of the signal.

Temporal pitch in normal listeners can also be elicited by amplitude modulating a noise signal (Burns & Viemeister, 1976). The modulation rate corresponds to the perceived pitch. Normal listeners can detect rate changes up to around 300 – 600 Hz (Patterson et al., 1978; Carlyon & Deeks, 2002). The temporal pitch mechanism can be explained by the first-order interval model. According to this model, the pitch is extracted from the weighted sum of the first-order intervals between pulses and neglecting the higher order intervals (Kaernbach & Demany, 1998; Kaernbach & Bering, 2001; Carlyon & Deeks, 2002; Moore, 2003).

Pitch can alternatively be explained by pattern recognition models (Goldstein, 1973; Wightman, 1973; Terhardt, 1974; Langner, 1992). According to this model, the harmonic complexes form specific patterns and these patterns are usually matched with the best internal template from the pool of templates to extract the pitch percept. In order to adequately process complex pitch information, the accurate tonotopic representation is vital (Oxenham et al., 2004).

3.4 Pitch Mechanisms in Cochlear Implants

In normal listeners, the place and temporal cues to pitch are interlinked and co-vary in real world scenarios. Conversely, cochlear implant users provide an excellent platform to investigate place pitch and temporal pitch independently (Tong et al., 1983; McKay et al., 2000). The pitch percept evoked by stimulating different electrodes inside the cochlea corresponds to place pitch. These intracochlear electrodes mimic the natural tonotopic organization of the cochlea, i.e., the basal electrodes correspond to high frequencies, while the apical electrodes are activated by low frequencies. Temporal pitch can be elicited by varying the pulse rate on an electrode (Simmons et al., 1965; Tong & Clark, 1985; Pijl & Schwarz, 1995b; Fearn et al., 1999; McKay et al., 2000; Zeng, 2002) or by amplitude modulating a carrier pulse train (Shannon, 1983; Tong et al., 1983; Busby et al., 1993; Mckay et al., 1994; Busby & Clark, 1997; Geurts & Wouters, 2001; Laneau et al., 2006). Temporal pitch usually fades off at around 300 Hz (Shannon, 1983; Blamey et al., 1984; Tong et al.,1985; Townshend et al., 1987; McKay et al., 2000; Zeng, 2002; Kong et al., 2009). In the following section, a review on temporal and place coding of pitch will be discussed.

3.4.1. Temporal Coding in Cochlear Implants

Temporal pitch in CI recipients can be conveyed either by varying the pulse rate or modulation frequency. CI recipients are happy to categorise these two percepts as *"pitch"*. Temporal pitch was first reported by Simmons et al.(1965). In this study, a cluster of six gross electrodes were implanted in the modular portion of the cochlea of a single post-lingually deafened patient under local anaesthesia. Assessments began after 1 week of post implantation. Pitch perception was among the several parameters evaluated in this recipient. The striking outcome of this study is that, *"the pitch is affected by both electrode selection and stimulus repetition rate, suggesting that two modes of "pitch" encoding are operative within one group of auditory fibres"*. The result showed that as the pulse rate increased from 50 Hz to 300 Hz with constant loudness, there was a steady increase in pitch. Moreover, they also reported about temporal pitch coding mediated by amplitude modulation, but were unable to explicitly explain the underlying mechanism.

Eddington (1978a,b) administered pitch scaling procedures in a single subject who was implanted with a six electrode device. Similar to the above study, Eddington et al. reported two mechanisms involved in processing pitch (Place & Rate) and reported the upper cut-off range for rate pitch as 300 Hz. The JNDs (Just Noticeable Difference) for pitch become extremely high for frequencies above 400 – 500 Hz. The above studies used fewer channel implant devices which provided coarse-grained information regarding speech (& pitch) and this may be attributed to the devices' slow integration time of 1 – 2 ms and the lack of spectral resolution (Shannon, 1983). In the early 1980s, research was spurred in the area of exploring multichannel CI devices.

An attempt was made by Tong et al. (1983) to explore the psychophysical capabilities of a multichannel CI device in a single CI subject. They evaluated two pulse patterns (single pulse per period (SPP) and multiple pulses per period (MPP)) with respect to current level, repetition rate, and electrode position. In the SPP pulse pattern, a single pulse is presented at every period. For the MPP pulse pattern, multiple pulses are presented in the first half of the period and no stimulation in the second half of the period. This stimulation pattern is called On-Off modulation. The repetition rate for the MPP pulse pattern corresponded to the pulse rate of the SPP pulse pattern. For a fixed duration stimulus, there is an increase in the number of pulses as the pulse rate for SPP increases. Conversely, the number of pulses remains constant irrespective of the increase in repetition rate for the MPP pulse pattern. In order to minimize the influence of loudness in the SPP pulse pattern, the MPP pulse pattern was proposed in this study. The results showed that, for constant current levels, there was a significant increase in loudness for SPP sequences with repetition rate ranging from 100 to 1000 pps. Conversely, the loudness growth in MPP was small and inconsistent. In the pitch estimation task, there was a steep increase in pitch as the repetition rate increases up to 300 pps. The result revealed that the dynamic range of the current levels (C & T levels), the pitch variations in repetition rate, and the DLs for repetition rate are similar across SPP and MPP pulse patterns. Additionally, in a multidimensional scaling procedure (two dimensions) they reported that the repetition rate and electrode position are independent parameters in electrical hearing. An extension of this study was carried out by Busby & Clark (1997) in a group of postlingual (N = 6) and prelingual (N = 8) CI recipients. The pitch estimate for both SPP and MPP pulse patterns increased with respect to increasing repetition rates. Interestingly, five out of eight were able to judge the pitch estimate for MPP pulse patterns as increasing as a result of an increasing repetition rate.

Swanson (2008 & 2010) investigated the temporal pitch perception using four pulse patterns (SPP, MPPU (Multiple Pulses per Period, Uniform-Sampling), MPPS (Multiple Pulses per Period, Synchronized), and MPPH (Multiple Pulses per Period, Half-wave). This study compared the pitch percept produced by a low-rate pulse train to pitch percepts produced by an amplitude modulated high-rate pulse train. The study aimed to investigate whether the pitch of the modulated high-rate pulse train depended on fundamental period or the shape of the modulating waveform. The pitch ranking abilities of six postlingually deafened CI recipients across the four pulse patterns showed interesting results. High performance was obtained for SPP and MPPH and the scores were comparatively less for modulation pulse patterns (MPPU and MPPS). Overall, the results provided clear evidence that the amplitude modulation of the high-rate pulse train evoked pitch information and the shape of the modulation waveform is crucial as it is evident from the performance of MPPH pulse pattern. Pulse time resolution seems to be vital for accurate temporal coding. Importantly, the temporal pitch results were consistent with the first-order inter spike interval model (Kaernbach & Bering, 2001; Carlyon & Deeks, 2002).

The above studies investigated temporal pitch sensitivity mainly by varying the pulse patterns. It is apparent that the amplitude modulation of the pulse train does evoke a sensation of pitch in CI users. The CI users are able to detect the low-pass slow varying temporal modulation and the temporal resolution can be examined by the temporal modulation transfer function (TMTF) (Viemeister, 1979). The TMTF is defined as the ability to just-detect the amplitude of the amplitude modulation as a function of modulation frequency. The TMTF in CI recipients was first reported by Shannon (1992) and the result obtained from 10 CI recipients suggest that the recipients were able to detect modulation frequencies up to 300 Hz and more sensitively in the range of 80 – 100 Hz. A similar finding was obtained by Busby et al.(1993) who tested seven Nucleus CI recipients (four postlinguals and three prelinguals). However, there was a large inter-subject variability seen especially in the prelingually deafened subjects and this is possibly attributed to the poor neuronal survival rates inside the cochlea and also the low-stimulation rate in the speech processor (SPEAK strategy). A recent study (Won et al., 2011) found a significant correlation between the modulation detection (TMTF) and speech recognition thresholds in 24 postlingually deafened CI recipients.

Earlier studies exploring only temporal pitch in normal listeners used sinusoidally amplitude modulated (SAM) noise in order to avoid the contribution of spectral cues (Burns, 1976, 1981; Moore & Rosen, 1979). These studies reveal that the SAM noise provides musical pitch information and the subjects were able to judge the musical interval using the SAM noise for modulation frequency up to 300 Hz – 500 Hz. The researchers argued that perhaps the pitch elicited by SAM noise is a weak pitch and anecdotal reports reveal that a few subjects perceived the modulations as roughness. Extrapolating the result from normal listeners, McKay et al.(1994) investigated different types of SAM pulse trains in six CI recipients. In CI speech processors, the output of the harmonic tones is analogous to the SAM pulse trains corresponding to the stimulus. The subjects were able to consistently rank the 150 – 200 Hz modulation frequency when the carrier rate was well above 800 Hz (e.g., see Figure 3.1). Performance deteriorated when the carrier frequency was less that 800 Hz because of the harmonic relationship between the carrier and the modulation frequency. Therefore, for precise ranking of modulation frequency, it was recommended that, the carrier frequency should be at least four times higher than the modulation frequency. The pitch evoked by amplitude modulated pulse train depends on the modulation frequency with sufficiently high carrier rate and large modulation depth (McKay et al., 1995).

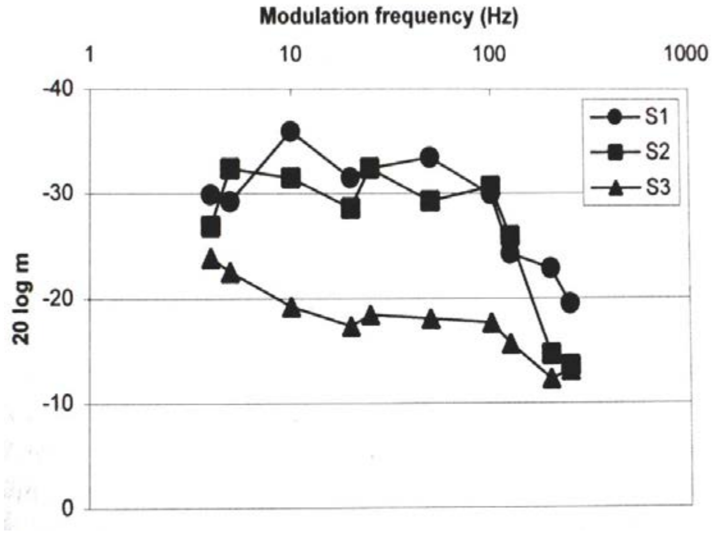


Figure 3.1: Amplitude modulation detection threshold for different modulation rate using 1000 Hz carrier pulse train (Busby et al., 1993).The modulation detection threshold decreases as the modulation frequency increases after 100 Hz. S1, S2, S3 are three CI recipients (Retrieved from: McKay (2004)).

McKay et al. (1995) investigated the importance of modulation depth using modulated and unmodulated pulse trains in four CI recipients. Result obtained from the 2-Interval forced choice (2IFC) pitch matching task showed that, CI recipients' were able to match a modulated pulse train to a similar unmodulated pulse train of equal rate, when the modulation depth was sufficiently large. Conversely, when the modulation depth was small, the perceived pitch was the collective function of modulation frequency, carrier rate, and modulation depth. A simple pitch model was developed to explain these results and the model predicted that the pitch matched rates are determined by the weighted average of the modulation and carrier pulse rates and the weighted sum is proportional to the number of neurons firing at respective frequencies. The model became unpredictable for carrier frequency greater than about 700 Hz. The study by Zhao & Liang (1996) showed a high level of phase locking for modulation frequencies in the range of 400 – 1200 Hz in the Dorsal Cochlear Nucleus (DCN) of guinea pigs. They reported that there might be a mechanism exclusively detecting modulation depths.

In a further study reported by McKay et al.(1999) using modulated and unmodulated pulse trains, the result obtained from a multidimensional scaling (MDS) analysis showed that there is a single dimension stimulus space for unmodulated rates (60 – 300 Hz) and two-dimensional stimulus space for modulated pulse trains. The two-dimensions of a modulated pulse train are carrier (140 – 300 Hz) and modulation rates (60 – 150 Hz). Therefore, the dual pitch elicited by the modulation pulse train corresponds to both carrier and modulation rates. Overall, they concluded that the central auditory system can extract two forms of temporal patterns specific to each location.

The above studies provided a clear understanding of how the amplitude modulated pulse trains evoke a pitch sensation (Temporal / Rate Pitch). The auditory nerve fires at a specific rate for an input signal and these neuronal excitation patterns are phase locked to the corresponding input signal. This mechanism is similar in both electrical and acoustic hearing. Interestingly, the nerve spikes are phase locked more synchronously in electrical hearing compared to acoustic hearing (Shepherd & Javel, 1997; Abbas & Miller, 2004). Experiments pertaining to rate pitch usually stimulate a single electrode by varying the pulse rate provided to that particular electrode. The pitch evoked by varying the pulse rate ranges from 50 pps to 300 pps. The percept below 50 pps is usually unpleasant, in literature several authors have described this percept in different ways, e.g., muffled by pillow or bee buzz (Simmons et al., 1965), flutter (Eddington et al., 1978), and rattle (Moore & Carlyon, 2005).

The studies related to rate pitch suggest that the upper limit of temporal pitch is around 300 pps (Figure 3.2). A few subjects performed exceptionally well and their upper limit extended up to 1000 pps (Townshend et al., 1987; Wilson et al., 1997), 500 pps (Fearn & Wolfe, 2000), and 900 pps (Kong & Carlyon, 2010). Interestingly, Moore & Carlyon (2005) analysed the rate discrimination results (N = 19) obtained from five earlier studies (Pfingst et al., 1994; van Hoesel & Clark, 1997; McKay et al., 1999, 2000; Zeng, 2002) and reported that the CI recipients are capable of detecting on average a 7.3 % increase in the rate of a 100 pps pulse train. The overall threshold ranged from less than 2 % to about 18%. The variation in the rate detection threshold was attributed to the difference in experimental procedures and a high percentage of inter-subject variability. The similarity among all these studies is that the rate pitch fades off at around 300 pps. The inter-subject variability among CI recipients may be due to several factors. They are etiology, type and duration of hearing loss, percentage of surviving neurons, and electrode insertion depth.

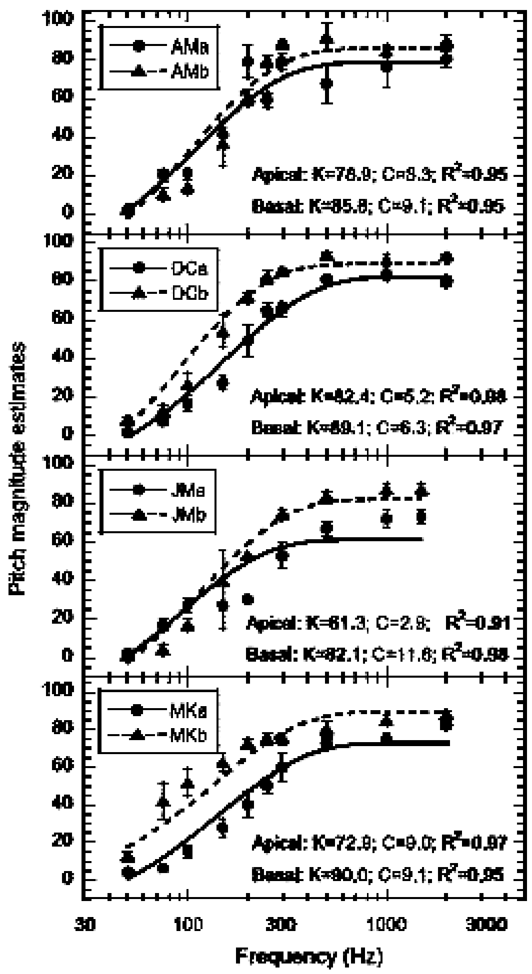


Figure 3.2: Pitch estimation scores for four CI recipients at (a) apical, (b) basal regions. The pitch estimation scores reaches a plateau for frequencies above 300 Hz (Zeng, 2002).

Carlyon & Deeks (2002) attempted to study the limits of temporal pitch in normal hearing. Earlier studies used amplitude-modulated (SAM) noise carrier signal to simulate electrical hearing, but Carlyon & Deeks argued that the inherent modulation present in the noise carrier may elevate the overall modulation more than by the slowly varying AM imposed on a signal. The inherent modulations may interfere with the high rate AM and reduce the depth of modulation. These factors urged the researchers to use a fixed bandpass filter to process harmonic complexes and the output is a filtered pulse train providing only temporal cues in normal hearing. The resolved components present in the harmonic complexes were removed by using the alternative phase complexes and sine phase complexes. The result obtained from the three normal hearing listeners revealed that at high band pass filtering (7800 – 10800 Hz) the subjects were able to detect the difference in F0 up to 712 pps rate for alternating phase complexes. The upper limit of 712 pps was far higher than the upper limit of 300 pps in CI recipients. Contrary to the findings in McKay & Carlyon (1999) study, this study suggests that the deficit seen in CI temporal processing is mediated by peripheral deficit and not attributed to the central coding of pitch. In an additional experiment, a low and high rate stimulus was presented simultaneously to a single ear and when an additional low rate stimulus was introduced in the opposite ear, the subject heard a single tone at the middle of the head. But when a high rate stimulus was introduced in the opposite ear, the subjects were able to discriminate from the single diffused low rate tone. Based on this result, they reported that for normal listeners, there is a central factor which determines the upper limit for rate discrimination at high overall rates. They concluded that there is "*temporal information present in the auditory nerve that is unavailable to the temporal mechanism, but which is accessible when a binaural cue is available*".

The commonly used psychophysical approach for determining DL (difference limen) in a rate discrimination task are; the method of constant stimuli (Fechner, 1966) and the adaptive procedure (Levitt, 1971). The adaptive procedure is used in most instances to reduce the testing timing. However, the pitch reversal phenomenon commonly seen in CI recipients will be unnoticed with the adaptive procedure (Laneau et al., 2004b; Kong et al., 2009; Kong & Carlyon, 2010). So, the procedures should be wisely picked depending on the goals of the study. Kong et al. (2009) used a novel forced choice method (2I-2AFC) to determine the limits of temporal pitch in CI recipients. In this study, rate discrimination was measured for five base rates as a function of large rate difference (ΔR) which remained fixed at 35 % between the test and the reference rates. The base rates were 100, 200, 300, 400 and 500 pps and the signal rates were 35 % higher than the base rates, which were 135, 270, 405, 540, and 675 pps respectively. Although the base rates and the signal rates were loudness balanced, roving was not used in this test. So, there might possibly be some contribution due to learning effect or involvement of non-pitch cues (loudness or brightness). In their first experiment, eight Med-El CI recipients scored well above chance for the five base rates demonstrating a non-monotonic performance. Maximum performance was obtained for medium rates (200 – 300 pps) compared to low (100 pps) and high rates (400 – 500 pps). These scores were compared with the Nucleus CI users and their scores deteriorated after 300 pps and approached to a chance level performance at high rates (Figure 3.3). The rate discrimination scores were compared with the scores obtained from SAM stimuli with a carrier rate of 5000 pps and the scores showed no significant difference between the two modes of eliciting temporal pitch and this coincided with the results obtained from earlier studies mentioned above. The rate discrimination scores were affected with respect to electrode location in four Med-El CI recipients, but the performance trend was not consistent across subject. So, this result does not satisfy the question of whether the accurate place-rate match is a prerequisite for appropriate coding of high rate pitch.

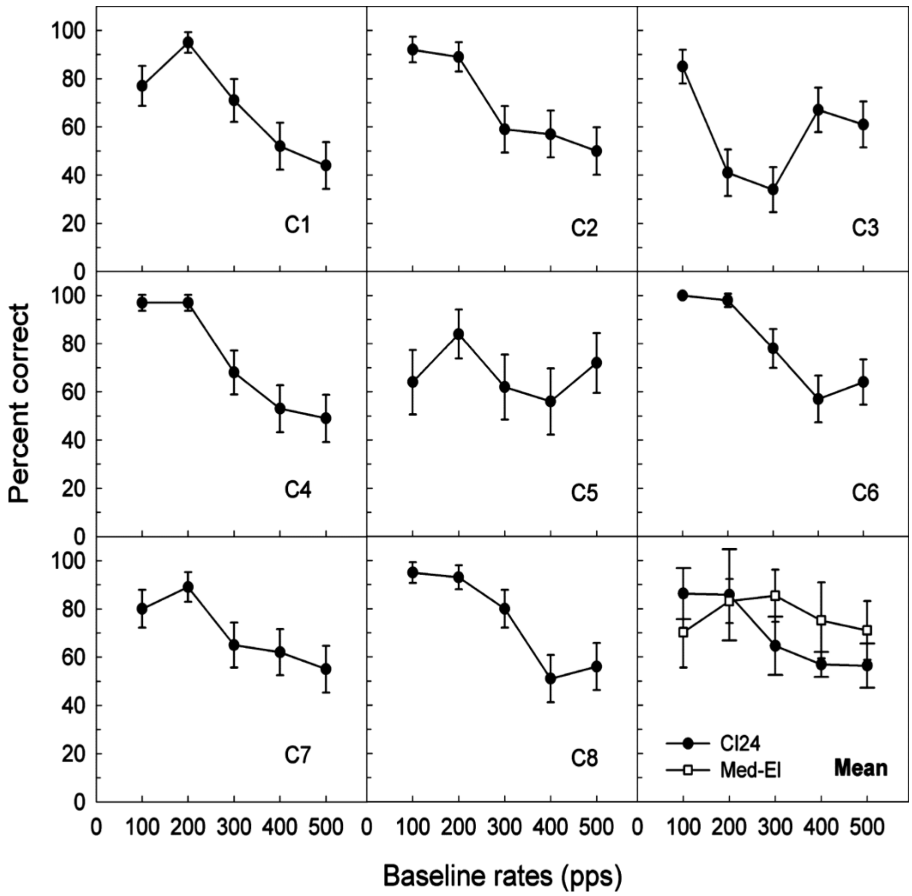


Figure 3.3: Rate discrimination scores for Eight Nucleus CI recipients. The mean scores of Nucleus CI recipients were compared with Med-El CI recipients for the rate discrimination task (Kong et al., 2009).

An extension study was reported by Kong & Carlyon (2010) to further investigate the upper limit of temporal pitch in CI recipients. Six Med-El subjects who participate in the earlier study (Kong et al., 2009) participated in this study as well. The rate discrimination task was similar to their previous study, but the only difference was that no feedback was provided in this study. In the pitch ranking task, they applied the method of midpoint comparison which was initially developed by Steinhaus (1950) and adapted by Long et al. (2005). In this task, the subject had to order the pitch of a stimulus in an orderly fashion. Listeners were asked to identify the higher stimulus when two pairs of stimuli are presented randomly. In the first trial, if the test stimulus is higher than the reference stimulus, then the algorithm will place the test stimulus at a higher rank order. In the subsequent trial, a high ranked stimulus will be compared with the new stimulus. If the stimulus was judged higher or lower, then it will be placed higher than the test stimulus or lower than the test stimulus. In this way, they can arrange the set of stimuli in an orderly manner. The results obtained from these two experiments reveal that the CI recipients can rank and order pitch well above 300 pps which is contradictory to the earlier studies (Shannon, 1983; Tong et al., 1983; McDermott & McKay, 1997; McKay et al., 2000; Zeng, 2002). Additionally, the two “Star” CI performers were able to detect temporal pitch changes up to 900 pps. Pitch reversals were also reported in a few instances at rates above or below the CI recipients’ upper cut off limits. The multidimensional scaling analysis reveals that the pulse rates (temporal pitch) and place of excitation (place pitch) produced independent and separate pitch percepts which are consistent with the earlier findings (Tong et al., 1983; McKay et al., 2000).

Carlyon et al. (2010) investigated the upper limits of temporal pitch by varying the signal duration of stimuli in the first experiment, adding a high rate (5000 pps) conditioning pulse to the existing base rates in the second experiment, and finding the effect of concurrent electrode stimulation as a function of rate in the third experiment. In the first experiment, the pulse rate was turned on abruptly for durations of 200 ms or 800 ms or 800 ms turned ‘ON’ and 300 ms turned ‘OFF’ ramps. The rate discrimination task, similar to the above study (Kong et al., 2009), measured performance at base rates of 100, 200, 300, 400, and 500 pps. The feedback was provided at the end of each trial. The scores for the six Nucleus CI recipients showed a traditional monotonic decline in performance as the rate increase above 300 pps for the three stimulus durations. Although there was a slight advantage for the shorter (200 ms) stimuli, the overall performance was similar across the three stimulus duration (Figure.3.4). The motivation for the second experiment was that the “alternating-amplitude” pattern of nerve spikes evoked by electrical hearing may impair pitch perception at high rates. Therefore, adding a background 5000 pps conditioning pulse abolishes the alternating-amplitude ECAP pattern (Rubinstein et al., 1999). The rate discrimination scores were determined with and without conditioning pulses at base rate from 100 – 500 pps. Six CI recipients (4 Nucleus CI users and 2 Med-El CI users) showed no superiority for the conditioner pulses compared to no-conditioner stimuli. In the third experiment, single and multiple electrode (E8 to E14) patterns as a function of rate, and spectral profile were evaluated using the midpoint comparison procedure (Long et al., 2005; Kong et al., 2009). The result showed that the performance was similar across single electrode, *flat* profile (Individual ‘C’ level was established for each electrode from E8 to E14, while stimulating all the 7 electrodes, a global correct factor was used to compensate for the combined loudness of 7 electrodes) and *peaked* profile (predetermined attenuation current levels were applied to the 7 electrodes and then overall ‘C’ level was established). Statistically, there was no significant difference among these profiles. The multiple electrode performance was not better than the single electrode performance. In all the three experiments there was an inter-subject variability seen across subjects. In conclusion, they suggest that these stimulation patterns did not affect the auditory nerve response evidently nor assisted in improving the rate discrimination at high rates. Therefore, these limitations of temporal pitch at high rates are not specific to temporal patterns of AN activity.

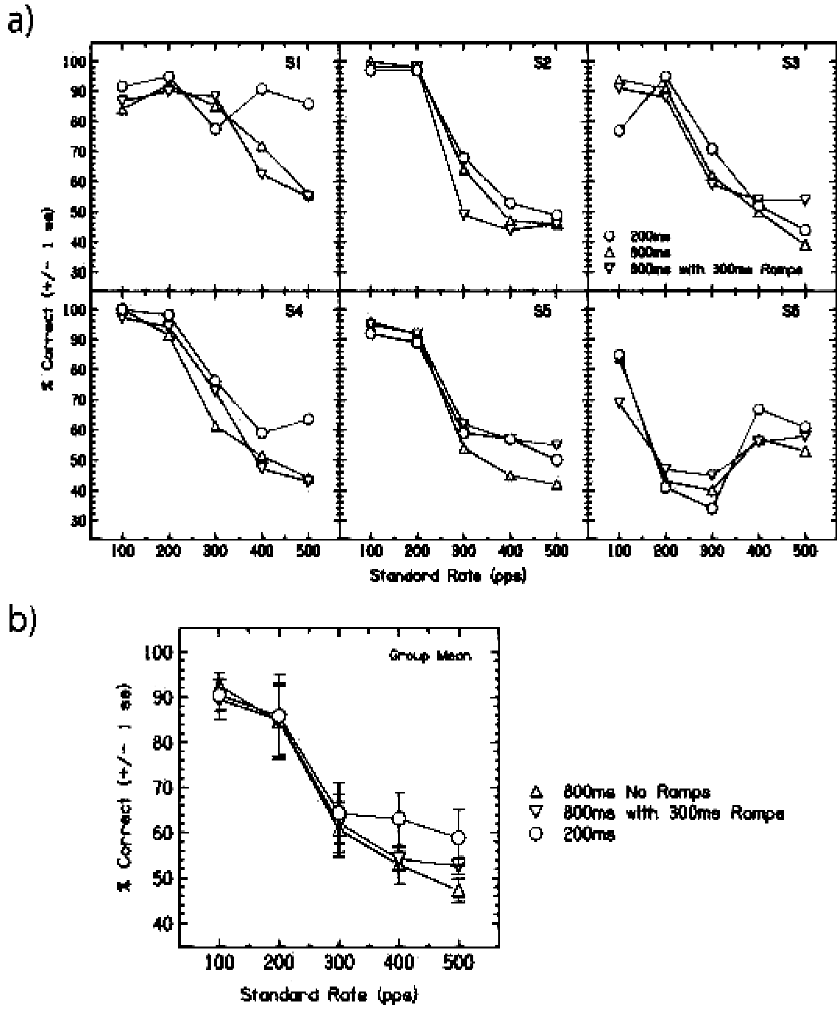


Figure.3.4: (a) Individual rate discrimination scores for eight CI recipients using three signal durations. (b) The mean rate scores for the three different pulse durations (Carlyon et al., 2010).

Recently attempts have been made to extend the limits of temporal and place pitch by using different kinds of bipolar stimulation (Macherey et al., 2011). Different asymmetrical pulses were tested in bipolar mode. One of the asymmetrical pulses was pseudomonophasic pulses (PSA). These PSA pulses have a short, high-amplitude anodic phase relative to the most apical electrode. The anodic short phase is salient, because it excites fibres very near to the apical electrodes compared to other electrodes. Earlier studies have reported that the short anodoic pulses are more effective than the long cathodic pulses (van Wieringen et al., 2008; Undurraga et al., 2010). Seven Advanced Bionics CI recipients were tested on several temporal and place pitch tasks. The subjects were able to significantly detect changes as high as 713 pps while using bipolar PSA pulses at apical region. The MDS results demonstrated that the temporal pitch at high rates were independent to the place of excitation.

The above studies can elucidate that for the majority of CI recipients, the upper limit for temporal pitch is around about 300 pps. Recent studies have also shown that some CI users were able to consistently detect changes at high rates above 300 pps (Kong et al., 2009). The study using pseudomonophasic pulse for bipolar stimulation especially in apical regions looked very promising for extending upper limits for temporal pitch. Extending the upper limits for temporal pitch depends on the number of “star” performers in the study and the mode of stimulation as in the case of PSA pulses. Interestingly, there was a significant difference in performance between Med-El & Nucleus CI users (Kong et al., 2009), but more studies are required to ascertain this finding. The extension of the upper limit for temporal pitch is crucial for processing fine structure information in CI recipients (Wilson et al., 2004; Green et al., 2005; Nie et al., 2005; Laneau et al., 2006).

Normal listeners are remarkably efficient in identifying the rate changes up to 600 – 700 pps which is much higher than the average CI counterparts (Carlyon & Deeks, 2002). This is consistent with the physiological data of the auditory nerve which showed highly synchronization up to 800 pps for electrical pulse train (Hartmann et al., 1984; van den Honert & Stypulkowski, 1984; Javel et al., 1987; Shepherd & Javel, 1997). Although the pulse trains are tightly phase locked in electrical hearing as opposed to acoustic hearing, this does not provide any kind of advantage for better coding of timing information in CI recipients and this paradox was clearly explained by Carlyon & Deeks (2002). The above studies explored temporal pitch using rate discrimination or ranking tasks. Alternatively, temporal pitch can also be explored in a musical context. The variation in pulse rate does indeed convey melody, therefore the next section will address the topic of exploring temporal pitch in a musical context.

3.4.1.1. Exploring temporal pitch mediated by musical pitch

Eddington et al. (1978) was the first study to demonstrate that CI temporal pitch can convey musical pitch information. Five commonly used melodies without rhythm cues were selected and played as varying pulse rates at a single electrode (E3). One subject was able to spontaneously recognise 3 out of 5 melodies ("Mary Had a Little Lamp", "Yankee Doodle", and 'Twinkle Twinkle Little Star'). The subject was unable to recognise the remaining two melodies and this was attributed to the mismatch between the pitch intervals mediated by pulse rates and the musical scale. In the subsequent year, Fourcin et al. (1979) used two approaches to test melodic pitch in CI users. In an informal test, the experimenter's voice was low-passed at 300 Hz and presented via a constant-current pulse train to an electrode. The commonly used melodies with minimal tempo cues were provided to a single subject. The subject was able to recognise all the melodies with ease and reported that the notes were in-tune. The subject described the quality of sound as "comb and paper". In a formal test, two melodies were selected which consists of equal numbered and equal duration tones to avoid rhythmic cues. The subject who participated in the previous informal testing participated in this test and performed well. However, there was a large intersubject variability seen across three unilateral subjects who performed this task. The overall conclusion was that the electrical pulse train does evoke a sensation of pitch and in some instances can convey musical pitch information. These earlier attempts provided evidence that temporal pitch in isolation can convey melodic pitch information.

In a widely quoted study, Pijl & Schwarz (1995b) investigated cochlear implant melody and interval recognition using a bipolar mode of stimulation. The melodies were presented as varying pulse train and the pulse rate corresponded to the fundamental frequency (F0) of each note in the melodies (F0 range = 75 – 600 pps). Seventeen subjects obtained a mean score of 44 % in an open set melody recognition task with rhythm intact. CI users found the open set paradigm extremely arduous. In a closed set melody recognition (no rhythm) task, subjects (N = 3) secured almost 100 % score at low pulse rates and the performance deteriorated as the pulse rate increased. The subjects were able to recognise the melodies up to 600 pps. Moreover, the performance was significantly superior ($p < 0.00001$) for apical electrodes compared to basal electrodes. Subjects reported that the melodies were more musical and pleasant when stimulated at the apical electrode. In the same study, the same three subjects were asked to label the musical intervals with respect to pulse rates on a single apical electrode (E18). Results showed than the subjects were able to accurately label the musical interval (Minor 3rd, 4th, 5th, and Major 6th) by using their memory. Investigators reported inter subject variability in this task and attributed this to the difference in electrically based pitch percept and also to the internal representation of musical intervals in the memory of these CI recipients. In an extension study, Pijl & Schwarz (1995a) tested three CI recipients who had a strong inclination towards music and one of them received violin training during his childhood. In one of the experiments, subjects were asked to reconstruct the melodic interval (5th, 4th, and minor 3rd) by adjusting the electrical pulse rate on a single electrode (E18). Subjects were able to tune the musical intervals abstracted from familiar melodies especially for low pulse rate. The pulse rate ratio was in close proximity to the actual frequency ratio of the musical intervals. In an additional experiment, two subjects were able to transpose the musical intervals (5th, 4th, and minor 3rd) either higher or lower pulse rates. Overall subjects were able to tune the musical intervals mediated by pulse rates and this was analogous to the acoustic musical intervals especially for low pulse rates. The work by Pijl and his colleagues (1995a, 1995b; 1997) elucidate that temporal pitch alone is quite sufficient in conveying adequate melodic pitch information in CI recipients.

An interesting single-subject study by McDermott & McKay (1997), examined a CI subject who was a piano tuner prior to implantation, to judge musical intervals based on method of adjustments. In the interval production procedure, the subject had to adjust the pitch pulse rate or modulation frequency on an electrode to a specified musical interval by using an unmarked knob. In the interval estimation procedure, the subject had to name the musical interval between the two tone sequence. The stimuli were presented at different electrode locations (Apical – E18, Middle – E12, Base – E5) using a bipolar mode of stimulation. The subject, to an extent, was capable of accurately determining the musical intervals. This result was analogous to the previous study and warrants that pulse rate or modulation frequency alone is capable of providing musical pitch information, but to a limited range of two octaves.

Recently, Swanson & McDermott (2010) used the Modified Melodies test to investigate cochlear implant place pitch and temporal pitch perception using direct stimulation. Three conditions were tested (Place C5 – Pure tones with octave starting from 523 Hz; Place C3 – Pure tones with octave starting from 131 Hz, and Rate C3 – Pulse rate equal to the fundamental frequency of each note, octave starting from 131 Hz). The Modified Melodies test supports several types of pitch modification, but in this study investigators used the nudge modification. In this modification, one note of the melody was shifted away from its correct pitch by a specified number of semitones. Blocks of trials were performed with shifts in the range 0.5 semitones to 7 semitones. Results showed that three CI users were able to recognise a correct version of a melody using temporal cues in isolation. The results obtained from the above studies warrant that temporal pitch in isolation can convey both melodic contour and interval size information of a melody in CI recipients.

3.4.2. Place Coding in Cochlear Implants

In a multichannel cochlear implant, each electrode evokes different pitch percepts and these percepts correspond to natural tonotopic organisation in the cochlea, i.e., the basal electrodes correspond to high frequency and the apical electrodes correspond to low frequency. Therefore, the cochlear implants mimic the natural tonotopic organisation of the cochlea. The pitch percepts associated with different locations along the cochlea was documented even in the early studies (Simmons et al., 1965; Eddington et al., 1978; Tong et al., 1982; Shannon, 1983; Twonshend et al., 1987) and it is well established that different intracochlear electrodes evoke different pitch percepts ranging from *low* to *high* (Figure 3.5).

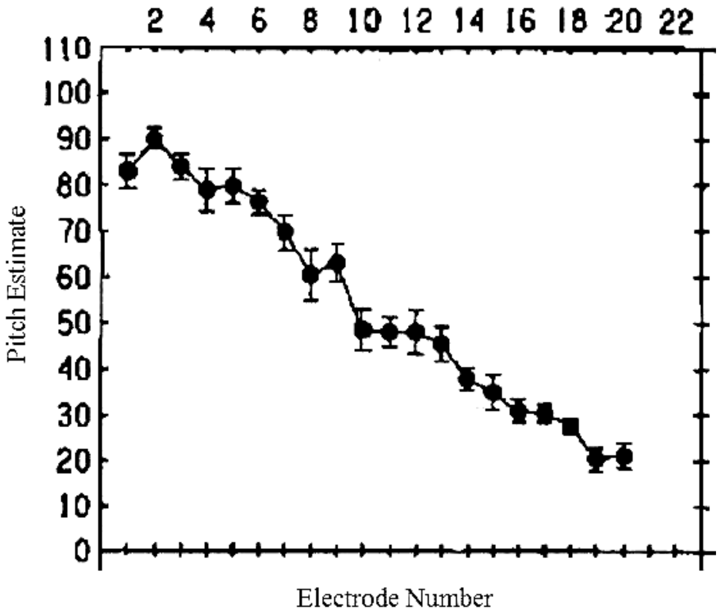


Figure 3.5: The pitch estimation performance of a single CI recipient in a pitch scaling experiment (Cohen et al., 1996b).

A classical study by Townshend et al. (1987) investigated the place pitch using pitch ranking procedure in three CI recipients and reported a mixed performance (low to high). Although there was a large inter-subject variability across subjects, overall performance revealed that the subjects were indeed listening to pitch across a single dimension (place pitch). Additionally, they investigated pitch perception as a function of dual electrode stimulation across these CI recipients. The intermediate pitch can be evoked by dual electrodes and produced only place pitch percepts across different locations along the electrode array. The intermediate pitches were varied depending on the ratio of current delivered to the two electrodes. The forward masking procedure was used to investigate the neural excitation pattern at different electrode locations (Cohen et al., 1996b; Chatterjee & Shannon, 1998) and reported different forward-masking patterns for different electrode locations.

Nelson et al. (1995) investigated fourteen CI recipients whose task was to pitch rank electrodes as a function of different locations along the cochlea. The subjects were presented with two sequences of 500 ms biphasic pulse train separated by a 500 ms silent interval. The subjects' task was to select the sequence with higher pitch. The electrode ranking procedure was tested for three spatially separated electrodes: E2 & E3 (0.75mm), E9 & E11 (1.5 mm), and E16 – E20 (3.0 mm). The correct scores were converted into d' per mm scores (refers to the distance between two electrodes) and a few subjects were able to perfectly pitch rank stimuli differences as small as 0.75 mm, while some subjects were only able to pitch rank stimuli when the electrodes were significantly apart (13mm). The pitch ranking scores were influenced by the locations inside the cochlea, i.e., ranking performance was better in the apical region compared to the basal region of the cochlea. The cumulative d' scores showed a monotonic increase in pitch from the basal to the apical region of the cochlea. Interestingly, there was an additive place pitch effect seen only in a few subjects who scored poorly in a ranking task. In an additive effect, the individual d' scores obtained from E5 & E6 and E6 & E7 was equal to the collective sum of d' scores obtained from E5 & E7. The additive effect does not hold true for high scoring subjects who attained maximum scores (ceiling effect). Although there was a large intersubject variability seen across subjects (0.12 d' per mm – 3.16 d' per mm), postlingually deafened subjects performed better than prelingually deafened subjects and pitch reversals were also reported on a few occasions. The results obtained from this study was consistent with the performance obtained from Busby et al. (1994) study who used another procedure (pitch estimation technique) and varied the electrode configuration (MP, BP, and CG). The results showed that the place pitch percept in electrical hearing was tonotopically organised from the apex to the base of the cochlea.

The Nucleus devices consist of 22 electrodes which produce 22 distinctive place pitches. Apart from the traditional way of stimulating individual electrodes to evoke place pitch, additional place pitches can also be elicited by simultaneous stimulation (Townshend et al., 1987; Wilson et al., 1993; Donaldson et al., 2005; Busby et al., 2008) or sequential stimulation (McDermott & McKay, 1994; Kwon & van den Honert, 2006) of adjacent electrodes. When a dual electrode pair was stimulated simultaneously, a single CI recipient using Ineraid implant was able to discriminate 25 % increments between electrodes which were 4 mm apart (Wilson et al., 2003).

Donaldson et al. (2005) investigated the place pitch sensitivity in six Clarion CI recipients (HiFocus or the HiFocus II electrode array) for single and dual electrode simultaneous stimulations. The dual electrode stimulation (Virtual channel stimulation) can be created by simultaneously delivering current to the two physical electrodes. Therefore, by adjusting the current weighting, the peak of the excitation pattern is steered at different sites between the two physical electrodes. The amount of current delivered to the basal electrode of a dual pair is represented as α. When α is 0, current is delivered to the more apical electrode. When α is 1, current is delivered to the more basal electrode. A two alternative forced choice task using a two down, one up adaptive procedure was used to estimate the place pitch discrimination thresholds. The results (d' scores) showed a monotonic increase in performance as α increases from 0 to 1. There was a good agreement between the results obtained from the psychometric function and the adaptive procedure. When the stimuli were presented at medium loud level, subjects reported to have better pitch percepts compared to presentation at medium soft level. Interestingly, when the current was steered between the two physical electrodes by means of linear interpolation, the loudness level for dual electrode stimulation was equal to the sum of the loudness levels obtained from the two individual electrodes. In conclusion, they reported that, the place pitch discrimination was apparent for 16 out of 17 electrode pairs and almost two to nine place pitch percepts were plausible with the dual electrode stimulation. In an extension study, Firszt et al. (2007) investigated 106 postlingually deafened CI recipients (115 Ears) from twelve centres across North America. The aim of this study was to investigate the CII or 90K cochlear implant users to hear additional spectral channels using current steering method. The method was similar to Donaldson et al. (2005) and they tested 3 electrode pairs (E2 & E3 (Apical), E8 & E9 (Middle), and E13 & E14 (Basal)) in different locations of the cochlea. Results showed that the subjects were able to perceive additional pitch percepts produced by current steering method. The number of discriminable pitch sensations for basal pair, middle pair, and apical pair were 3.8, 6.0, and 5.3. Thus, the number of discriminable pitch sensations varied as a function of electrode location. Overall, the potential number of discriminable pitch sensations across the entire array ranged from 8 to 451 with a mean score of 63. The mean score was four times greater than the number of physical electrodes present in the implant array and they noticed large intersubject variability among CI users. These two studies provide clear evidence that simultaneous stimulation of dual electrodes can evoke spectral pitch cues in CI recipients. A recent electrophysiological study (Snel-Bongers et al., 2012) has shown that the spread of excitation (SOE) and channel interaction was similar across single electrode stimulation and simultaneous stimulation of dual electrode and good correlation was seen across the dual electrode stimulation and channel interaction in twelve HiRes 90K implant users.

In an investigation by Luo et al.(2010), seven postlingually deafened CI recipients (8 Ears) identified pitch contours as a function of time varying virtual channels. In a pitch contour identification (PCI) task, nine pitch contours were created by steering the current between the two physical electrodes or halfway (a virtual channel) between the two physical electrodes. The first pulse was presented either to the apical part of the electrode pair ($\alpha = 0$) or the basal part ($\alpha = 1$) or a virtual channel halfway between the two electrodes ($\alpha = 0.5$). In the first PCI experiment, the three electrode pairs (apical, medial, and basal) were tested at five stimulus durations (100, 200, 300, 500, and 1000 ms). The PCI task is a 9 alternative forced choice procedure, where a subject listens to each stimulus and clicks on the appropriate pitch contour sign on the screen. In the second PCI experiment, only three contours were used (flat, falling, and rising). Although, the 9-contour PCI (9 AFC) task was strenuous as compared to the 3-contour PCI task, the performance was similar across the two PCI tasks and there was no significant difference among these tasks. The scores were better for high stimulus duration and there was no effect of electrode locations. The subjects were able to discriminate VC (virtual channels) and the cumulative d' scores ranged from 0.5 – 4.4. The VC discrimination scores varied greatly among the subjects. Interestingly, there was a strong correlation between the cumulative d' scores and the 9-contour PCI scores, but it is also noteworthy that there was a significant correlation between cumulative d' scores and 100, 200, and 500 ms 3-contour and 9 contours PCI scores. These results add to the existing literature confirming that the simultaneous stimulation of the virtual channels does indeed convey pitch information.

The simultaneous stimulations of the dual electrodes encounter potential problems of channel interaction when the current fields are summated inside the cochlea. The channel interaction can smear the spectral resolvability and reduce the performance in CI recipients (Fu & Nogaki, 2005). The channel interaction can be reduced by sequential stimulation which reduces the current field summation and interactions usually take place at the neural level. An additional way of reducing the channel interaction is by using focused electrode stimulation like tripolar (TPS) or partial tripolar (pTP) stimulation (Bierer, 2007; Landsberger et al., 2012).

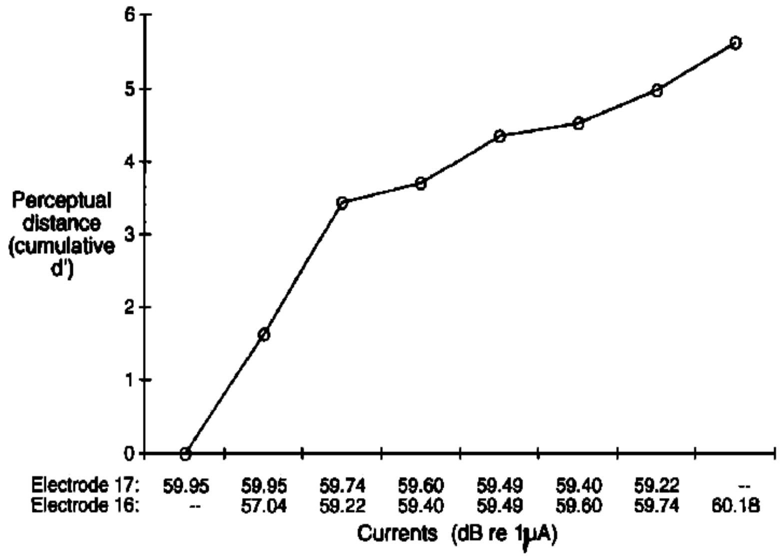


Figure 3.6: Six dual-electrode stimuli were created between E17 & E16 with relative current levels in a single Nucleus recipient. The pitch discrimination scores (2-AFC) were plotted as a cumulative d' perceptual distance measure. There is a monotonic increase in scores as the current levels vary between the two electrodes (McDermott & McKay, 1994)

In a classical study, McDermott & McKay (1994) reported non-simultaneous dual electrode stimulation in five experienced postlingually deafened Nucleus recipients. All pulse trains had a duration of 4 ms and for dual electrode stimulation there was a 0.4 ms gap between the two biphasic pulses. The result obtained from the 2-AFC pitch ranking procedure (N = 4) demonstrates that the pitch elicited by the dual electrode stimuli was ordered intermediately between the two adjacent electrodes. The intermediate pitch of a dual electrode can be varied in accordance with the proportion of current (α) delivered to these physical electrodes (Figure 3.6). In some instances, they reported pitch reversal. The intermediate pitch is created by the centroid (geometric centre) of the combined spatial distribution between the two adjacent electrodes and probably creates a single, broader neural excitation pattern. This is supported by the multidimensional result obtained by McKay et al.(1996), who reported that, the dual electrode stimuli in the 'stimulus space' depends on the distance between the two adjacent electrodes along the electrode array. The MDS results showed that the dual electrode stimuli have two dimensions. Firstly, increasing the width of current creates more overlapping areas and hence it is not plausible to stimulate distinct set of neurons for transmitting information. Secondly, the variations in the temporal delay between the two electrode pair. For sequential dual electrode stimulation, the timing between the two biphasic pulses is very crucial for accurate coding of information.

Kwon & van den Honert (2006) investigated pitch discrimination as a function of dual electrode sequential stimulation in Nucleus CI24 users. Eleven CI recipients (12 ears) were instructed to discriminate pitch between the three electrode pairs (E19 & E18, E19 & E18, and E4 & E3). The mean pitch sensitivity scores d' was 7.26, 3.97, and 2.97 for apical (E19 & E18), middle (E19 & E18), and basal (E4 & E3) electrode pairs respectively. The overall d' scores ranged from 0.7 to 9.6. The performance varied largely across individuals, nonetheless there was better pitch sensitivity at the apical region which was also reported by McDermott & McKay (1994).

Busby & Plant (2005) reported that in dual electrode sequential stimulation (N = 8), intermediate pitch percept of around 43 channels were plausible from 22 physical electrodes in Nucleus freedom implants (CI24RE). Interestingly, the SOE (spread of excitation) function for single and double electrode stimulation were very similar using ECAP (Evoked Compound Action Potentials) across different electrode locations and there was no significant difference between the single and double electrodes in pitch ranking task (Busby et al., 2008). Saoji et al.(2009) reported that the SOE was similar across the simultaneous dual electrode stimulation and the intermediate physical electrodes. Conversely, the SOE was not similar for sequential dual electrode and the intermediate physical electrode and this relied on the pulse sequence order which was seen in seven Advanced Bionics CI users.

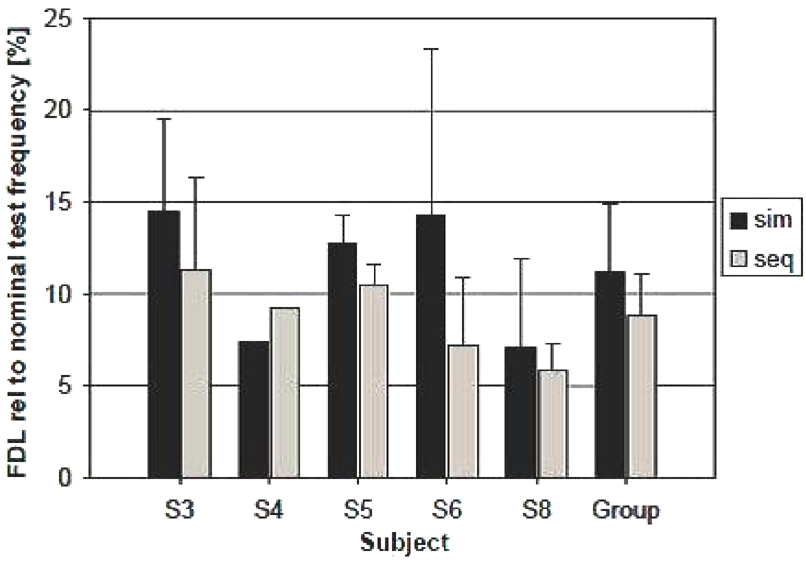


Figure 3.7: Frequency difference limen (FDL) scores using simultaneous and sequential stimulation in Med-El CI recipients. The error bars represent standard deviation of the mean (Nobbe et al., 2007).

Nobbe et al.(2007) investigated eight Med-El postlingually deafened CI recipients in a frequency discrimination task using both sequential and simultaneous stimulation patterns. The frequency difference limen (FDL) for sinusoids was obtained using three –

down one-up two-interval two-alternative adaptive forced choice procedure. Pitch reversals become unnoticed when an adaptive procedure is used, so a non-adaptive procedure should always be preferred and it also depends on the research question. The mean JNDs of 8.8%, 11.2% were obtained for sequential and simultaneous stimulation respectively. However, no significant difference was seen across the two modes of stimulation and there was a larger intersubject variability seen across subjects and also among electrode pairs (Figure 3.7). The intermediate pitch can be elicited by amplitude weighting between the two adjacent electrode pairs using simultaneous or sequential stimulation.

The discrimination ability of seven Advanced Bionics CI recipients using simultaneous and sequential virtual channels was investigated recently by Landsberger & Galvin (2011). The discrimination was measured for both monopolar (MP) and bipolar (BP + 1) modes. The interpulse interval (IPI) varied between 0.0 to 1.8 ms for sequential virtual channels. In the virtual channel discrimination task (3IFC), subjects were able to discriminate between both sequential and simulation VCs and the scores were well above chance. It should be noted that the scores were not influenced by either the IPI or the stimulation mode (MP or BP +1). The discrimination result was correlated with previous studies and concluded that the discrimination of simultaneous or sequential virtual channels was influenced by subjects' spatial selectivity. The difference in scores obtained from the sequential and simultaneous VCs may be attributed to the SOE and this was apparent from the previous study (Saoji et al., 2009) which reported a relative shift in peak for forward masking functions in sequential VCs compared to simultaneous VCs. Moreover, the direction of shift depends on which electrode was stimulated first in sequential VCs. The above mentioned studies provide clear evidence that both simultaneous and sequential stimulation does evoke intermediate pitches in CI recipients.

A different approach to evaluate the place pitch sensitivity was introduced by Laneau & Wouters (2004a). They investigated the place pitch sensitivity as a function of number of active channels in four postlingually deafened CI recipients. Pitch ranking task (2I-2AFC) was measured at two different locations (E17 and E13) for all the subjects except one (E6 was additionally tested) and there were six active channels (1, 2, 3, 4, 5, and 8). In each trial, two stimuli were presented and the subject's task was to identify the stimulus with higher pitch (excluding the loudness cue). The reference stimulus was centered on the reference electrode and the variant stimulus was shifted 1 or 2 electrodes towards the apex or base. Prior training was provided before the commencement of the test. The result showed excellent place pitch sensitivity in monopolar mode and the average JNDs (pitch ranking task) expressed in electrodes ranged from 0.34 to 0.61 which was significantly smaller than the Nucleus physical electrode space (0.75mm apart). Most importantly, the place pitch sensitivity was not dependent on the number of active channels. The large spectral overlap with respect to active channels was found to have no influence on the place pitch sensitivity. The results were consistent with the centroid model of place pitch described by Laneau (2005). According to this model, the place pitch is determined by the centroid (centroid of the gravity) of the stimulation pattern. Additionally, the result obtained from Cohen et al. (1996a) supported the centroid place pitch model and demonstrated that the centroid of the forward masking distribution varied systematically with respect to pitch estimation threshold.

Earlier studies have reported that place pitch in CI is in close proximity to the brightness aspect of timbre as opposed to pitch (McDermott & McKay, 1997; McDermott, 2004; Moore & Carlyon, 2005; Swanson et al., 2009). The brightness of a harmonic tone depends on the centroid of spectral profile, i.e., tones having strong high harmonics sound brighter (Plomp, 1976; Schubert & Wolfe, 2006). The ability to rank stimuli on a low-to-high scale does not differentiate between brightness and pitch. An operational definition of pitch is that variations in pitch can convey a melody (Moore & Carlyon, 2005), therefore a musical (Melody perception tasks) approach was desirable to explore CI pitch perception skills (this will be discussed in more detail in the following chapter).

Swanson et al. (2009) investigated cochlear implant place pitch using the Modified Melodies test (4.2.2.3). The melodies were presented as pure tones with frequency ranging from C5 – C6 (532 -1046 Hz) through the ACE strategy on a freedom processor. The ACE strategy with quadrature envelope detection produces very little ripples for pure tones, thereby minimizing the amplitude modulation and therefore, the processed signal consists of place pitch information only. Using place-pitch cues alone, all six subjects were able to recognise an incorrect melodic contour. Four out of six subjects were able to detect a large (five-semitone) error in the size of a musical interval, and two subjects were able to detect a smaller (two-semitone) error. The results from the two studies (Swanson et al., 2009; Swanson & McDermott, 2010) suggested that place-pitch cues can provide melodic pitch. Swanson et al. argued that the subjects were indeed listening to pitch changes, but the contribution of brightness could not be completely ruled out and this was the main need of chapter 6 which investigated the role of brightness in different pitch perception tasks and addresses the issue of whether the brightness percept has the potential to convey melody or not.

3.4.3. Interaction between Place Pitch and Temporal Pitch Perception

Cochlear Implants provide an excellent platform to investigate the individual contribution of place and temporal pitch in CI recipient. The independent nature of place and temporal pitch was first reported by Tong et al.(1983) using a multidimensional scaling (MDS) method. In this method, the multiple pulse per second (MPP) was used to create nine perceptually different stimuli consisting of three electrode locations (E2, E3, and E6) and three pulse rates (83, 125, and 250 pps). All nine stimuli were loudness balanced and the subject had to identify the least similar stimuli presented in an AAB sequence. The dissimilarity matrix was analysed using a MDS procedure and revealed a two dimensional representation specific to electrode location and pulse rate (Figure 3.8). The result demonstrated that the two dimensions (place and rate) are independent and perceived differently.

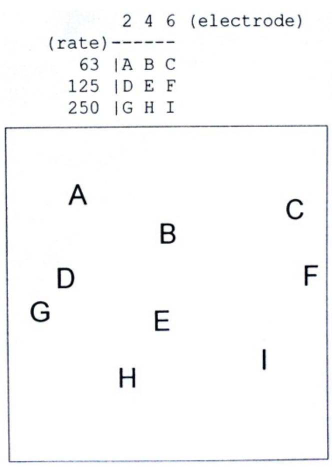


Figure 3.8: Multidimensional scaling results for the MPP (multiple pulses per period) stimuli (Tong et al., 1983) (Retrieved from: McKay (2004)).

McKay et al. (2000) reconfirmed the above finding by using a different approach. In their first experiment, four CI recipients participated in a discrimination procedure (3I-2AFC) consisting of eight test stimuli (increase and decrease in rate of stimulation, apical and basal place change only, basal change with increase or decrease in pulse rate, and apical place change with increase or decrease in pulse rate). The place pitch percept was generated by sequential dual electrode stimulation. The subjects' performance was plotted as a ratio between the predicted d' and obtained d'. The mean score for the four conditions (Basal/Rate Up, Basal/Rate Down, Apical/Rate Up, and Apical/Rate Down) did not vary more than d' = 1.0. The result showed that the performance of consistent cues (Basal/Rate Up, and Apical/Rate Down) was better than the performance of inconsistent cues (Basal/Rate Down, and Apical/Rate Up). This suggests that CI recipients were able to combine the consistent rate and place cues effectively. However, statistically, the scores failed to reach a significant level. In the second experiment, high rates especially at basal electrodes were investigated using consistent and inconsistent cues. The d' scores for the five conditions (Basal alone, Basal 500 Hz, Basal 600 Hz, Basal/Rate up (500 to 600 Hz), and Basal/Rate Down (600 to 500 Hz)) demonstrated that, there was no specific advantage for consistent cues (Basal/Rate up (500 to 600 Hz)) over the inconsistent cue (Basal/Rate Down (600 to 500 Hz)). The subjects were able to effectively combine the concurrent place of excitation and pulse rate cues. Finally, these two experiments suggest that the place of stimulation and pulse rate are two independent perceptual dimensions.

Although these two dimensions (place and rate cues) are perceptually independent in CI recipients, there are a few instances where they can combine these dimensions effectively. For e.g., CI recipients may perceive the high pulse rate tone provided to the apical electrodes (low frequency region) similar to the low pulse rates tone delivered to the basal electrodes (high frequency region). Therefore, the overall pitch perceived may be an integration of both place and temporal cues to pitch. In order to establish adequate pitch perception in normal hearing, temporal information should be presented in the appropriate tonotopic region of the cochlea (Oxenham et al., 2004).

A recent study (Luo et al., 2012) explored the interaction of place and temporal pitch in a PCI (pitch contour identification) experiment. The PCI was evaluated in three modules. Firstly, the place pitch alone was investigated by steering the current between the two adjacent electrodes. Twelve pitch contours (six rising & six falling contours) were created with current steering (α ranges from 0.2 – 0.8). They used single or simultaneous dual electrode stimulation. Five postlingually deafened Advanced Bionics recipients were tested using monopolar mode for an apical electrode pair (2 -3) or a middle electrode pair (7 – 8). Secondly, temporal pitch alone was investigated by using the time-varying AM frequencies with 30 % AM depth. Similar to place pitch contours, there were twelve temporal pitch contours created by varying the AM pulse train. In a PCI (2AFC) task, the percentage of raising pitch contour responses were recorded with respect to current steering ratio in case of place pitch or AM frequency in case of temporal pitch. The scores improved as a function of Δα (place pitch) or AM frequency changes and there was no effect of electrode location (Apical or Middle) in either place or temporal tasks alone. The slopes of the psychometric functions illustrated that there was a strong correlation between Δα and AM frequency changes. Thirdly, the PCI was evaluated as a function of consistent and inconsistent place pitch and temporal pitch and reported a significant performance for consistent cues rather than the inconsistent cues. This finding was similar to the previously mentioned study (McKay et al., 2000). Finally, results showed a significant integration between the place pitch and the temporal pitch in a dynamic pitch perception task. They also recommended better coordination of current steering and AM frequency changes for better representation of pitch. It would be interesting to compare the PCI results obtained from single and dual electrode stimulation with the multielectrode stimulation.

The aforementioned studies reported the relationship between the place and temporal pitch in CI recipients. Anecdotally, when the same information is conveyed either as place pitch or temporal pitch, CI recipients may prefer temporal cues over place cues or vice versa. For example, when a melody is presented first as exclusive temporal pitch information, CI recipients may perceptually prefer melodies with temporal cues over place cues. It is also plausible that CI recipients may prefer exclusive place cues over temporal cues as well. Therefore, it is vital to deliver both the cues to the CI recipients accurately. Moreover, it is pivotal to understand the independent nature of these two percepts, but on the other hand, certain constrain (music perception, speech in noise, etc.) could possibly be surmounted by adequate integration of place pitch and temporal pitch in CI recipients.

3.5 Loudness Perception

Loudness is a psychological correlate of physical strength (amplitude) and defined as “that attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud” (ANSI, 1973). In electrical hearing, as the current amplitude increases, there is a steady increase in the loudness perceived by the CI recipients. The physiological studies provide evidence that the loudness of the sound is correlated with the number of neurons activated in the auditory nerve (Abbas & Miller, 2004). As soon as the initial ‘Switch ON’ is made in the CI recipients, one of the clinical protocols is to establish the C (comfort) level and T (Threshold) level. The difference between the C & T level is called the dynamic range (DR) usually represented in clinical units (roughly about 8 dB between C & T levels). In electrical hearing, the loudness grows much steeper than the acoustic hearing and the DR is substantially reduced in electric hearing as compared to acoustic hearing (DR: Approximately 120 dB in Normal subjects). This may be attributed to the compression function of cochlea seen in normal hearing and absent in CI recipients (Zeng & Shannon, 1994). The variation in the electrical dynamic range depends on intersubject variability, electrodes, and stimulus parameters (pulse duration, pulse rate, and number of active electrodes). The Monopolar (MP) mode is widely used in commercial CI devices because of its ability to produce lower C & T levels as compared to Bipolar (BP) or Tripolar (TP) modes.

Loudness in electrical hearing depends on three physical dimensions, such as current (expressed in microamps), phase width (expressed in microseconds), and pulse rate (expressed in Hz). In the current CI devices, the phase width and pulse rate are constant, but current changes. The results obtained from the earlier study (Fu & Shannon, 1998) on loudness estimation in CI recipients were consistent with the Stevens Power Law. The loudness (L) in electrical hearing is directly proportional to the current (I) with a mean exponent $\beta = 2.7$. So, the equation is $L \propto I^{2.7}$. In the Nucleus 24 device, current is specified by 8 bit current level (c) and the current (*i* in microamps) is an exponential function of current level (c) in clinical units, $i = i_0 e^{rc}$, where i_0 = 10µA (equal to 0 clinical units or current level (c)) and i = 1750 µA (equal to 255 clinical units (c)) and r = 0.0203 (constant value). Technically, for each clinical unit, there is a 2 % (0.176 dB) increase in current.

McKay et al.(2001) investigated loudness in multichannel CI users using sequential stimulation. The results were consistent with the loudness prediction model and they found that the loudness increases as a function of spatial separation between electrode pairs up to approx. 3mm. They concluded that the loudness summation was independent of the channel spacing. According to this model, the loudness of individual pulses are estimated and the loudness of all the pulses are summated in a 2ms time window. Hence, the resultant loudness is a good predictor of current levels. The result obtained from this study shows that the log loudness was linearly represented at low current levels, but loudness grew more steeply at high current levels. Additionally, Zeng & Shannon (1992; 1995) also reported that the loudness is exponentially related to the stimulus frequency above 300 Hz and a power function above 300 Hz.

.

3.6 Timbre Perception

Timbre is a multidimensional attribute defined as "the attribute of auditory sensation whereby a listener can judge two sounds are dissimilar using any criteria other than pitch, loudness and duration" (Pratt & Doak, 1976). Timbre is the perceptual quality that distinguishes between two tones that have the same pitch, loudness, and duration (Plomp, 1976; Moore, 2003). It is the quality of sound that differs between two musical instruments such as an oboe and a flute. In acoustic hearing, timbre is represented by three physical dimensions, the two salient dimensions are log-attack time and the spectral centroid (spectral centre of gravity), the less salient dimension is the spectral flux (spectral fluctuations over time).

Timbre perception in CI users can be explored in three ways, timbre recognition or discrimination tasks (ability to recognise the musical instrument), timbre appraisal (perceptual evaluation of the quality of a sound), and musical timbre space (determining the multidimensional timbre space and its psychophysical capabilities). The CI users find it extremely arduous to perform this task. The CI subjects score almost at chance (40 % - 60 %) in a timbre recognition task (Gfeller et al., 1998; Gfeller et al., 2002; McDermott, 2004; Kang et al., 2009). The CI subjects' rating in the timbre appraisal tasks was comparatively lower than the normal hearing counterparts (Gfeller et al., 1991; Gfeller at al., 2002; Looi et al., 2007). The spectral space in CI users was recently evaluated by Kong et al. (2011). Unlike the 3-dimensional spacing for normal subjects, the CI recipients had a 2-dimensional timbre space related to robust temporal envelope and less salient spectral envelope. There was a strong correlation observed between CI and NH subjects for log-attack time and a weak correlation for spectral envelope. Finally, they reported that the temporal envelope cues are more reliable cues as opposed to spectral envelope cues in CI users.

Brightness, a perceptual attribute of the spectral centroid, depends on the energy concentration across the harmonic structure of a stimulus. It is determined by the amplitude structure of all the partials and does not directly rely on the fundamental frequency (F0) (Marozeau & de Cheveigne, 2007). Therefore, the overall brightness (spectral profile) is determined by the centroid of excitation pattern (Anantharaman et al., 1993; Dai et al., 1996). A tone with strong high harmonics may sound brighter than a tone having a strong energy concentration at the lower harmonics (Lichte, 1941). Brightness (spectral profile) constantly varies in the outside environment, therefore, the overall profiles of spectral shape permit listeners to differentiate between different musical instruments as well as differentiating speech from non-speech sounds. Variations in the spectral envelope mainly affect the brightness attribute of timbre, but this variation does have a slight influence on pitch (Marozeau & de Cheveigne, 2007). There are two prominent approaches for studying pitch & timbre interaction. The first approach is a psychophysical method using synthetic tones or musical instruments. In this approach, the researchers have direct control over these dimensions, i.e., loudness and duration can be kept constant while pitch or timbre is varied (Demany & Semal, 1993; Pitt, 1994; Russo & Thompson, 2005; McDermott et al., 2008). Russo & Thompson (2005) observed that when participants were asked to judge changes in pitch, their judgements were influenced by changes in the spectral centroids of the component tones.

The second approach involves neurophysiological studies using MMN (Mismatch Negativity) (Caclin et al., 2006). Psychophysical studies have reported that the two dimensions (pitch & timbre) are independently stored in auditory short-term memory (Semal & Demany, 1991) and a few others have found some degree of interaction between these dimensions (Caclin et al., 2007). The neurophysiological studies supported that these dimensions are individually represented in the sensory memory (Caclin et al., 2006). The reason for the independent or interactive performance of the auditory attributes is because, during the initial stages (peripheral level), the auditory system processes all the dimensions interactively and later they are represented individually especially in the sensory memory (Caclin et al., 2007).

Demany et al., (1993) made an attempt to study the relationship between the two perceptual dimensions (pitch and brightness). He investigated whether a physical shift in the fundamental frequency (F0) or the brightness aspects of timbre (Spectral centroid (Fc)) had a perceptual influence on each other. His question was "are these two types of physical shifts perceptually independent?" The conditions which were tested are a shift in F0 with Fc fixed, a shift in Fc with F0 fixed, and combined shifts in both F0 and Fc. Their multidimensional analysis supports the hypothesis that the pitch and the brightness aspects of timbre are perceptually independent.

An important study by Pitt (1994) investigated the role of pitch and timbre in musically trained and untrained individuals. He reported a detailed description of how the two groups perceived varying pitch and varying timbre. He used four combinations of pitch-timbre conditions (same pitch same timbre; same pitch different timbre; different pitch same timbre and different pitch different timbre). He used four tonal stimuli, two played on a trumpet and the other two played on a piano. He conducted two experiments; a categorization task and a speeded classification task (Garner Interference). The results indicated that the timbre is more salient than pitch for non-musicians and the cross-talk phenomenon is much stronger from timbre to pitch thus exhibiting an asymmetrical interference. Conversely, musicians are trained to perceive only pitch changes and the cross-talk is bidirectional with equal strength in both directions and this leads to symmetrical interference. Generally, the acoustic stimuli in the real world consist of harmonic tones and its complexes. When the musically untrained subjects perceive different stimulus types (brightness and noise band sequences), they are easily distracted by these stimuli, concealing the true identity of pitch. Conversely, the musicians are conditioned to concentrate only on the pitch changes. So, the distraction effects are minimal in these participants. The results from the above studies suggest that the auditory dimensions are processed differently by these two groups of participants (Musicians & Non-Musicians).

Recent studies on CI place pitch reported that the place of excitation cue is in closer proximity to brightness aspect of timbre than to pitch (McDermott & McKay, 1997; Laneau & Wouters, 2004a; Moore & Carlyon, 2005; Swanson et al., 2009). In chapter 7, the brightness mechanism is explored using various psychophysical procedures and the outcomes are discussed with relevance to CI place pitch perception.

3.7 Music Perception

Music is one of the oldest art forms mediated by sounds, generally produced by musical instruments or vocal tones in a systematic and continuous manner. Music comprises of four major perceptual dimensions (pitch, loudness, timbre, and duration) (Krumhansl & Iverson, 1992). Musical listening is affected, when any of these dimensions are improperly coded in the auditory system.

In western music, when two musical notes are presented in a frequency ratio of 2:1, then the normal listeners will perceive these tones as similar and this relationship is called as an *octave*. For example, a 440 Hz tone, when doubled (880 Hz) or reduced by half (220 Hz), the resulting tone will sound similar to the 440 Hz. In an equal temperament tuning, each octave is divided into a series of twelve equal steps called the twelve-tone equal temperament (12TET). These twelve tones are referred to as *semitones.* Semitones are usually the smallest musical interval in western music. In simple terms, a semitone is an interval between two adjacent notes in a twelve tone scale. The size of the musical interval can be established based on the ratio of their frequency. The frequency ratios of the musical interval in an octave are, 1:1 (unison), 2:1 (octave), 3:2 (perfect fifth), 4:3 (perfect fourth), 5:4 (major third), and 6:5 (minor third). When two notes are separated by a small ratio (E.g., unison), the normal listeners report a pleasant sensation of consonance. The frequency ratios are usually calculated with respect to a reference frequency of A4 (440 Hz). In a 12TET, a single semitone is represented as 1.06:1 and the remaining musical intervals can be calculated based on this relation. For *Major second – 'D Note'* = $(1.06)^2$:1, *Major third – 'E Note'* = $(1.06)^4$:1, *Perfect 4th – 'F Note'* = $(1.06)^5$:1, *Major 6th – 'A Note'* = $(1.06)^9$:1, *Octave* = 2:1.

Musical scales around the world are centred on pitch (E.g., Diatonic and pentatonic) and that is one of the reasons why pitch has a special status in music. Pitch compared to other dimensions (loudness and timbre) is most strongly associated with the melody. Melody can be defined as the organisation of successive musical sounds with respect to pitch (Tovey, 1975). Pitch relations among successive musical notes are more important than the absolute pitch of those notes (Dowling & Fujitani, 1970). The two components that constitute a melody are contour and interval size. Pitch contour refers to the direction in which the successive notes vary in a melody, whereas pitch interval refers to the distance between two successive notes. Melody recognition will be affected when either of these components are not adequately conveyed to listeners (Dowling & Fujitani, 1970). In this thesis, two kinds of pitch modification were used in melody recognition tasks. Firstly, the *backward modification* in which the entire contour of the melody is altered. Secondly, the *warp modification* in which the interval size of the melody is modified based on a linear input-output function, but the melodic contour remains unaltered.

Normal listeners effortlessly perceive music and on some instances, music is perceived as a background task. This is not the case with CI users (McKinney, 2010). Musical listening remains an uphill task for the majority of CI recipients (for excellent reviews on music perception in CI see McDermott (2004, 2012) & Looi (2008)). It is well documented that CI recipients are on par with the normal hearing counterparts in *rhythm* discrimination tasks. The CI recipients are quite efficient in identifying the variation in the temporal patterns. Conversely, their performance often drops to chance level in pitch perception tasks. Normal listeners have a remarkable ability to detect a fraction of a semitone; perhaps only “Star CI Performers” can potentially detect changes as low as two semitones. The majority of CI recipients can discriminate pitch changes ranging from two – seven semitones, but there is a substantial amount of intersubject variability seen across CI recipients (Gfeller et al., 2002; Sucher & McDermott, 2007; Cooper et al., 2008). The plausible reasons are poor neural survival rate inside the cochlea, electrode insertion depth, and auditory deprivation prior to implantation.

The most widely used test battery for evaluating music perception in CI recipients is the CAMP test (Clinical Assessment of Music Perception) (Nimmons et al., 2008) which assesses the pitch discrimination and direction of complex tones, melody recognition, and timbre identification. The pitch discrimination is an adaptive test and it is important to note that pitch reversals may be unidentified when an adaptive test is used. Therefore, one needs to be aware of the pros and cons of the test battery and how it should be used in accordance to the needs of the research question. Interestingly, Cooper et al.(2008) tested normal and CI subjects using the Montreal Battery of Evaluation of Amusia (MBEA) which was initially developed by Peretz et al.(2003). Based on the performance of CI recipients, they argued that MBEA is a suitable test for evaluating the music perception in CI recipients. The Melodic contour identification (MCI) test developed by Galvin et al.(2007) is used to assess the ability of CI recipients to recognise melodies and is additionally used as a training tool for improving melody recognition in CI users. Swanson (2008) developed the Modified Melodies test to evaluate pitch perception abilities in CI recipients. The Modified Melodies test supports several types of pitch modification and this will be explained in more detail in the following chapter.

In summary, this chapter discussed the pitch processing mechanisms in CI recipients. There are two pitch percepts corresponding to CI pitch mechanism. *Place pitch* percept can be evoked by stimulating different electrodes inside the cochlea, thereby mimicking the natural tonotopic organisation of the normal cochlea. *Temporal pitch* can be elicited by varying the pulse rate. The upper limit for temporal pitch is approximately round 300 pps, but there are a few exceptions seen with respect to “star CI performers” in different studies. Preliminary reports have shown promising results in extending the upper limits of temporal pitch in CI recipients (Kong & Carlyon, 2010; Macherey et al., 2011), but additional studies need to affirm these findings. The place pitch and temporal pitch are two independent percepts, which can be manipulated independently and separately in CI recipients. The next chapter will investigate the role of pitch sequences in normal hearing individual using various pitch perception tasks.

Chapter 4: Investigating Pitch Perception in Normal Hearing Individuals

Abstract

The aim of the present study was to investigate the performance of normal hearing individuals on various pitch perception tasks. Eighteen normal-hearing adults participated in four experimental procedures: 4 AFC Discrimination, 2 AFC Ranking, and 2 AFC Modified Melodies test (Backwards and Warp modifications) using pitch sequences (harmonic tones varying in pitch, with constant brightness). Results revealed that the subjects' performance was high in all the four experimental procedures. d' scores revealed that the discrimination scores were better than the ranking scores. In the Modified Melodies test, subjects were able to detect a melody contour change and the performance reached ceiling. Additionally, when the melody contour was preserved but the musical intervals were changed, subjects were able to detect a musical interval change with a high level of accuracy using pitch sequences.

4.1. Introduction

The human auditory system is efficient in processing the three perceptual dimensions of pitch, loudness, and timbre simultaneously (Pitt, 1994). Accurate perception of music requires precise coding of pitch in normal hearing individuals and also in hearing aid or CI users (McDermott, 2004; Looi, 2008). Pitch is a fundamental attribute in both channels of communication (speech and music). However, whereas speech tends to involve coarse-grained changes in pitch (e.g., prosody and intonation), music is typically characterised by fine-grained changes in pitch (Wolfe, 2002).

This chapter aims to determine how the C4 harmonic tones (pitch sequences) affect the performance of normal hearing individuals. We will examine four experimental procedures to determine which of these procedures were sensitive for evaluating pitch perception. Normal hearing performance in this chapter acts as a baseline and comparisons will be made with CI recipients in later chapters. An operational definition of pitch is that the variation in pitch can convey a melody (Moore & Carlyon, 2005), thus a test involving melody was utilised. This thesis utilised four experimental procedures: discrimination, ranking, the Modified Melodies test using backward modification, and the Modified Melodies test using warp modification. The four procedures are described in the following section. The baseline stimuli used in this chapter was pitch sequences (Harmonics tones varying in pitch).

4.2. Method

4.2.1. Subjects

Eighteen normal hearing adults with age ranging from 18 – 45 years (M = 25.83, S.D = 6.46, Males = 5) participated in the study. Sixteen participants had no formal musical experience, and two were musically trained (having played an instrument for at least 3 years). The study was approved by the Human Research Ethics Committee (HREC) - Macquarie University.

4.2.2. Experimental Procedures

The four experimental procedures are described below in terms of the baseline pitch sequences. The musical notes used in these tasks were harmonic tones with varying pitch (fundamental frequency) and constant spectral profile (Brightness). . Feedback was provided only at the end of each block of trials and no trial-by-trial feedback was provided. Prior to the start of each task, subjects were presented with practise trials until they were familiarised with the tasks.

4.2.2.1. Note Discrimination Task

The ability to detect changes in pitch was assessed in this task. The four musical notes used in this procedure were C4 = 262 Hz, D4 = 294 Hz, G4 = 392 Hz, and A4 = 440

Hz, and these notes were paired among each other to form six pairs [C-D; G-A; D-G; C-G; D-A; C-A]. A four-alternative forced choice method was used. In each trial, there were three presentations of the reference note and one presentation of the variant note, in randomised order (Figure 4.1). The subject's task was to identify the interval containing the variant note. For example, C C C A note sequences were presented and the subject's task was to identify the variant note which is A. A block of trials comprised of 48 trials (6 note pairs x 8 trials per pair).

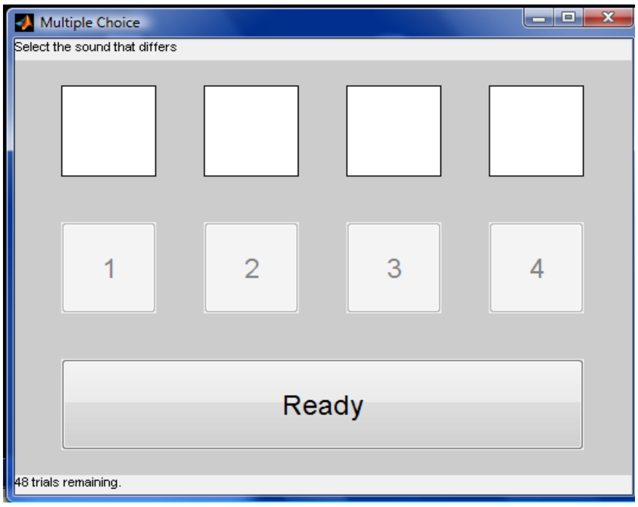


Figure 4.1: Snapshot of discrimination procedure.

4.2.2.2. Note Ranking Task

The ability to rank the notes in the correct order was assessed in this task. The set of notes and pairings used in this task was the same as in the discrimination task. In each trial, a pair of notes was presented in an XXY sequence, i.e., the first note was presented twice, followed by the second note. The subject's task was to identify whether the final note was either "rising" or "falling" (a two alternative forced choice procedure) with respect to the first two notes (Figure 4.2). As in discrimination procedure, the ranking task had 6 note pairs and 8 trials per pair. Overall there were 48 trials per block. The Nucleus

Matlab Toolbox (NMT) (Swanson, 2008) was used to administer both discrimination and ranking procedures.

Figure 4.2: Snapshot of ranking procedure.

4.2.2.3. Modified Melodies Test

The Modified Melodies Test was developed by Swanson (2008). In each trial, the name of the melody was first displayed, followed by the presentation of the initial phrase of the melody played twice (two-alternative forced choice procedure). The subject's task was to identify the correct version of the melody (Figure 4.3). The pitch of the melody was modified based on the type of modifications used. The rhythm was left intact. In this thesis, only two familiar melodies, 'Old MacDonald Had a Farm' and 'Twinkle Twinkle Little Star' were used. Each melody was defined by a sequence of notes. The pitch of each note was specified as the number of semitones above a base frequency. The notes were transposed up by 0, 1, 2, or 3 semitones (randomly selected on each trial) to minimise the contribution of non-pitch cues and also to suppress the learning effects. The duration of each note was specified as a number of beats. A beat was 300 ms in duration.

Old Macdonald was defined as:

```
Notes: [5,5,5,0, 2,2,0, 9,9,7,7, 5]
Beats: [1,1,1,1, 1,1,2, 1,1,1,1, 2]
```

Twinkle Twinkle Little Star was defined as:

```
Notes: [0,0,7,7, 9,9,7, 5,5,4,4, 2,2,0]
Beats: [1,1,1,1, 1,1,2, 1,1,1,1, 1,1,2]
```

Figure 4.3: Snapshot of the Modified Melodies test.

4.2.2.3.1. Backward Modification

In this modification, the melody notes were reversed in time. To preserve the exact rhythm of the original melody, repeated notes were firstly replaced by a single note, to give a merged note sequence, as illustrated in Table 2. The merged note sequence was then reversed in time, and finally the appropriate notes were split into repeated notes. The rhythm was included in this test because the subjects feel more familiar with the melodies and removing the rhythm will make these melodies unrealistic and less familiar. The Figures 4.4 and 4.5 illustrates the result for Old Mac Donald and Twinkle Twinkle Little Star respectively. It is important to note that the backward modification altered the contour of the melodies. Each block contained 32 trials.

Table 2: Backward Modification Note Sequence.

Melody	Original Note Sequence	Backward Note Sequence
Old MacDonald	FCDCAGF	FGACDCF
Twinkle Twinkle	CGAGFEDC	CDEFGAFC

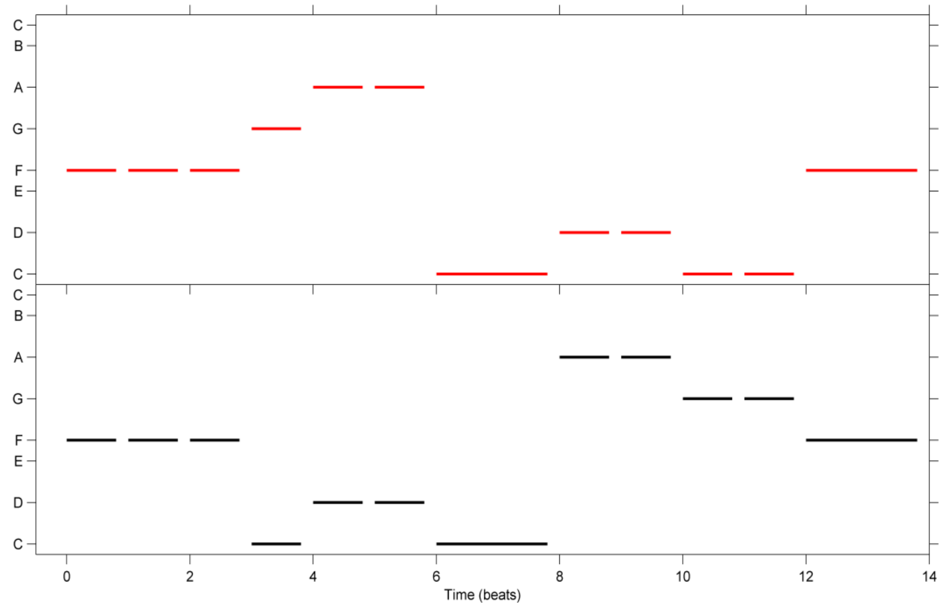


Figure 4.4: Old MacDonald [Black-Bottom] original version and Backward Modification [red-top].

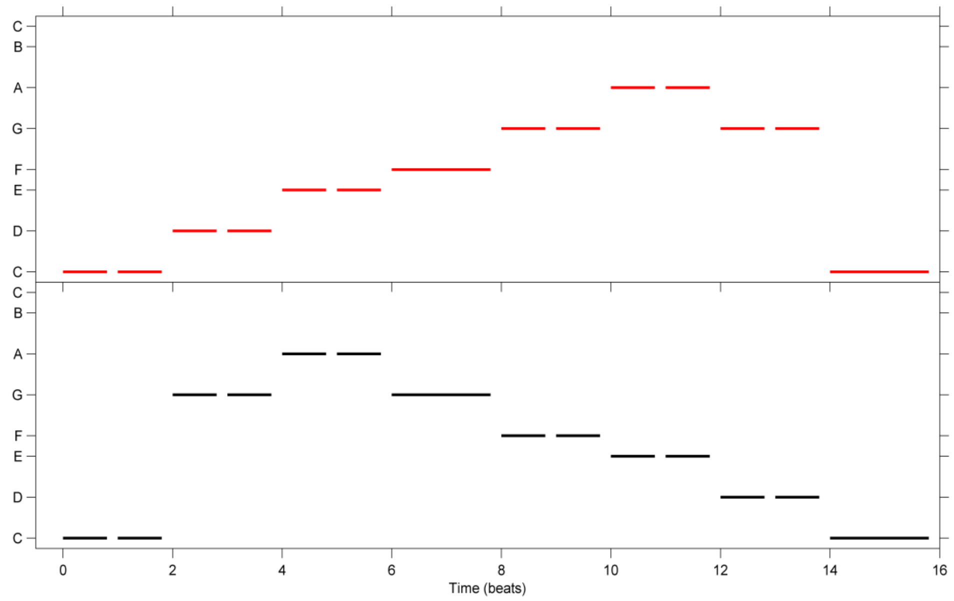


Figure 4.5: Twinkle Twinkle [Black-Bottom] original version and Backward Modification [red-top].

4.2.2.3.2. Warp Modification

In the warp modification, each note in the melody was expressed as the number of semitones above the base frequency, and was shifted in pitch by applying a piece-wise linear input-output function, as shown in Figure 4.6. The amount of pitch shift was controlled by a warp factor, which specified the slope of the initial segment of the input-output function. In the present chapter, warp factors 0.75 and 1.33 were used. The highest and lowest notes in the melody were unchanged, and the frequency range of the modified melody was same as the original melody. For warp 0.75, the intermediate notes were shifted downwards in pitch. Conversely, for warp 1.33, the intermediate notes were shifted upwards in pitch. The modified melodies contained mistuned notes lying between the notes of the musical scale, which cannot be represented in standard musical notation. The main feature of the warp modification is that, it changes the size of the musical intervals while keeping the melodic contour intact. Thus the modified melodies will have mistuned notes embedded within the melody with unaltered melodic contour. Melody recognition is affected when melodic contour or interval size of the melody is affected (Dowling & Fujitani, 1970). Each block comprised of 32 trials (16 trials x 2 warp factors) and the stimuli were presented in a randomised order.

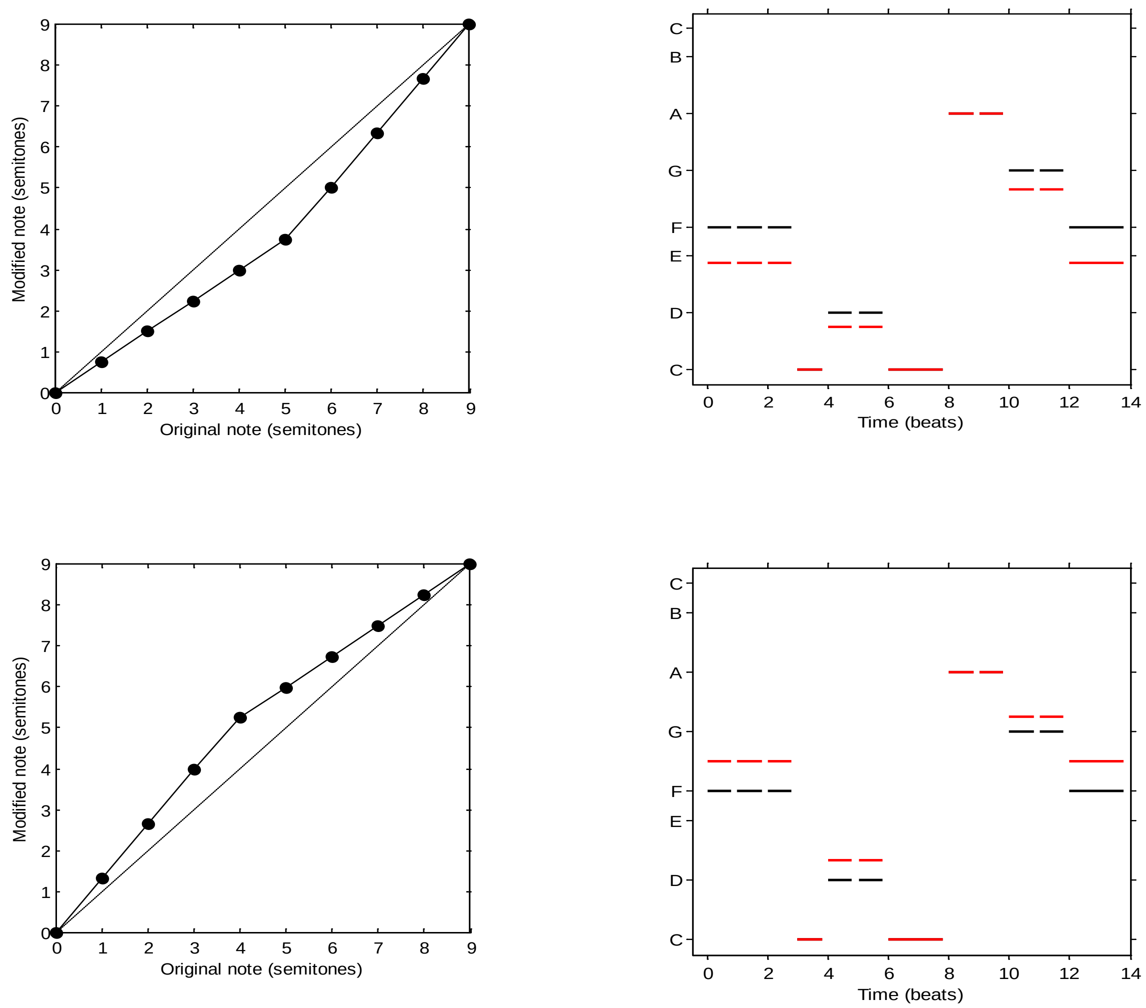


Figure 4.6: Pitch input-output relationship for the Warp modification (Left side). Top: Warp 0.75; bottom: Warp 1.33 . Original (black) and modified (red) versions of Old MacDonald (Right side). Each note is shown as a horizontal line, with length indicating duration, and vertical position indicating pitch. The vertical axis is linear in semitones (Right). Top: Warp 0.75; bottom: Warp 1.33.

4.2.3. Stimulus Description

All the stimuli were synthesised on a PC at a sampling rate of 16 kHz, and presented binaurally via a PC sound card and headphones (Sennheiser HD280 Pro) in a sound-treated room. Each frequency step (consisting of one long 300 ms note or two short 150 ms notes) was 300 ms in duration, with a smooth (sinusoidal-shaped) rise and fall time of 50 ms. All the stimuli had the same spectral profile (i.e., same timbre). During the practise trial, the subjects were instructed to set a comfortable loudness level which was kept unaltered throughout the experimental procedures. The amplitude of all the stimuli were normalised from -1 to +1 in order to maintain equal loudness level across subjects. . The detail of the baseline stimulus type is described below.

4.2.3.1. Pitch sequences: Harmonic tones varying in pitch

Each note was a harmonic tone. The harmonic tones had fundamental frequencies in the range of C4 – A4 (262 – 440 Hz), and all had the same spectral profile, i.e. the same brightness. They were generated by summing a number of harmonics. Harmonics having frequencies below C6 (1047 Hz) had 0 dB amplitude (i.e., the reference amplitude). Harmonics having higher frequencies were reduced by 3 dB per semitone above C6, up to an upper frequency of C7 (2093 Hz). No harmonics beyond C7 were included. Thus, the tones contained only low-numbered harmonics which would be resolved in the normal auditory system. The spectral profile of the four notes C4, D4, G4, and A4 are shown in Figure 4.7**.**

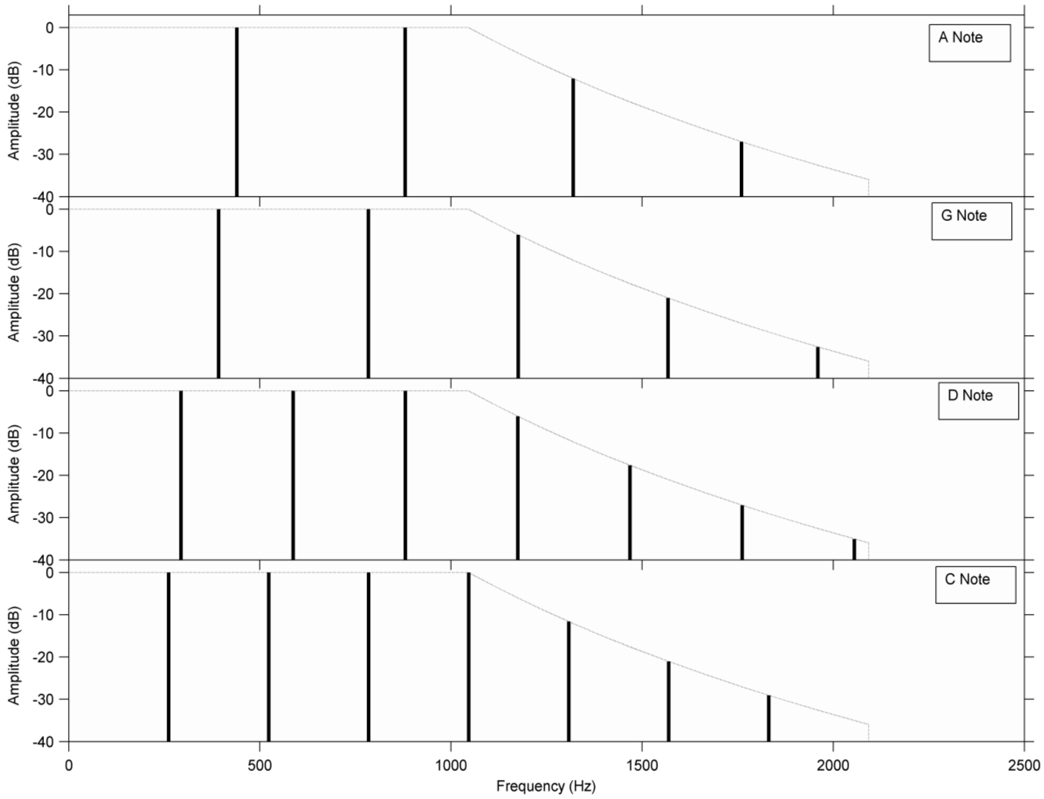


Figure 4.7: Spectral profile of harmonic tones C4, D4, G4, A4.

4.3. Results and Discussion

The capacity of normal subjects to recognise musical patterns based on pitch sequences were assessed in this chapter. The percentage correct scores for discrimination and raking were obtained by summing across note pairs and for modified melodies warp modification, the scores were summed across the two warp factors (0.75 and 1.33). The percentage correct scores for discrimination, ranking, modified melodies backward, and modified melodies warp were 92 %, 89 %, 100 %, and 98 % respectively.

The percentage correct scores across the four procedures cannot be compared directly, because the discrimination procedure was a 4-AFC method and the other procedures were 2-AFC methods. In order to compare these procedures, the scores were converted into d' scores based on table A5.7 of Macmillan & Creelman (2005). The group mean d' scores for the four procedures are illustrated in Figure 4.8 and the statistical analysis was administered on these d' scores.

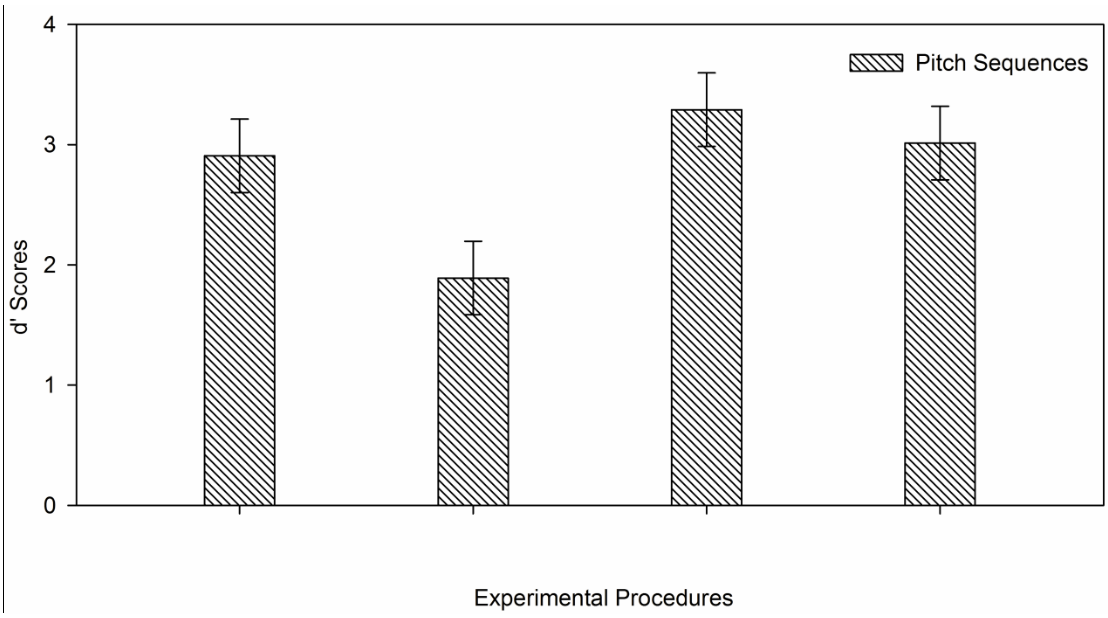


Figure 4.8: The group mean d' scores for the three stimulus types across four procedures. A d' score of zero indicates performance at chance. The error bars represent the standard errors.

A one-way ANOVA was computed between the four experimental procedures and the results showed a high significant effect of the experimental procedures [$F(3,68) = 15.79$, $p < 0.001$]. The Bonferroni post hoc comparison showed a significant difference between discrimination and ranking ($p < 0.001$), ranking and modified melodies backward ($p < 0.001$), and ranking and modified melodies warp ($p < 0.001$). There was no significant difference among other pairs ($p > 0.05$). The normal hearing subjects performed well above chance in all the four experimental procedures. The maximum scores were seen for the modified melodies backward (Mean (M) = 3.29) and the least performance was seen in ranking task (M = 1.89). The intermediate scores were obtained for modified melodies warp (M = 3.01) and discrimination (M = 2.90). Pearson's correlation showed a high positive correlation only between discrimination and ranking procedures ($r = 0.61$, $p = 0.006$) and there was no correlations among other procedures.

The scores obtained from the modified melodies backward modification suggests that the participants had no difficulty in identifying pitch contour changes and the subjects were capable of accurately judging the size of the musical intervals in the modified melodies warp modification.

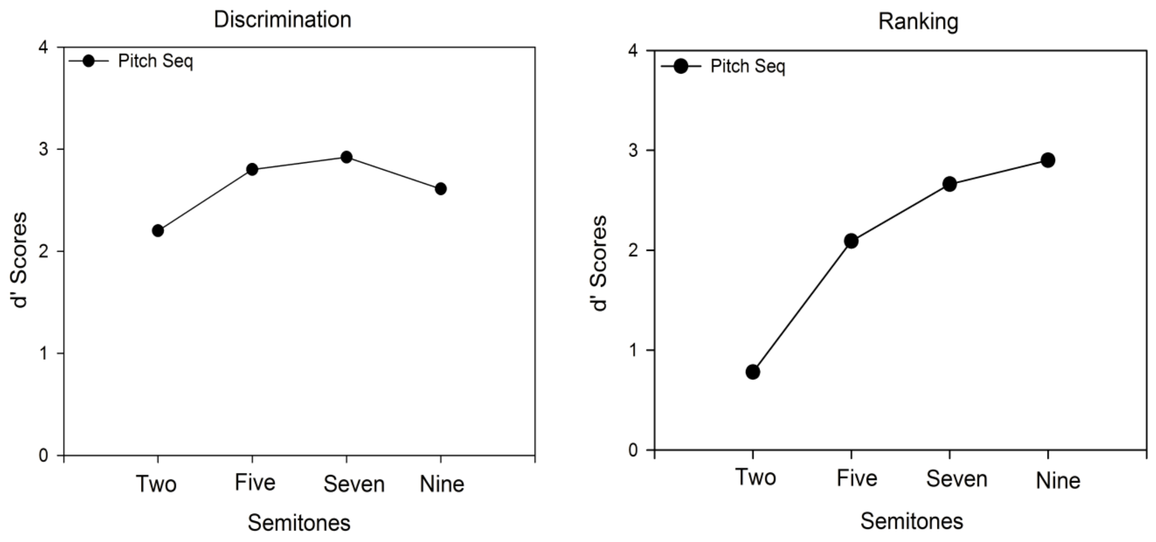


Figure 4.9: d' scores as a function of semitone difference for (a) Discrimination and (b) Ranking.

The discrimination and ranking scores in Figure 4.8were averaged across all note pairs. In both these procedures, six pairs of notes were used. These notes were 2 semitones apart (C-D; G-A), 5 semitones apart (D-G), 7 semitones apart (C-G; D-A), and 9 semitones apart (C-A). The scores are plotted as a function of semitone difference in Figure 4.9.

The discrimination results were analysed using One-way ANOVA to investigate whether there was any significant difference among the semitones. The results showed that there was no significant difference among the semitones [$F (3, 68) = 1.58$, $p > 0.05$]. Conversely, in ranking task, One-way ANOVA revealed a significant difference among semitones [$F (3, 68) = 17.3$, $p < 0.001$]. Bonferroni post hoc analysis revealed that the two semitone scores were significantly different from five, seven, and nine semitones ($p < 0.001$) and there was no significant difference among other semitone pairs. Ranking scores showed a monotonic increase in performance from two semitones to nine semitones. Two musicians scored 100 % and 91 % in discrimination and ranking tasks respectively as compared to 78 % and 72 % by non-musicians. Additional musician subjects would be required to do a detail analysis.

4.4. Conclusion

This chapter investigated the performance of eighteen normal subjects on four experimental procedures using pitch sequences (harmonic tone varying in pitch). The normal subjects performed extremely well on all the four procedures and the performance reached ceiling in the Modified Melodies test using backward modification. The normal hearing performance behaves like a baseline and it is compared with the CI group in the following chapters.

Chapter 5: Exploring Temporal Pitch Sensitivity as a Function of Electrode Stimulation Patterns and Pulse Rate

Abstract

Six cochlear implant (CI) users participated in a study which investigated (a) the temporal pitch sensitivity using different pitch perception tasks; (b) the CI performance as a function of electrode stimulation patterns and base pulse rates. Four stimulation patterns were tested using three experimental procedures. The stimulation patterns were single electrode stimulation (apical (Electrode E22) or middle (E12)), dual-electrode stimulation (E22 & E12), and multiple electrode stimulation (E22 to E12). The three procedures were: Ranking, Modified Melodies Test - Backward modification, and Modified Melodies Test - Warp modification. The stimuli were presented as pulses at a base rate of 131 pulses per second (pps) (C3 range). Additionally, a high pulse rate of 262 pps (C4 range) was also tested in two stimulation modes apical (E22) and middle (E12) across the three procedures. The first hypothesis tested was that the performance would be better at the apical (E22) electrode compared to middle electrode (E12). The second hypothesis was that the performance would be better for multiple electrodes (E22 to E12) compared to single or dual electrode stimulation. The third hypothesis was that the performance would drop as the base rate increased from C3 (131 pps) to C4 (262 pps). Surprisingly, the results showed no significant difference among the stimulation patterns for the three procedures. The performance at C3 (131 pps) base rate was significantly better than C4 (262 pps) base rate, but above chance scores were obtained for the C4 (262 pps) pulse rate. Results showed a good correlation between Modified Melodies Backward and Modified Melodies Warp modification. Overall, the results did not support the first two hypotheses, but did support the third hypothesis which revealed that the performance decreased as the base rate increased from C3 (131 pps) to C4 (262 pps). Overall the results suggested that the CI recipients received same information from different stimulation pattern and were unable to combine cues from different places in the cochlea to give a stronger cue.

5.1. Introduction

Cochlear implant devices convert the incoming acoustic signals into a series of pulse trains, which stimulate multiple electrodes simultaneously or sequentially interleaved. Pulses delivered to the intra cochlear electrodes evoke two kinds of pitch sensation. Firstly, the place pitch corresponds to different electrode positions along the cochlea. As the electrode position moves from apex to base, there is an increase in place pitch percept (Nelson et al., 1995). Cochlear implants mimic the natural tonotopic organisation of the cochlea, i.e., the low frequencies are represented at the apex of cochlea, while the high frequencies are represented at the base. Nucleus devices consist of 22 physical electrodes which produce 22 distinctive place pitches. Additional place pitch can be elicited by sequential stimulation of adjacent electrodes (McDermott & McKay, 1994; Kwon & Van den Honert, 2006) or simultaneous stimulation (Townshend et al., 1987; Donaldson et al., 2005; Busby et al., 2008). The temporal pitch can be produced by varying the pulse rate on an electrode (Pijl & Schwarz, 1995b; Fearn et al., 1999; McKay et al., 2000; Zeng, 2002) or by amplitude modulating the carrier pulse train (Mckay et al., 1994; Geurts & Wouters, 2001; Laneau et al., 2006). CI recipients usually label these two patterns of temporal pitch as pitch. Although there is a large inter subject variability in the upper cut-off range for temporal pitch, it is reported that the temporal pitch usually fades off at around 300 pulse per seconds (pps) (Shannon, 1983; Tong et al.,1985; Zeng, 2002; Kong et al., 2009, Milczynski et al., 2009), but interestingly a high range was also observed by Townshend et al. (1987) - 1000 pps and Kong & Carlyon (2010) – 900 pps.

Earlier studies have provided evidence that the place pitch and temporal pitch can be studied independently in CI recipients (Tong at al., 1983; McKay et al., 2000). Studies on temporal pitch have furnished sufficient evidence that temporal pitch alone is capable of conveying melodic information in CI users. The two classical studies (Pijl & Schwarz, 1995b; McDermott & McKay, 1997) on temporal pitch reported that CI users were able to identify the melodies and label the musical intervals when presented as varying pulse rates on an electrode.

In order to further understand the saliency of temporal pitch especially in musical context, the present study was designed to investigate temporal pitch sensitivity on various stimulation patterns (Single, Dual, and Multiple Electrodes) at two base pulse rates (C3 – 131 pps and C4 – 262 pps). The study hypothesised that, (a) the performance would be better at the apical (E22) electrode compared to middle electrode (E12), because of a better place-rate match. (b) When more electrodes are stimulated, more temporal information can be carried by a larger number of auditory nerve fibres and better scores can be anticipated. So, it was hypothesised that the performance would be better for multiple electrodes (E22 to E12) compared to single or dual electrode stimulation. (c) The performance would drop as the base rate increased from C3 (131 pps) to C4 (262 pps) as seen in previous studies (Zeng, 2002; Kong et al., 2009). These hypotheses were tested using three experimental procedures (Ranking, Modified Melodies test – Backward modification, and Modified Melodies test – Warp modification).

5.2. Method

5.2.1. Subjects

Six post-lingually-deafened CI recipients, with at least one year of implant usage, participated in the study. The study was approved by Macquarie University human research ethics committee (HREC) and additional approval was obtained from the Sydney South West Area Health Service (SSWAHS). The detailed description regarding the implant age, etiology, implant type, pulse rate for CI stimulation, processor type, and processing strategies for each recipient was mentioned in Table 3.

Table 3: Cochlear Implant Recipients' Details.

Subjects ID	**CI-1**	**CI-2**	**CI 03**	**CI 04**	**CI 05**	**CI -6**
Age	71/M	70/F	65/M	65/F	37/F	**78/M**
Implant Age	R: 4.5 yrs	R: 4 yrs	R : 7 yrs	R: 12 yrs	R : 7 yrs	R: 6 yrs
Etiology	Progressive	Progressive	Progressive	Sudden (Ear Infection)	Ototoxicity	Progressive
Implant Type	CI24 RE (ST)	CI24 RE (ST)	CI24 RE (CA)	CI24M	CI24M	CI24 RE (CA)
Pulse Rate	900	500	900	900	900	900
Processor Type	Nucleus 5 [CP 810]	Nucleus 5 [CP 810]	Freedom	Freedom	Freedom	Nucleus 5 [CP 810]
Processing Strategy	ACE	ACE	ACE	ACE	ACE	ACE

5.2.2. Experimental Procedure

The three experimental procedures (Ranking, Modified Melodies test – Backward modification, and Modified Melodies test – Warp modification) were investigated as a function of different stimulation patterns at two base pulse rates. In these procedures, no trial-by-trial feedback was provided, but feedback was provided only at the end of each block. Prior to the start of each task, subjects were presented with practise trials until they were familiarised with the tasks. The entire experimental procedure for CI recipients was conducted in six sessions. Each session lasted for 2 hrs with approximately 10 minutes break after an hour.

The note ranking task (4.2.2.1) and the Modified Melodies test both backward and warp modification (4.2.2.3) are explained in great detail in the previous chapter. In this chapter, the melodies were presented as varying pulse rate at different stimulation patterns. In the Modified Melodies warp modification several warp factors were tested. But warp factors 0.10 & 10.00 were summed and compared with other experimental procedures, because all the subjects were able to perform only on these two warp factors (0.10 & 10.00). The discrimination procedure was not included in this study because pitch reversals cannot be detected in this procedure and also due to time constrains. Before the start of each procedure, loudness was balanced separately for each stimulation pattern (Apical, Middle, Dual and Multiple electrode stimulation) at two base pulse rates [C3 (131 pulse per second (pps)) and C4 (262 pps)]. Loudness balancing was done for four notes (C, D, G, and A) in each stimulation pattern. Each individuals' MAP was used as a reference and each electrode was stimulated at a gain (Current Levels (CL)) below the C (comfortable) levels. The gain was increased at 2 CL steps until the recipient says that the note is at a comfortable loudness level. Once this was established for a single note, a similar procedure was followed for the remaining notes in each stimulation pattern. Once comfortable loudness levels were established for all the four notes, a complete sweep was done and the subjects were asked which note needs to be increased or decreased in loudness until all the four notes had the same loudness level.

5.2.3. Stimulus Description

The pulses were presented as biphasic, monopolar (MP1+2) pulse trains at a base pulse rate of C3 (131 pps) and C4 (262 pps). The pulse rates have a 9 semitone range for both C3 and C4 base rate. The participants' latest MAPs were used to set the initial presentation level for these experimental procedures. Loudness-balanced comfortable levels for each musical note (C, D, G, and A) at different stimulation patterns were established for two base pulse rates. The same current levels were maintained across the three experimental procedures. The current levels varied relatively with respect to different stimulation patterns and the pulse width was maintained at 25µsec/phase, with an interphase gap of 8 µsec. Subject CI-1 reported difficult in listening to these procedures, more specifically, he reported that the sound was fading off at the end of the each stimulus. Therefore, to get an optimal loudness level and to avoid increasing the current levels to maximum (out of compliance issues), the pulse width alone was increased to 50 µsec/phase. Finally, the pulse sequences were delivered to the intra cochlear electrode using a clinical pod via a loaner Freedom processor for all the participants. Communications with the CI and recording of recipients' responses were accomplished by using NIC (Nucleus Implant Communicator). NIC developed by Cochlear Ltd is used to define and deliver stimulation patterns and other operations. The ranking and Modified Melodies test was implemented in Matlab using Nucleus Matlab Toolbox (NMT) and python respectively. In the ranking procedure, the currents were roved by 2 CL, i.e., either increase or decrease the presentation level by 2 CL. The significance of doing roving is to minimise the contribution of loudness cues and also to avoid the involvement of any non-pitch cues. The details of the three stimulation patterns are described below.

5.2.3.1.Single Electrode Stimulation

The musical notes were played as a pulse train stimulating the apical electrode (E22) or middle electrode (E12). In CI-4, the middle electrode (E11) was used. The three procedures were administered on a single electrode (Figure 5.1). This mode of stimulation provided exclusive temporal pitch information. Except CI-2, the remaining subjects participated in the additional testing of a base pulse rate of C4 (262 pps) at apical (E22) and middle electrodes (E12).

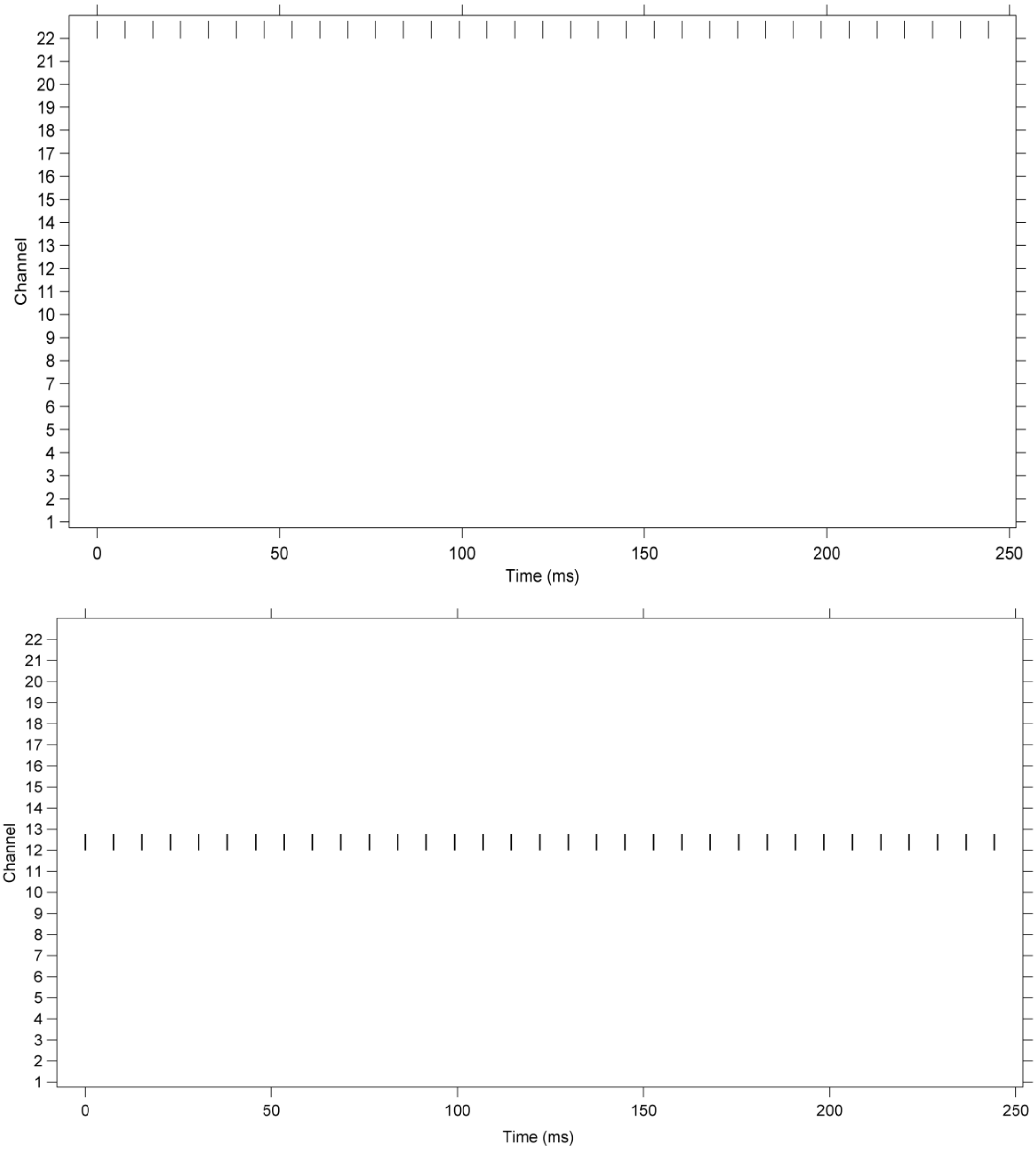


Figure 5.1: Single electrode stimulation. Apical - E22 (Up) and Middle – E12 (Below). The electrodes 1 to 22 are numbered from basal to apical portion of the cochlea.

5.2.3.2.Dual Electrode Stimulation

The musical notes were played as a pulse train stimulating both the apical (E22) and middle electrode (E12) sequentially. All the procedures were administered on the dual electrodes (E12 & E22) at a comfortable level (Figure 5.2).

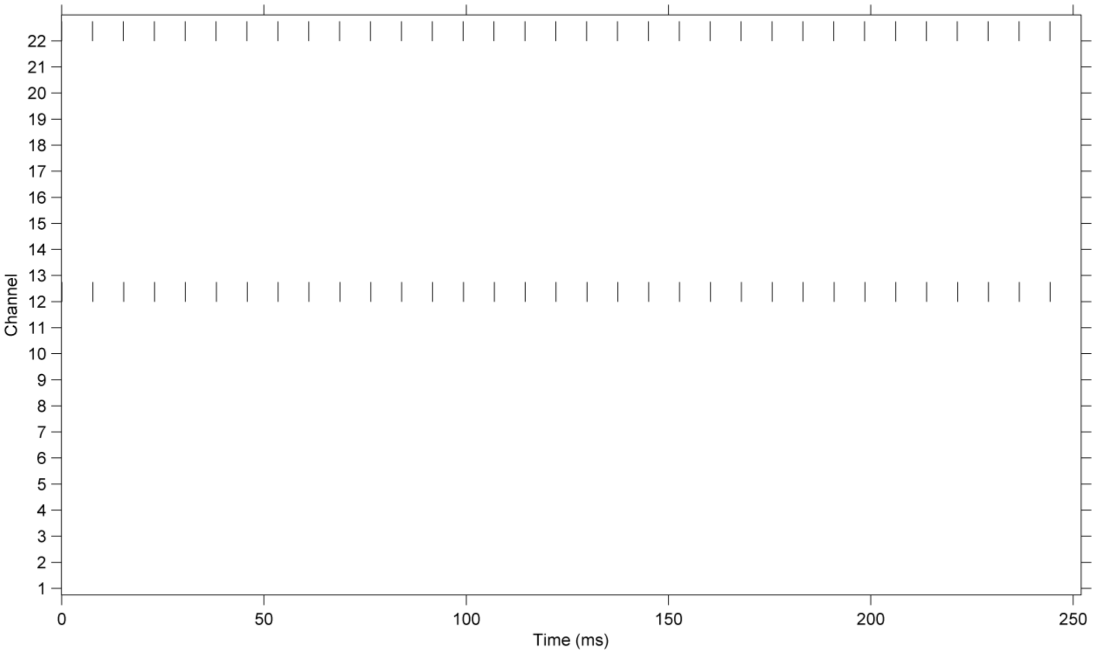


Figure 5.2: Dual electrode stimulation (E22 and E12). The electrodes 1 to 22 are numbered from basal to apical portion of the cochlea.

5.2.3.3.Multiple Electrode Stimulation

In this condition, the notes were played as a pulse train stimulating the apical half of the electrode array, i.e., stimulating eleven electrodes from apical (E22) to middle electrode (E12) sequentially (Figure 5.3). Eleven electrodes were stimulated sequentially with an inter-pulse interval of 70 µsec.

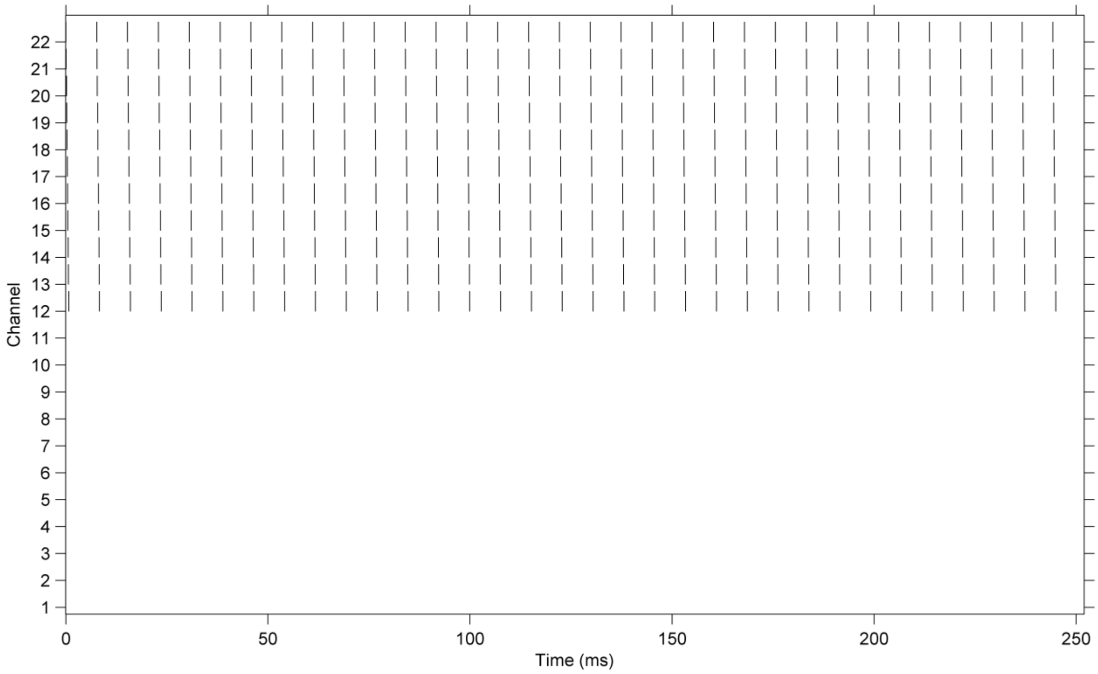


Figure 5.3: Multiple electrode stimulation (E22 to E12). The electrodes 1 to 22 are numbered from basal to apical portion of the cochlea.

5.3. Results

The study investigated the temporal pitch sensitivity in CI recipients using various pitch perception tasks. A single score was obtained for each procedure across two base rates by summing note pairs in ranking, and summing warp factors 0.10 and 10.00 in the modified melodies warp. The proportion correct scores are tabulated in Table4.

Table 4: Percentage correct scores for the three experimental procedures.

Experimental procedures* Vs *Stimulation patterns	***131 PPS*** *(C3 Range)*				***262 PPS*** *(C4 Range)*	
	E12	*E22*	*E22&E12*	*E22toE12*	*E12*	*E22*
Ranking	88	86	86	87	92	90
Modified Melody Backward	90	90	93	85	82	75
Modified Melody Warp	87	90	88	86	86	82

The proportion correct scores obtained from these procedures were converted into d' scores based on table A5.7 of Macmillan & Creelman (2005). This would enable us to compare the present scores with the scores obtained from the similar procedures using loudspeakers (Chapter 6). The group mean d' scores for the three procedures across different stimulation patterns are illustrated in Figure 5.4 and the statistical analysis was administered on these d' scores.

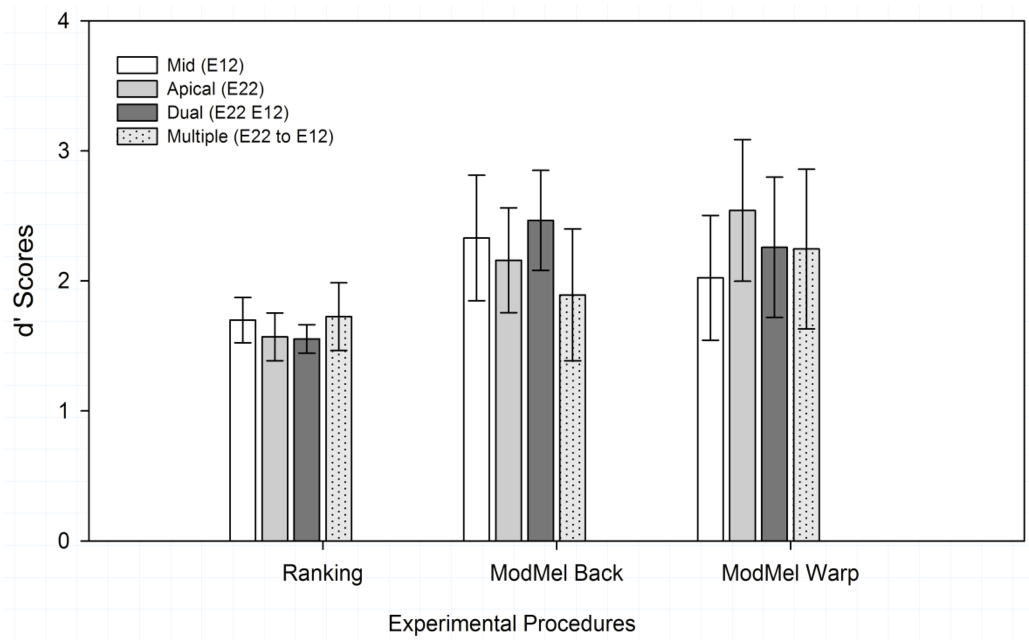


Figure 5.4: d' scores for the three experimental procedures across different stimulation patterns at C3 - 131 pps (N = 6). A d' score of 0 indicates that the performance was at chance. The maximum d' score is 4. The error bars represent the standard errors.

A repeated measures analysis of variance (ANOVA) with the factors, experimental procedures (ranking, modified melodies backward, and modified melodies warp) and stimulation patterns (middle, apical, dual, and multiple) were computed. The ANOVA revealed no main effect of experimental procedures [$F(2, 10) = 1.84$, $p > 0.05$]. The mean scores for modified melodies warp, modified melodies backward, and ranking were 2.26 ($SE = 0.49$), 2.21 ($SE = 0.37$), and 1.63 ($SE = 0.16$) respectively. Bonferroni corrected pairwise comparison showed no significant difference among the experimental procedures. Additionally, there was no main effect of stimulation patterns [$F(3, 15) = 0.18$, $p > 0.05$]. The mean scores for four stimulation patterns were very similar, i.e., the mean scores were 2.09 ($SE = 0.30$), 2.09 ($SE = 0.32$), 2.01 ($SE = 0.32$), 1.95 ($SE = 0.37$) for dual, apical, middle, and multiple stimulations respectively. Similarly, the pairwise comparison revealed no significant difference among these stimulation patterns ($p > 0.05$). There was no interaction between experimental procedures and stimulation patterns [$F(6, 30) = 0.30$, $p > 0.05$].

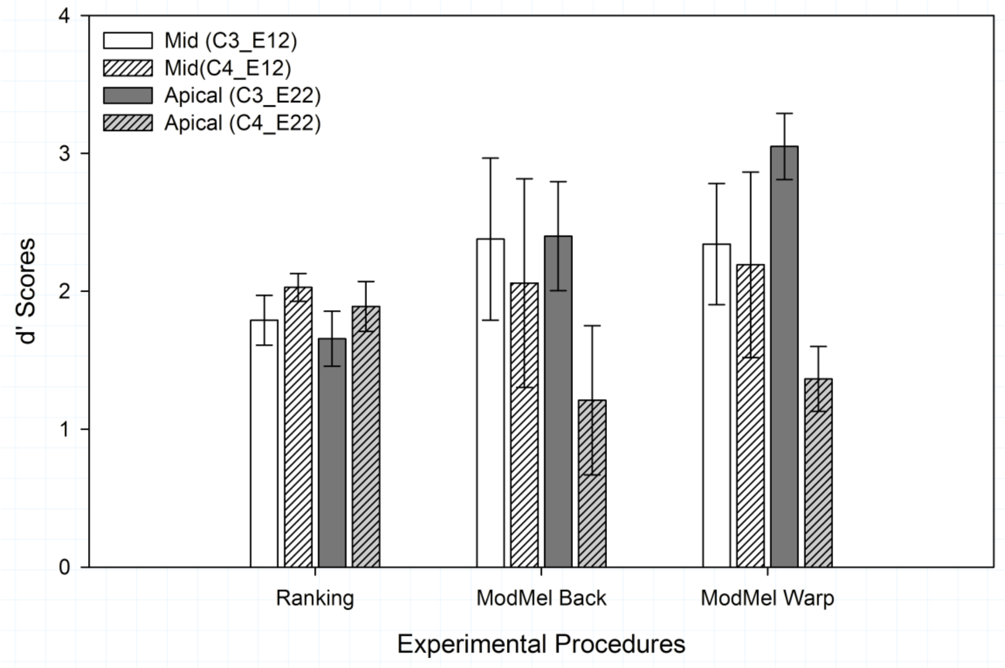


Figure 5.5: d' scores for the three experimental procedures across different stimulation patterns at C4 - 262 pps (N = 5). The d' score of 0 indicates that the performance was at chance. The maximum d' score is 4. The error bars represent the standard errors.

A repeated measures ANOVA was computed for the three experimental procedures at a high base rate of C4 (262 pps). The factors considered were experimental procedures and stimulation pattern (C4_E12 and C4_E22). The analysis was similar to the performance using the C3 (131 pps) rate, i.e., there was no main effect of the experimental procedures and also no main effect of the stimulation patterns. Pairwise comparisons indicated that there was no significant difference among the experimental procedures and stimulation patterns. Similarly, there was no interaction between the two factors ($p > 0.05$). The mean scores for ranking, modified melodies backward, and modified melodies warp were 1.75 (SE = 0.10), 1.63 (SE = 0.58), and 1.77 (SE = 0.38) respectively. The mean scores for middle (C4_E12) and apical (C4_E22) electrode stimulation patterns were 2.09 (SE = 0.49) and 1.48 (SE = 0.27) respectively.

In order to understand the effect of base rate (C3 – 131 pps and C4 – 262 pps) on different experimental procedures and electrodes, a repeated measures ANOVA with factors experimental procedures, electrodes (middle (E12), and apical (E22)), and rates (C3

– 131 pps and C4 – 262 pps) were computed (Figure 5.5). The results showed no main effect of experimental procedures and electrodes. A pairwise comparison showed no significant difference across procedures and electrodes. As expected, there was a main effect of rate [$F(1, 4) = 10.06$, $p = 0.03$], and also a main effect of experimental procedures and rate [$F(2, 8) = 4.19$, $p = 0.05$]. The pairwise comparison between the two rates yielded a significant difference ($p < 0.05$) and produced a maximum score for C3 – 131 pps rate ($M = 2.26$, $SE = 0.24$) as compared to C4 – 262 pps ($M = 1.79$, $SE = 0.33$). In order to understand the interaction effect of experimental procedures and rate, series of one-way ANOVA was computed. Bonferroni corrected pairwise comparison showed no significant difference between the two base rates in ranking and Modified Melodies backward tasks. But apical electrode (E22) performance was significantly affected ($p < 0.01$) by increasing the base rate from C3 (131 pps) to C4 (262 pps) in Modified Melodies warp task.

This analysis suggests that the performance for 131 pps was substantially better than the performance with respect to 262 pps (high) rate. Interestingly, there was no advantage of different stimulation patterns on these experimental procedures. A paired t test between the pooled middle and apical electrodes as a function of base rate revealed that the performance of the apical electrode dropped significantly ($p < 0.01$) as the rate increased from C3 – 131 pps to C4 – 262 pps. The decrease in performance especially in apical region may be attributed to the mismatch between the rate and place. Pearson-Correlation revealed high positive correlation between modified melodies backward and modified melodies warp ($r = 0.73$, $p < 0.001$). There was no significant correlation between modified melodies warp and ranking or modified melodies backward and ranking ($p > 0.05$).

There was a large intersubject variability seen across different stimulation patterns in ranking (Figure 5.6), modified melodies backward modification (Figure 5.7), and modified melodies warp modification (Figure 5.8).

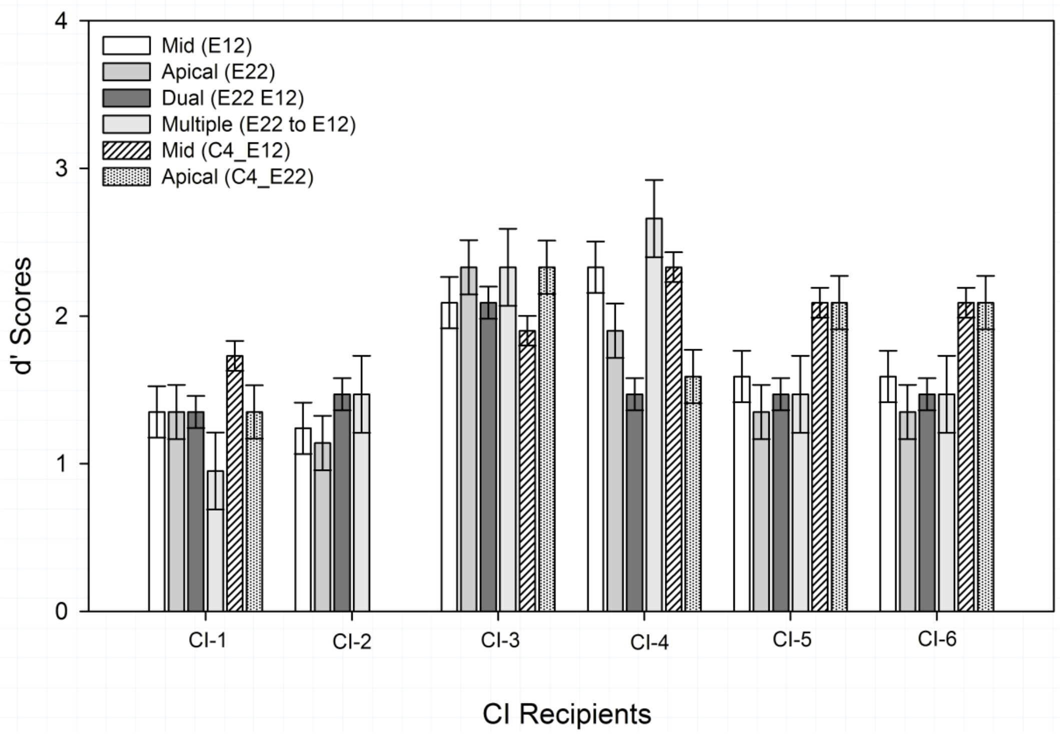


Figure 5.6: The individual subjects' d′ scores obtained from ranking procedure for all the stimulation patterns. CI-2 did not participate in the high rate stimulation pattern. The error bars represent the standard errors of the means.

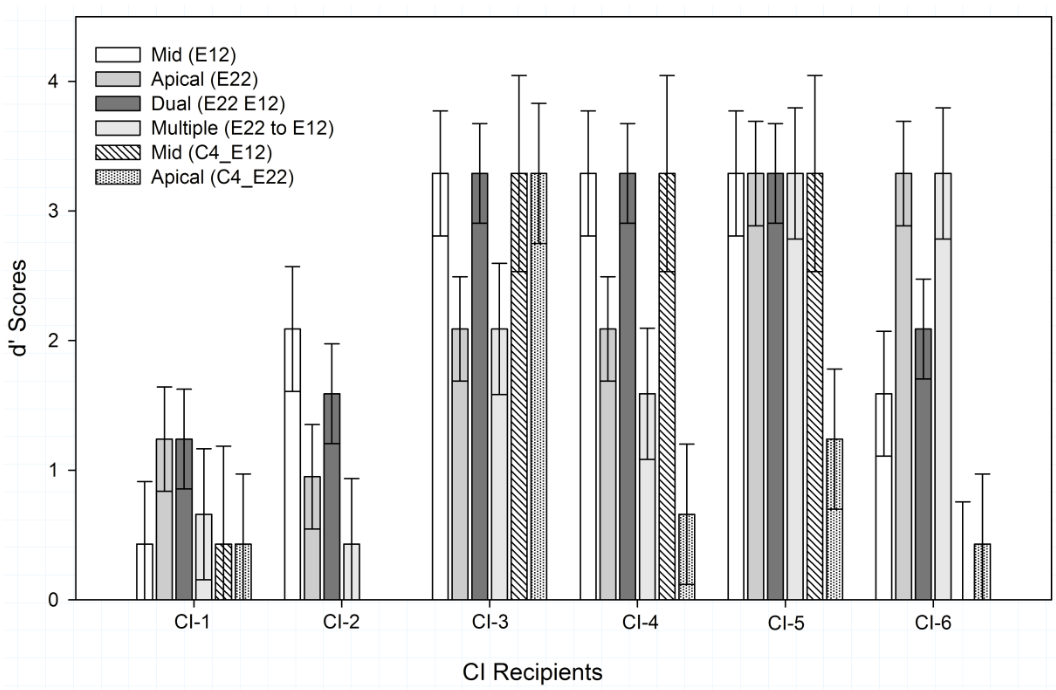


Figure 5.7: The individual subjects' d' scores obtained from modified melodies backward modification. CI-2 did not participate in the high rate stimulation pattern. The error bars represent the standard errors of the means.

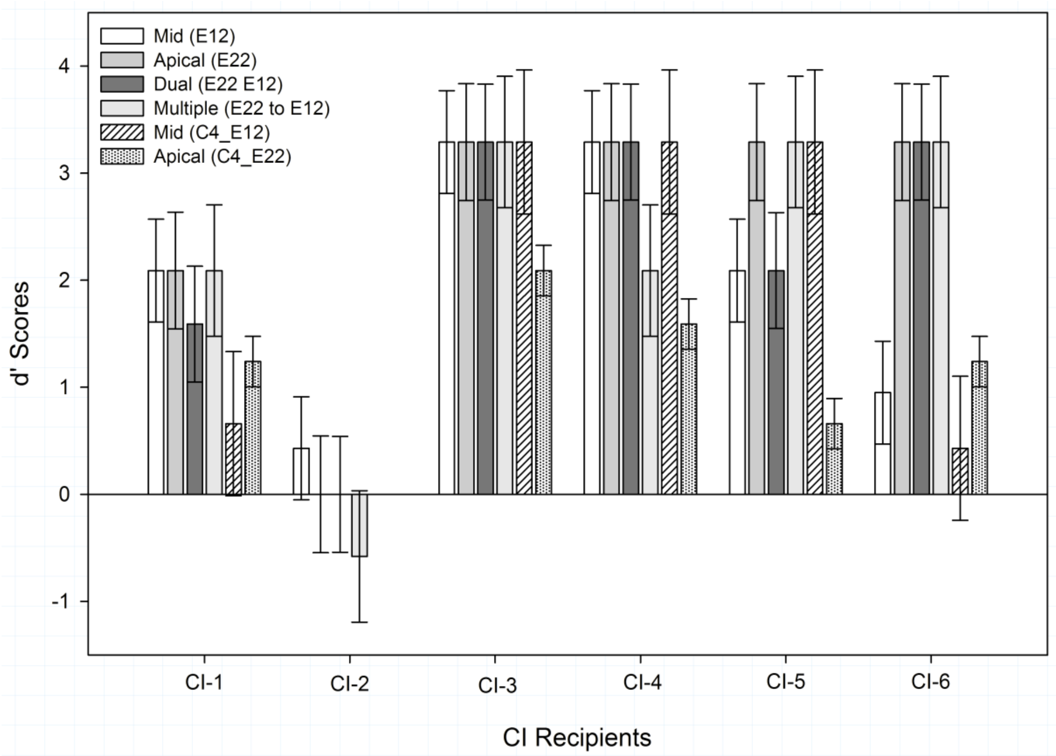


Figure 5.8: The individual subjects' d' scores obtained from modified melodies warp modification. CI-2 did not participate in the high rate stimulation pattern. The error bars represent the standard errors of the means.

An attempt was made to investigate the performance of CI recipients in the Modified Melodies test using different warp factors. This modification alters the interval size of the melody and varying the warp factor affects the interval shift between the notes of a melody (4.2.2.3.2). As the warp factor approaches 1.0, it becomes harder to detect the mistuned melody. The warp factors were tested in pairs (0.10 & 10.00, 0.25 & 4, 0.50 & 2, and 0.75 & 1.33). In this study, warp factor 0.10 & 10.00 were summed and compared with other experimental procedures (Figure 5.4), because all the subjects were able to perform only on these two warp factors (0.10 & 10.00). It should be noted that all the CI users did not perform on all the warp factors because of the increasing level of difficulty. The individual scores for all the warp factors are illustrated in Figure 5.9. When a subject scored at chance for a particular warp factor, the test was terminated and harder warp factors were not tested. As the warp factor decreases from 10.00 to 2.00 and increases from 0.1 to 0.50, the CI performance drops and this is evident in the Figure 5.9. The CI performances as a function of warp factors show a large intersubject variability.

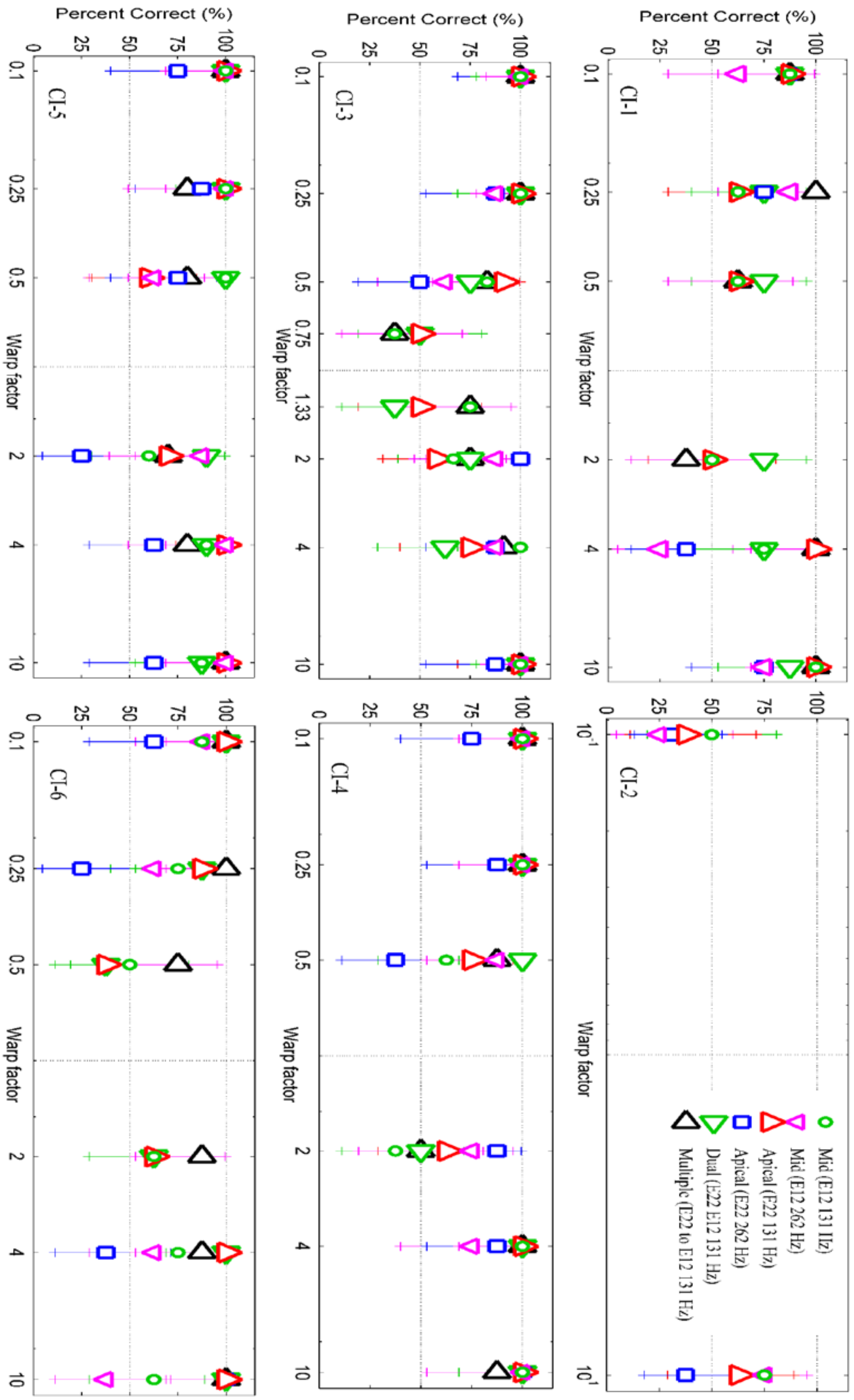


Figure 5.9: The individual modified melodies warp scores for different stimulation patterns. The 90% binomial confidence intervals are marked on the error bars, so that if the lower bound is above 50% it means that the score was significantly above chance according to a one-sided binomial test (p < 0.05).

In the Modified Melodies test, the subjects were able to detect the contour changes in the backward modification task and were able to obtain scores above chance for interval size judgements (warp modification). The overall performances across different electrode stimulation patterns were similar across the three experimental procedures. The performance on the apical electrode was significantly affected when the rates increased from C3 (131 pps) to C4 (262 pps), conversely, this was not seen in middle electrode. The overall performance using C3 – 131 pps rate was significantly better than the C4 – 262 pps rate especially in these pitch perception tasks.

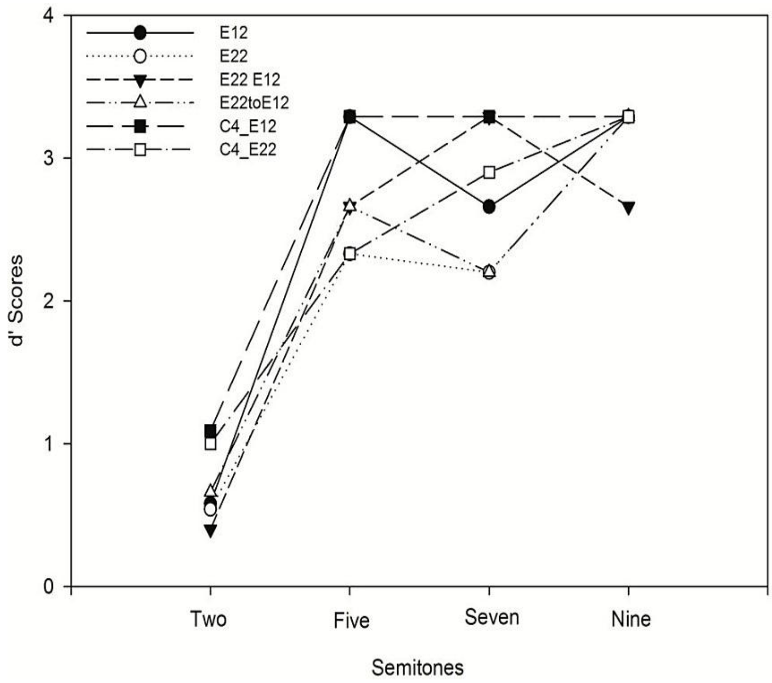


Figure 5.10: CI Ranking scores for six stimulation patterns as a function of semitone difference.

An attempt was made to study the performance of ranking with respect to semitone difference for different electrode stimulation patterns. A repeated measure ANOVA with factors, stimulation pattern and semitones were computed. Results suggested that there was no main effect of stimulation patterns [$F(3, 15) = 2.00$, $p = 0.15$] and pairwise comparison showed no significant difference among the electrode stimulation patterns. Conversely, there was a main effect of semitone difference [F (3, 15) = 63.45, p = 0.001]. The Bonferroni corrected pairwise comparison showed significant difference between two and five semitones, two and seven semitones, and two and nine semitones (p < 0.01). The remaining pairs were not significant (p > 0.05). There was no interaction between the electrode stimulation pattern and the semitone difference for the 131 pps base rate. An additional repeated measures ANOVA was computed for high rates (262 pps) at two stimulation patterns (N = 5). The factors were stimulation patterns (C4_E12 and C4_E22) and semitone difference. The analysis revealed no main effect of stimulation patterns, but there was a main effect of semitone difference [F (3, 12 = 113.17, p = 0.001]. Bonferroni pairwise analysis showed that the two semitones were significantly different from the other semitones (p < 0.01) and there was no significant difference among other semitone pairs.

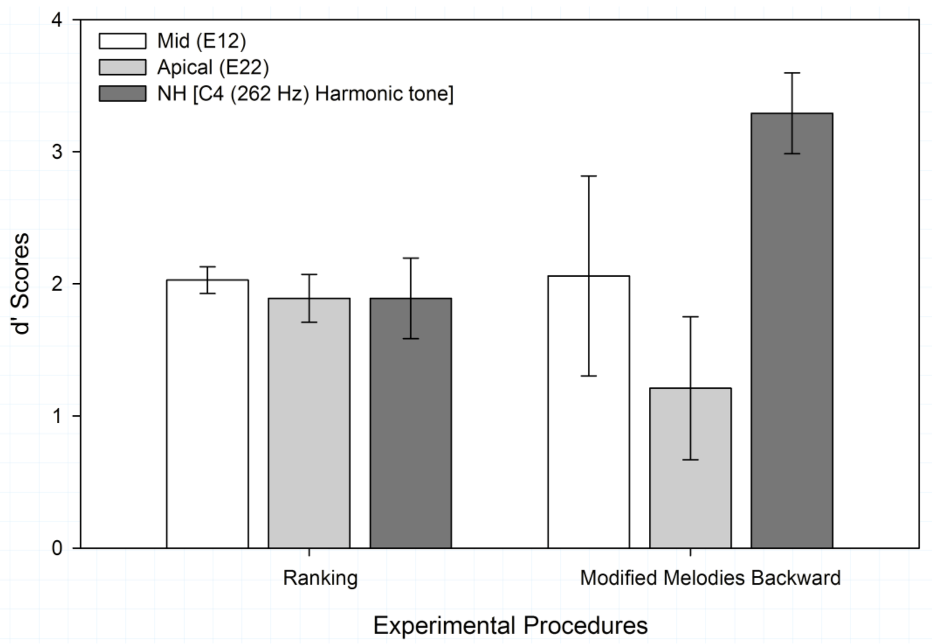


Figure 5.11: d' scores for the two procedure comparing NH subjects and CI recipients. The NH performance on harmonic tones (262 Hz) was compared with CI subjects' performance on middle (E12) and apical (E22) electrode using high base rate of 262 pps. The error bars represent the standard errors of the means.

An attempt was made to compare the results obtained from normal hearing subjects in the previous chapter (4.3) with the results obtained from CI subjects using high rate pulse train (Figure 5.11). The baseline stimulus (C4 (262 Hz) harmonic tones) in normal hearing subjects were compared with CI recipients at base rate of C4 (262 pps) on middle (E12) and apical (E22) electrodes. Interestingly, ANOVA revealed no significant difference among the three condition in ranking procedure [$F(2, 25) = 0.08$, $p > 0.05$]. However, in modified melodies backward [$F(2, 25) = 14.03$, $p < 0.001$] performance was highly significant among the conditions. The post hoc (Bonferroni corrected) analysis revealed that the scores for middle and apical electrodes were not significantly different ($p > 0.05$). But the scores between middle electrode and NH subjects ($p < 0.05$) and apical electrode and normal subjects ($p < 0.001$) were significant. The modified melodies warp was not analysed because NH subjects used harder warp (warp factors 0.75 and 1.33) as compared to CI subjects (warp factors 0.10 and 10.0). Overall, NH subjects performed significantly better than the CI counterparts. Although the middle electrode performance was slightly better than apical electrode, statistically, it failed to reach a significant level ($p > 0.05$).

5.4. Discussion

The study investigated the role of temporal pitch processing in CI users using different stimulation patterns. Six CI users participated in the experimental procedures using a base pulse rate of C3 (131 pps) and except CI-2, the remaining subjects participated in the same procedure using a high rate of C4 (262 pps). Surprisingly, there was no significant difference between the scores for the different stimulation patterns and the performance indicated that the CI users obtained the same information across the three stimulation patterns (Single, Dual, and Multiple Electrodes). This study investigated three hypotheses. Firstly, the performance would be better at the apical (E22) electrode compared to middle electrode (E12). Secondly, the performance would be better for multiple electrodes (E22toE12) compared to single or dual electrode stimulation. Thirdly, the performance would drop as the base rate increased from C3 (131 pps) to C4 (262 pps). The study did not support the first and the second hypotheses, but supported the third hypothesis which revealed that the performance was affected as a function of base pulse rate.

Comparison between Apical Vs Middle Electrode

The results obtained from the present study show little effect of electrode location on temporal pitch sensitivity. The performance was similar across apical (E22) and middle (E12) electrodes. The CI performance was affected significantly at apical electrode only for high base rates (C4 – 262 pps) compared to middle electrode. Interestingly, Pijl et al. (1995b) found an advantage of apical electrode ($p < 0.00001$) compared to basal electrode, but there was no significant advantage between apical and middle or basal and middle electrode. In their study, CI users reported that the apical stimulation sounded more musical and pleasing than the other electrode location. Similar to the present study, there was no advantage of apical electrode compared to middle electrode.

The CI performance in middle and apical electrode using C4 (262 pps) base rate was compared with normal hearing performance using C4 (262 Hz) harmonic tones (Figure 5.11). Result showed that the normal subjects were superior to CI counterparts in modified melodies backward and modified melodies warp modification. Interestingly, in ranking procedures, there was no significant difference between normal and CI subjects. This implies that the Modified Melody test is measuring something different to the ranking procedure, and perhaps the Modified Melodies test is a more sensitive test of musical pitch. The CI performance was reported only for five subjects. It would be worthwhile to test more CI recipients and to see their performance trend in this task.

Kong et al. (2009) reported no superiority for the apical (E4 or E3) region compared to the middle (E7 or E8) or basal (E11) in Med-El cochlear implant users. The performance was measured as a factor of increasing pulse rate for the different electrode locations. The results showed no significant effect of the electrode location and argued that there might be a mismatch between place and temporal pitch across individual subjects.

Comparison between Single, Dual, and Multiple Stimulations

In the present study, a single electrode (apical (E22) or middle (E12)), dual electrode (E22 & E12), and multiple (E22 to E12) electrode stimulations yielded good scores in the three procedures. The subjects were able to rank the musical notes well above chance and produced high scores for large semitone differences (Five, Seven & Nine semitone difference). Despite the low scores at two semitones, the performance was still above chance.

The contour and interval size are the two important constituents of a melody. The melody recognition is impaired when either one of these parameters are altered (Dowling & Fujitani, 1970). In the modified melodies backward, subjects were able to recognise the correct contour of the melody when presented as pulse sequences. Additionally, subjects were able to judge the interval size of a melody in the modified melodies warp procedure. The CI performance was well above chance in the Modified Melodies test.

The CI users were able to rank musical notes, identify the correct contour, and judge the interval size of a melody, when stimulated with changes in pulse rate on a single (apical (E22) or middle (E12)), dual (E22 & E12), and multiple electrode stimulations (E22 to E12). Swanson & McDermott (2010) investigated the melody recognition abilities of 3 CI users using base rate of C3-131 pps on a single apical electrode. In this study, the researchers used nudge modification, where a single note was shifted up or down in pitch by a specified number of semitones. Results were similar to the present study that the temporal cue alone is sufficient to convey melodic information.

Pijl and his colleagues tested 3 CI users (Pijl & Schwarz, 1995b) and 2 CI users (Pijl, 1997) on their ability to label the intervals of common melodies (initial phrase) as “in tune”, “flat” or “Sharp” . The melodies were presented as varying pulse rates at apical electrode (E18). Although subjects were able to correctly identify the “in tune” melodies, high performance was seen for minor 3rd (137 pps) compared to major 5th (163 pps) and major 6th (127). Similar results were observed in the present study using C3 (131 pps) base rate, where the subjects were able to rank and recognise the correct version of melody using pulse rate sequences. Additionally, stimulating the intracochlear electrodes by varying the pulse rate yielded substantially better scores than presenting the signal via the loudspeaker (Pijl, 1997). One of the plausible reasons is that the phase locking is more synchronized in the directly stimulated auditory nerve compared to the acoustically stimulated auditory nerve (Abbas & Miller, 2004).

Comparison between C3 (131 pps) Vs C4 (262 pps) base rates

The temporal pitch was investigated with respect to base pulse rates C3 (131 pps) and C4 (262 pps). In ranking and Modified Melodies backward procedures, there was no significant difference between the base rates at mid and apical locations. Conversely, Modified Melodies warp at apical location showed a significant decline in performance when base rate was increased from C3 (131 pps) to C4 (262 pps). This implies that the Modified Melody warp test is measuring something different to the ranking procedure, and perhaps is a more sensitive test of musical pitch. Zeng (2004) reported that the CI recipients can detect temporal pitch cues up to 300 pps. Conversely, the results obtained from the ranking procedure provide significant evidence that temporal pitch can be elicited at higher base rate of 262 pps. Similar to this study, Kong et al. (2009) reported a higher cut-off range for temporal pitch in CI recipients.

The overall performance at C4 (262 pps) base rate was significantly less than the performance at C3 (131 pps). Hence, the result supports the hypothesis that the performance deteriorates as a function of pulse rate. The subjects' performance was well above chance for the three procedures using both pulse rates. Performance on the apical electrode was more sensitive to pulse rate changes compared to the middle electrode. The CI performance significantly declined for high pulse rate C4 (262 pps) at apical electrode, but a similar decline was not evident in middle electrode. The apical region may not correspond to this particular characteristic frequency of C4 (262 pps) and this base pulse rate may be out-of-range for the apical electrode (E22). This could possibly explain the performance decline in CI recipients. Additionally, the apical region have better neural survival rate (Fayad & Linthicum, 2006; Bierer, 2007) and be more sensitive to much lower frequency range (C3 (131 pps) range) than the middle electrode (E12) region. The overall performance may also be attributed to the CI rate-place mismatch. The CI rate-place mismatch was reported in the earlier studies in terms of consistent and inconsistent cues (McKay et al., 2000; Luo et al., 2012).

Temporal pitch in CI recipients usually fades of at around 300 Hz (Zeng, 2002) and studies by Kong (2009 & 2010) provide evidence that temporal pitch can be extended beyond 300 Hz and reported high intersubject variability at high rates. The current study indicated that the recipients were able to rank notes well above chance for high base rate of C4 ranging from 262 Hz to 440 Hz. This suggests that the temporal pitch sensitivity does not significantly fade off at 300 Hz and that experimental procedure does have an influence on the performance as well. Ranking the notes is a primitive step towards music recognition and involves less mental abilities than melody recognition tasks. This is probably one of the reasons why the CI recipients could perform well on ranking task at a high base rate of C4 (262 pps). But the CI performance reduced for Modified Melodies test which requires higher mental abilities and performance was significantly affected for Modified Melodies warp. The Modified Melodies warp scores were significantly affected when base rate increased from C3 (131 pps) to C4 (262 pps) at the apical electrode. Therefore, temporal pitch sensitivity does reduce at 300 Hz but the temporal pitch range does vary depending on the tasks involved.

Implication for pitch processing in Normal hearing

Pitch processing mechanisms in cochlear implants deviate from that of the normal hearing individuals. Unlike normal listeners, the place and timing cues to pitch can be studied independently and separately in CI recipients (McKay et al., 2000). Even when a fundamental frequency is removed, still the pitch information is conveyed by the harmonic complexes (Missing Fundamentals) in normal listeners. Earlier studies have reported that the frequency discrimination score for harmonic tones were better than the pure tones (Zeitlin, 1964; Platt & Racine, 1985; Moore & Peters, 1992). In musical interval judgement tasks, the performance improves in normal hearing individuals when additional harmonics were added to the musical notes (Houtsma & Smurzynski, 1990). The performance significantly improved when the low numbered harmonics increased from 7 to 13 and performance remained static after 13th harmonics. Interval judgments were possible at the high harmonics range of 20th – 30th harmonics. Similarly, Moore & Peters (1992) showed that the frequency difference limens were lowest (best performance) for tones containing harmonics from 1 – 12, moderate for tones having 1 – 5 harmonics, and highest (worst performance) for pure tone. Their result suggests that the performance improved when additional harmonics were added to the signal.

In the present study, when pitch information was delivered to a single electrode (E12 or E22) or dual electrodes (E22 & E12) or multiple electrodes (E22 to E12), the performance was not affected as a function of stimulation patterns nor as a function of electrode location. It was hypothesized that adding an electrode may provide an extra cue for performing these pitch perception tasks. Surprisingly, there was no influence of different stimulation patterns and recipients obtained similar information from these stimulation patterns. It suggests that CI recipients are unable to combine cues from different places in the cochlea to give a “stronger” cue. This seems different to NH subjects, where extra harmonics (1 – 12 harmonics) in a signal improves the performance as opposed to pure tones (single harmonic) (Moore et al., 1985).

5.5. Conclusion

The present study aimed to investigate the role of temporal pitch in electrical hearing as a function of stimulation patterns, base pulse rate, and electrode location. The three experimental procedures were: ranking, Modified Melodies test – backward, and

Modified Melodies test – warp. Results indicated that subjects were able to rank, identify the melodic contour in backward modification, and judge the interval size in warp modification. The CI performance was well above chance and statistically there was no significant difference between the stimulation patterns [Single (E12 or E22), Dual (E22 & E12), and Multiple (E22 to E12)] across the three procedures. Interestingly, there was no apical advantage seen in this study, but there was a significant decline in performance for high base rate (C4 – 262 pps) and this was attributed to the rate-place mismatch. Although there was a high intersubject variability, the result revealed a high correlation between modified melodies backward and modified melodies warp modifications. This finding did not support the hypothesis that the stimulation patterns had an effect on the pitch perception tasks, but supported the hypothesis that the performance was affected as a function of base pulse rate. The present study along with the results obtained from the previous studies (Pijl & Schwarz, 1995b; McDermott & McKay, 1997; Pijl, 1997; Swanson & McDermott, 2010) provided sufficient evidence that temporal pitch alone is sufficient for conveying musical information. However results showed no specific advantage for different stimulation pattern for the experimental procedures. Results suggest that CI recipients received same information from different stimulation pattern and failed to combine cues from different places in the cochlea to give a "stronger" cue.

Chapter 6: Investigating Cochlear Implant Place Pitch and Temporal Pitch Perception

Abstract

Place and temporal cues to pitch can be studied independently in cochlear implant (CI) recipients. Temporal pitch has been shown to convey melody in CI recipients, but few studies have investigated place pitch alone. The present study investigated the role of place pitch and temporal pitch using different pitch perception tasks. Six post-lingually-deafened CI recipients, with at least one year of implant usage, participated in the study. The four experimental procedures were: 4AFC Discrimination, 2AFC Ranking, and 2AFC Modified Melodies test (backward and warp modification). The three stimulus types were: (1) Pure tones with base frequency of C5 (523 Hz), providing place cues only; (2) Harmonic tones with base frequency of C3 (131 Hz), providing temporal cues only; (3) Harmonic tones with base frequency of C4 (262 Hz). The stimuli were presented via loudspeaker at a comfortable loudness level. The recipients used their own sound processor. The overall scores for discrimination and ranking were high for all three stimulus types, however, three subjects showed pitch reversals in ranking the C4 harmonic tones. In the Modified Melodies test, scores were similar for C5 pure tones and C3 harmonic tones, while scores using C4 harmonic tones were worse and mostly near chance. The C4 harmonic tones potentially offered temporal and place cues. However, their fundamental frequency range of 262 to 523 Hz was probably above the upper frequency limit of temporal pitch for most subjects, and additionally place pitch cues may have been ambiguous. This may explain the observed pitch reversals, which prevented good performance in a melody task. Scores with C5 pure tones were as good as those with C3 harmonic tones. These results suggest that CI place pitch may convey melodic pitch information, but the contribution of brightness cannot be completely ruled out.

6.1. Introduction

Music has been an integral part of every culture around the world. This is evident from a simple alarm clock tune to complex symphony; almost every single event in the real world can be associated with music. Thus, music plays a significant role in everyday life. Musical listening is effortless in normal hearing individuals. On the other hand, it is extremely effortful for people with cochlear implants (CI). Although there is a steep rise in the number of people opting for cochlear implantation, it remains an irony that the majority of them do not appreciate music. The contemporary CI devices convey adequate speech information in quiet, but the performance deteriorates drastically when CI users perform music perception tasks. One of the underlying factors is that the pitch is poorly coded in these implantable devices (McDermott, 2004; Looi, 2008; Nimmons et al., 2008). Therefore, the reduced pitch perception skills substantially affect their musical ability and restrict them from enjoying real world music.

The pitch sensation in cochlear implants can be evoked by either stimulating different electrodes (place cue) or by varying the pulse rate in an electrode (temporal cue). Cochlear implant technology provides a suitable platform for studying place pitch and temporal pitch separately and independently (McKay et al., 2000). These two cues can be investigated either by the direct stimulation method (Dorman et al., 1994; Pijl & Schwarz, 1995a,b; McDermott & McKay, 1997; McKay et al., 2000; Swanson & McDermott, 2010) or the traditional method of presenting the acoustic output from the soundcard to the speech processor via direct audio input or delivering the signal via loudspeakers. The traditional method has been incorporated by several researchers for understanding pitch perception in CI recipients (Gfeller et al., 2002; Kong et al., 2004; Looi et al., 2004; Nimmons et al., 2008; Galvin et al., 2009). Direct stimulation studies provide a suitable platform to better control the amount of cues provided to CI recipients, but do not really test a realistic environment. However, a widely quoted direct stimulation study by Pijl & Schwarz (1995b) investigated cochlear implant melody and interval recognition using a bipolar mode of stimulation. The melodies were presented as varying pulses and the pulse rate corresponds to the fundamental frequency (F0) of each note in the melodies (F0 range = 75 – 600 pps). Results showed that the CI users were able to recognise melody and judge interval size when the notes were presented by varying the pulse rate on a single electrode and elucidated that temporal pitch alone is quite efficient in delivering melodic pitch information. Unfortunately, the direct stimulation method does not replicate the real world scenario (Gfeller et al., 2002).

A recent study by Singh et al. (2009) investigated whether cochlear implant melody recognition was affected by varying melody frequency range, harmonicity, and number of electrodes. In the first experiment (N = 11), CI recipients were asked to recognise melodies played in three frequency ranges (Low = 104-262 Hz, Middle = 207-523 Hz, and High = 414-1046 Hz). The procedure had 12 isochronous melodies administered as a close-set forced choice procedure. The melodies were presented directly (direct audio input) into their speech processor. CI users obtained high scores in high frequency ranges compared to chance level performance in middle frequency range. The low frequency range produced scores below chance level. In the second experiment (N = 4), the cochlear implant melody recognition was investigated using pure tones in the three frequency ranges. The results reveal that the scores were better using pure tones compared to complex tones in experiment 1. There was no correlation between phoneme recognition and melody recognition and they reported that these two tasks involve different processes for coding acoustic signals (Speech and Music). Overall, they concluded that current cochlear implants do not encode either temporal pitch or place pitch cues adequately. They emphasised different signal processing strategies for speech and music.

Swanson et al.(2009) used the Modified Melodies test to investigate cochlear implant place pitch perception. In their study, several types of pitch modifications were tested. The melodies were presented as pure tone with frequency ranging from C5 – C6 (532 -1046 Hz) through the ACE strategy on a freedom speech processor. Using place-pitch cues alone, all six subjects were able to identify the correct contour of a melody in modified melodies backward modification. In modified melodies nudge modification, one note of the melody was shifted away from its correct pitch by a specified number of semitones either two or five semitones. The majority of them (four out of six) were able to detect a large semitone difference (five-semitones) and only two subjects were able to detect the two semitone shift in the modified version of the melody. Additionally, Swanson & McDermott (2010) used the Modified Melodies test to investigate cochlear implant place pitch and temporal pitch perception using direct stimulation. Three conditions were tested (Place C5 – Pure tones with octave starting from 523 Hz; Place C3 – Pure tones with octave starting from 131 Hz, and Rate C3 – Pulse rate equal to the fundamental frequency of each note, octave starting from 131 Hz). In their study, modified melodies nudge modification was used. Blocks of trials were performed with shifts in the range 0.5 semitones to 7 semitones. Results showed that three CI users were able to recognise a correct version of a melody using temporal cues in isolation. Interestingly, they found that all subjects (n = 6) were able to recognise melodies based on place cues alone. The results from these two studies suggest that both temporal pitch and place pitch alone can convey melodic pitch information. The CI performance in the Modified Melodies test suggests that the recipients were listening only to pitch changes while performing this task, but the contribution of brightness cannot be completely ruled out (Swanson et al., 2009).

Aforementioned studies reveal that temporal pitch alone is capable of providing melodic pitch information. Earlier studies investigating place pitch were centred on psychophysical procedures, such as, discrimination (McKay et al., 2000; Geurts & Wouters, 2001 & 2004), and ranking (Laneau & Wouters, 2004). Surprisingly, not many studies have explored place pitch in a musical context.. Therefore, the aim of the present study was to investigate the role of place pitch and temporal pitch in CI recipients using different pitch perception tasks. The study consisted of four experimental procedures, which were discrimination, ranking, modified melody test using backward modification, and modified melody test using warp modification. The three types of stimuli used in the experimental procedures were, (i) C3 (131 Hz) harmonic tones; (ii) C4 (262 Hz) harmonic tones; and (iii) C5 (523 Hz) Pure tones.

Previous research on temporal pitch reveals that the upper cut-off range for temporal pitch is around 300 Hz (Zeng, 2002; Kong et al., 2009). Therefore, using stimuli at the range of 131 to 220 Hz will evoke temporal pitch cues in CI recipients. Thus the C3 harmonic tone range was implemented in this study. Additionally, the quadrature envelope in the sound processor produce minimal ripple for pure tones which suggests a minimal temporal cue and maximum place-of-excitation cues. The pure tones used in this study had a range (C5 range - 523Hz – 880 Hz) which is well above the upper cut-off limit for temporal pitch. Thus, the C3 harmonic tones and C5 pure tones provided temporal and place pitch in isolation respectively. Conversely, C4 harmonic tones with a range of 262 Hz to 440 Hz provide a combination of both place and temporal pitch cues. Using these stimulus types with a specific frequency range would provide an in-depth knowledge of how the temporal and place pitch in isolation are coded and how effectively these cues are utilised by CI recipients’ in various pitch perception tasks.

6.2. Method

6.2.1. Subjects

Six CI recipients who participated in the previous experiment participated in this study as well. The details of the CI recipients are mentioned in the previous chapter (Table 3).

6.2.2. Experimental Procedure

The four experimental procedures (Discrimination, Ranking, Modified Melodies test – backward modification, and Modified Melodies test – warp modification) were used to evaluate the role of place pitch and temporal pitch in CI recipients. The details of each procedure are described in chapter 4 (4.2.2).

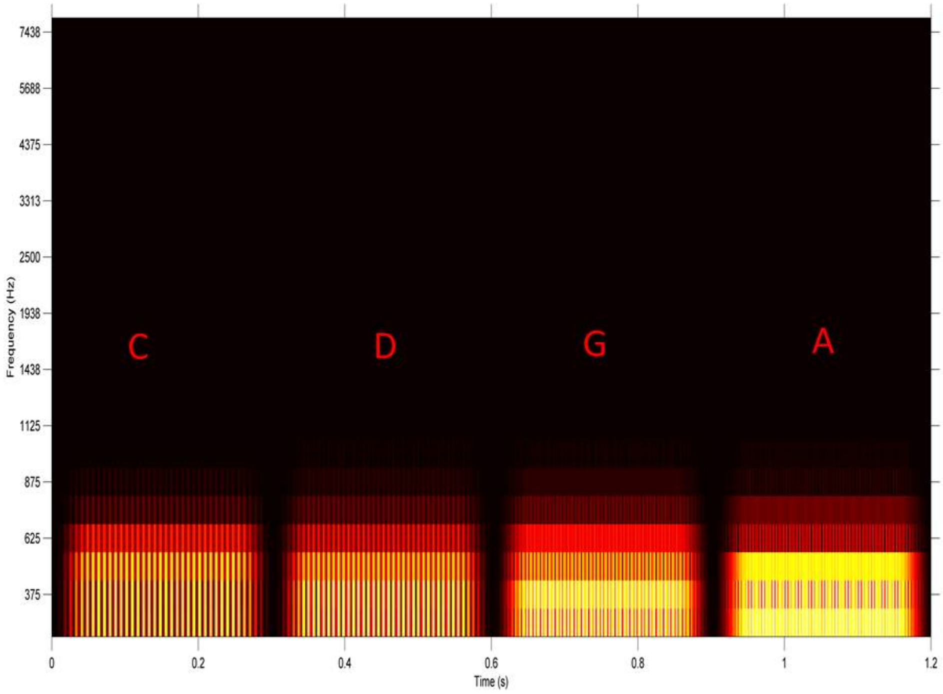
6.2.3. Stimulus Description

All the stimuli were synthesised on a PC at a sampling rate of 16 kHz, and presented via loudspeaker in a sound-treated room. During the testing, all subjects used their own sound processor which implemented ACE processing strategy. Each note was 300 ms in duration with a smooth (sinusoidal-shaped) rise and fall time of 50 ms. The stimuli were presented at a comfortable loudness level for all the CI recipients. The details of the three stimulus types are described below.

6.2.3.1. Harmonic Tone with Base Frequency C3 (131 Hz)

The musical notes were played using harmonic tones with frequency ranging from C3 (131 Hz) to A3 (220 Hz). The resulting note patterns stimulate apical electrodes which correspond to low frequencies (E17 – E22). There are two ways of visualising the pattern of stimulation of the intracochlear electrodes. Firstly, figure 6.1(above) shows the filter bank output for C3 harmonic tones. The representation is similar to the spectrogram with

22 filter banks. The Y axis represents the centre frequency of filters which drive the corresponding electrode. Black colour represents zero amplitude, while white represents maximum amplitude. Secondly, the electrodogram (Figure 6.1 (below)) shows the pulse pattern of stimulation for C3 harmonic tones. Each vertical line represents a stimulation pulse and the height of the pulse represents the amplitude. A zoomed in version of musical note “C” was provided in Figure 6.2 which provides a clear view of pulse sequences with varying amplitude across different electrodes. Thus, the C3 harmonic tones modulate the carrier pulse train and the amplitude fluctuations of the pulse train correspond to the fundamental frequency of the incoming musical signal. The amplitude modulations (temporal cues) are represented by vertical line stripes in Figure 6.1 (above) & 6.2. These electrodogram provides clear evidence that the C3 harmonic tones convey exclusive temporal pitch information to the CI recipients. All the stimuli had the same spectral profile (i.e., same timbre).



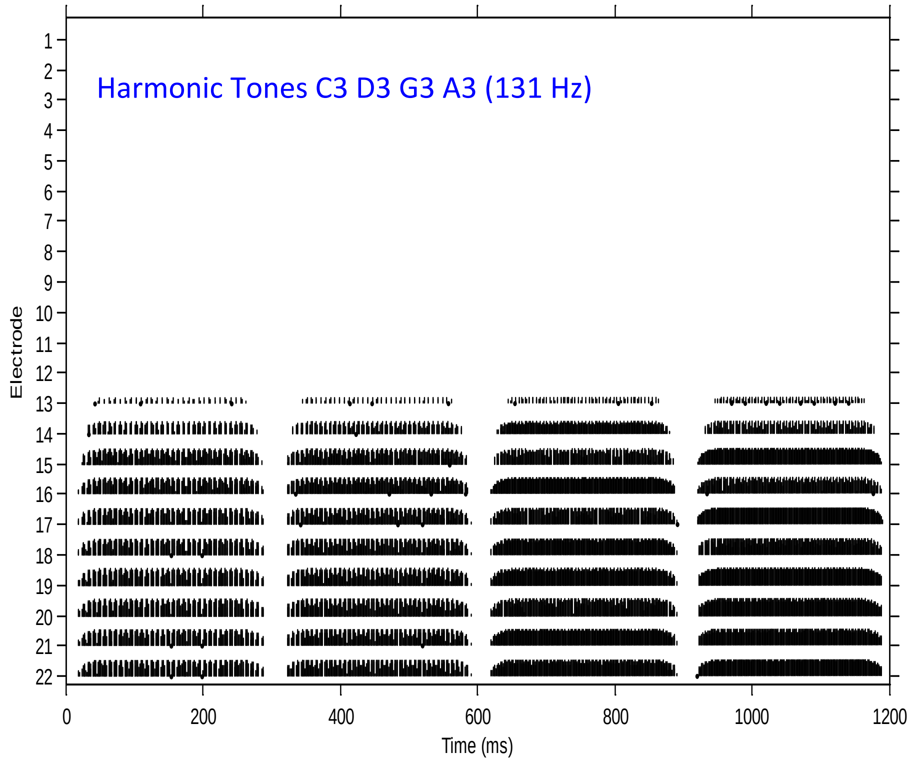


Figure 6.1(above): Electrodogram (filter bank output) of 4 musical notes played at C3 – 131 Hz base frequency. Y axis denotes frequency and x axis denotes time. Figure 6.1(below) Electrodogram of 4 musical notes. Y axis denotes electrode number and x axis denotes time.

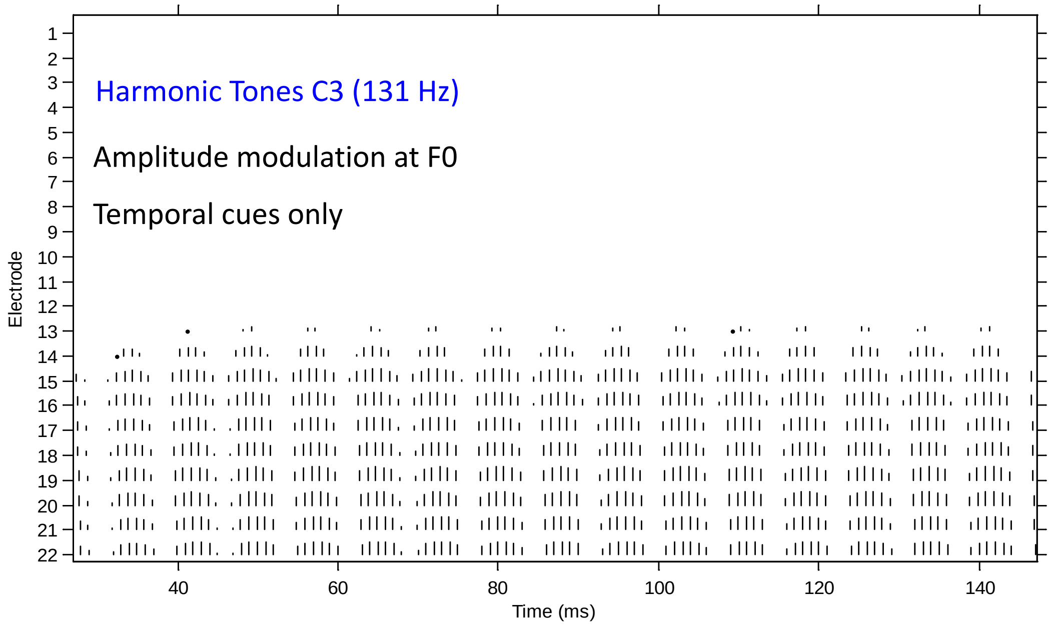


Figure 6.2: Electrodogram of "C" musical notes played at C3 – 131 Hz (Harmonic tone) base frequency. Y axis denotes electrode number and x axis denotes time.

6.2.3.2. Pure Tone with Base Frequency C5 (523 Hz)

In this condition, the notes were played using pure tones with frequency ranging from C5 (523 Hz) to A5 (880 Hz). For pure tones, the quadrature envelope has very little ripples, thereby minimizing the amplitude modulation, therefore the processed signal consists only of place pitch information (Swanson et al., 2007). The horizontal striations schematised in the Figure 6.3 demonstrate that these pure tone signals (C5 to A5) clearly stimulate distinct electrodes thereby providing exclusive place pitch information. Figure 6.3 demonstrates that for a single musical note, specific electrodes are being activated. Bright white colour (Figure 6.3 (above)) illustrates that a single electrode is being stimulated and bright white light fades off as the distance from the stimulation electrode increases. For example, for musical note C, E20 is being maximally stimulated (bright white colour) and fades off (dull black colour) at around E18. Even the pulse pattern in Figure 6.3 (below)) shows that distinct electrodes are being stimulated for different pure tone stimuli. Each pure tone activates multiple electrodes due to broad analysis filters in the processor. A few notes such as C5 and D5 activated the same set of electrodes with different amplitude; under these circumstances the place cue depends on the centroid of the stimulation pattern (Laneau & Wouters, 2004a; McDermott, 2004; Swanson, 2008).

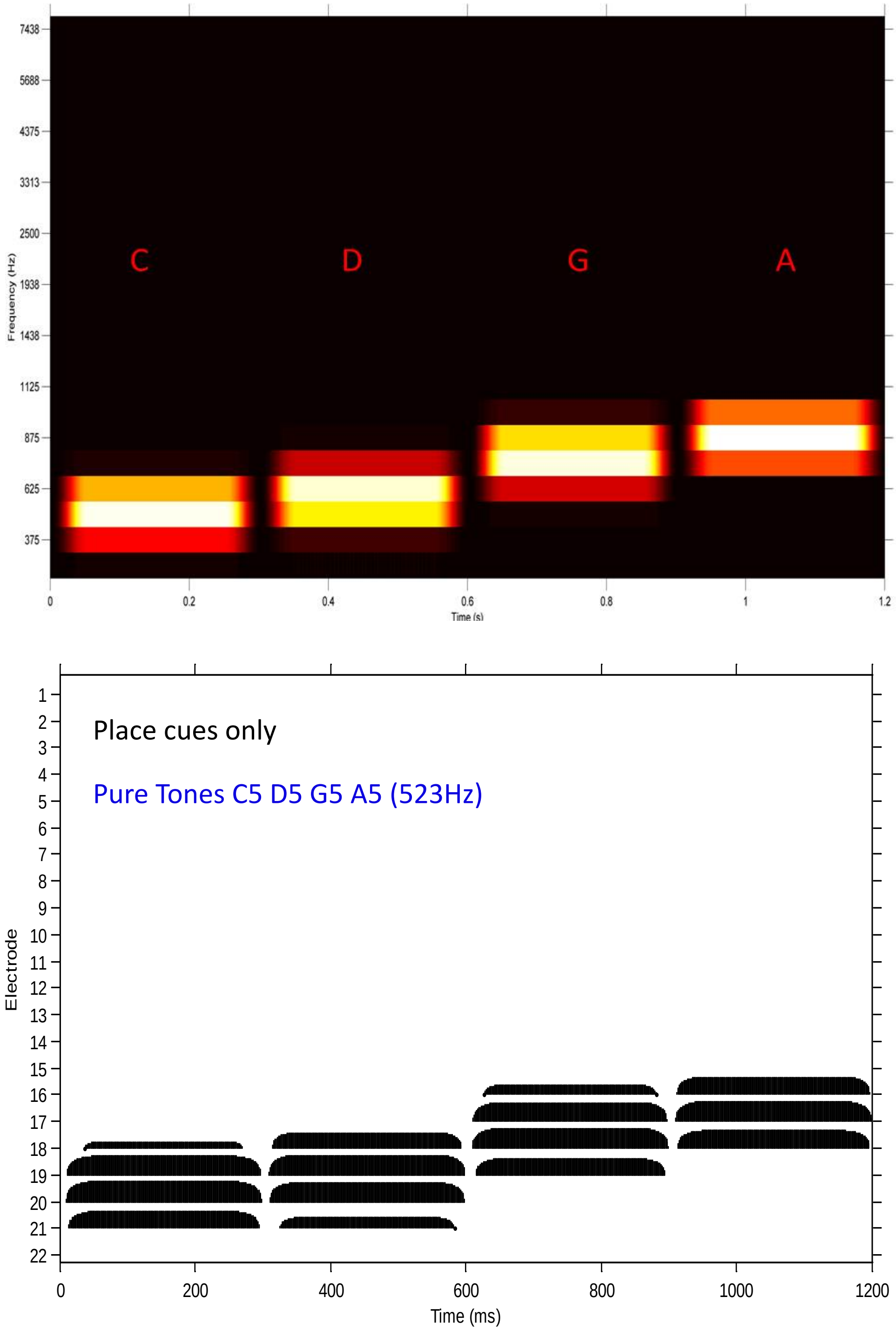
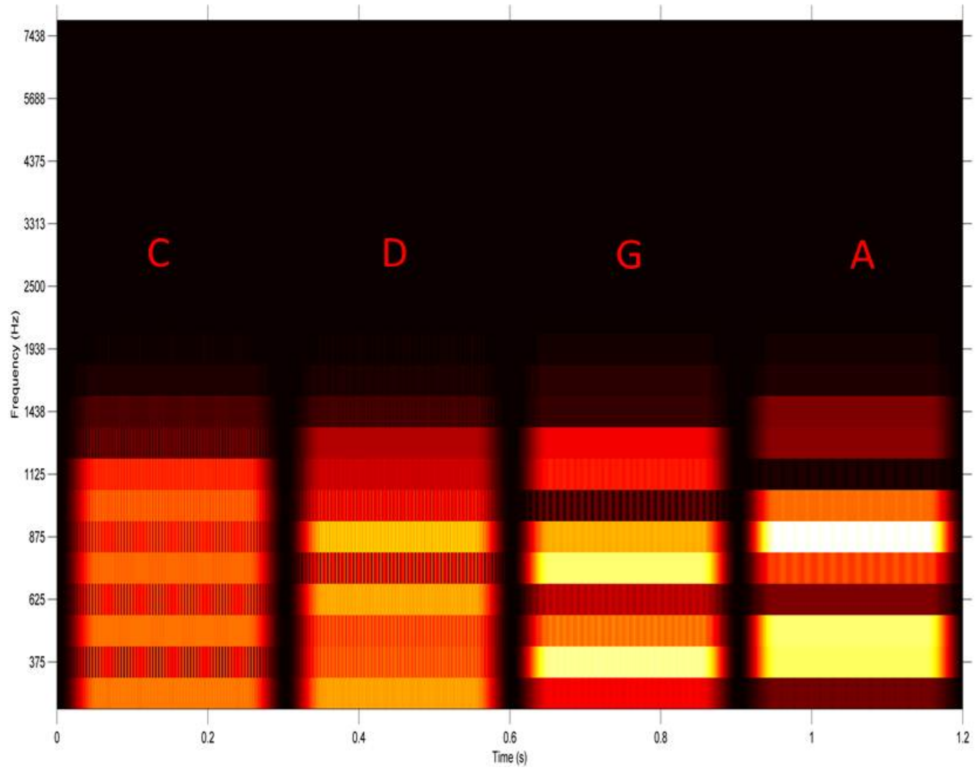


Figure 6.3: (above): Electrodogram (filter bank output) of 4 musical notes played at C5 – 523 Hz base frequency. Y axis denotes frequency and x axis denotes time. Figure 6.3(below) Electrodogram of 4 musical notes. Y axis denotes electrode number and x axis denotes time.

6.2.3.3. Harmonic Tone with Base Frequency C4 (262 Hz)

The musical notes were played using harmonic tones with frequency ranging from C4 (262 Hz) to A4 (440 Hz). There was a 9 semitone range for both the melodies (Old MacDonald & Twinkle Twinkle). The C4 Harmonic tones stimulate multiple electrodes with varying modulation rates (Figure 6.4). Indeed, the C4 harmonic tones may provide both temporal and place cue information to the CI recipients and these are the most common harmonic complexes present in the real world. The electrodogram (Figure 6.4) shows that both temporal (vertical line stripes) and place cues (bright white light) are not clearly represented for the four musical notes (C4, D4, G4, and A4). But for C3 (131 Hz) harmonic tones, the temporal cues (vertical line stripes) are clearly represented for the four musical notes and for C5 (523 Hz) pure tones, the place cues (bright white light) activate distinct set of electrodes for different musical notes, thereby providing exclusive place pitch information. All the stimuli had the same spectral profile (i.e., same timbre).



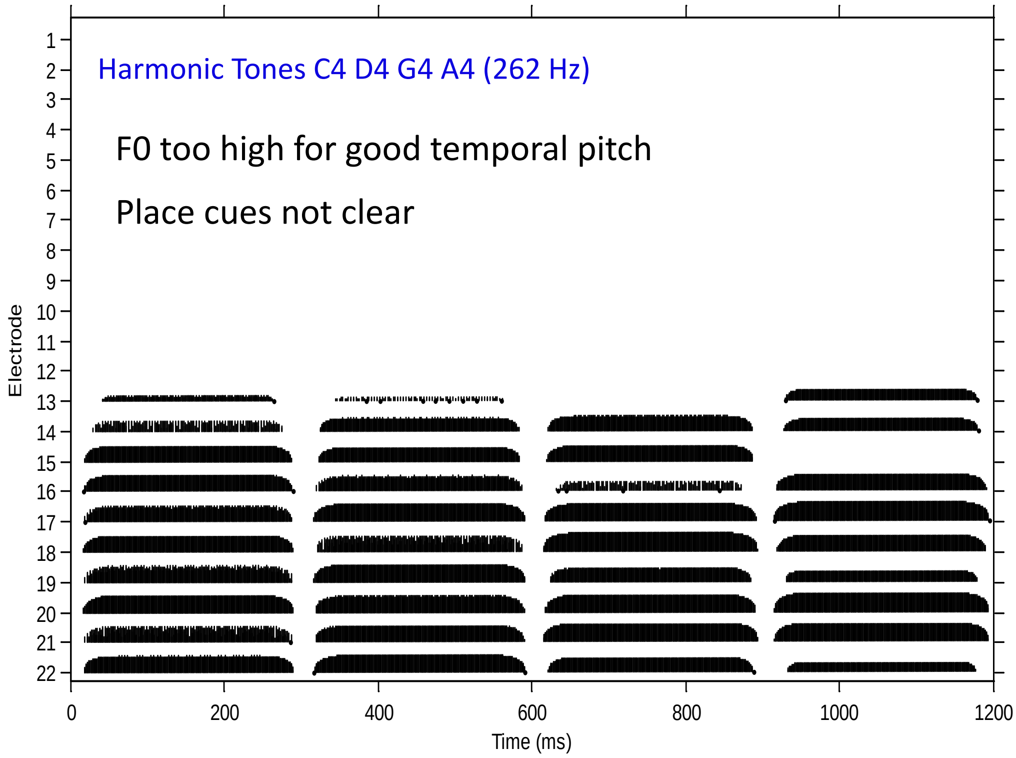


Figure 6.4: (above): Electrodogram (filter bank output) of 4 musical notes played at C4 – 262 Hz base frequency. Y axis denotes frequency and x axis denotes time. Figure 6.4(below) Electrodogram of 4 musical notes. Y axis denotes electrode number and x axis denotes time.

6.3. Results

The present study investigated the role of place pitch and temporal pitch using three stimulus types in CI recipients. A single score was obtained for each procedure and stimulus type by summing across note pairs in discrimination and ranking, and summing across warp factors 0.10 and 10.00 in modified melodies warp. The group mean scores are shown in Table5.

Table 5: Percentage correct scores for the four experimental procedures.

Stimulus Types	**Discrimination** (4-AFC)	**Ranking** (2-AFC)	**Modified Melody Backward** (2-AFC)	**Modified Melody Warp** (2-AFC)
C3 Harmonic Tone	81	80	84	70
C4 Harmonic Tone	75	72	63	49
C5 Pure Tone	90	91	89	75

The discrimination task was a 4-AFC method, whereas the other procedures were 2-AFC methods, the proportion correct scores could not be directly compared. Instead, the scores were converted into d' scores based on table A5.7 of Macmillan & Creelman, (2005). The group mean d' scores for the three stimulus types and four procedures are illustrated in Figure 6.5 and the statistical analysis was administered on these d' scores. There was a large intersubject variability across the four experimental procedures, but this was more prominent in the Modified Melodies test. The individual scores of discrimination (Figure 6.6), ranking (Figure 6.7), modified melodies backward (Figure 6.8), and modified melodies warp (Figure 6.9) are illustrated in the figures below.

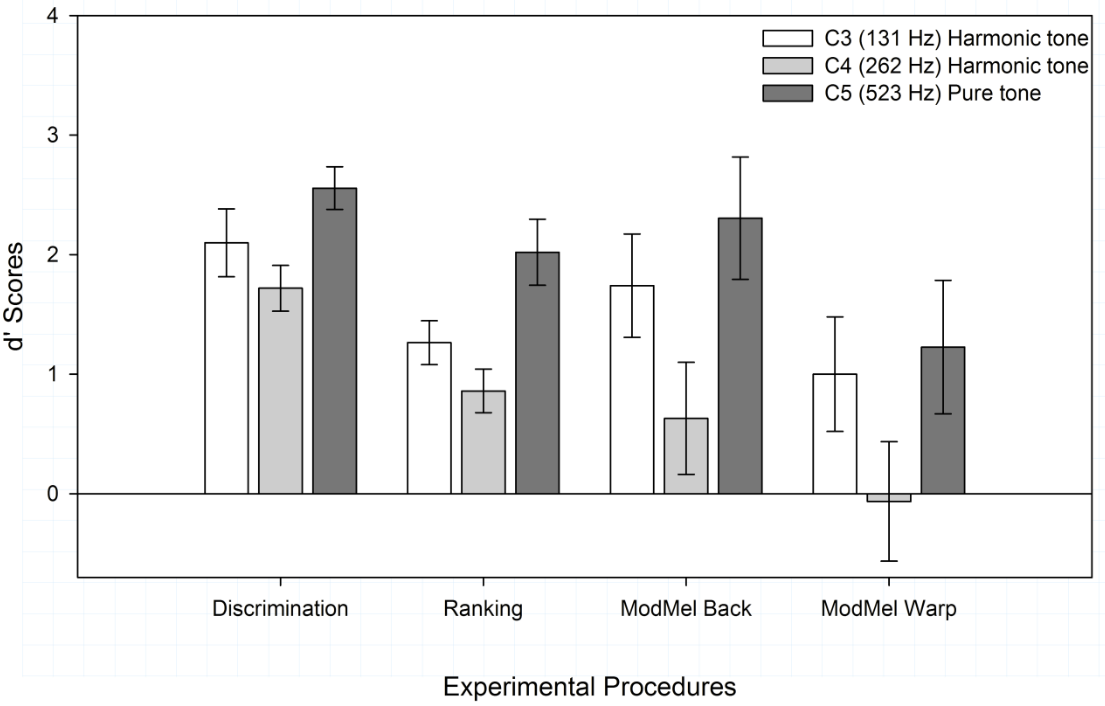


Figure 6.5: d' scores for the three base frequencies across four experimental procedures. A d' score of 0 indicates that the performance was at chance. The maximum d' score is 4. The error bars represent the standard errors.

A repeated measures analysis of variance (ANOVA) with the factors, experimental procedures (discrimination, ranking, modified melodies backward, and modified melodies warp) and stimuli (C3 harmonic tones, C4 harmonic tones, and C5 pure tones) were computed. The ANOVA revealed a main effect of experimental procedures [$F(3, 15) = 5.67$, $p = 0.08$]. The Bonferroni corrected pairwise comparisons showed that the discrimination scores ($M = 2.12$, $SE = 0.17$) were significantly higher than the ranking scores ($M = 1.38$, $SE = 0.13$, $p < 0.05$). Interestingly, the pairwise comparisons for other experimental procedures were not significant. The mean scores for modified melodies backward ($M = 1.55$, $SE = 0.40$) were higher than the mean scores for modified melodies warp ($M = 0.72$, $SE = 0.43$). Additionally, there was a main effect of stimuli [$F(2, 10) = 9.54$, $p = 0.005$]. However, the pairwise comparisons showed no significant difference among the three stimulus types ($p > 0.05$). Overall, high mean scores were obtained for C5 pure tones (M = 2.02, SE = 0.28) compared to C3 harmonic tones (M = 1.52, SE = 0.27). The least mean scores were obtained for C4 harmonic tones (M = 0.78, SE = 0.30). The interaction between experimental procedures and stimuli were not significant [F (6, 30) = 1.40, p = 0.24].

Pearson-Correlation revealed high positive correlation between discrimination and ranking (r = 0.74, p < 0.001). High correlation was observed between modified melodies backward and discrimination (r = 0.69, p < 0.001), modified melodies backward and ranking (r = 68, p < 0.005), and modified melodies backward and modified melodies warp (r = 0.74, p < 0.001). There was no significant correlation among other pairs.

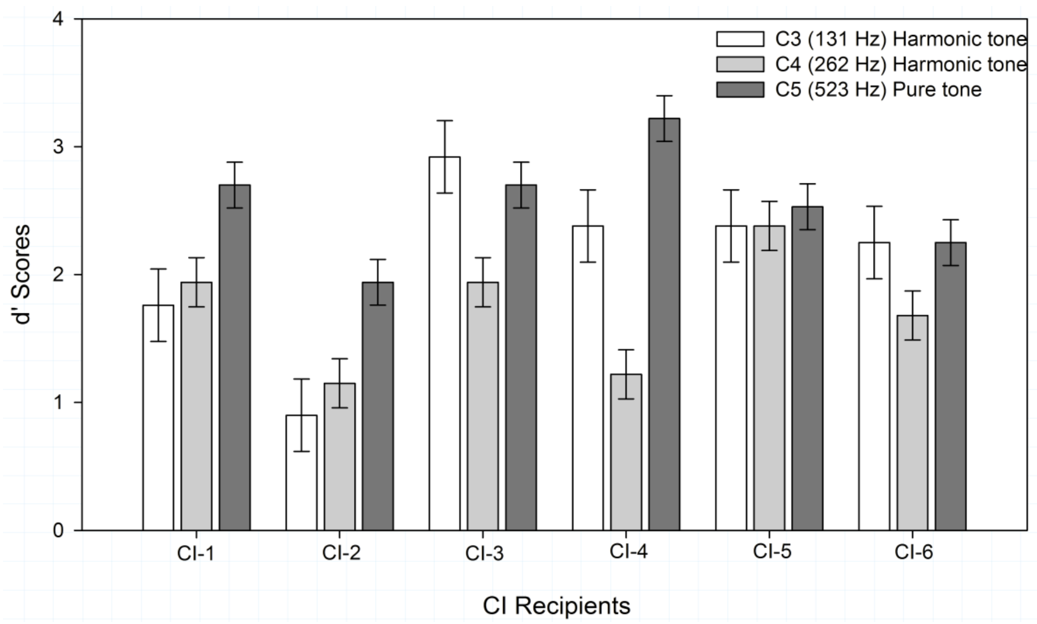


Figure 6.6: Individual d' scores for the discrimination procedure across the three base frequencies. The error bars represent the standard errors.

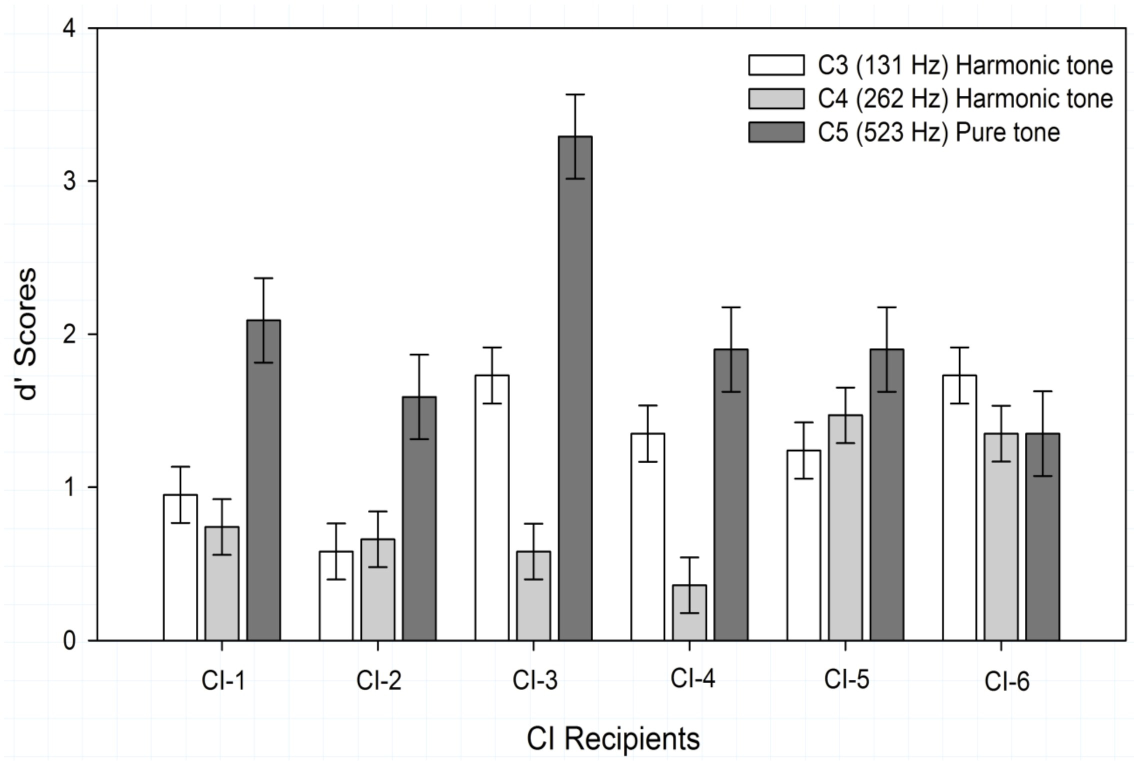


Figure 6.7: Individual d' scores for the ranking procedure across the three base frequencies. The error bars represent the standard errors.

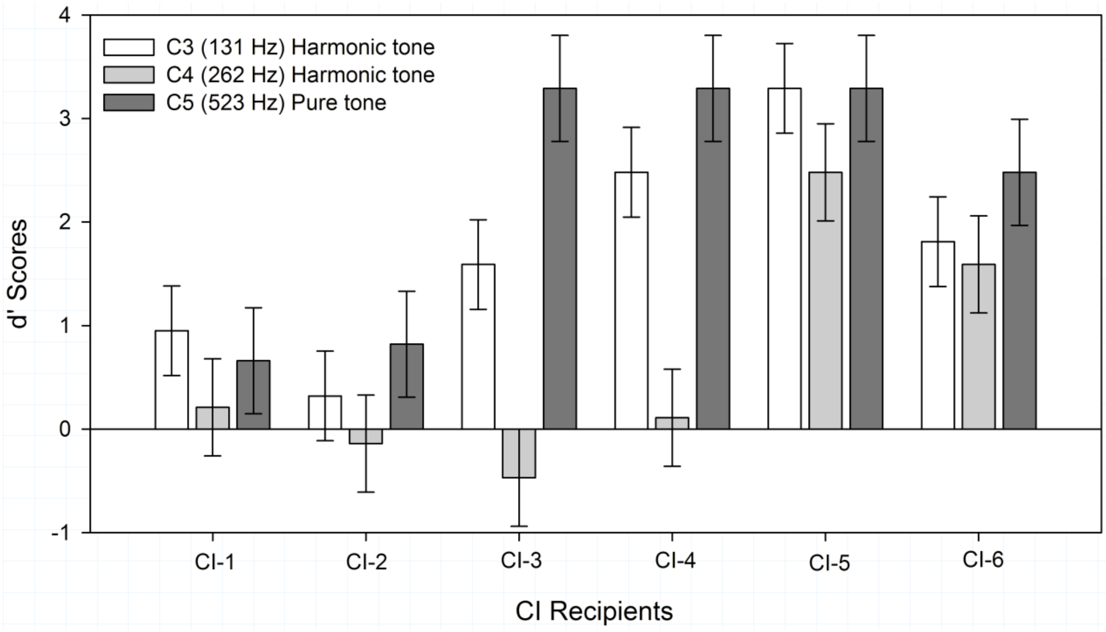


Figure 6.8: Individual d' scores for the modified melodies backward modification across the three base frequencies. The error bars represent the standard errors.

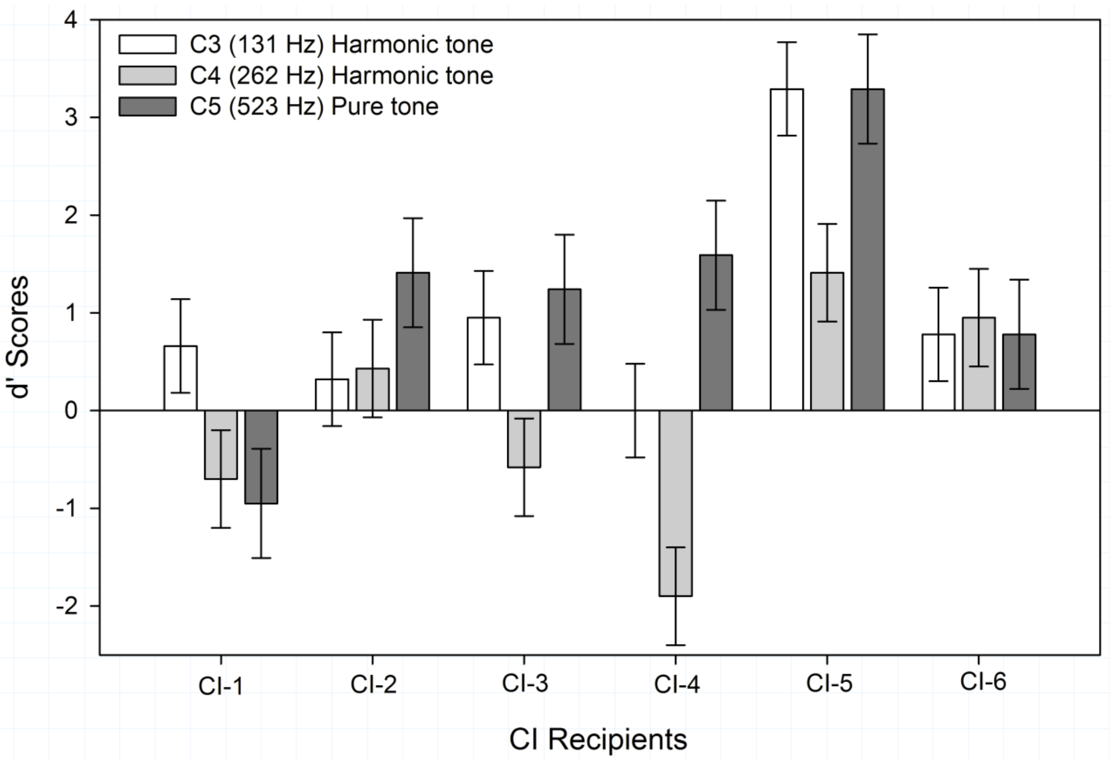


Figure 6.9: Individual d' scores for the modified melodies warp modification (warp factors: 0.10 & 10.00) across the three base frequencies. The error bars represent the standard errors.

An attempt was made to investigate the performance of CI recipients in Modified Melodies test using different warp factors. In warp modification, the interval size of the melody is altered, but the melodic contour remains intact (4.2.2.3.2). The warp factors were tested in pairs (0.10 & 10.00, 0.25 & 4, 0.50 & 2, and 0.75 & 1.33). As the warp factor decreases, the test becomes harders and it becomes difficult to detect the mistuned melody.

In this study, warp factor 0.10 & 10.00 were summed and compared with other experimental procedures (Figure 6.5), because all the subjects were able to perform only on these two warp factors (0.10 & 10.00). It should be noted that all the CI users did not perform on all the warp factors because of the increasing level of difficulty. The individual scores for all the warp factors are illustrated in Figure 6.10. When a subject scored at chance for a particular warp factor, the test was terminated and harder warp factors were not tested.

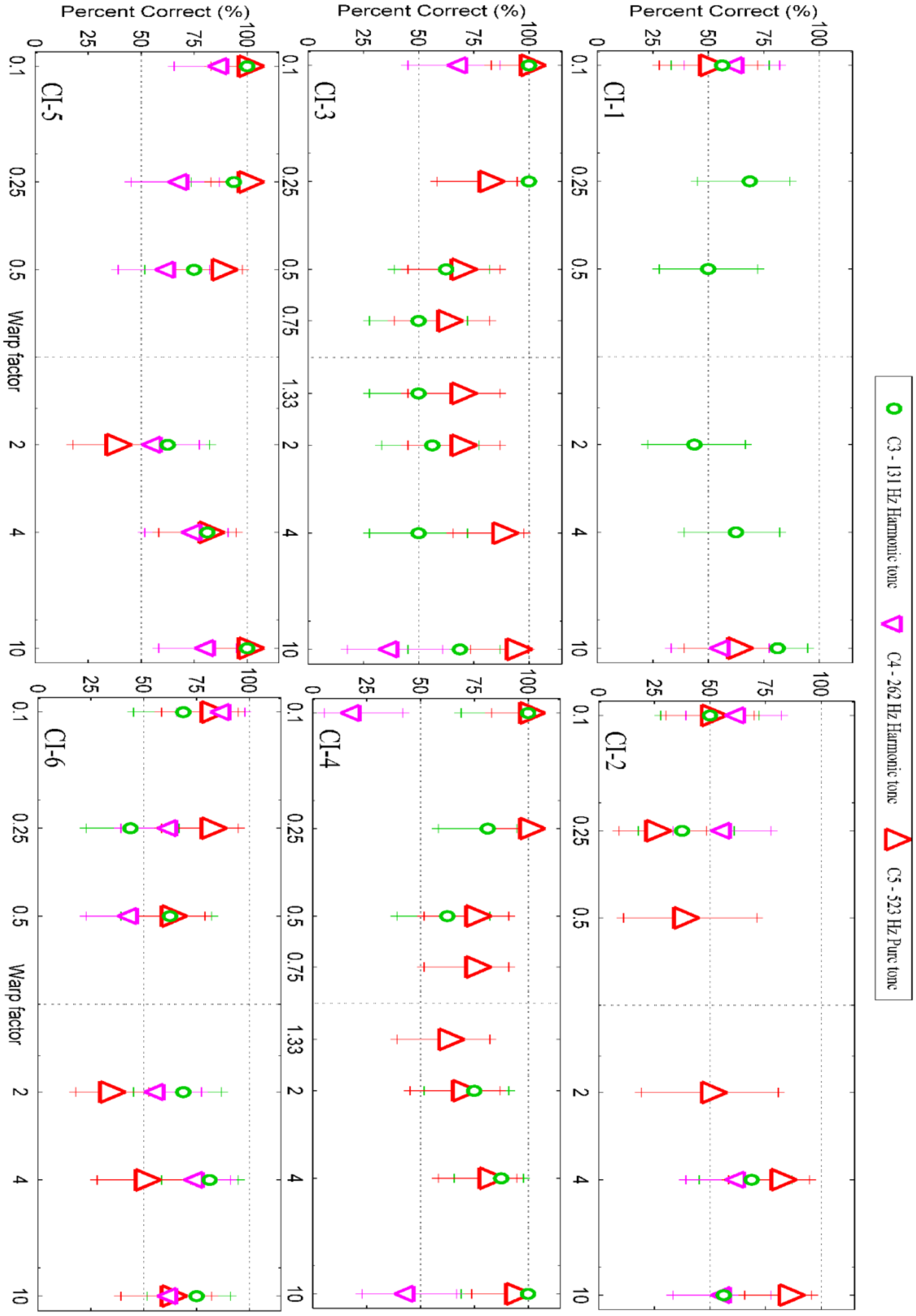
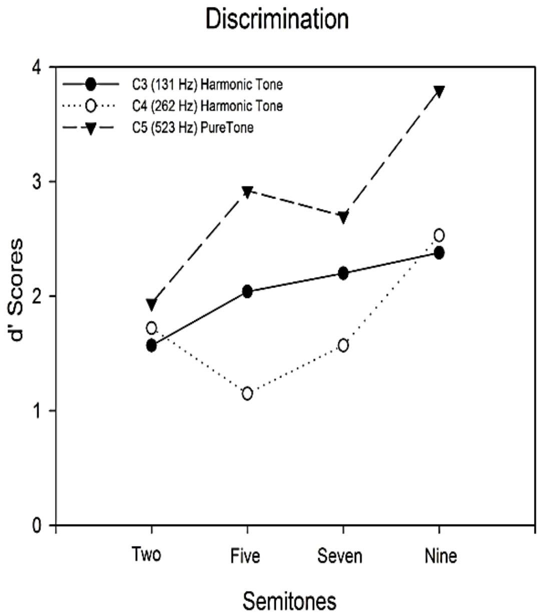


Figure 6.10: The individual modified melodies warp scores for different stimulus types. The 90% binomial confidence intervals are marked on the error bars, so that if the lower bound is above 50% it means that the score was significantly above chance according to a one-sided binomial test ($p < 0.05$).



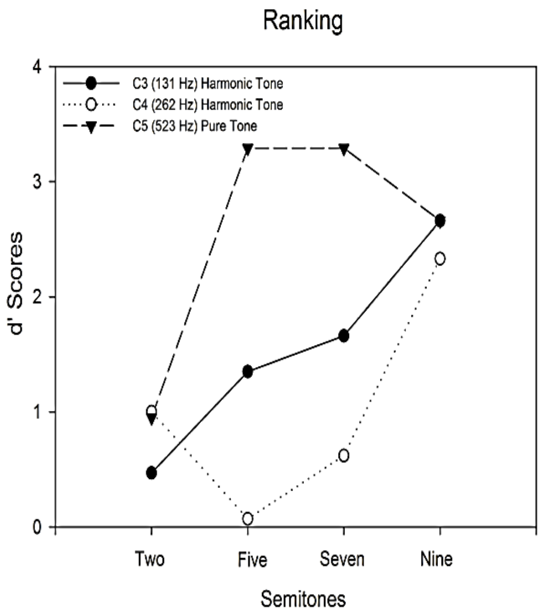


Figure 6.11: CI Discrimination scores as a function of semitone difference (Left). CI Ranking scores as a function of semitone difference (Right).

The discrimination and ranking scores in Figure 6.5 represents the cumulative score across all note pairs. In both these procedures, six note pairs were used and plotted as a function of semitone difference. These notes were 2 semitones apart (C-D; G-A), 5 semitones apart (D-G), 7 semitones apart (C-G; D-A), and 9 semitones apart (C-A).

An attempt was made to investigate whether the performance increased with respect to semitone differences. A repeated measures of ANOVA with the factors, stimuli (C3 Harmonic Tones, C4 Harmonic Tones, and C5 Pure tones) and semitones (Two, Five, Seven, and Nine) were computed. In discrimination, ANOVA revealed a main effect of stimuli [$F(2, 10) = 6.27$, $p = 0.01$]. The Bonferroni corrected pairwise comparisons showed that the C5 pure tones ($M = 3.06$, $SE = 0.16$) were significantly higher than C4 harmonic tones ($M = 2.03$, $SE = 0.30$, $p < 0.05$). The scores for C3 harmonic tones ($M = 2.50$, $SE = 0.37$) were not significant. Additionally, there was a main effect of semitones

[F (3, 15) = 7.92, p = 0.002]. The Bonferroni corrected pairwise comparison also revealed that the nine semitone scores were significantly higher than the two semitone scores ($p < 0.01$). Overall, there was no interaction between stimuli and semitones [F (6, 30) = 1.27, p = 0.29].

In Ranking, ANOVA revealed a main effect of stimuli [F (2, 10) = 7.58, p = 0.01]. The Bonferroni corrected pairwise comparison showed that C5 pure tones (M = 2.69, SD = 0.14) were significantly higher than C4 harmonic tones (M = 1.34, SE = 0.33). The C3 harmonic tones (M = 1.99, SE = 0.32) were not significant. There was a main effect of semitones [F (3, 15) = 7.28, p = 0.003]. The Bonferroni corrected pairwise comparison revealed that the nine semitone scores were significantly higher than the two semitone scores ($p < 0.05$). Overall, there was a significant interaction between stimuli and semitones [F (6, 30) = 2.75, p = 0.03]. In order to understand the interaction effect, a series of one-way ANOVA were conducted for stimuli. The analysis revealed a significant main effect of semitone for C3 harmonic tones [F (3, 3) = 18.84, p = 0.019] and C5 pure tones [F (2, 4) = 8.57, p = 0.03]. Post Hoc comparisons for C3 harmonic tones showed that the two semitone scores were significantly lower than the nine semitone scores ($p < 0.01$) and for C5 pure tones, the two semitone scores were significantly lower than the five and the seven semitone scores ($p < 0.05$). None of the other comparisons were significant. The analysis revealed that the performance improved as a function of semitone difference in discrimination and ranking tasks. The highest performance was seen for the C5 pure tones and lowest performance was obtained for the C4 harmonic tones.

An attempt was made to plot the individual CI performance as a function of note pairs for C4 harmonic tones (Figure 6.12). The results show a large intersubject variability among CI subjects. Three CI users found it extremely hard to rank the C4 harmonic tones and exhibited pitch reversals (CI 2, CI3, and CI4). This is a phenomenon where the CI users consistently rank the pitch in the opposite order. The results show that the mean scores between CD and GA (2 semitones) and also between CG and DA (7 semitones) were not similar. For E.g., the musical notes CG and DA were seven semitones apart, but the CI performance varied drastically. The CI subjects probably obtained temporal cues for ranking C4 (262 Hz) – G4 (392 Hz) and this cue could have probably faded for ranking D4 (294 Hz) – A4 (440 Hz). The place cues (centre of gravity) seem to be ambiguous (See C4 Harmonic tones Electrodogram- Figure 6.4) and this could probably explain the obtained results.

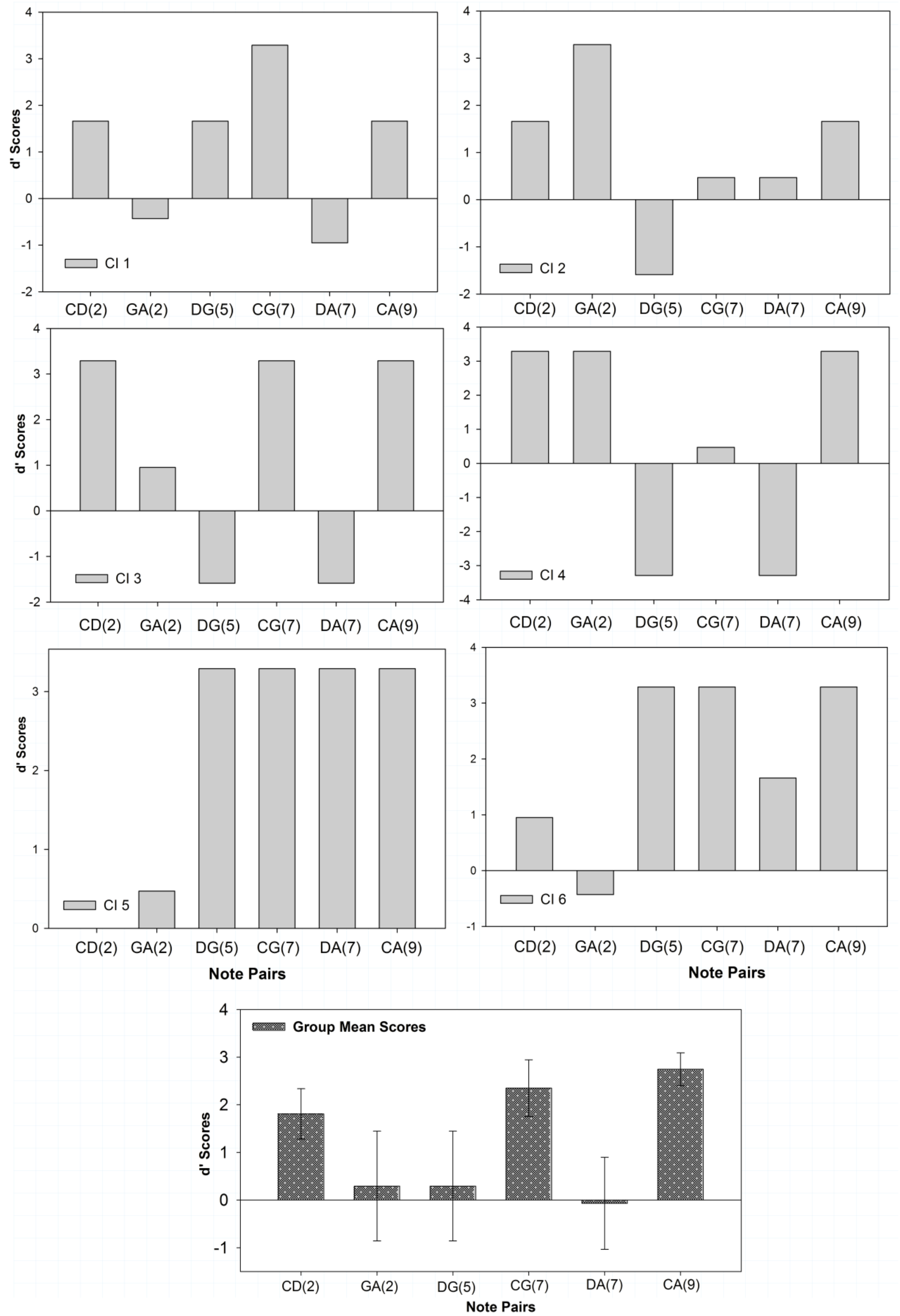


Figure 6.12: CI Ranking scores as a function of note pairs for C4 (262 Hz) harmonic tones. The note pairs are two semitones apart (CD and GA), five semitones apart (DG), seven semitones apart (CG and DA), and nine semitones apart (CA). The subjects CI 2, CI 3, and CI 4 exhibited pitch reversals. The error bars represent the standard errors.

Comparing CI subjects Vs NH Subjects

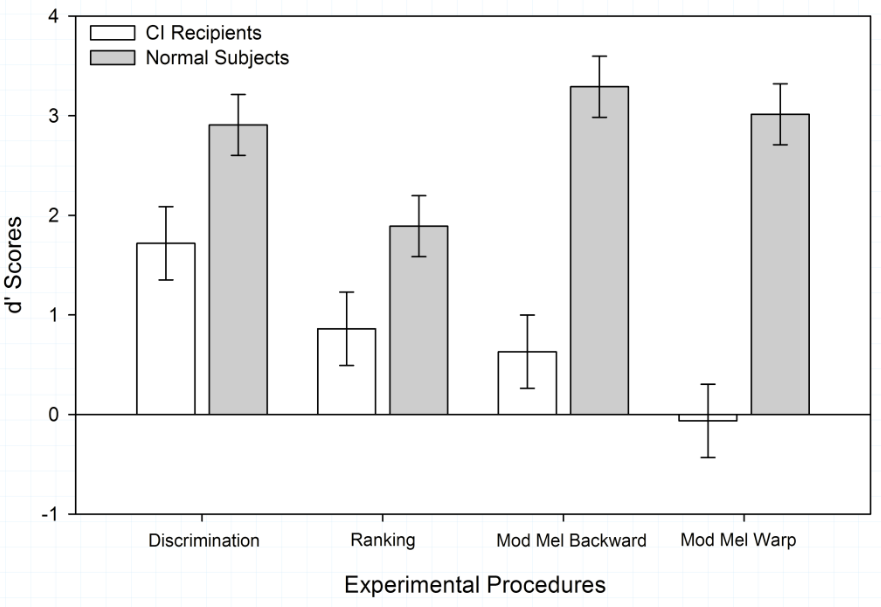


Figure 6.13: The performance of CI recipients on C4 (262 Hz) harmonic tones was compared with the performance of normal hearing individuals across the four experimental procedures. The error bars represent the standard errors.

Figure 6.13 schematises the performance of CI and NH subjects for the four experimental procedures using the baseline stimulus, i.e., C4 (262 Hz) harmonic tone. The normal subjects performed better than the CI counterparts in all the procedures. ANOVA revealed that the performance of NH subjects were significantly ($p < 0.01$) better than the CI subjects. The CI subjects found the Modified Melodies test extremely arduous. In warp modification, the subjects were required to judge the interval size of the melody and CI performance dropped to chance level for this condition. There was a large intersubject variability among CI recipients in these pitch perception tasks.

Comparing CI temporal pitch processed by ACE strategy Vs Direct stimulation

The C3 (131 Hz) harmonic tone which conveyed exclusive temporal pitch information was compared with the other ways (Direction Stimulation – Chapter 5) of conveying temporal pitch information in the six CI recipients (Figure 6.14).

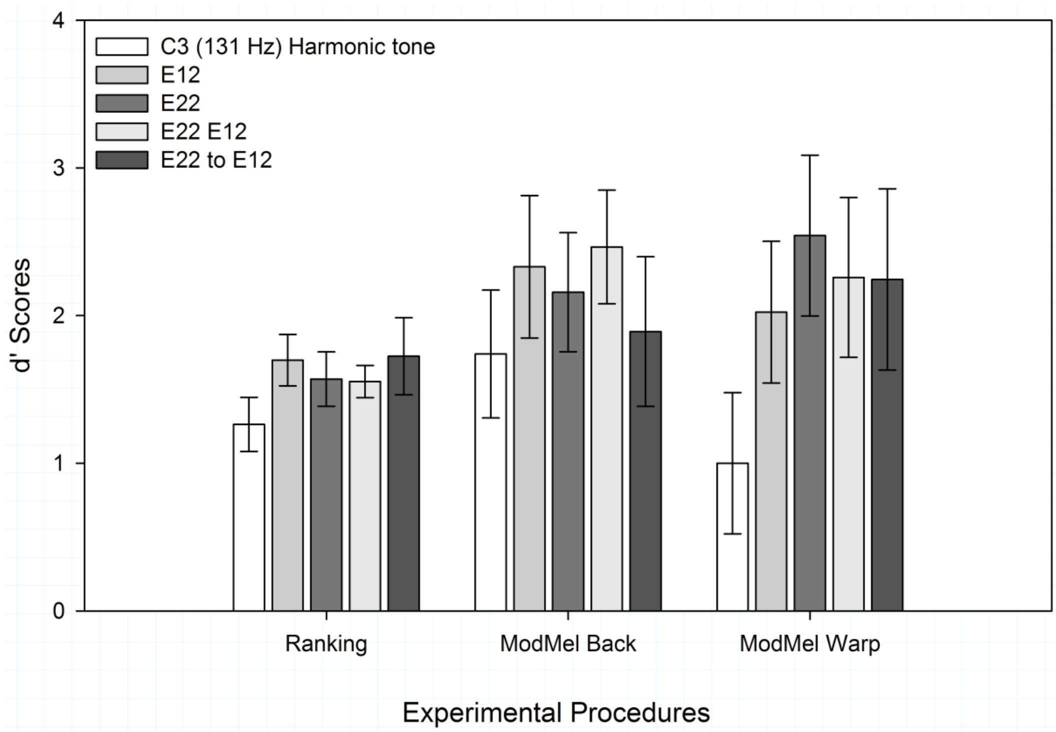


Figure 6.14: Comparing the performance of CI recipients for C3 (131 Hz) harmonic tone with the performance of same CI recipients at C3 (131 pps) base pulse rate. The error bars represent the standard errors.

The C3 (131 Hz) harmonic tones processed by ACE strategy (amplitude modulation cues) and delivered via loudspeaker was compared with temporal cues conveyed via direct stimulation of electrode. The three electrode stimulation were single (apical – E22 and middle – E12), dual (E22 & E12), and multiple (E22 to E12). The details of these stimulations are explained in chapter 5 (5.2.3). ANOVA revealed no significant difference between the temporal pitch percepts mediated by amplitude modulation (C3 (131 Hz) harmonic tone), and pulse rate variation on single or dual or multiple electrodes for the three procedures. Even in modified melodies warp modification, the C3 (131 Hz) harmonic tone failed to reach a significant level ($p > 0.05$).

In summary, the scores obtained from the modified melodies backward suggest that the participants were able to detect the contour patterns using C5 pure tones and C3 harmonic tones. Conversely, the C4 harmonic tone scores in the modified melodies warp was at chance level. The scores reveal that the C4 harmonic tones were incapable of providing adequate interval size information and only C3 and C5 tones provided this quite adequately. There was a large intersubject variability seen across CI recipients and in some instances pitch reversals were seen in the ranking procedure for C4 harmonic tones. The overall performance reveals that the CI users exhibited difficulty in discrimination, ranking, and identifying melodies using the C4 harmonic tones. Good performance was seen only for C5 pure tones and C3 harmonic tones in all the procedures.

6.4. Discussion

An attempt was made to investigate the individual contribution of place pitch and temporal pitch in various pitch perception tasks. The three stimulus types used in this study elicited pitch sensation using different mechanisms (Figure 6.1, 6.2, 6.3 and 6.4). The quadrature envelope detection for pure tones produce very little ripples, thereby minimizing the amplitude modulation and therefore, the processed signal consists of place pitch information only. Similarly, the C3 harmonic tones modulate the carrier pulse train and these amplitude modulations correspond to the fundamental frequency of the incoming musical notes, thereby providing temporal pitch information. Finally, the C4 harmonic tones stimulate multiple electrodes with varying modulation rates thereby providing both temporal and place cue information. The C4 harmonic tones are the most realistic tones present in the real world environment. The significance of choosing C3 harmonic tones is because this range of 131 Hz to 220 Hz is particularly sensitive for CI recipients to retrieve the temporal pitch cues. Previous studies have documented that up to 300 Hz the temporal pitch is perceivable by CI recipients (Zeng, 2002; Kong et al., 2009). In order to ensure only place pitch was conveyed, pure tones of C5 range (523 Hz to 880 Hz) were selected. Firstly, the C5 range of 523 Hz to 880 Hz is well above the cut-off limit of temporal pitch.

Secondly, the quadrature envelope detection in the filter banks of a sound processor detects very little ripples for pure tones suggesting no amplitude modulation and implies that the pure tones stimulate different intra cochlear electrode providing only place cue information (Swanson et al., 2007). Thirdly, the broad analysis filters in the sound processor activates multiple electrodes for pure tones. This is evident in the electrodogram (Figure 6.2) which indicates that different musical notes stimulate different intracochlear sites.

In discrimination, CI users obtained relatively good scores across the three base frequencies. There was a significant difference only between C4 harmonic tones and C5 pure tones. Subjects obtained high score using pure tones as compared to harmonic tones. A similar finding was observed by Gfeller et al. (2002). Their study investigated the factors which influence the melody recognition in CI recipients. The study consisted of three experiments, familiar melody recognition, complex harmonic tone discrimination, and pure tone discrimination. All the test stimuli were presented via loudspeakers. The complex harmonic tone discrimination was a two alternative forced-choice (2 AFC) adaptive procedure. The mean scores for NH adults (1.13 semitones; range 1-2 semitone) were significantly better than CI recipients (7.56 Semitones; range 1-12 semitones, S.D = 5.18). The pure tone discrimination was a four alternative forced-choice adaptive procedure. The NH adults displayed frequency difference limens of less than 0.01 for this test. Conversely, CI recipients showed considerable variability in this task. However, fewer (approximately 6 %) subjects were able to discriminate frequency differences smaller than 0.02 – 0.03 at both high and low levels. However, the CI recipients' mean scores were significantly poorer than the NH adults' scores and coincide with the results obtained from the present study (Figure 6.13). Cross comparison between pure tone and complex tone reveal that CI users obtained high scores using pure tones compared to complex tones.

For ranking, relatively good performance was seen across all the three base frequencies. Similar to discrimination, high scores were obtained for C5 pure tones and C3 harmonic tones and poor scores were observed for C4 harmonic tones. CI subjects exhibited a problem in ranking two semitones compared to seven and nine semitones. A few CI users exhibited pitch reversals using C4 harmonic tones. The ranking performance improved with respect to semitone difference and a similar finding was observed in Laneau & Wouters (2006) study. The study compared the three speech processing schemes (F0mod, ACE, and ACE512) using different music perception experiments. In one of their experiments (F0 Discrimination for Musical Tones), ranking ability was measured for notes of five different instruments (grand piano, clarinet, guitar, and synthetic voice). The three reference F0 values were C3 (130.8), F3# (185.8 Hz), and F4# (370.0 Hz). The C3 (130.8 Hz) base frequency produced scores which improved from 1 semitone to 4 semitones for the ACE processing scheme. Gfeller et al. (2007) studied the accuracy of direction of pitch change as a function of base frequency (131, 164, 208, 262, 330, 417, 524, 663, and 831 Hz) and interval size (1, 2, 3, and 4 Semitones). The stimuli used in their task were pure tones and the pitch ranking procedure was similar to the present study. The GLMM (generalized linear mixed model) analysis revealed that the C5 (523 Hz) base frequency performance was similar to the present study where the performance improved as a function of semitone difference. As an alternative to traditional studies, Sucher & McDermott (2007) investigated the pitch ranking abilities of CI recipients on 'real-world' stimuli of sung vowels (1 semitone and 6 semitone). CI subjects obtained scores of 60.2 % and 40.0 % for 6 and 1 semitone respectively. The results suggested that CI subjects frequently confused the direction of pitch change. It is important to note that the ranking procedure in the current study is more sensitive than the adaptive procedures used in the previous studies (Gfeller et al., 2002; Nimmons et al., 2008; Pretorius & Hanekom, 2008), because it is sensitive in finding the pitch reversals seen in CI recipients. Pitch reversals exhibited by CI recipients can shed some light on better understanding the pitch (tonotopic) organisation in these recipients.

It has been argued that both contour and interval size can contribute to melody recognition (Dowling & Fujitani, 1970). The recognition of melody is affected when either of these components is altered. The Modified Melodies test alters the melody in two different ways. In backward modification the contour of the melody is altered, whereas in warp modification the interval size of the melody is altered, but keeping the melodic contour unaltered. The scores obtained for backward modification reveal that the CI users were able to recognise the contour of the melody when the melody contained either place cue only (C5 pure tones) or temporal cue only (C3 harmonic tones). Performance reduced when the melody consisted of both place and temporal cue (C4 harmonic tones). Similarly, in warp modification the CI users were able to judge the musical interval size when the melodies contained place cues only (C5 pure tones) or temporal cues only (C3 harmonic tones). CI users obtained scores below chance using C4 harmonic tones. Anecdotal reports indicated that most CI users found it extremely hard to judge interval size using C4 harmonic tone (both place and temporal cues).

There was a large intersubject variability seen in ranking, Modified Melodies backward, and Modified Melodies warp. In some instances, scores were below chance in these tasks.The ranking task was sensitive enough to detect pitch reversals but not plausible in the discrimination task. In the discrimination task, subjects can rely on cues other than pitch (maybe the brightness cue, addressed in detail in Chapter 7) to perform this task, so discrimination may not be as sensitive as ranking. There was a high correlation between Modified Melodies backward and other procedures which suggests that Modified Melodies backward is a sensitive test investigating melodic pitch in CI recipients.

Good performance using pure tones was seen in Singh et al. (2009) study. The CI users performed well using the high base frequency (414 – 1046 Hz) compared to low (104 -262 Hz) and middle (207 – 523 Hz) frequency ranges in a familiar melody recognition task. Results showed that the melody scores for pure tones were significantly better than the complex harmonic tones. Cooper et al.(2008) used the Montreal Battery for Evaluation of Amusia (MBEA) for assessing the music perception abilities of CI users. The MBEA comprised of six tests (Scales, Contour, Interval, Rhythm, Meter and Melody Memory). The musical stimuli were presented in a free field condition (stimuli presented via loudspeakers) and the frequency ranged from B3 (247 Hz) to B5 (988 Hz). CI users were on par with their NH counterparts in rhythm and meter (temporal-based task). However, the scores for scales, contour, and interval (pitch-based tasks) were near chance.

Donnelly et al. (2009) used a novel pitch separation task to investigate the perception of polyphony (or harmony) in CI recipients. All the stimuli had F0 ranging from C4 (262 Hz) to C5 (523 Hz) and the mode of presentation was using loudspeaker. The experiment consisted of three kinds of stimuli: Single-pitch stimuli consisted of either pure tones or piano tones from C4 – B4 (12 unique pitches, 24 total stimuli), Two-pitch stimuli consisted of either pure tone from C4 – C5 (1-12 semitones interval distance, 24 total stimuli), and Three-pitch stimuli consisted of either pure tones or piano tone representation of 6 unique symmetric chords (equal interval spacing between lower/middle and middle/higher pitches) within the range of C4 – C5. The stimuli were presented randomly using 3-alternative, single interval, forced choice procedure. The subjects' task was to choose whether the presented stimuli consisted of one, two or three pitches. Twelve CI subjects obtained significantly poor scores when asked to distinguish single and multiple acoustic stimuli (Two-pitch and Three-pitch stimuli). The CI performance was near chance for two-pitch and three-pitch stimuli. They concluded that the CI recipients demonstrated frequency perceptual fusion of multiple-pitch stimuli as single-pitch units. Aforementioned studies used a wide variety of approaches to study pitch and music perception in CI users and these studies concur that pitch is significantly affected in CI users and likely to impact on music perception. Overall, the plausible reasons for the intersubject variability across subjects may be attributed to poor neural survival rate inside the cochlea, electrode insertion depth, and auditory deprivation prior to implantation.

Implication for CI Place Pitch and CI Temporal Pitch

The results obtained from C3 harmonic tones suggest that the temporal pitch alone is sufficient for conveying adequate melody information and this was evident from the earlier study by Pijl et al.(1995b). In the present study, good scores were obtained for C5 pure tones and the overall performance of C5 pure tones supports the notion that CI place pitch alone can support musical pitch. Results showed that the CI users were able to discriminate the C4 harmonic tones quite effectively, but a few CI users were unable to order the musical notes in the correct order and exhibited pitch reversals. Ranking or Ordering the pitch in correct order is the fundamental ability in the perception of melodic contour (Gfeller et al., 2002). Results showed that there is indeed a high correlation between ranking and modified melodies backward procedures.

CI recipients were able to recognise the correct melodic contour in backward modification and scores were well above chance. However, when required to judge the musical interval size using C4 harmonic tones, recipients' performance was below chance. The C4 fundamental frequency range of 262 to 523 Hz was probably above the upper frequency limit of temporal pitch for most subjects and additionally the place pitch cues may have been ambiguous. This may reflect the limitations of place-pitch discrimination. Additionally, the frequency allocated by the speech processor to an electrode does not match the characteristic frequency corresponding to that electrode position. This frequency-to-electrode mismatch is likely to distort the representation of musical intervals (Skinner et al., 2002; Boëx et al., 2006). Overall results show that place pitch may convey musical pitch information and this is an enigma not explained by current models of pitch processing (Moore & Carlyon, 2005). Interestingly, CI users find it hard to process pitch information when both place and temporal cues are available in a signal. One of the solutions may be to increase the number of channels at the fundamental frequency range (Laneau et al., 2004b) and also to improve the spatial resolution (van den Honert & Kelsall, 2007; Frijns et al., 2011).

6.5. Conclusion

In summary, this study investigated the contribution of place cues and temporal cues using different pitch perception tasks. The results indicated that ranking is a better procedure than discrimination. High correlation was observed between modified melodies backward and the other procedures and indicated that modified melodies backward is a sensitive test compared to modified melodies warp. The CI recipients obtained good scores when they relied on a single cue, either a place (C5 Pure tones) or a temporal cue (C3 harmonic tones). However, the CI performance decreased or dropped to chance level when the stimulus consists of both place and temporal cues (C4 harmonic tones). Finally, the results suggest that CI place pitch may convey melodic pitch information, but the role of brightness cannot be completed ruled out and this issue is addressed in the following chapter.

Chapter 7: Role of Brightness in Pitch Perception Tasks - Implication for Cochlear Implant Place Pitch

Abstract

Researchers have speculated that cochlear implant (CI) place pitch is more closely related to brightness aspect of timbre than to pitch. As brightness can be ordered on a low-to-high scale, it can provide high scores on ranking and discrimination tests. Therefore, the aim of the present study was to investigate the role of brightness in different pitch perception tasks in normal hearing individuals. Eighteen normal-hearing adults participated in four experimental procedures: 4 AFC Discrimination, 2 AFC Ranking, and 2 AFC Modified Melodies test (Backwards and Warp modifications), using three stimulus conditions: (i) Pitch sequences: harmonic tones varying in pitch, with constant brightness; (ii) Brightness sequences: harmonic tones varying in brightness, with constant pitch; (iii) Noise sequences: Low-pass noise bands, varying in cut-off frequency. The scores for discrimination and ranking were high for all three stimulus types, and d' analysis revealed that the subjects' performance was better in discrimination than ranking. In the Modified Melodies test, when subjects were required to detect a melody contour change, scores were high for all three stimulus conditions. Conversely, when the melody contour was preserved but the musical intervals were changed, scores using pitch sequences were high, while scores using brightness or noise sequences were at chance-level. Thus subjects were able to discriminate and rank brightness, and were able to detect brightness contour changes, but were unable to make judgements of musical intervals for brightness. These results suggest that the cochlear implant recipients in the previous chapter may have perceived place cues as brightness rather than pitch sequences.

7.1. Introduction

The human auditory system is efficient in processing the three perceptual dimensions of pitch, loudness and timbre simultaneously and coherently (Pitt, 1994). Timbre is a multidimensional perceptual quality that distinguishes between two tones that have the same pitch, loudness, and duration (Plomp, 1976; Moore, 2003). It allows the musical instrument that played a particular note to be identified. Timbre depends on many physical properties of a tone, but the most prominent is the spectral profile, i.e., the amplitudes of the harmonics. The brightness of a harmonic tone depends on the centroid of the spectral profile (Plomp, 1976; Schubert & Wolfe, 2006), i.e., tones having strong high harmonics sound brighter. Like pitch, brightness can also be ordered on a low-to-high scale. The background information regarding timbre perception in normal hearing subjects was addressed in chapter 3 (3.6).

The place pitch corresponds to the pitch percept which gradually rises as the place of stimulation moves from apex to base of the cochlea, thus mimicking the natural tonotopic organisation (3.4.2). Researchers have speculated that cochlear implant place-pitch (place-of-excitation) is in closer proximity to the brightness aspect of timbre than to pitch (McDermott, 2004; Moore & Carlyon, 2005; Swanson, 2009). The ability to rank stimuli on a low-to-high scale does not differentiate between brightness and pitch. An operational definition of pitch is that variations in pitch can convey a melody; thus a test involving melody perception was preferred.

In a study of place-pitch perception cues in cochlear implants, Swanson et al. (2008) presented CI recipients with target and comparison melodies and asked them to identify the correct version of the melody (Modified Melodies test). Using place-pitch cues alone, all seven subjects were able to detect changes to the melodic contour of target melodies. Four out of seven subjects were able to detect a large (five-semitone) error in the size of a musical interval, and one CI recipient was able to detect a smaller (two-semitone) error. The results supported the hypothesis that place-pitch cues can provide melodic pitch. Although anecdotal reports from subjects suggested that they were indeed hearing a melody based on pitch changes, the possibility that subjects were recognising patterns of brightness changes could not be completely ruled out.

The purpose of the present study was to determine whether good scores can be obtained in the absence of a true perception of pitch. Therefore, the present study measured the performance of normal hearing subjects on stimuli containing pattern of brightness changes compared to pattern of pitch changes. The study utilised four procedures: discrimination, ranking, the Modified Melodies test using backward modification, and the Modified Melodies test using warp modification. The three stimuli used were (i) *Pitch sequences:* harmonic tones varying in pitch, with constant brightness; (ii) *Brightness sequences:* harmonic tones varying in brightness, with constant pitch; (iii) *Noise sequences:* Low-pass noise bands, varying in cut-off frequency. The result obtained from NH subjects will be compared with CI subjects to determine whether CI subjects in previous chapter perceived place pitch as pitch or as a pattern of brightness changes.

7.2. Method

7.2.1. Subjects

Eighteen normal hearing adults who participated in the previous experiment (Chapter 4) using pitch sequences participated in this study as well.

7.2.2. Experimental Procedures

The four experimental procedures are described in terms of the baseline condition, where each note was a harmonic tone, and pitch (fundamental frequency) was varied. The details of the experimental procedure are described in chapter 4 (4.2.2). The four procedures used in this study were note discrimination (4 AFC), note ranking (2 AFC), modified melodies backward modification (2 AFC), and modified melodies warp modification (2 AFC). It should be noted that for NH subjects, warp factors of 0.75 and 1.33 were used, whereas for CI recipients warp factors of 0.10 and 10.00 were used. As the warp factors decreased, the Modified Melodies test became difficult and it becomes harder to detect the original version of the melody. The warp factors 0.10 and 10.00 were extremely easy for NH subjects, but that was not the case with CI recipients. Therefore, NH subjects were able to perform at much more difficult warp factors (0.75 and 1.33), unlike CI recipients who reach below chance scores for much higher warp factors (10.00 and 0.10).

7.2.3. Stimulus Description

All the stimuli were synthesised on a PC at a sampling rate of 16 kHz, and presented via a PC sound card and headphones (Sennheiser HD280 Pro) in a sound-treated room. Each note was 300 ms in duration, with a smooth (sinusoidal-shaped) rise and fall time of 50 ms. The stimuli were presented at the comfortable loudness level for all the subjects. The details of the three stimulus types are described below.

7.2.3.1. Pitch sequences: Harmonic tones varying in pitch

This is a baseline condition for the four experimental procedures and described in chapter 4 (4.2.3.1). The NH performance using the baseline pitch sequences were compared with brightness and noise sequences. The NH scores using pitch sequences were obtained from Chapter 4 and compared with the remaining sequences.

7.2.3.2. Brightness sequences: Harmonic tones varying in brightness

In this condition, a digital pulse train with fundamental frequency of C4 (262 Hz) was generated. This created a series of harmonics with equal amplitude. Each time the experimental procedures required a note at a specified frequency, the note was synthesised by passing the C4 pulse train through a low-pass filter that had a cut-off frequency equal to the specified frequency. MATLAB firrcos function was used to design each filter which had a 64-tap FIR filter with a 1000 Hz wide raised cosine transition band. The nominal note frequencies (i.e., the filter cut-off frequencies) were in the range C6 – A6 (1047 – 1760 Hz). The spectral profile of the four “notes” C, D, G, and A are shown in Figure 7.1. The figure shows that the spectral profile varied as a function of each note. For example, a larger spectral profile was seen for ‘A Note’ compared to ‘C Note’. Thus, in this condition, the spectral profiles (and hence the brightness) varied in a pattern that depended on the notes, although the pitch was constant. In order to maintain a constant loudness, the amplitude of all the stimuli were normalised from -1 to +1. Therefore all the stimuli had equal amplitude with constant loudness.

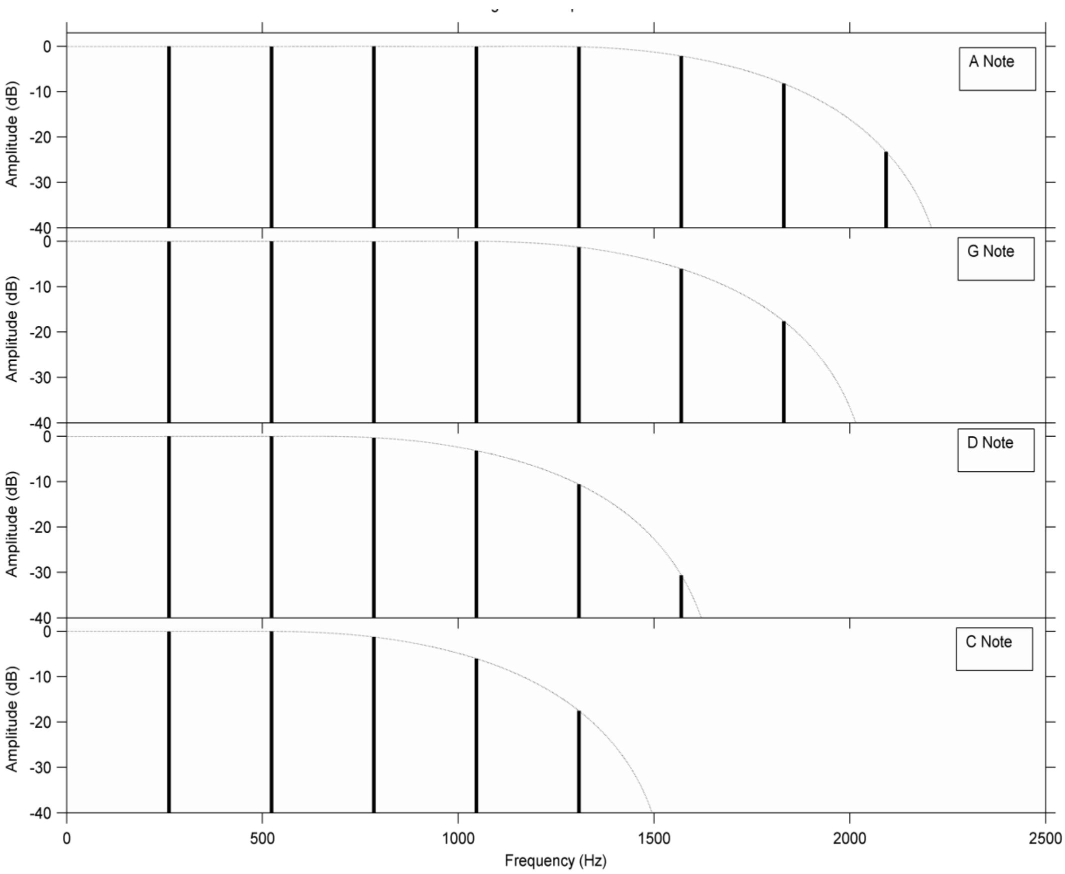


Figure 7.1: Spectral profile of brightness sequences as a function of four notes.

7.2.3.3. Noise sequences: Noise bands varying in cut-off frequency

This condition was similar to the Brightness condition, except that the input to the low-pass filter was a burst of white noise. The same set of low-pass filters were used, i.e., the filter cut-off frequencies were in the range C6 – A6 (1047 – 1760 Hz). The spectral profile of the four "notes" C, D, G, and A are shown in Figure 7.2. Thus, in this condition, the spectral profiles varied in a pattern that depended on the notes. In order to maintain a constant loudness, all amplitude of all the stimuli were normalised to be in the range of -1 to +1. Therefore all the stimuli had equal amplitude with constant loudness.

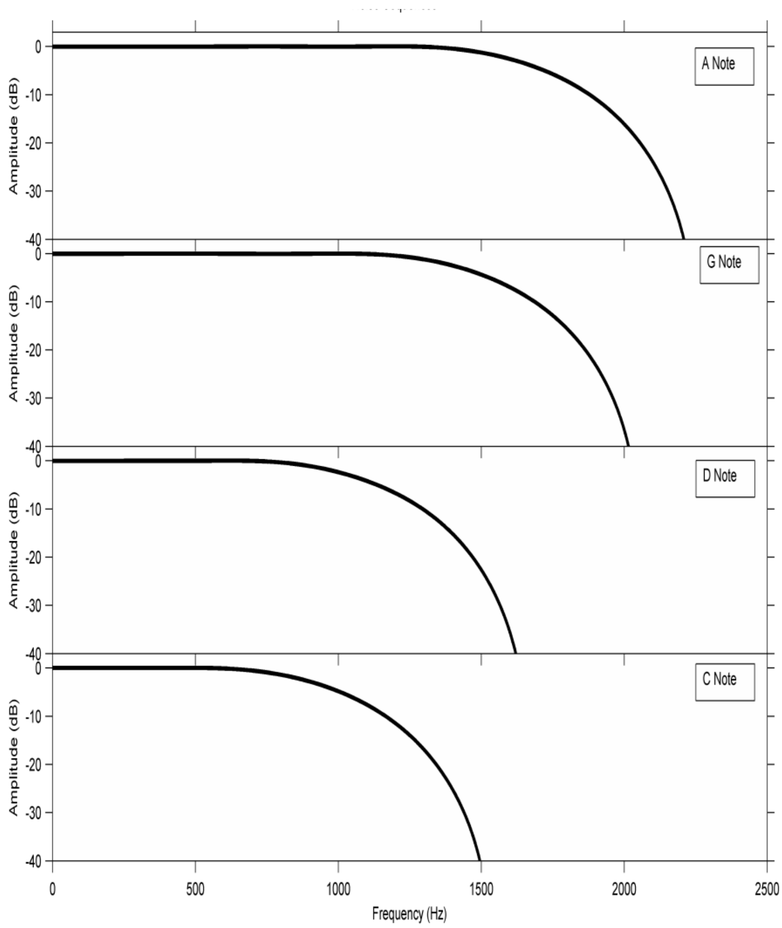


Figure 7.2: Spectral profile of noise sequences as a function of four notes

.

7.3. Results

The present study investigated the role of brightness in different pitch perception tasks and also to answer the question of whether the CI recipients in previous experiments perceived place pitch as pitch and not as a pattern of brightness changes. A single score was obtained for each procedure and stimulus type by summing across note pairs in discrimination and ranking, and summing across warp factors 0.75 and 1.33 in modified melodies warp. It should be noted that for NH subjects warp factor of 0.75 and 1.33 were used, whereas for CI recipients warp factors of 0.10 and 10.00 were used. As the warp factors decreased, the Modified Melodies test became difficult and it becomes harder to detect the original version of the melody. The warp factors 0.10 and 10.00 were extremely easy for NH subjects, but that was not the case with CI recipients. The percentage correct scores for the four experimental procedures are tabulated in Table6. The pitch sequences scores were used as a baseline scores in this study. The percentage correct scores across the four procedures cannot be compared directly, because the discrimination procedure was a 4-AFC method and the other procedures were 2-AFC methods. In order to compare these procedures, the scores were converted into d' scores based on table A5.7 of Macmillan & Creelman (2005).

Table 6: Percentage correct scores for the four experimental procedures.

Stimulus Types	Discrimination	Ranking	Modified Melody Backward	Modified Melody Warp
Pitch sequences	92	89	100	98
Noise sequences	90	88	97	56
Brightness Sequences	86	88	92	54

The group mean d' scores for the four procedures are illustrated in Figure 7.3 and the statistical analysis were administered on these d' scores.

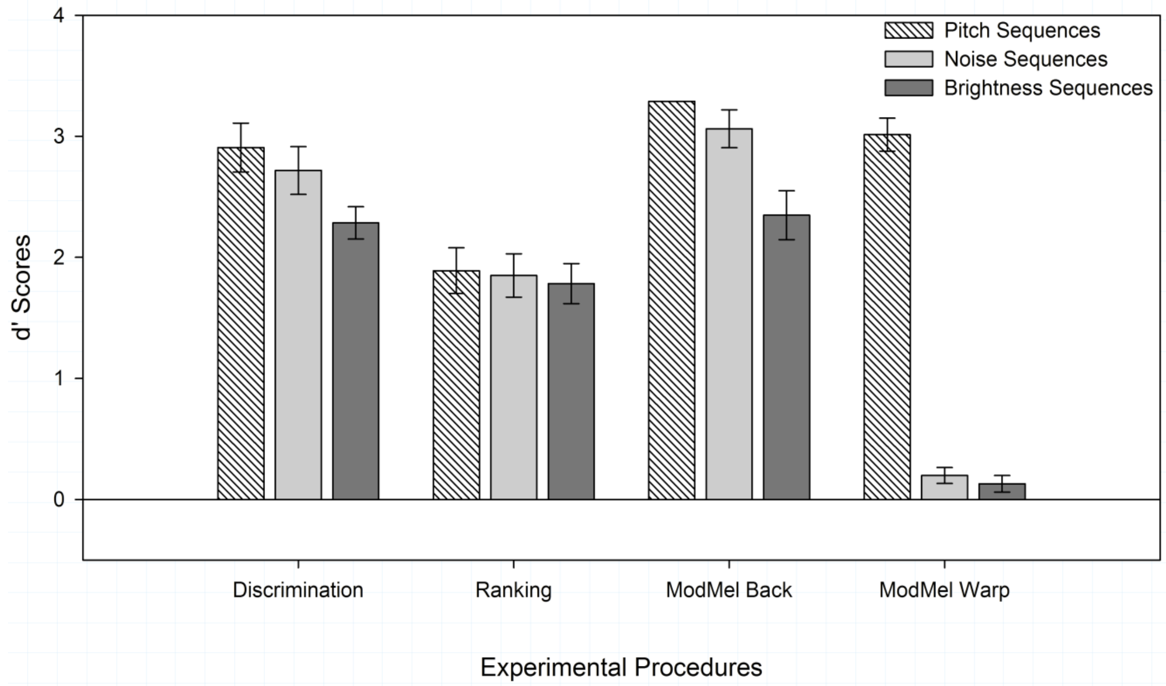


Figure 7.3: The group mean d' scores for the three stimulus types across four procedures. A d' score of zero indicates performance at chance. The error bars represent the standard errors.

A repeated measures analysis of variance (ANOVA) with the factors, experimental procedures (discrimination, ranking, modified melodies backward, and modified melodies warp) and stimuli (pitch, noise, and brightness sequences) were computed. The ANOVA revealed a main effect of both condition [$F(3, 51) = 73.28$, $p < 0.01$] and stimuli [$F(2, 34) = 148.81$, $p < 0.01$]. Pairwise comparisons revealed that the three stimuli were significantly different from each other ($p < 0.01$, Bonferroni corrected, pitch estimates M = 2.77, SE = 0.10; noise estimates M = 1.95, SE = 0.11; brightness estimates M = 1.63, SE = 0.09). Finally, there was a significant interaction observed between conditions and stimuli [$F(6,102) = 43.93$, $p < 0.01$].

In order to understand the interaction, a series of one-way ANOVA was computed. In discrimination task, pairwise comparison showed a significant difference between pitch and brightness ($p < 0.01$), and noise and brightness sequences ($p < 0.05$). There was no significant difference for the remaining pairs. In ranking task, there was no significant difference between the three stimulus types. In modified melodies backward, there was a significant difference between pitch and brightness ($p < 0.01$), and noise and brightness sequences ($p < 0.05$). In modified melodies warp, there was a significant difference between pitch and noise ($p < 0.0001$), and pitch and brightness sequences ($p < 0.0001$). There was no significant difference between the remaining pairs ($p > 0.05$).

The maximum scores were seen for the modified melodies backward (M = 2.90; SE = 0.09) and the least scores were obtained for the modified melodies warp (M = 1.11; SE = 0.05). Intermediate scores were obtained for discrimination (M = 2.63, SE = 0.15) and ranking (M = 1.84, SE = 0.17). The two musicians yielded slightly higher scores than the non-musicians, but more musician subjects were required to demonstrate this advantage. There was a high positive correlation (Pearson's correlation) seen between discrimination and ranking ($r = 0.62$, $p < 0.001$) and between discrimination and modified melodies warp ($r = 0.29$, $p < 0.05$). Unfortunately, there was no correlation among other pairs.

The scores obtained from the modified melodies backward modification suggests that the participants had no difficulty in identifying a contour change across the three stimulus types. On the contrary, the modified melodies warp scores were high only for pitch sequences and chance level scores for brightness and noise sequences. This suggests that the participants were able to accurately judge the size of musical intervals only for pitch sequences and other dimensions such as brightness and noise sequences failed to convey this information.

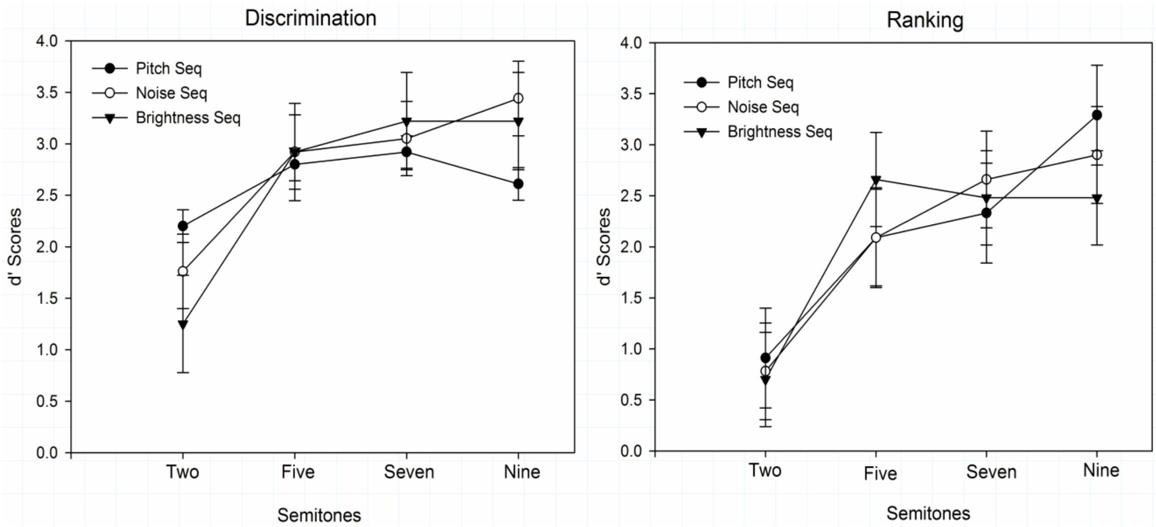


Figure 7.4: d′ scores as a function of semitone difference for (a) Discrimination and (b) Ranking.

The discrimination and ranking scores in Figure 7.3 were averaged across all note pairs. In both these procedures, six pairs of notes were used. These notes were 2 semitones apart (C-D; G-A), 5 semitones apart (D-G), 7 semitones apart (C-G; D-A), and 9 semitones apart (C-A). The scores are plotted as a function of semitone difference in Figure 7.4.

The discrimination results were analysed using ANOVA with factors stimulus types and semitones were computed. The results showed no main effect of stimulus types [$F (2, 34) = 0.73$, $p > 0.05$], but there was a main effect of semitones [$F (3, 51) = 31.86$, $p = 0.001$] and an interaction effect was seen between stimulus types and semitones [$F (6, 102) = 8.86$, $p = 0.001$]. The pitch performance was similar and reached ceiling for the four semitones and there was no significant difference across these semitones ($p > 0.05$). The noise scores at two semitones were significantly different from the seven and nine semitones scores ($p < 0.01$), but there was no significant difference between two and five semitones. Similarly, brightness performance at two semitones was significantly less than five, seven, and nine semitones ($p < 0.01$). At two semitones, the brightness scores were significantly lower than pitch and noise scores ($p < 0.05$) and there was no significant difference among other pairs. As expected, there was an increase in performance from two semitones to nine semitones especially for brightness and noise sequences.

In ranking, similar to discrimination, ANOVA showed no main effect of stimulus types [F (2, 34) = 0.12, p > 0.05], but there was a main effect of semitones [F (3, 51) = 46.70, p < 0.001] and an interaction between stimulus types and semitones were also seen [F (6, 102) = 2.20, p < 0.05]. Bonferroni pairwise comparisons reveal that the two semitones were significantly different from five, seven, and nine semitones and there was no significant difference among other pairs. The results showed a monotonic increase in performance from two semitones to nine semitones across the three stimulus types.

Comparison between CI Place pitch (C5 Pure tones) Vs NH Brightness sequences

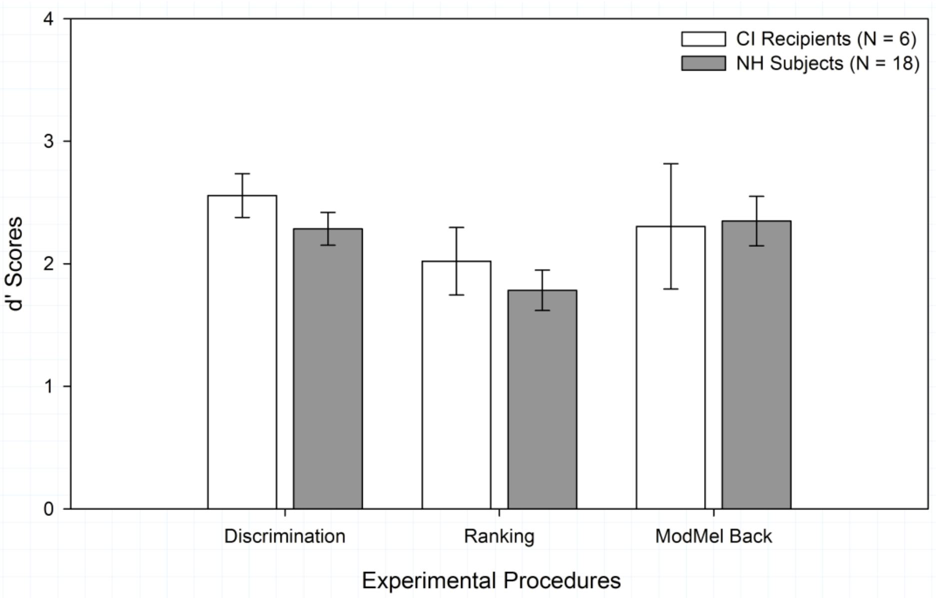


Figure 7.5: Comparison of performance between CI subjects using C5 (523 Hz) pure tone and NH subjects using brightness sequences. The error bars represents standard errors.

An attempt was made to compare the performance of NH subjects using brightness sequences with the performance of CI recipients using C5 (523 Hz) pure tone base frequency. ANOVA revealed no significant difference among the three procedures for CI subjects [$F(2, 15) = 0.58$, $p > 0.05$], but there was a significant difference for NH subjects [$F(2, 51) = 3.36$, $p = 0.04$]. Interestingly, Bonferroni post hoc comparison revealed no significant difference between the three procedures among NH and CI groups. Additionally, the comparison between NH and CI subjects among the three procedures showed no significant difference ($p > 0.05$). The modified melodies warp was not analysed because NH subjects used harder warps (warp factors - 0.75 and 1.33) compared to CI subjects (warp factors - 0.10 and 10.00). These results showed similar performance trend between CI subjects using C5 (523 Hz) pure tones and NH subjects using brightness sequences.

7.4. Discussion

This study investigated the performance of brightness and noise sequences on various pitch perception tasks in NH subjects. Overall, subjects obtained high scores for all three stimuli in discrimination and ranking tasks. Pitch sequences yielded significantly better scores than brightness sequences for the discrimination, modified melodies backwards, and modified melodies warp procedures. Ranking performance suggests that the subjects can rank these stimulus types with high level of confidence and as mentioned earlier, ability to rank the stimuli from low-to-high does not differentiate these stimulus types. Therefore, results showed no significant difference among stimulus types in the ranking task. Overall, the present study confirms better performance for pitch sequences as compared to brightness sequences.

Earlier findings of McDermott et al. (2010) may provide a basis for interpreting our results. Their stimuli consisted of intervals of pitch, brightness, and loudness. Pairs of these intervals were presented and participants indicated which of the two intervals was larger. Interval acuity was calculated for all three attributes and was no better for pitch than for timbre or loudness when expressed as the number of just-noticeable differences (JNDs). However, JNDs were considerably smaller for pitch than for brightness or loudness. These findings suggest that better performance for pitch sequences in the present study may have arisen because JNDs were smaller for pitch than for brightness or loudness.

The two important components which help in recognizing a melody are melodic contour and interval size (Dowling & Fujitani, 1970). The recognition of melody is affected when either of these components is altered. In the modified melodies backwards procedure, subjects were required to detect a change in the contour of the melody. The three percepts (pitch, brightness, and noise sequences) provided adequate information about the contour and thus good scores were evident in the modified melodies backwards procedure. These results are consistent with another study investigating perception of contours in pitch and brightness (McDermott et al., 2008). In Experiment 2, they found that the subjects were able to judge whether a brightness contour matched a pitch contour. In Experiment 4, they found that the subjects were able to identify a familiar melody when it was played as a pattern of brightness changes.

In the modified melodies warp procedure, subjects were required to detect a change in the interval sizes of the melody. High scores were obtained only for pitch sequences, with chance level scores for brightness and noise sequences. Thus, only pitch sequences conveyed sufficient interval size information; the brightness and noise sequences were incapable of providing this information. These results can be compared to those of experiment 4 in McDermott et al. (2008). In that study, the scores for recognising familiar melodies decreased when the pitch intervals were stretched and this was not observed in brightness interval stretching. This performance underscores the good sensitivity of pitch for interval size and also demonstrates the poor sensitivity of brightness for brightness interval size.

The noise sequences consisted of white noise filtered through low pass noise bands varying in the upper cut-off frequencies. On perceptual evaluation, one can faintly recognise that the stimulus contains pitch information embedded beneath the noisy component. A similar kind of stimulus (noise sequence) was used by Spahr et al. (2008) to simulate the shape of the electrical excitation in cochlear implants. In that study, the narrow band stimuli (fuzzy tones) were generated by varying the upper and lower cut off frequencies of the set of filters. Therefore, the shape of the slopes varied on either side of the noise spectrum. The pure tone and noise band stimuli (fuzzy tones) were used to create familiar melodies varying in different noise band indexes (r). The melodies were played either as pure tones (r = 0) or fuzzy tones (noise index (r) = 0.5, 1.0 and 1.5). The r value of 0, 0.5, 1.0, and 1.5 corresponds to bandwidth of 0.01, 0.19, 0.79, and 2.5 octaves respectively. In one of their experiments (musical tuning), the normal hearing subjects were required to judge the version of the melody having correct musical scale. The interval size of these melodies were either stretched or reduced, so that the subjects had to identify the in tune melody. The scores deteriorated as the noise index (r) increased from 0.5 and the performance was at chance for r =1.0 & 1.5. They concluded that melody recognition requires gross resolution of the pitch, while the musical tuning required fine resolution of pitch.

Music recognition is a complex perceptual skill, which requires a person to understand the relationship, sequences, and exact step size between the notes for correct recognition of a melody (Dowling & Fujitani, 1970). Discrimination and ranking are among the primitive skills required for melody recognition. An operational definition of pitch is that the variation in pitch can convey melody (Moore & Carlyon, 2005). This does not hold true for brightness sequences. The variation in brightness can convey gross information about the contour, but is certainly not capable of providing the finer melodic cues such as interval size judgements (McDermott et al., 2008 & 2010).

Implication for Cochlear Implant Place Pitch

The findings from the present study can address some issues related to the processing of pitch in cochlear implant (CI) recipients. The place-of-excitation in the cochlea depends on the spectral shape of the incoming electric signal and these variations lead to changes in the perceived pitch. It is reported that these variations consequently affect the perceived timbre (McDermott & McKay, 1997; McDermott, 2004; Moore & Carlyon, 2005). An operational definition of pitch is that variation in pitch can convey a melody (Moore & Carlyon, 2005). Therefore, only pitch variation does indeed convey a melody, but not many studies have investigated the place pitch (place of excitation) in a musical context.

Previous studies on cochlear implants have postulated that the place of excitation cue is in closer proximity to the brightness aspect of timbre than to pitch (McDermott & McKay, 1997; McDermott, 2004). In chapter 6, three kinds of stimuli were used to investigate the role of place and temporal pitch in CI users using loudspeakers. The stimuli were C3 (131 Hz) harmonic tones, C4 (262 Hz) harmonic tones, and C5 (523 Hz) pure tones. The quadrature envelope detection in a CI sound processor produces very little ripples for pure tones and stimulates specific corresponding intracochlear electrodes. Thus, a pure tone provides exclusive place pitch information and acts as an excellent signal for investigating place pitch in CI users. The CI pitch perception results show same level of performance for both place pitch and temporal pitch and demonstrate that both cues (in isolation) are capable of conveying adequate melodic pitch information.

The anecdotal reports from the previous chapter (6) reveals that the CI subjects were listening to pitch changes and not to brightness changes. However, to ascertain this notion, a series of pitch perception procedures were tested on normal hearing individuals using brightness and noise sequences. Performance with brightness in the present study reveals that the participants were quite efficient in discriminating, ranking, and also detecting contour changes in brightness. Brightness does indeed convey adequate cues about the contour, but not about interval size judgements. Interval size is crucial for recognising familiar melodies (Dowling & Fujitani, 1970) and is unique to pitch and not to any other perceptual dimensions (McDermott et al., 2010). Interestingly, the brightness performance (NH subjects) was similar to C5 (523 Hz) pure tones (CI subjects) performance. This suggests that place pitch in CI subjects and brightness in NH subjects can convey some aspects of musical pitch. The overall performance for brightness in the present study suggests that the CI recipients in chapter 6 and in Swanson et al.(2009) study may actually have perceived place pitch more similarly to the brightness attribute of timbre, and additional studies are needed to confirm these findings.

7.5. Conclusion

This chapter aimed to investigate the role of brightness and noise sequences on various pitch perception tasks in NH subjects. The performances with brightness sequences (harmonic tones varying in brightness and constant pitch) and noise sequences (noise band varying in cut-off frequency) were compared with the baseline pitch sequences (harmonic tones varying in pitch and constant brightness). Results showed that normal subjects were able to discriminate, rank and identify the correct contour of the melody using all the three stimulus types. However, they were unable to make judgments regarding musical intervals for brightness. This implies that the CI users in chapter 6 and Swanson et al. (2009) study may actually have perceived place pitch as a pattern of brightness changes and not as pitch and more studies are needed to confirm this findings.

Chapter 8: Summary and Conclusion

Cochlear Implant technology has improved so tremendously that a majority of CI recipients secure high scores in speech perception tasks in quiet. This remarkable improvement is confined to speech perception tasks in quiet, but fail to facilitate improvement in pitch perception tasks. This challenging topic has attracted a great deal of attention over the past few years. The final chapter summarises the findings from all the experiments, their limitations, and future directions for this study. Additionally, how these findings would contribute to the pool of knowledge on CI pitch perception is also discussed.

8.1. Summary of the experiments

The initial three chapters focused on the introduction and the literature review of CI pitch processing. In chapter 4, the performance of NH subjects on various pitch perception tasks using pitch sequences (C4 harmonic tones (262 Hz)) were investigated. The pitch sequences acted as a baseline condition and were compared with other perceptual dimensions (brightness and noise sequence) and also with the performance of CI recipients. This chapter aims to determine how the C4 harmonic tones (pitch sequences) affect the performance of normal hearing individuals. Additionally, what exactly the four experimental procedures were measuring and which among these procedures were sensitive for evaluating pitch perception was also investigated.

Four experimental procedures used in this study were discrimination (4 AFC), ranking (2 AFC), Modified Melodies test – backward modification (2 AFC), and Modified Melodies test - warp modification (2 AFC). Results showed that eighteen NH subjects were extremely efficient in detecting the contours for modified melodies backward and performance reached ceiling. Additionally, high scores were obtained for modified melodies warp and discrimination tasks. Ranking scores were comparatively less than the remaining procedures, but the scores were well above chance level. The NH subjects showed high correlation between the discrimination and the ranking procedures. Two musicians scored better than non-musicians especially in discrimination and ranking tasks. Additional musician subjects would be required to investigate this further.

Chapter 5 investigated the role of temporal pitch sensitivity as a function of different stimulation patterns (single, dual, and multiple) and base pulse rates (C3 – 131 pps, C4- 262 pps). Four stimulation patterns were tested using three experimental procedures. The stimulation patterns were single electrode stimulation (apical (Electrode E22) or middle (E12)), dual-electrode stimulation (E22 & E12), and multiple electrode stimulation (E22 to E12). The three procedures were: Ranking, Modified Melodies Test - Backward modification, and Modified Melodies Test - Warp modification. Each stimulus was a pulse train delivered on either single or multiple electrodes. The stimuli were presented as pulses at a base rate of 131 pulses per second (pps) (C3 range) and additional pulse rate of 262 pps (C4 range) was also used. The study hypothesised that, (a) the performance would be better at the apical (E22) electrode compared to middle electrode (E12), because of a better place-rate match. (b) When more electrodes are stimulated, more temporal information can be carried by a large number of auditory nerve fibres in CI recipients and better scores can be anticipated. So, it was hypothesised that the performance would be better for multiple electrodes (E22 to E12) compared to single or dual electrode stimulation. (c) The performance would drop as the base rate increased from C3 (131 pps) to C4 (262 pps) as seen in previous studies (Zeng, 2002; Kong et al., 2009).

At C3 - 131 pps base pulse rate, there was no significant difference in performance between middle (E12), apical (E22), dual (E22 & E12), and multiple (E22 to E12) electrode stimulation patterns. Interestingly, for C3-131 pps base rate there was no significant difference between the performance of the apical (E22) and the middle (E12) electrode. This performance reveals that the CI subjects were unable to combine temporal information from different places in the cochlear to give a stronger pitch cue. . At high base rate (C4 – 262 pps), overall performance dropped, but this was more evident in the Modified Melodies test (backward and warp). The apical (E22) performance was predominantly affected as the base rate increased from C3 (131 pps) to C4 (262 pps). This result supported the salience of correct rate-place match for accurate coding of pitch. Results showed a good correlation between modified melody backward and modified melody warp. The group mean score reveal that there was no significant difference among pitch ranking scores among different conditions. However, there were some significant differences between the conditions for the Modified Melodies test, implying that they were more sensitive. Finally, the CI recipients obtained similar information from these (single, dual, and multiple electrode) stimulation patterns. This implies that CI recipients were unable to combine temporal cues from different places in the cochlea to give a “stronger” cue. Thus, this study does not support the first two hypotheses and supports the third hypothesis, i.e., performance decreased when base rate increased from C3-131 pps to C4-262 pps.

Chapter 6 investigated the role of place pitch and temporal pitch in CI recipients using three stimulus types. The three stimuli were, C3 (131 Hz) harmonic tones (which provided exclusive temporal pitch information – see Figure 6.1 & 6.2); C5 (523 Hz) pure tones (which provided place pitch information – see Figure 6.3); and C4 (262 Hz) harmonic tones (which provided both place pitch and temporal pitch information). The stimuli were presented via loudspeaker at a comfortable loudness level. The recipients used their own sound processor. The performance of CI subjects in the four experimental procedures (discrimination, ranking, modified melodies backward, and modified melodies warp) using the three stimulus types were analysed. Results showed that good scores occurred only when the subjects were provided with a single cue either temporal pitch (C3 (131 Hz) harmonic tones) or place pitch (C5 (523 Hz) pure tones). The scores for C4 (262 Hz) harmonic tones were comparatively less than the scores obtained using other stimuli and scores dropped to chance for modified melodies warp. The performances of CI subjects in the four experimental procedures using C4 (262 Hz) harmonic tones (pitch sequences) were significantly poorer than the NH counterparts in chapter 4. The C4 harmonic tones potentially offered both temporal and place cues which can be evident from the electrodogram in Figure 6.4. However, their fundamental frequency range of 262 Hz – 523 Hz is probably above the upper cut-off limit of temporal pitch for most CI recipients and additionally place cues may have been ambiguous. This may explain the poor performance of CI recipients using C4 (262 Hz) harmonics tones and may also explain the pitch reversals seen in these subjects. Although there was a large intersubject variability among CI recipients, the results showed a high correlation between discrimination and ranking, ranking and modified melodies backward, and modified melodies backward and modified melodies warp. The modified melodies backward was found to be a sensitive test compared to the modified melodies warp. Finally, the CI performance in this study suggests that CI place pitch may convey melodic pitch information, but the contribution of brightness cannot be completely ruled out.

Chapter 7 addressed the research question about whether CI recipients in the previous chapter perceived place pitch as pitch and not as a pattern of brightness changes. As brightness can be ordered on a low-to-high scale, it can provide high scores on ranking and discrimination tests. In order to explore this issue, brightness sequences were created and the role of brightness was investigated in various pitch perception tasks in NH subjects. The brightness sequences (harmonic tones varying in brightness but with constant pitch) and noise sequences (noise band varying in cut-off frequency) were compared with the baseline pitch sequences (harmonic tones varying in pitch and constant brightness). The four experimental procedures using the three stimuli were analysed. The scores for discrimination and ranking were high for all three stimulus types, and d' analysis revealed that the subjects' performance was better in discrimination than ranking. In the Modified Melodies test, when subjects were required to detect a melody contour change, scores were high for all three stimulus conditions. Conversely, when the melody contour was preserved but the musical intervals were changed, scores using pitch sequences were high, while scores using brightness or noise sequences were at chance-level. Thus subjects were able to discriminate and rank brightness, and were able to detect brightness contour changes, but were unable to make judgements of musical intervals for brightness.. This implies that the CI users in previous studies (chapter 6 and Swanson et al. (2009)) may have perceived place pitch as a pattern of brightness changes and not as pitch and additional studies are needed to confirm these findings.

8.2. Temporal Pitch, Place Pitch, and Brightness

Temporal pitch and place pitch are two cues to pitch in CI recipients. These two pitch percepts can be studied independently only in CI recipients (Tong et al., 1983; McKay et al., 2000). In the real world scenario, it is hard to unlock the interdependency of these two percepts in normal hearing individuals. An operational definition of pitch is that the variation in pitch can convey a melody (Moore & Carlyon, 2005). Thus in this thesis, Modified Melodies test along with other pitch perception procedures were used to assess the pitch perception abilities of NH and CI subjects.

Temporal pitch can be conveyed either by varying the pulse rate on an electrode (5.2.3) or amplitude modulating the carrier pulse train (e.g., C3 (131 Hz) harmonic tones (6.2.3.1)). The results obtained from varying the pulse rates (single, dual, and multiple) were similar to the results obtained from C3 (131 Hz) harmonic tones processed by the ACE processing strategy (Figure 6.14). This clearly demonstrates that temporal pitch can be mediated by both pulse rate variation on an electrode and amplitude modulation of the carrier pulse train. Previous studies on CI temporal pitch provided evidence that temporal pitch in isolation can convey adequate melodic pitch (Pijl & Schwarz, 1995b; McDermott & McKay, 1997). Similarly, in this thesis, when only temporal cues were provided, CI subjects were able to discriminate musical notes, rank musical notes, identify the melodic contours, and judge the interval size of the melodies. Interesting finding in this thesis was that when different pattern of stimulation (single, dual, and multiple) were used CI performance did not vary across the pitch perception tasks. Conversely, in normal hearing subjects, performance increases when additional harmonic was added to the stimulus (Houtsma & Smurzynski, 1990; Moore & Peters, 1992). This normal hearing trend was not observed in CI recipients indicating that these recipients were unable to combine the temporal cues from different places in the cochlear to give a strong pitch cue. The Modified Melodies scores at C4 base rate were significantly worse than C3 base rate for the apical electrodes compared to the mid electrode. The scores from this study suggests that the interaction between rate and place are more complex than suggested by the first and the third hypothesis.

Earlier studies and chapter 6 clearly demonstrated that temporal pitch alone is efficient in conveying melodic pitch information to CI recipients. There are only a few studies which have attempted to investigate the place pitch cues in CI recipients. In Chapter 6, three stimulus types which evoked different set of cues in CI recipients were used. The C3 (131 Hz) harmonic tones which had a range of 131 Hz to 220 Hz provided exclusive temporal pitch information and this is evident in the electrodogram (Figure 6.1 & 6.2). The electrodogram clearly demonstrate the amplitude modulation cue which helps in conveying temporal pitch information in CI recipients. Not many researchers have used C5 (523 Hz) pure tones to explore pitch cues in CI recipients. The significance of using C5 pure tones was that the range of 523 Hz to 880 Hz was well above the temporal pitch range and the output of the quadrature envelope in a filter bank of speech processor produce very little ripples, i.e., very minimal amplitude modulation, therefore providing only place of stimulation cue to CI recipients (Swanson et al., 2007). This experiment provides clear evidence that when a single cue (either Temporal – C3 harmonic tones or Place – C5 pure tones) is provided CI recipients obtained good scores and performance dropped to chance when the stimulus contained both place and temporal cues (C4 harmonic tones). The C4 fundamental frequency range of 262 Hz to 523 Hz was probably above the upper frequency limit of temporal pitch for most subjects and additional place pitch cues may have been ambiguous.

Earlier studies have postulated that the place of excitation cue in CI corresponds to the brightness aspect of timbre than to pitch (Laneau & Wouters, 2004; Moore & Carlyon, 2005). A few CI subjects self-reported that they were able to associate the three stimulus types (C3 (131 Hz) harmonic tones, C4 (262 Hz) harmonic tones, and C5 (523 Hz) pure tones) with an appropriate musical instrument class. For example, C5 (523 Hz) pure tones were associated with a flute. A majority of them at least got the instrument class correct. Informally, this does indeed suggests that the CI recipients were relying on pitch cues. In order to ascertain whether CI recipients were listening to pitch changes and not to brightness changes, a series of pitch perception procedures were tested on normal hearing individuals.

Brightness, similar to pitch, can be ordered on a scale from low to high. Pitch depends on the fundamental frequency of the stimulus, while brightness depends on the spectral profile (amplitude of the harmonics) of the signal (3.6). Thus, tasks such as discrimination and ranking do not answer the research question of whether place cue in CI is either pitch or brightness. In these two tasks, both pitch and brightness sequences yielded high scores and this narrows down to the fundamental definition of pitch, i.e., variation in pitch does convey a melody. Thus, the Modified Melodies test was utilised and the results reveal that brightness variation may convey contour information, but is not capable of providing interval size judgements and this is evident in chapter 7 (7.3). Therefore, the results suggest that the CI subjects may actually have perceived place pitch as brightness changes rather than pitch.

A similar performance trend was observed between CI place pitch (C5 (523 Hz) pure tones) and NH brightness. Looking solely at this result, one can draw a conclusion that CI place pitch may be analogous to brightness changes rather than pitch. Surprisingly, the results obtained from chapter 6 revealed that CI performance was similar for both temporal pitch and place pitch stimuli in various pitch perception tasks. The conclusion drawn from this chapter suggests that CI place pitch does indeed convey melodic pitch information. These two major findings in some way contradict each other and hinge on the definition of pitch. As mentioned earlier, an operational definition of pitch is that the variation in pitch can convey a melody (Moore & Carlyon, 2005). This typically suggests that CI place pitch performance may be pitch, however the NH subjects were able to discriminate, rank, and identify the contour of brightness sequences and showed a similar performance trend as for CI place pitch. This result along with the results obtained from all the chapters suggest that the CI recipients may have perceived place pitch as brightness and not as pitch. Although CI subjects were able to perform these pitch perception tasks using place or temporal cues in isolation, there still remains a significant performance gap between NH and CI subjects. It should be noted that the NH subjects had a mean age of 25.83 yrs as compared to the CI subjects who had a mean age of 64.3 yrs. Indeed, hearing sensitivity deteriorates as the age progresses (usually seen above 60 years – presbycusis or age-related hearing loss). However, these CI recipients had normal or near normal aided thresholds at speech frequencies. However, the performance of NH subjects was far superior to CI subjects.

8.3. Limitation of this thesis

The NH subjects who participated in this study were mostly non-musicians, but only few were musicians (N = 2). Their performance was better than the non-musicians. It would be interesting to test the performance of musicians using similar protocols. Previous studies have shown significant differences between musician and non-musician on various pitch perception tasks (Pitt, 1994; McDermott et al., 2010).

In the real world, normally hearing listeners are conditioned to listen to sounds which vary in pitch and constant brightness. Therefore, NH subjects did not face any problem performing the experimental procedures using pitch sequences. However, NH subjects incurred problems in orienting themselves to brightness sequences and it would be interesting to see the influence of training on these tasks.

Finally, in this thesis, only six CI recipients were tested for the period of six sessions. It would have been better if more subjects had participated in this study.

8.4. Future Directions

The behavioural results from this thesis demonstrated that CI place of excitation cues (place pitch) may correspond to the brightness aspect of timbre rather than to pitch.

An extension study using objective measurement such as EEG (Electroencephalography), fMRI (Functional magnetic resonance imaging), and cochlear implant enabled MEG (Magnetoencephalography) would enable the researchers to visualise the accurate localisation of these two perceptual dimensions in CI recipients' brains.

Musicians are trained to rely on the pitch dimension and are taught to understand the melodic contours and judge musical interval sizes. Musicians outperformed non-musicians in the majority of the pitch perception tasks. Therefore, it would be worth investigating the performance of musicians in these pitch perception tasks using pitch, brightness, and noise sequences.

Researchers have showed that auditory training does indeed improve pitch processing in CI recipients (Galvin et al., 2007; Fu & Galvin, 2012). Therefore, the influence of training CI recipients on these pitch perception tasks could also be investigated.

Currently researchers have developed focused spatial stimulation to improve pitch perception in CI recipients (2.3.2.4). The partial tripolar (pTP) and phased array (PA) stimulation provided promising results. Temporal pitch sensitivity can also be investigated as a function of focused stimulation, electrode stimulation patters (single, dual, and multiple), and base pulse rate. It would also be worthwhile to investigate the CI place pitch perception using focused intracochlear electric stimulation.

References

Abbas, P. J., & Miller, C. A. (2004). Biophysics and Physiology. In F. -G. Zeng, A. N. Popper & R. R. Fay (Eds.), Cochlear Implants: Auditory Prostheses and Electric Hearing: (Vol. 20). New York: Springer.

Anantharaman, J. N., Krishnamurthy, A. K., & Feth, L. L. (1993). Intensity-weighted average of instantaneous frequency as a model for frequency discrimination. The Journal of the Acoustical Society of America, 94(2), 723-729.

ANSI. (1973). Psychoacostical terminology. ANSI S3.20. New York: American National Standards Institute.

ANSI. (1994). American National Standard Acoustic Terminology. New York: American National Standards Institute.

Arnd, P., Staller, S., Arcaroli, J., Hines, A., & Ebinger, K. (1999). Within-subjects comparison of advanced coding strategies in Nucleus 24 cochlear implant. Technical Report, Cochlear Corporation, Englewood, Colorado.

Arnoldner, C., Riss, D., Brunner, M., Durisin, M., Baumgartner, W.-D., & Hamzavi, J.-S. (2007). Speech and music perception with the new fine structure speech coding strategy: preliminary results. Acta Oto-laryngologica, 127(12), 1298-1303.

ASA. (1960). Acoustical Terminology SI, 1-1960. New York: American Standards Association.

Berenstein, C. K., Mens, L. H. M., Mulder, J. J. S., & Vanpoucke, F. J. (2008). Current steering and current focusing in cochlear implants: comparison of monopolar, tripolar, and virtual channel electrode configurations. Ear Hear, 29(2), 250-260.

Bernstein, J. G., & Oxenham, A. J. (2003). Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number? The Journal of the Acoustical Society of America, 113(6), 3323-3334.

Bierer, J. A. (2007). Threshold and channel interaction in cochlear implant users: Evaluation of the tripolar electrode configuration. The Journal of the Acoustical Society of America, 121(3), 1642-1653.

Bilger, R., Black, F., Hopkinson, N., Myers, E., Payne, J., Stenson, N., et al. (1977). Evaluation of subjects presently fitted with implanted auditory prostheses. Annals of Otology, Rhinology & Laryngology, Suppl 86, 3 - 10.

Blamey, P. J., Dowell, R. C., Tong, Y. C., & Clark, G. M. (1984). An acoustic model of a multiple-channel cochlear implant. The Journal of the Acoustical Society of America, 76(1), 97-103.

Boëx, C., Baud, L., Cosendai, G., Sigrist, A., Kós, M.-I., & Pelizzone, M. (2006). Acoustic to Electric Pitch Comparisons in Cochlear Implant Subjects with Residual Hearing. Journal of the Association for Research in Otolaryngology, 7(2), 110-124.

Bonham, B. H., & Litvak, L. M. (2008). Current focusing and steering: Modeling, physiology, and psychophysics. Hearing Research, 242(1–2), 141-153.

Burns, E. M., & Viemeister, N. F. (1976). Nonspectral pitch. The Journal of the Acoustical Society of America, 60(4), 863-869.

Burns, E. M., & Viemeister, N. F. (1981). Played-again SAM: Further observations on the pitch of amplitude-modulated noise. The Journal of the Acoustical Society of America, 70(6), 1655-1660.

Busby, P. A., Battmer, R. D., & Pesch, J. (2008). Electrophysiological Spread of Excitation and Pitch Perception for Dual and Single Electrodes Using the Nucleus Freedom Cochlear Implant. Ear & Hearing, 29(6), 853-864.

Busby, P. A., & Clark, G. M. (1997). Pitch and loudness estimation for single and multiple pulse per period electric pulse rates by cochlear implant patients. The Journal of the Acoustical Society of America, 101(3), 1687-1695.

Busby, P. A., & Plant, K. L. (2005). Dual Electrode Stimulation Using the Nucleus CI24RE Cochlear Implant: Electrode Impedance and Pitch Ranking Studies. Ear & Hearing, 26(5), 504-511.

Busby, P. A., Tong, Y. C., & Clark, G. M. (1993). The perception of temporal modulations by cochlear implant patients. The Journal of the Acoustical Society of America, 94(1), 124-131.

Busby, P. A., Whitford, L. A., Blamey, P. J., Richardson, L. M., & Clark, G. M. (1994). Pitch perception for different modes of stimulation using the Cochlear multiple-electrode prosthesis. The Journal of the Acoustical Society of America, 95(5), 2658-2669.

Caclin, A., Brattico, E., Tervaniemi, M., Naatanen, R., Morlet, D., Giard, M. H., et al. (2006). Separate neural processing of timbre dimensions in auditory sensory memory. Journal of Cognitive Neuroscience, 18(12), 1959-1972.

Caclin, A., Giard, M.-H., Smith, B. K., & McAdams, S. (2007). Interactive processing of timbre dimensions: A Garner interference study. Brain Research, 1138, 159-170.

Cariani, P. A., & Delgutte, B. (1996). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. Journal of Neurophysiology, 76(3), 1698-1716.

Carlyon, R. P., & Deeks, J. M. (2002). Limitations on rate discrimination. The Journal of the Acoustical Society of America, 112(3), 1009-1025.

Carlyon, R. P., Deeks, J. M., & McKay, C. M. (2010). The upper limit of temporal pitch for cochlear-implant listeners: Stimulus duration, conditioner pulses, and the number of electrodes stimulated. The Journal of the Acoustical Society of America, 127(3), 1469-1478.

Chittka, L., & Brockmann, A. (2005). Perception Space—The Final Frontier. PLoS Biology, 3(4), e137.

Choi, C. T. M., & Lee, Y.-H. (2012). A Review of Stimulating Strategies for Cochlear Implants. In C. Umat & A. R. Tange (Eds.), Cochlear Implant Research Updates. Shanghai: InTech.

Clark, G. (2003). Cochlear Implants: Fundamentals and Applications. New York: Springer.

Cohen, L. T., Busby, P. A., & Clark, G. M. (1996a). Cochlear Implant Place Psychophysics. Audiology and Neurotology, 1(5), 278-292.

Cohen, L. T., Busby, P. A., Whitford, L. A., & Clark, G. M. (1996b). Cochlear Implant Place Psychophysics. Audiology and Neurotology, 1(5), 265-277.

Cooper, H., & Craddock, L. (2006). Cochlear Implants: A Practical Guide. New York: John Wiley & Sons.

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 & Hearing, 29(4), 618-626.

Dai, H., Nguyen, Q., Kidd, J. G., Feth, L. L., & Green, D. M. (1996). Phase independence of pitch produced by narrow-band sounds. The Journal of the Acoustical Society of America, 100(4), 2349-2351.

Dallos, P. (1992). The active cochlea. The Journal of Neuroscience, 12(12), 4575-4585.

Demany, L., & Semal, C. (1993). Pitch versus Brightness of Timbre: Detecting Combined Shifts in Fundamental and Formant Frequency. Music Perception: An Interdisciplinary Journal, 11(1), 1-13.

Djourno, A., & Eyries, C. (1957). Auditory prothesis by means of a distant electrical stimulation of the sensory nerve with the use of an indwelt coiling. La Presse Médicale, 65(63), 1417.

Djourno, A., Eyries, C., & Vallancien, P. (1957a). Electric excitation of the cochlear nerve in man by induction at a distance with the aid of micro-coil included in the fixture. Comptes Rendus des Seances de la Societe de Biologie et des ses Filiales, 151, 423 - 425.

Djourno, A., Eyries, C., & Vallancien, P. (1957b). Preliminary attempts of electrical excitation of the auditory nerve in man, by permanently inserted micro-apparatus. Bulletin of the National Academy of Medicine, 141, 481 - 483.

Donaldson, G. S., Kreft, H. A., & Litvak, L. (2005). Place-pitch discrimination of single- versus dual-electrode stimuli by cochlear implant users. The Journal of the Acoustical Society of America, 118(2), 623-626.

Donnelly, P. J., Guo, B. Z., & Limb, C. J. (2009). Perceptual fusion of polyphonic pitch in cochlear implant users. The Journal of the Acoustical Society of America, 126(5), EL128-EL133.

Dorman, M. F., Loizou, P. C., & Rainey, D. (1997). Speech intelligibility as a function of the number of channels of stimulation for signal processors using sine-wave and noise-band outputs. The Journal of the Acoustical Society of America, 102(4), 2403-2411.

Dorman, M. F., Smith, M., Smith, L., & Parkin, J. L. (1994). The pitch of electrically presented sinusoids. The Journal of the Acoustical Society of America, 95(3), 1677-1679.

Dowling, W. J., & Fujitani, D. S. (1970). Contours, interval, and pitch recognition in memory for melodies. The Journal of the Acoustical Society of America, 49, 524 - 531.

Eddington, D. (1980). Speech discrimination in deaf subjects with cochlear implants. The Journal of the Acoustical Society of America, 68(3), 885-891.

Eddington, D., Dobelle, W., Brackmann, D., Mladejovsky, M., & Parkin, J. (1978a). Auditory prothesis research with multiple channel intracochlear stimulation in man. Annals of Otology, Rhinology & Laryngology, 87, 1 - 39.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., & Parkin, J. L. (1978b). Place and periodicity pitch by stimulation of multiple scala tympani electrodes in deaf volunteers. Transactions - American Society for Artificial Internal Organs, 24, 1-5.

Emadi, G., Richter, C.-P., & Dallos, P. (2004). Stiffness of the Gerbil Basilar Membrane: Radial and Longitudinal Variations. Journal of Neurophysiology, 91(1), 474-488.

Fayad, J. N., & Linthicum, F. H. (2006). Multichannel Cochlear Implants: Relation of Histopathology to Performance. The Laryngoscope, 116(8), 1310-1320.

Fearn, R., Carter, P., & Wolfe, J. (1999). The perception of pitch by users of cochlear implants: possible significance for rate and place theories of pitch. Acoustics Australia, 27(2), 41-43.

Fearn, R., & Wolfe, J. (2000). The relative importance of rate and place: Experiments using pitch scaling technique with cochlear implantees. Annals of Otology, Rhinology & Laryngology, 109 (Supp 185) (12), 51-53.

Fechner, G. (1966). Elements of psychophysics, translated by H.E. Adler (Vol. I, pp. 402). New York: Holt, Rinehart & Winston.

Firszt, J. B., Koch, D. B., Downing, M., & Litvak, L. (2007). Current Steering Creates Additional Pitch Percepts in Adult Cochlear Implant Recipients. Otology & Neurotology, 28(5), 629-636.

Fourcin, A. J., Rosen, S. M., Moore, B. C. J., Douek, E. E., Clarke, G. P., Dodson, H., et al. (1979). External Electrical Stimulation of the Cochlea: Clinical, Psychophysical, Speech-Perceptual and Histological Findings. British Journal of Audiology, 13(3), 85-107.

Frijns, J. H. M., Briaire, J. J., Zarowski, A., Verbist, B. M., & Kuzma, J. (2004). Concept and initial testing of a new, basally perimodiolar electrode design. International Congress Series, 1273, 105-108.

Frijns, J. H. M., Dekker, D. M. T., & Briaire, J. J. (2011). Neural excitation patterns induced by phased-array stimulation in the implanted human cochlea. Acta Oto-laryngologica, 131(4), 362-370.

Fu, Q.-J., & Galvin, J. J. (2012). Auditory Training for Cochlear Implant Patients. In F.-G. Zeng, A. N. Popper & R. R. Fay (Eds.), Auditory Prostheses (Vol. 39, pp. 257-278). New York: Springer.

Fu, Q.-J., & Nogaki, G. (2005). Noise Susceptibility of Cochlear Implant Users: The Role of Spectral Resolution and Smearing. Journal of the Association for Research in Otolaryngology, 6(1), 19-27.

Fu, Q.-J., & Shannon, R. V. (1998). Effects of amplitude nonlinearity on phoneme recognition by cochlear implant users and normal-hearing listeners. The Journal of the Acoustical Society of America, 104(5), 2570-2577.

Galvin, J., Fu, Q.-J., & Shannon, R. (2009). Melodic Contour Identification and Music Perception by Cochlear Implant Users. Annals of the New York Academy of Sciences, 1169 (The Neurosciences and Music III Disorders and Plasticity), 518-533.

Galvin, J., Fu, Q.-J., & Nogaki, G. (2007). Melodic Contour Identification by Cochlear Implant Listeners. Ear and Hearing, 28(3), 302-319.

Geurts, L., & Wouters, J. (2001). Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. The Journal of the Acoustical Society of America, 109(2), 713-726.

Geurts, L., & Wouters, J. (2004). Better place-coding of the fundamental frequency in cochlear implants. The Journal of the Acoustical Society of America, 115(2), 844-852.

Gfeller, K., Knutson, J. F., Woodworth, G., Witt, S., & DeBus, B. (1998). Timbral recognition and appraisal by adult cochlear implant users and normal-hearing adults. Journal of the American Academy of Audiology, 9(1), 1-19.

Gfeller, K., & Lansing, C. R. (1991). Melodic, Rhythmic, and Timbral Perception of Adult Cochlear Implant Users. Journal of Speech and Hearing Research, 34(4), 916-920.

Gfeller, K., Turner, C., Mehr, M., Woodworth, G., Fearn, R., Knutson, J., et al. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. Cochlear Implants International, 3(1), 29-53.

Gfeller, K., Turner, C., Oleson, J., Zhang, X., Gantz, B., Froman, R., et al. (2007). Accuracy of Cochlear Implant Recipients on Pitch Perception, Melody Recognition, and Speech Reception in Noise. Ear & Hearing, 28(3), 412-423.

Goldstein, J. L. (1973). An optimum processor theory for the central formation of the pitch of complex tones. The Journal of the Acoustical Society of America, 54(6), 1496-1516.

Gray, H. (1918). Anatomy of the human body (20 ed.). Philadelphia: Lea & Febiger.

Grayden, D. B., Burkitt, A. N., Kenny, O. P., Clarey, J. C., Paolini, A. G., & Clark, G. M. (2004). A cochlear implant speech processing strategy based on an auditory model. Paper presented at the Intelligent Sensors, Sensor Networks and Information Processing Conference.

Green, T., Faulkner, A., Rosen, S., & Macherey, O. (2005). Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification. The Journal of the Acoustical Society of America, 118(1), 375-385.

Hartmann, R., Topp, G., & Klinke, R. (1984). Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea. Hearing Research, 13(1), 47-62.

Helms, J., Müller, J., Schön, F., Winkler, F., Moser, L., Shehata-Dieler, W., et al. (2001). Comparison of the TEMPO+ Ear-Level Speech Processor and the CIS PRO+ Body-Worn Processor in Adult MED-EL Cochlear Implant Users. Journal for Oto-Rhino-Laryngology, Head and Neck Surgery, 63(1), 31-40.

Hochmair-Desoyer, I., Hochmair, E., Burian, K., & Fischer, R. (1981). Four years of experience with cochlear prostheses. Medical Progress Through Technology, 8(3), 107 - 119.

Houtsma, A. J. M., & Smurzynski, J. (1990). Pitch identification and discrimination for complex tones with many harmonics. The Journal of the Acoustical Society of America, 87(1), 304-310.

Javel, E., Tong, Y. C., Shepherd, R. K., & Clark, G. M. (1987). Responses of cat auditory nerve fibers to biphasic electrical current pulses. Annals of Otology, Rhinology & Laryngology, 96 (suppl.128)(1), 26-30.

Kaernbach, C., & Bering, C. (2001). Exploring the temporal mechanism involved in the pitch of unresolved harmonics. The Journal of the Acoustical Society of America, 110(2), 1039-1048.

Kaernbach, C., & Demany, L. (1998). Psychophysical evidence against the autocorrelation theory of auditory temporal processing. The Journal of the Acoustical Society of America, 104(4), 2298-2306.

Kang, R., Nimmons, G. L., Drennan, W., Longnion, J., Ruffin, C., Nie, K., et al. (2009). Development and Validation of the University of Washington Clinical Assessment of Music Perception Test. Ear and Hearing, 30(4), 411-418.

Katz, J., Burkard, R. F., Medwetsky, L., & Hood, L. (2010). Handbook of Clinical Audiology. Philadelphia: Lippincott Williams & Wilkins.

Kennedy, H. J., Evans, M. G., Crawford, A. C., & Fettiplace, R. (2006). Depolarization of Cochlear Outer Hair Cells Evokes Active Hair Bundle Motion by Two Mechanisms. The Journal of Neuroscience, 26(10), 2757-2766.

Koch, D. B., Osberger, M. J., Segel, P., & Kessler, D. (2004). HiResolution and Conventional Sound Processing in the HiResolution Bionic Ear: Using Appropriate Outcome Measures to Assess Speech Recognition Ability. Audiology and Neurotology, 9(4), 214-223.

Kong, Y.-Y., & Carlyon, R. P. (2010). Temporal pitch perception at high rates in cochlear implants. The Journal of the Acoustical Society of America, 127(5), 3114-3123.

Kong, Y.-Y., Deeks, J. M., Axon, P. R., & Carlyon, R. P. (2009). Limits of temporal pitch in cochlear implants. The Journal of the Acoustical Society of America, 125(3), 1649-1657.

Kong, Y.-Y., Mullangi, A., Marozeau, J., & Epstein, M. (2011). Temporal and Spectral Cues for Musical Timbre Perception in Electric Hearing. Journal of Speech, Language, and Hearing Research, 54(3), 981-994.

Kong, Y.-Y., Cruz, R., Jones, J., & Zeng, F.-G. (2004). Music perception with temporal cues in acoustic and electric hearing. Ear & Hearing, 25(2), 13.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual Interactions Between Musical Pitch and Timbre. Journal of Experimental Psychology: Human Perception & Performance, 18(3), 739-751.

Kwon, B. J., & Van den Honert, C. (2006). Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users. The Journal of the Acoustical Society of America, 120(1), EL1-EL6.

Landsberger, D., & Galvin, J. J. (2011). Discrimination between sequential and simultaneous virtual channels with electrical hearing. The Journal of the Acoustical Society of America, 130(3), 1559-1566.

Landsberger, D. M., Padilla, M., & Srinivasan, A. G. (2012). Reducing current spread using current focusing in cochlear implant users. Hearing Research, 284(1–2), 16-24.

Landsberger, D. M., & Srinivasan, A. G. (2009). Virtual channel discrimination is improved by current focusing in cochlear implant recipients. Hearing Research, 254(1–2), 34-41.

Laneau, J. (2005). When The Deaf Listen to Music Pitch Perception with Cochlear Implants. PhD Thesis, Katholieke Universiteit Leuven, Leuven.

Laneau, J., & Wouters, J. (2004a). Multichannel Place Pitch Sensitivity in Cochlear Implant Recipients. Journal of the Association for Research in Otolaryngology, 5(3), 285-294.

Laneau, J., Wouters, J., & Moonen, M. (2004b). Relative contributions of temporal and place pitch cues to fundamental frequency discrimination in cochlear implantees. The Journal of the Acoustical Society of America, 116(6), 3606-3619.

Laneau, J., Wouters, J., & Moonen, M. (2006). Improved Music Perception with Explicit Pitch Coding in Cochlear Implants. Audiology & Neuro-Otology, 11(1), 38-52.

Langner, G. (1992). Periodicity coding in the auditory system. Hearing Research, 60(2), 115-142.

Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. The Journal of the Acoustical Society of America, 49(2B), 467-477.

Lichte, W. H. (1941). Attributes of complex tones. Journal of Experimental Psychology, 28(6), 455-480.

Licklider, J. (1951). A duplex theory of pitch perception. Experientia, 7(4), 128-134.

Litvak, L. M., Spahr, A. J., & Emadi, G. (2007). Loudness growth observed under partially tripolar stimulation: Model and data from cochlear implant listeners. The Journal of the Acoustical Society of America, 122(2), 967-981.

Long, C. J., Nimmo-Smith, I., Baguley, D. M., O'Driscoll, M., Ramsden, R., Otto, S. R., et al. (2005). Optimizing the Clinical Fit of Auditory Brain Stem Implants. Ear and Hearing, 26(3), 251-262.

Looi, V. (2008). The Effect of Cochlear Implantation of Music Perception: A Review. Otorinolaringologia, 58(4), 169-190.

Looi, V., McDermott, H., McKay, C. M., & Hickson, L. (2004). Pitch discrimination and melody recognition by cochlear implant users. International Congress Series, 1273, 197-200.

Looi, V., McDermott, H., McKay, C. M., & Hickson, L. (2007). Comparisons of Quality Ratings for Music by Cochlear Implant and Hearing Aid Users. Ear and Hearing, 28(2), 59S-61S.

Luo, X., Landsberger, D. M., Padilla, M., & Srinivasan, A. G. (2010). Encoding pitch contours using current steering. The Journal of the Acoustical Society of America, 128(3), 1215-1223.

Luo, X., Padilla, M., & Landsberger, D. M. (2012). Pitch contour identification with combined place and temporal cues using cochlear implants. The Journal of the Acoustical Society of America, 131(2), 1325-1336.

Macherey, O., Deeks, J., & Carlyon, R. (2011). Extending the Limits of Place and Temporal Pitch Perception in Cochlear Implant Users. Journal of the Association for Research in Otolaryngology, 12(2), 233-251.

Macmillan, A. N., & Creelman, C. D. (2005). Detection theory : A User's Guide (Second Edition ed.): Psychology Press.

Marozeau, J., & de Cheveigne, A. (2007). The effect of fundamental frequency on the brightness dimension of timbre. The Journal of the Acoustical Society of America, 121(1), 383-387.

McDermott, H. J. (2012). Music Perception. In F.-G. Zeng, A. N. Popper & R. R. Fay (Eds.), Auditory Prostheses (Vol. 39, pp. 305-339). New York: Springer

McDermott, H. J. (2004). Music Perception with Cochlear Implants: A Review. Trends in Amplification, 8(2), 49-82.

McDermott, H. J., & McKay, C. M. (1994). Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea. The Journal of the Acoustical Society of America, 96(1), 155-162.

McDermott, H. J., & McKay, C. M. (1997). Musical pitch perception with electrical stimulation of the cochlea. The Journal of the Acoustical Society of America, 101(3), 1622-1631.

McDermott, J. H., Keebler, M. V., Micheyl, C., & Oxenham, A. J. (2010). Musical intervals and relative pitch: Frequency resolution, not interval resolution, is special. The Journal of the Acoustical Society of America, 128(4), 1943-1951.

McDermott, J. H., Lehr, A. J., & Oxenham, A. J. (2008). Is Relative Pitch Specific to Pitch? Psychological Science, 19(12), 1263-1271.

McKay, C. M. (2004). Psychophysics and Electrical Stimulation. In F.-G. Zeng, A. N. Popper & R. R. Fay (Eds.), Cochlear Implants: Auditory Prostheses and Electric Hearing (Vol. 20, pp. 286-333). New York: Springer.

McKay, C. M., McDermott, H., & Carlyon, R. (2000). Place and temporal cues in pitch perception: are they truly independent? Acoustics Research Letters Online, 1(1), 25-30.

McKay, C. M., McDermott, H., & Clark, G. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. The Journal of the Acoustical Society of America, 96(5), 2664-2673.

McKay, C. M., & Carlyon, R. P. (1999). Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains. The Journal of the Acoustical Society of America, 105(1), 347-357.

McKay, C. M., McDermott, H. J., & Clark, G. M. (1995). Pitch matching of amplitude-modulated current pulse trains by cochlear implantees: The effect of modulation depth. The Journal of the Acoustical Society of America, 97(3), 1777-1785.

McKay, C. M., McDermott, H. J., & Clark, G. M. (1996). The perceptual dimensions of single-electrode and nonsimultaneous dual-electrode stimuli in cochlear implantees. The Journal of the Acoustical Society of America, 99(2), 1079-1090.

McKay, C. M., Remine, M. D., & McDermott, H. J. (2001). Loudness summation for pulsatile electrical stimulation of the cochlea: Effects of rate, electrode separation, level, and mode of stimulation. The Journal of the Acoustical Society of America, 110(3), 1514-1524.

McKinney, M. (2010). Music through hearing aids: Perception and modeling. Paper presented at the 159th Meeting of the Acoustical Society of America - Lay Language Papers, Baltimore, Maryland.

Meddis, R., & Hewitt, M. J. (1991). Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification. The Journal of the Acoustical Society of America, 89(6), 2866-2882.

Meddis, R., & O'Mard, L. (1997). A unitary model of pitch perception. The Journal of the Acoustical Society of America, 102(3), 1811-1820.

Mens, L. H. M., & Berenstein, C. K. (2005). Speech Perception with Mono- and Quadrupolar Electrode Configurations: A Crossover Study. Otology & Neurotology, 26(5), 957-964.

Michelson, R. (1971). The result of electrical stimulation of the cochlea in human sensory deafness. Annals of Otology, Rhinology & Laryngology, 80, 914 - 919.

Milczynski, M., Wouters, J., & van Wieringen, A. (2009). Improved fundamental frequency coding in cochlear implant signal processing. The Journal of the Acoustical Society of America, 125(4), 2260-2271.

Moller, A. (2006). Hearing: Anatomy, Physiology, and Disorders of the Auditory System (2 ed.). San Diego: Academic Press.

Moore, B. (2003). An Introduction to the Psychology of Hearing (5 ed.). London: Academic Press.

Moore, B. C. J., & Carlyon, R. P. (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In C. Plack, R. Fay, A. Oxenham & A. Popper (Eds.), Pitch: Neural Coding and Perception (Vol. 24, pp. 234-277). New York: Springer.

Moore, B. C. J., Glasberg, B. R., & Peters, R. W. (1985). Relative dominance of individual partials in determining the pitch of complex tones. The Journal of the Acoustical Society of America, 77(5), 1853-1860.

Moore, B. C. J., & Peters, R. W. (1992). Pitch discrimination and phase sensitivity in young and elderly subjects and its relationship to frequency selectivity. The Journal of the Acoustical Society of America, 91(5), 2881-2893.

Moore, B. C. J., & Rosen, S. M. (1979). Tune recognition with reduced pitch and interval information. Quarterly Journal of Experimental Psychology, 31(2), 229 - 240.

Nelson, D. A., Van Tasell, D. J., Schroder, A. C., Soli, S., & Levine, S. (1995). Electrode ranking of "place" pitch and speech recognition in electrical hearing. The Journal of the Acoustical Society of America, 98(4), 1987-1999.

Nie, K., Stickney, G., & Zeng, F.-G. (2005). Encoding Frequency Modulation to Improve Cochlear Implant Performance in Noise. IEEE Transactions on Biomedical Engineering, 52(1), 64-73.

Nimmons, G. L., Kang, R. S., Drennan, W. R., Longnion, J., Ruffin, C., Worman, T., et al. (2008). Clinical Assessment of Music Perception in Cochlear Implant Listeners. Otology & Neurotology, 29(2), 149-155.

Niparko, J. K. (2009). Cochlear Implants: Principles & Practices. Philadelphia: Lippincott Williams & Wilkins.

Nobbe, A., Schleich, P., Zierhofer, C., & Nopp, P. (2007). Frequency discrimination with sequential or simultaneous stimulation in MED-EL cochlear implants. Acta Oto-laryngologica, 127(12), 1266-1272.

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(5), 1421-1425.

Patterson, R. D., Johnson-Davies, D., & Milroy, R. (1978). Amplitude-modulated noise: The detection of modulation versus the detection of modulation rate. The Journal of the Acoustical Society of America, 63(6), 1904-1911.

Peeters, I., Marquet, J., & Offeciers, E. (1987). The Laura cochlear prosthesis development description. In P. Banfai (Ed.), Cochlear Implant: Current situations. International Cochlear Implant Symposium. Duren, Germany.

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(1), 58-75.

Pfingst, B. E., Holloway, L. A., Poopat, N., Subramanya, A. R., Warren, M. F., & Zwolan, T. A. (1994). Effects of stimulus level on nonspectral frequency discrimination by human subjects. Hearing Research, 78(2), 197-209.

Pickles, J. O. (2008). An Introduction to the Physiology of Hearing (3 ed.). Bingley: Emerald.

Pijl, S. (1997). Labeling of Musical Interval Size by Cochlear Implant Patients and Normally Hearing Subjects. Ear & Hearing, 18(5), 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(1-2), 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. The Journal of the Acoustical Society of America, 98(2), 886-895.

Pitt, M. A. (1994). Perception of Pitch and Timbre by Musically Trained and Untrained Listeners. Journal of Experimental Psychology: Human Perception & Performance, 20(5), 976-986.

Plack, C., & Oxenham, A. (2005a). Overview: The Present and Future of Pitch. In C. Plack, R. Fay, A. Oxenham & A. Popper (Eds.), Pitch: Neural Coding and Perception (Vol. 24, pp. 1-6). New York: Springer.

Plack, C., & Oxenham, A. (2005b). The Psychophysics of Pitch. In C. Plack, R. Fay, A. Oxenham & A. Popper (Eds.), Pitch: Neural Coding and Perception (Vol. 24, pp. 7-55). New York: Springer.

Platt, J., & Racine, R. (1985). Effect of frequency, timbre, experience, and feedback on musical tuning skills. Attention, Perception, & Psychophysics, 38(6), 543-553.

Plomp, R. (1967). Pitch of Complex Tones. The Journal of the Acoustical Society of America, 41(6), 1526-1533.

Plomp, R. (1976). Aspects of tone sensation. London: Academic Press.

Pratt, R. L., & Doak, P. E. (1976). A subjective rating scale for timbre. Journal of Sound and Vibration, 45(3), 317 - 326.

Pretorius, L. L., & Hanekom, J. J. (2008). Free field frequency discrimination abilities of cochlear implant users. Hearing Research, 244(1–2), 77-84.

Ropshkow, O. (2010). Cochlea-crosssection [Online image]. Retrieved March 14, 2013 from http://en.wikipedia.org/wiki/File:Cochlea-crosssection.svg.

Rubinstein, J. T., Wilson, B. S., Finley, C. C., & Abbas, P. J. (1999). Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. Hearing Research, 127(1–2), 108-118.

Russo, F., & Thompson, W. (2005). An interval size illusion: The influence of timbre on the perceived size of melodic intervals. Attention, Perception, & Psychophysics, 67(4), 559-568.

Saoji, A. A., Litvak, L. M., & Hughes, M. L. (2009). Excitation Patterns of Simultaneous and Sequential Dual-Electrode Stimulation in Cochlear Implant Recipients. Ear & Hearing, 30(5), 559-567.

Schubert, E., & Wolfe, J. (2006). Does Timbral Brightness Scale with Frequency and Spectral Centroid? Acta Acustica united with Acustica, 92, 820-825.

Seligman, P., & McDermott, H. J. (1995). Architecture of the Spectra 22 speech processor. Annals of Otology, Rhinology & Laryngology, 104(suppl.166), 139-141.

Semal, C., & Demany, L. (1991). Dissociation of pitch from timbre in auditory short-term memory. The Journal of the Acoustical Society of America, 89(5), 2404-2410.

Shannon, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. Hearing Research, 11(2), 157-189.

Shannon, R. V. (1992). Temporal modulation transfer functions in patients with cochlear implants. The Journal of the Acoustical Society of America, 91(4), 2156-2164.

Shannon, R. V., Zeng, F.-G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech Recognition with Primarily Temporal Cues. Science, 270(5234), 303-304.

Shepherd, R. K., & Javel, E. (1997). Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status. Hearing Research, 108(1–2), 112-144.

Simmons, F. B., Epley, J. M., Lummis, R. C., Guttman, N., Frishkopf, L. S., Harmon, L. D., et al. (1965). Auditory Nerve: Electrical Stimulation in Man. Science, 148(3666), 104-106.

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 & Hearing, 30(2), 160-168

Skinner, M. W., Ketten, D. R., Holden, L. K., Harding, G. W., Smith, P. G., Gates, G. A., et al. (2002). CT-Derived Estimation of Cochlear Morphology and Electrode Array Position in Relation to Word Recognition in Nucleus-22 Recipients. Journal of the Association for Research in Otolaryngology, 3(3), 332-350.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. Nature, 416(6876), 87-90.

Snel-Bongers, J., Briaire, J. J., Vanpoucke, F. J., & Frijns, J. H. M. (2012). Spread of Excitation and Channel Interaction in Single- and Dual-Electrode Cochlear Implant Stimulation. Ear & Hearing, 33(3), 367-376.

Spahr, A. J., Litvak, L. M., Dorman, M. F., Bohanan, A. R., & Mishra, L. N. (2008). Simulating the Effects of Spread of Electric Excitation on Musical Tuning and Melody Identification With a Cochlear Implant. Journal of Speech, Language, and Hearing Research, 51(6), 1599-1606.

Steinhaus, H. (1950). Mathematical Snapshots. New York: Oxford University Press.

Stevens, S. S. (1937). On Hearing by Electrical Stimulation. The Journal of the Acoustical Society of America, 8(3), 191-195.

Sucher, C. M., & McDermott, H. J. (2007). Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. Hearing Research, 230(1–2), 80-87.

Swanson, B. (2008). Pitch Perception with Cochlear Implants. PhD Thesis, University of Melbourne, Melbourne.

Swanson, B., Dawson, P., & McDermott, H. J. (2009). Investigating cochlear implant place-pitch perception with the Modified Melodies test. Cochlear Implants International, 10(S1), 100-104.

Swanson, B., & McDermott, H. J. (2010). What cochlear implants can tell us about pitch perception in normal hearing. Paper presented at the 20th International Congress on Acoustics, Sydney.

Swanson, B., Van Baelen, E., Janssens, M., Goorevich, M., Nygard, T., & Van Herck, K. (2007). Cochlear Implant Signal Processing ICs. Paper presented at the IEEE Custom Integrated Circuits Conference (CICC), San Jose, California.

Terhardt, E. (1974). Pitch, consonance, and harmony. The Journal of the Acoustical Society of America, 55(5), 1061-1069.

Tong, Y. C., Blamey, P. J., Dowell, R. C., & Clark, G. M. (1983). Psychophysical studies evaluating the feasibility of a speech processing strategy for a multiple channel cochlear implant. The Journal of the Acoustical Society of America, 74(1), 73-80.

Tong, Y. C., & Clark, G. M. (1985). Absolute Identification of electric pulse rates and electrode positions by cochlear implant patients. The Journal of the Acoustical Society of America, 77(5), 1881-1888.

Tong, Y. C., Clark, G. M., Blamey, P. J., Busby, P. A., & Dowell, R. C. (1982). Psychophysical studies for two multiple-channel cochlear implant patients. The Journal of the Acoustical Society of America, 71(1), 153-160.

Tovey, D. F. (1975). The Forms of Music. Cleveland: Meridian Books.

Townshend, B., Cotter, N., Compernolle, D. V., & White, R. L. (1987). Pitch perception by cochlear implant subjects. The Journal of the Acoustical Society of America, 82(1), 106-115.

Undurraga, J. A., van Wieringen, A., Carlyon, R. P., Macherey, O., & Wouters, J. (2010). Polarity effects on neural responses of the electrically stimulated auditory nerve at different cochlear sites. Hearing Research, 269(1–2), 146-161.

van den Honert, C., & Kelsall, D. C. (2007). Focused intracochlear electric stimulation with phased array channels. The Journal of the Acoustical Society of America, 121(6), 3703-3716.

van den Honert, C., & Stypulkowski, P. H. (1984). Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings. Hearing Research, 14(3), 225-243.

van Hoesel, R. J. M., & Clark, G. M. (1997). Psychophysical studies with two binaural cochlear implant subjects. The Journal of the Acoustical Society of America, 102(1), 495-507.

van Wieringen, A., Macherey, O., Carlyon, R. P., Deeks, J. M., & Wouters, J. (2008). Alternative pulse shapes in electrical hearing. Hearing Research, 242(1–2), 154-163.

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., & McDermott, H. J. (2005). Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies. The Journal of the Acoustical Society of America, 117(5), 3126-3138.

Vandali, A. E., & van Hoesel, R. J. M. (2011). Development of a temporal fundamental frequency coding strategy for cochlear implants. The Journal of the Acoustical Society of America, 129(6), 4023-4036.

Vandali, A. E., Whitford, L. A., Plant, K. L., Clark, G. M. (2000). Speech Perception as a Function of Electrical Stimulation Rate: Using the Nucleus 24 Cochlear Implant System. Ear and Hearing, 21(6), 608-624.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. The Journal of the Acoustical Society of America, 66(5), 1364-1380.

Waltzman, S. B., & Roland, J. T. (2006). Cochlear Implants. New York: Thieme Medical Publishers.

Warren, R. M. (2008). Auditory Perception: An Analysis and Synthesis (3 ed.). Cambridge: Cambridge University Press.

Wightman, F. L. (1973). The pattern-transformation model of pitch. The Journal of the Acoustical Society of America, 54(2), 407-416.

Wilson, B., Zerbi, M., Finley, C., Lawson, D., & van den Honert, C. (1997). Speech processors for auditory protheses (Eighth Quaterly Progress Report). National Institutes of Health.

Wilson, B. S. (2004). Engineering Design of Cochlear Implants. In F.-G. Zeng, A. N. Popper & R. R. Fay (Eds.), Cochlear Implants: Auditory Prostheses and Electric Hearing (Vol. 20, pp. 14-52). New York: Springer.

Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: Current designs and future possibilities. Journal of Rehabilitation Research & Development, 45(5), 695-730.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. Nature, 352(6332), 236-238.

Wilson, B. S., Sun, X., Schatzer, R., & Wolford, R. D. (2004). Representation of fine structure or fine frequency information with cochlear implants. International Congress Series, 1273, 3-6.

Wilson, B. S., Wolford, R., Schatzer, R., Sun, X., & Lawson, D. T. (2003). Speech processors for auditory protheses (Seventh Quaterly Progress Report). National Institutes of Health.

Wilson, B. S., Zerbi, M., & Lawson, D. T. (1993). Speech processors for auditory protheses (Third Quaterly Progress Report). National Institutes of Health.

Wolfe, J. (2002). Speech and Music, Acoustic and Coding, and What Music Might Be 'For'. Paper presented at the International Conference on Music Perpcetion and Cognition, Sydney.

Won, J. H., Drennan, W. R., Nie, K., Jameyson, E. M., & Rubinstein, J. T. (2011). Acoustic temporal modulation detection and speech perception in cochlear implant listeners. The Journal of the Acoustical Society of America, 130(1), 376-388.

Zeitlin, L. R. (1964). Frequency Discrimination of Pure and Complex Tones. The Journal of the Acoustical Society of America, 36(5), 1027-1027.

Zeng, F.-G., Rebscher, S., Harrison, W., Xiaoan, S., & Haihong, F. (2008). Cochlear Implants: System Design, Integration, and Evaluation. Biomedical Engineering, IEEE Reviews in, 1, 115-142.

Zeng, F.-G. (2002). Temporal pitch in electric hearing. Hearing Research, 174(1-2), 101-106.

Zeng, F.-G. (2004). Trends in Cochlear Implants. Trends in Amplification, 8(1), 1-34.

Zeng, F.-G., & Shannon, R. (1995). Loudness of simple and complex stimuli in electric hearing. Annals of Otology, Rhinology & Laryngology, 104(Supp.166), 235-238.

Zeng, F.-G., & Shannon, R. V. (1992). Loudness balance between electric and acoustic stimulation. Hearing Research, 60(2), 231-235.

Zeng, F.-G., & Shannon, R. (1994). Loudness-coding mechanisms inferred from electric stimulation of the human auditory system. Science, 264(5158), 564-566.

Zhao, H.-B., & Liang, Z.-A. (1996). Processing of modulation frequency in the dorsal cochlear nucleus of the guinea pig: sinusoidal frequency-modulated tones. Hearing Research, 95(1–2), 120-134.

Zierhofer, C. (2002). Electrical nerve stimulation based on channel specific sampling sequences. MED-EL. US Patent 10/303568.

Zwolan, T. A. (2009). Cochlear Implants. In J. Katz, R. F. Burkard, L. Medwetsky & L. Hood (Eds.), Handbook of Clinical Audiology (5 ed.). Philadelphia: Lippincott Williams & Wilkins.

MACQUARIE UNIVERSITY

Research Office
Research Hub, Building C5C East
MACQUARIE UNIVERSITY NSW 2109

Phone **+61 (0)2 9850 8612**
Fax +61 (0)2 9850 4465
Email ro@vc.mq.edu.au

Ethics
Phone **+61 (0)2 9850 6848**
Email ethics.secretariat@vc.mq.edu.au

13 August 2009

Mr Vijay Marimuthu
59/192 Vimiera Road
Marsfield
NSW 2122

Reference: HE31JUL2009-D00040

Dear Mr Marimuthu,

FINAL APPROVAL

Title of project: Effect of stimulating the spread of Electric Excitation on pitch and music perception in CI

Thank you for your recent correspondence. Your response has addressed the issues raised by the Ethics Review Committee (Human Research) and you may now commence your research.

Please note the following standard requirements of approval:

1. The approval of this project is **conditional** upon your continuing compliance with the *National Statement on Ethical Conduct in Human Research (2007)*.

2. Approval will be for a period of five (5 years) subject to the provision of annual reports. **Your first progress report is due on 01 August 2010**

If you complete the work earlier than you had planned you must submit a Final Report as soon as the work is completed. If the project has been discontinued or not commenced for any reason, you are also required to submit a Final Report on the project.

Progress Reports and Final Reports are available at the following website:
http://www.research.mq.edu.au/researchers/ethics/human_ethics/forms

3. If the project has run for more than five (5) years you cannot renew approval for the project. You will need to complete and submit a Final Report and submit a new application for the project. (The five year limit on renewal of approvals allows the Committee to fully re-review research in an environment where legislation, guidelines and requirements are continually changing, for example, new child protection and privacy laws).

4. Please notify the Committee of any amendment to the project.

5. Please notify the Committee immediately in the event of any adverse effects on participants or of any unforeseen events that might affect continued ethical acceptability of the project.

6. At all times you are responsible for the ethical conduct of your research in accordance with the guidelines established by the University. This information is available at: http://www.research.mq.edu.au/policy

If you will be applying for or have applied for internal or external funding for the above project it is your responsibility to provide Macquarie University's Research Grants Officer with a copy of this letter as soon as possible. The Research Grants Officer will not inform external funding agencies that you have final approval for your project and funds will not be released until the Research Grants Officer has received a copy of this final approval letter.

Yours sincerely

Dr Karolyn White
Director of Research Ethics
Chair, Ethics Review Committee (Human Research)

Cc: Dr Robert Manell, Department of Linguistics

Address for all correspondence
RESEARCH DEVELOPMENT OFFICE
ROYAL PRINCE ALFRED HOSPITAL
CAMPERDOWN NSW 2050

TELEPHONE: (02) 9515 6764
FACSIMILE: (02) 9515 7176
EMAIL: geetha.govender@email.cs.nsw.gov.au
REFERENCE: X10-0281 & HREC/10/RPAH/497

NSW GOVERNMENT | Health
Sydney
Local Health District

10 October 2011

Mr V Marimuthu
Department of Linguistics
C5A Building, Audiology Section
Macquarie University
NORTH RYDE NSW 2109

Dear Mr Marimuthu,

Re: Protocol No X10-0281 & HREC/10/RPAH/497 - "Cochlear implant pitch perception and channel interaction"

I refer to the above research study which was approved by the Ethics Review Committee in October 2010 and the requirement by the National Health and Medical Research Council that research be monitored on at least an annual basis.

Enclosed is an Annual Progress Report form which you should complete appropriately and have signed. Following the introduction of the NSW Department of Health's *Model for Single Ethical and Scientific Review of Multi-centre Research*, the scope of your report will depend on the location(s) at which the study is being undertaken.

For single site studies within the RPAH Zone / multicentre studies which are only being conducted in NSW at an RPAH Zone site: Complete the report for the RPAH Zone site only.

For multicentre studies for which you are the Co-ordinating Investigator: Provide information on the study's progress at each site in NSW at which it is being conducted (please refer to the *Model*'s Standard Operating Procedure 024).

The report should be returned to the Research Development Office at the above address by **Wednesday 26 October 2011** for inclusion in the agenda of the next meeting.

Please note that failure to fulfil this annual reporting requirement may seriously jeopardise consideration of future submissions to the Ethics Review Committee.

Thank you for your assistance in this matter.

Yours sincerely,

Lesley Townsend
Executive Officer
Ethics Review Committee

encl.

HERC\PROG RPT\CORRES\11-11

General Correspondence
PO Box M30
Missenden Road, NSW, 2050
Email: slhn.esu@sswahs.nsw.gov.au
Website: www.health.nsw.gov.au/sydlhn/

Sydney Local Health District
ABN 17 520 269 052
Level 11 North, King George V Building
83 Missenden Rd
CAMPERDOWN, NSW, 2050
Tel 612 9515 9600 Fax 612 9515 9610