MUSIC PERCEPTION OF COCHLEAR IMPLANT USERS

VALERIE WEI LOOI

SUBMITTED IN TOTAL FULFILMENT OF THE REQUIREMENTS OF THE DEGREE

DOCTOR OF PHILOSOPHY

2006

DEPARTMENT OF OTOLARYNGOLOGY
THE UNIVERSITY OF MELBOURNE

Produced on archival quality paper

ABSTRACT

Objectives:

This study investigated the music perception skills of cochlear implant (CI) users, compared to hearing aid (HA) users who met the audiological criteria for a CI. Further, to eliminate some of the inter-subject variability that arises when making such between-group comparisons, a group of patients on the waiting list for a CI were tested prior to implantation whilst utilising their HA, and then again post-surgery, with the CI. It was hypothesised:

1) That experienced CI users (CI subject group) would score lower than HA users (HA subject group) on the pitch, instrument identification, and melody tests, but not the rhythm test;
2) That subjects on the waiting list for a CI (WL subject group) would score higher on the pitch, instrument identification, and melody tests when tested with their HA pre-implantation than post-surgery with their CI; and
3) That subjects utilising a HA (i.e., both the HA subject group and the WL subject group when tested with their HA pre-implantation) would rate music to sound more pleasant than the subjects utilising a CI (i.e., the CI subject group and the WL subject group when tested post-implantation).

Method:

Fifteen postlingually deafened adults utilising a Nucleus CI (i.e. the CI group) were compared to 15 postlingually deafened adults using a HA (i.e. the HA group); both of these subject groups had at least one year's experience with their respective devices. Of the CI subject group, 8 subjects used the CI24 device with the ACE speech-processing strategy, and 7 subjects used the CI22 device with the SPEAK speech-processing strategy. All of the HA subjects met the audiological criteria for a CI in terms of hearing thresholds and speech perception scores. Further, 9 subjects on the waiting list (WL) for a cochlear implant were also tested pre-surgery with their HAs, and subsequently 3

months post switch-on of their CI24 device, implemented with the ACE speech-processing strategy.

A series of music tests were developed for this research incorporating four major tasks: (i) discrimination of 38 pairs of rhythms; (ii) pitch ranking of sung vowels, one-octave, half-octave, and a quarter-octave apart; (iii) instrument recognition and appraisal involving three subtests, each comprising 12 different instruments or ensembles; and (iv) recognition of 10 familiar melodies where the pitch and rhythm cues were preserved. The tests were initially verified with a group of normally hearing subjects. Stimuli were presented at comfortable presentation levels either via direct audio input or through a neck loop system activated by the telecoil on a subject's HA. The test battery was administered to each subject on two separate occasions.

Results:

The results of the assessments partially supported the first two hypotheses, but not the third. For the first hypothesis, as expected there was no significant difference between the CI and HA subjects on the rhythm test. The CI group scored significantly lower on the pitch and melody tests ($p < 0.001$ for both comparisons), but equivalent to the HA subjects on the instrument recognition tests. The second hypothesis only held true for the one-octave and quarter-octave subtests of the pitch task ($p = 0.007$, and $p < 0.001$, respectively), with lower pitch-ranking scores obtained post-surgery with the CI than pre-surgery. There were no significant differences between the pre- and post-surgery test scores for the rhythm, instrument identification, or melody tests. The third hypothesis was not supported by the findings of this research with the subjects utilising a CI rating the music stimuli to sound more pleasant than the subjects utilising a HA (WL subjects: $p = 0.005$ for subtest 2, and $p = 0.009$ for subtest 3).

Conclusions:

Regardless of their experience with the device, CI users at best only scored equivalently to, and for the pitch tasks, significantly worse than, HA users. However, the HA users did not score as well as what may be expected with normally hearing listeners. These findings suggest that despite the two devices using contrasting modes of auditory stimulation, and providing different perceptual cues for the listener, neither the CI nor the HA enabled the subjects in this study to achieve satisfactory or effective music perception.

DECLARATION and PREFACE

This is to certify that:

i) The thesis comprises only my original work towards the PhD, except where indicated in the Preface;

ii) Due acknowledgement has been made in the text to all other material used;

iii) The thesis is less than 100 000 words in length, exclusive of tables, maps, bibliographies and appendices.

The software program, *MACarena*, utilised in this study was written by Dr Wai Kong Lai, from the Department of Otorhinolaryngology Head and Neck Surgery, University Hospital, Zurich, Switzerland. The questionnaires designed for this research were based around the *Iowa Musical Background Questionnaire*, developed by Professor Kate Gfeller and colleagues, at Iowa University, Iowa, USA. Permission was obtained for the use of these resources. Assistance with statistical analyses and advice was obtained from Associate Professor Ian Gordon from the Melbourne Statistical Consulting Centre, along with Dr Ross Darnell from the University of Queensland. Numerous of the figures included in the thesis were developed by colleagues in the Department – namely, Dr Peter Seligman, Mr David Tsang, Mr Brett Swanson, as well as from promotional material published by Cochlear Limited. Ethical approval for this study was obtained from the RVEEH's Human Research Ethics Committee. Finally, mention must be made of my three supervisors – Associate Professor Hugh McDermott, Professor Colette McKay, and Associate Professor Louise Hickson, for advice during the course of the study with respect to experimental design, data collection methodology, data analyses, and overall interpretation of the results.

The following is a list of publications that have arisen from the work undertaken as part of this thesis.

Looi, V., McDermott H. J., McKay, C. M., & Hickson, L. M. (2006). Pre-to-post cochlear implantation comparisons for music. *Proceedings of the 9th International Conference on Cochlear Implants and Related Sciences*, Vienna, Austria. Wiener Medizinische Wochenschrift, Vol 156 (Supplementum 119), 104.

Looi, V., McDermott H. J., McKay, C. M., & Hickson, L. M. (2005). Music perception of cochlear implant users compared to hearing aid users. *Abstracts of the 2005 Conference on Implantable Auditory Prostheses,* Asilomar, Pacific Grove, California, USA. July 30-Aug 4 2005.

Looi, V., McDermott H. J., McKay, C. M., & Hickson, L. M. (2004). Pitch discrimination and melody recognition by cochlear implant users. *Proceedings of the VIII International Cochlear Implant Conference,* Indianapolis, Indiana USA, Elsevier. Vol 1273C, 197-200.

Looi, V., Sucher, C, McDermott, H. J. (2003). Melodies familiar to the Australian population across a range of hearing abilities. *The Australian and New Zealand Journal of Audiology*, 25(2), 75-83.

Looi, V. & McDermott, H. J. (2002). Music perception of cochlear implant users. *Abstracts of The Garnett Passe and Rodney Williams Memorial Foundation Conference, Frontiers in Otorhinolaryngology*, Noosa Heads, July 31-August 2, 2002.

McDermott, H. J. & Looi, V. (2004). Perception of complex signals, including musical sounds, with cochlear implants. *Proceedings of the VIII International Cochlear Implant Conference*, Indianapolis, Indiana USA, Elsevier. Vol 1273C, 201-204.

Looi, V., McDermott H. J., McKay, C. M., & Hickson, L. M. (accepted). Comparisons of quality ratings for music by cochlear implant and hearing aid users. *Ear and Hearing.*

Valerie Wei Looi

August 2006

The University of Melbourne, Australia.

ACKNOWLEDGEMENTS

There have been a host of people who have contributed to this undertaking in one way or other. To anyone who has been a part of this enriching experience, regardless of how small or large your contribution, I thank you. There are, however, numerous who warrant special mention.

Firstly, and foremostly, I am indebted to my main supervisor, Associate Professor Hugh McDermott, as well as my associate supervisors, Professor Colette McKay, and Associate Professor Louise Hickson. A sincere and profound thank you to Hugh for everything - your teaching, explanations, advice, guidance, suggestions, comments, support, time, and effort, to name a few. I do not have an engineering bone in my body, and your patience in explaining, and re-explaining the technicalities associated with this research until it 'sunk in' was not only greatly appreciated, but unequivocally necessary. I will be first in line to purchase that "Acoustic Engineering For Dummies" book I've suggested on several occasions! To Colette, my sincere thanks for your guidance, explanations, advice, suggestions, and time – particularly with respect to the perceptual and psychophysical side of this study. To say that I have learnt a lot would be more than an understatement. To Louise, my sincere thanks also, not just for your invaluable academic contribution and advice, but also for the more personal support, and guidance. Thank you for being a sounding-board for my concerns, and for your constant reassurance that things were on the right track, and more importantly, that doing this PhD was the right thing. I also want to particularly thank you for your time and effort – I realise that as you aren't associated with the University of Melbourne, your involvement with this PhD has been entirely out of your own time, and generosity, for which I am most grateful and appreciative.

This study was primarily funded by the National Health and Medical Research Council of Australia, via a Dora Lush Biomedical Postgraduate Research Scholarship (Grant ID: 263814). A mention must be made of Kate Drummond, who was of great assistance with the initial scholarship application, as well as to Loretta Gibson, Kimberley Nunes, and the Melbourne Scholarships Office, for managing and administering the budget associated with this scholarship. This study was also funded by the Co-operative Research Centre (CRC) for Cochlear Implant and Hearing Aid Innovation. To the CRC and particularly Dr Robert Cowan, the Director of the CRC, thank you for your continued financial support, and assistance over the course of the PhD. Your assistance with travel expenses, conference costs, and other equipment and research funding have made the PhD path a lot easier, less stressful, and more rewarding for me. Thank you also to Monica Chapman, the Executive Officer of the CRC. Funding for this study was also indirectly provided by the Garnett Passe and Rodney Williams Memorial Foundation, whom provide the funding for my principal supervisor. This PhD was undertaken at the Department of Otolaryngology at the University of Melbourne, and my thanks go to Professor Richard Dowell, the head of the department, along with the staff of the Department, and in particular - Paul Quilty, the librarian, and Stas Suriowiecki and Martin Wojak for the (all-too-frequent!) IT support. My thanks must also be extended to the University of Queensland. With this research being conducted over two states, the generous logistical and practical support I received was greatly appreciated.

To the clinicians and staff of the Cochlear Implant Clinic at the Royal Victorian Eye and Ear Hospital, Queensland Hearing, and Australian Hearing – thank you for your invaluable assistance with both recruiting subjects, and clinical support. Particular thanks must go to Rod Hollow, Karen Pedley, Andrea Simpson, Catherine Sucher, Sylvia Tari, Jennifer Bergman, Sally Smith, and Greg Angus. Thank you also to Cochlear Limited, and in particular Dr Peter Seligman, Nik Kydas, and John Heasman for equipment and technical assistance. I also must acknowledge the invaluable statistical assistance and advice I received from Associate Professor Ian Gordon of the Melbourne Statistical Consulting Centre, as well as Dr Ross Darnell, from the University of Queensland. Along with not having an engineering background, my statistical knowledge was sadly lacking – so your advice, explanations, and suggestions were much needed, and most appreciated. I learnt a lot! A big thank you also goes to Debbie Emerson from the Percy Baxter Collaborative Learning Centre at the University of Melbourne, for her invaluable extended assistance with the intricacies of master document and Microsoft Word formatting styles. Thank you for being so approachable and happy to help.

To my colleagues, and more correctly, friends, in Ardee, and at the Bionic Ear Institute – thank you. You have all contributed in some way: to Andrea Simpson and Cathy Sucher – thank you for your friendship, support, practical help, and subject recruitment assistance. Good luck to both of you with your current studies. To Adam Hershbach, and Justin Zakis, thank you for the technical expertise. To Cila, Carrie, Andy, Luke, and Mansze, thank you for your friendship. To my colleagues at the Audiology Department of the Royal Children's Hospital in Brisbane, and particularly Lisa Shannon, and Lisa Rainbow – my sincere gratitude for your flexibility in accommodating my 'here one week, gone the next' needs. I have appreciated, and thoroughly enjoyed the chance to continually keep up the clinical work throughout the PhD. Recognition and thanks must also be extended to Mike Robb and the Department of Communication Disorders at the University of Canterbury, Christchurch, New Zealand whom have been most generous in allowing me the time and resources to finish writing the thesis whilst commencing a new job. A separate thank you is also extended to many of you who helped proofread parts of the thesis.

On a more personal note, to my family – my parents and my sister, thank you from the bottom of my heart for your support. All of the little things (and big things!) that you did to make my life as easy, convenient, and less stressed as you could, were (and still are!) most appreciated. I couldn't have done it without your support and help, and I love you very much. To my friends, particularly Heather, as well as Melissa K, Jenny, Juliet, and Melissa F – thank you for keeping me sane (OK, vaguely resembling sane), and for keeping in touch even though I've been interstate. To my BJJ training club, thanks for the opportunity for frustration release!

Finally, but by no means least, I wish to express my sincere thanks and gratitude to all of my research subjects. I truly appreciate your time, effort, concentration, and interest. I definitely could not have done this without you, and hopefully one day that piece of music you know won't sound like it's being played by tone-deaf 'musicians'!

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

LIST OF ABBREVIATIONS

The following is a list of some of the major abbreviations used throughout this thesis. The abbreviations in *italics* indicate those relating to the name of a specific speech processing strategy.

ACE	*Advanced Combination Encoder*	**Hz**	Hertz
ADRO	Adaptive Dynamic Range Optimisation	**ICRA**	International Collegium of Rehabilitative Audiology
ANOVA	Analysis of Variance	**i.e.**	That is
BP	Bipolar stimulation	**kHz**	Kilohertz
CA	*Compressed Analogue*	**Ltd.**	Limited
CG	Common ground stimulation	**MLEQ**	Music Listening and Enjoyment Questionnaire
CI	Cochlear implant	**mm**	Millimetre
CIS	*Continuous Interleaved Sampling*	**MP**	Monopolar stimulation
cm	Centimetre	***MPEAK***	*Multipeak Strategy*
CUNY	City University of New York (sentence test)	**ms**	Millisecond
DAI	Direct audio input	**MTEQ**	Music Training and Experience Questionnaire
dB	Decibels	**NH**	Normal hearing
dBHL	Decibels hearing level	**PMMA**	Primary Measures of Music Audiation (test)
dBSPL	Decibels sound pressure level	**p./pp.**	Page/pages
e.g.	For example	**pps**	Pulses per second
F0	Fundamental frequency	**RMS**	Root-mean-squared
F1	1st Formant	***SAS***	*Simultaneous Analogue Stimulation*
F2	2nd Formant	**SD**	Standard deviation
FDA	Food and Drug Administration	***SMSP***	*Spectral Maxima Sound Processor*
Fig.	Figure	***SPEAK***	*Spectral Peak Strategy*
FM	Frequency modulated	**WL**	Waiting list
HA	Hearing aid	**yr**	Year

CHAPTER 1: INTRODUCTION

"The principal objective in the development of cochlear implants is to restore useful hearing to profoundly deaf subjects…To accomplish this, basic discoveries have been and continue to be necessary." (Fravel, 1986, p. xi). Developments in the field of cochlear implants have expanded at an expeditious rate, particularly in the last two decades. Current devices, technological innovations, and patient performances would have exceeded the initial expectations of Djourno & Eyries (1957) when they experimented with implanting a single electrode onto the eighth nerve of a patient undergoing facial nerve surgery (Eisen, 2003). Since then, single-channel implants have been superseded by the multi-channel cochlear implant (CI), along with a host of constantly updated speech processors and processing strategies. Current CIs offer the patient the potential of improved speech recognition with and without visual cues, the ability to use the telephone, a better quality of life, and an introduction to a new world of sounds. More cosmetically appealing ear-level speech processors now supplement the body-level processors, and a multitude of devices, sound processing options, hardware choices, and supplementary accessories are available to both the clinician and the patient.

Along with the new technology, improved results, increased marketing, and a larger patient population comes the inevitable increase in patients' expectations and aspirations. Despite the ongoing research and promotions by manufacturers regarding 'the latest' functions, additional benefits, and the superiority of their device(s) over others, ultimately it is the user who determines the true benefit, effectiveness, and thus success of the product. As many recipients can now achieve highly satisfactory speech recognition performance, they are looking toward improved perception of other aural stimuli such as musical and environmental sounds; hence the more recent expansion of research interests into non-speech stimuli. This study investigated the perception of music with the Nucleus CI via two methods. Firstly, experienced subjects utilising a CI were compared to hearing aid (HA) users who met the audiological criteria for a CI, on music tests assessing rhythm, pitch, instrument (timbre), and melody perception. The incorporation of CI and HA users with more-equitable unaided hearing thresholds allowed the assessment of electric versus acoustic stimulation for music perception, whilst accounting for some of the physiological differences present when comparing a normally hearing ear to one with a significant cochlear hearing loss. Secondly, in order to eliminate some of the inter-subject variability that arises when making such between-group comparisons, a group of patients on the waiting list (WL) for a CI were tested prior to implantation whilst utilising their HA, and then again post-surgery, with the CI. These two methods allowed the assessment of:

1) The music perception skills of CI users when compared to HA users with a moderately-severe to profound bilateral hearing loss;
2) The effect of cochlear implantation on an individual's music perception skills; and
3) Other subject factors that may affect a CI or HA user's perception of music.

Chapter 2 provides an overview of the fundamentals of sound, hearing, hearing loss, and music stimuli as applicable to this study. Once these fundamentals relating to sound and its perception through acoustic stimulation are outlined, the thesis progresses onto considering the other form of hearing stimulation relevant to this thesis – electrical stimulation via the CI, in Chapter 3. A comparison between HAs and CIs will also be made in this chapter. The information in these two chapters is fundamental to the review and interpretation of the literature directly related to this research, as presented in Chapter 4. Whereas the latter part of Chapter 3 concentrates on the more theoretically focused psychoacoustic research, Chapter 4 presents literature investigating the performance of CI users on general musical tests, involving musical stimuli presented to subjects utilising their speech processors in more realistic listening situations. For example, whilst Chapter 3 discusses the pitch cues used by CI users for tightly controlled stimuli presented to a single electrode, Chapter 4 focuses on the ability of CI users to undertake pitch-ranking tasks whilst listening with their speech processor to stimuli presented via a loudspeaker. This latter type of research is more comparable to the current study. A range of studies are reviewed in order to provide an understanding of both the general consensus with respect to CI users' performance on music listening tasks, as well as to establish 'the bigger picture' with respect to the current status of research in this area. This enables the reader to put this study into context, and ascertain how it addresses some currently unanswered questions. The literature review also assists in the later interpretation of the results and findings.

As Chapter 4 will verify, it has been established that there is no significant difference, on average, between normally hearing subjects and CI subjects on measures of rhythmic perception. However, the performance of CI subjects on frequency-based tasks including instrument identification, melody recognition, and higher/lower pitch discrimination, is significantly compromised. Existing studies have not shown reliable or definitive correlations between overall music perception skills and a range of subject variables such as implant type, age, speech processing strategy, length of time with the implant, or level of music experience. In comparison to the recognition of speech, research into the ability of CI users to perceive music, as well as the factors impacting upon their perception of such sounds, is not as well understood, with many areas as yet unaddressed. Research involving current day speech processing strategies, or comparing CI to HA users, is limited, and there has been no research comparing the same subjects pre- to post-implant surgery on music tasks.

In cognisance of the existing literature, the following hypotheses were derived:

1) That experienced CI users (CI subject group) would score lower than HA users (HA subject group) on the pitch, instrument identification, and melody tests, but not the rhythm test;
2) That subjects on the waiting list for a CI (WL subject group) would score higher on the pitch, instrument identification, and melody tests when tested with their HA pre-implantation than post-surgery with their CI; and
3) That subjects utilising a HA (i.e., both the HA subject group and the WL subject group when tested with their HA pre-implantation) would rate music to sound more pleasant than the subjects utilising a CI (i.e., the CI subject group and the WL subject group when tested post-implantation).

These are described more comprehensively in Chapter 5 along with the rationale and justification underlying this research.

To test these hypotheses, as well as to investigate other factors that may affect a CI or HA user's perception of music, a music test battery was developed for administration with three separate subject groups. The test stimuli for the battery were largely recorded and compiled by the researcher, as described in Chapter 6, with the music tests being subsequently verified with a group of normally hearing subjects. Two case-history questionnaires developed by the researcher for the purpose of this study are also described in Chapter 6. These materials were administered to three different research subject groups as described in Chapter 7. In the first part of the research, the performance of the CI and HA subject groups on the music test battery was compared. The subjects for the HA group were required to meet the current cochlear implantation criteria in terms of hearing levels and speech perception scores. In the second part of the research, patients on the waiting list for a CI (WL subject group) were tested prior to their operation whilst utilising HAs, and then again approximately 3 months post-implantation. The results of the CI and HA subject groups from part one of the study also served as control-group comparisons for the WL subject group's testing in part two. The music test battery was administered to the two former subject groups on two occasions, approximately 4 months apart, to account for the potential of learning effects leading to misleading comparisons and conclusions.

Due to the quantity of data, and multiple cross-group comparisons required for this research, the results of the study are presented over two chapters. The raw data for each subject group are presented in Chapter 8, with Chapter 9 providing the analysis of data via graphical and statistical comparisons within, and between, the subject groups. Relevant correlations are also calculated in this chapter. A discussion of the pertinent findings arising from these analyses, comparisons, and correlations is presented in Chapter 10. The thesis is culminated in Chapter 11 by summarising the findings of both this current study, as well as current and ongoing research developments related to this study's findings. It is hoped that the results of this study will help to increase the understanding of factors impacting upon the perception of musical stimuli by CI users, and in the longer term, to assist in improving the quality of music listening experiences for this population.

CHAPTER 2: OVERVIEW OF SOUND, HEARING & MUSIC

This chapter serves to provide an overview of the fundamental concepts relevant to this study – the nature and perception of sound in general, as well as music. It commences with an introduction to the physical components, production, and propagation of sound (section 2.1), followed by an overview of the hearing mechanism and hearing loss in section 2.2. The study of perception, integration, and interpretation of the elements comprising a sound is psychoacoustics; the ensuing section (section 2.3) focuses upon three major psychoacoustic percepts relevant to this study – loudness (section 2.3.1), pitch (section 2.3.2), and timbre (section 2.3.3). As the perception of these elements varies between a person with normal hearing and one with a cochlear hearing loss, section 2.3.4 highlights some of the potential consequences for a person with a hearing loss. Pitch, timbre, and loudness are fundamental percepts of the main stimuli used in this research – music. Consequently, the final parts of this chapter (section 2.4) delineate music per se, and its perception, as well as making a rudimentary comparison between music and speech signals (section 2.5).

2.1 SOUND

The creation of acoustic musical sound can be divided into three key steps. Firstly, an external energy source such as plucking, hitting, blowing, or bowing, creates a vibration which is transferred to a sound body, such as a sound board. The vibration of this body, in turn, creates air pressure variations (i.e. sound waves) which are then propagated to the listener via condensation and rarefaction of the air particles. This results in alternating areas of high and low pressure where the air molecules are close together and then far apart, successively. The ear detects, amplifies and transforms these pressure variations into signals interpretable by the brain. The intensity of sounds depends upon the degree of compression of the air molecules with increased pressure creating greater displacement of the molecules. The resulting sound can be modified by changes in one or all of these steps. For example, varying the force or method used to create the vibration will change the amplitude of the harmonics (Glossary), whilst modifying the shape, material, or size of the sound body will change the resulting vibration pattern. The propagation of the sound wave can also be affected through muting or modifying the room's acoustics (Handel, 1989). It is worth mentioning at this point that music can also be created by electronic means (electro-acoustic music), for example via the use of a Musical Instrument Digital Interface (MIDI). It is beyond the scope of this overview to discuss electro-acoustic music, however once created, electro-acoustic music is propagated to the listener in the same manner as for acoustical music sounds.

Harmonic motion, the simplest form of vibration, is created by exerting a force onto an elastic medium that is subsequently moved from equilibrium to a point of maximum displacement in one direction, through equilibrium to the point of maximal displacement in the contrary direction before returning to equilibrium. The restoring force is proportional to the force of displacement. Such vibration results in the simplest type of soundwave, a sinusoid, the description of which encompasses three features – the frequency (in Hertz – Hz), the amplitude, and the phase. The former specifies the number of repetitions (or cycles) per second and is dependent on the mass and stiffness of the system – for example, a stiffer vibrating source will result in a higher frequency sound, whereas a larger mass will decrease the frequency. 'Amplitude' details the size of the pressure variation, and is often measured in decibels (dB), with 'phase' referring to the stage of the cycle a waveform is at for a precise moment in time. Most of the sounds encountered in everyday life comprise complex tones or waveforms whereby two or more simple tones or sinusoids are combined. The most basic of these are periodic in nature, having a regular repetition rate corresponding to the frequency of the fundamental - the fundamental frequency (F0) (Glossary). For periodic complex sounds, the frequencies of the higher components, or harmonics, are multiples of the F0. Any complex wave can be analysed via 'Fourier Analysis' into a series of sinusoids, each with its own frequency, amplitude, and phase (Handel, 1989; Moore, 2003b). Durrant & Lovrinic (1995) provide a good overview of the physical parameters associated with sound per se, along with definitions of the relevant terminology.

2.2 HEARING AND HEARING LOSS

The ability of a person to detect and interpret sounds varies considerably, partially pertaining to the individual's auditory system. The following is a rudimentary outline of the mechanisms, processes, and variables related to hearing, aiming to give an overview of the basic issues fundamental to this study. There are a host of resources which provide more detailed information on this topic (e.g., Durrant & Lovrinic, 1995; Gelfand, 1998).

The human peripheral auditory system (Figure 2.1) can be subdivided into three sections – the outer, middle, and inner ears. The former two parts act as conductive mechanisms, detecting and transforming sound energy into mechanical vibrations for conduction to the inner ear. The inner ear contains the sensory organ of hearing, serving to convert these vibrations into neural impulses transmitted to the brain. An impairment in any of these areas can result in a hearing loss. A conductive hearing loss (Glossary) refers to a problem in the outer or middle ear affecting the conduction of sound through to the inner ear. A sensorineural hearing loss (Glossary) involves an impairment to one or more components in the inner ear and/or a malfunction of higher-order hearing mechanisms. Of most relevance to this study is a sensorineural hearing loss arising from damage to the hair cells in the cochlea which hinders the transformation of mechanical vibrations into neural impulses. The extent of damage and the resultant hearing impairment can be classified based on a person's hearing thresholds (Figure 2.2).

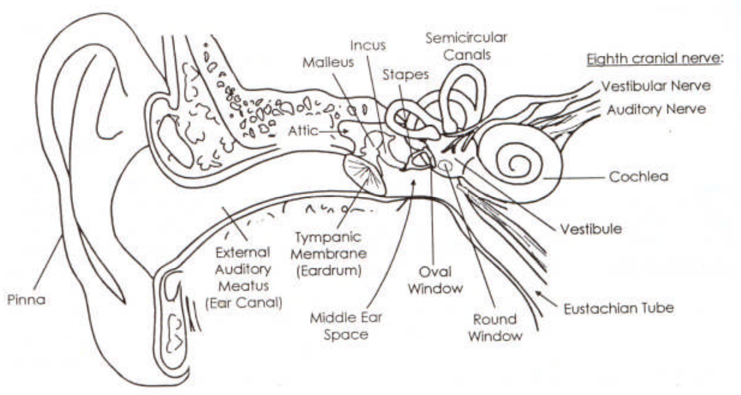


Figure 2.1: The human peripheral auditory system
Outer Ear: pinna, external auditory meatus, tympanic membrane. Middle Ear: malleus, incus, stapes. Inner Ear: cochlea. (Image from: Gelfand (1998))

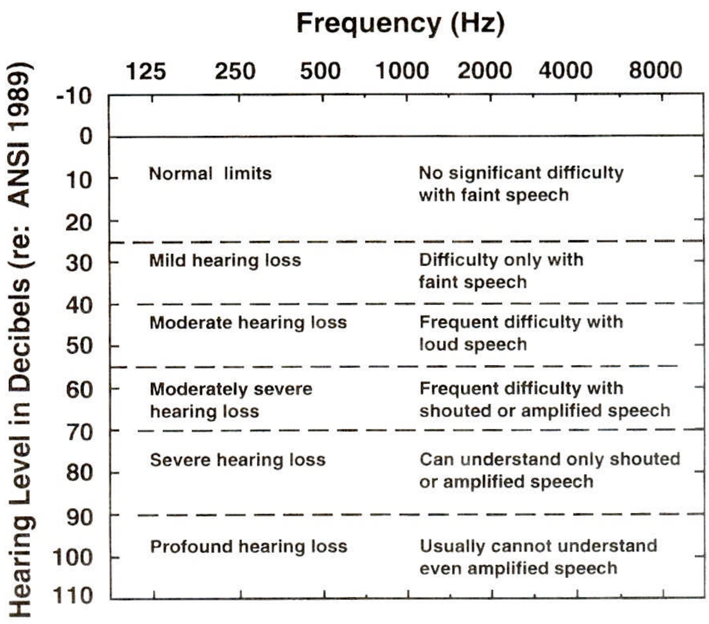


Figure 2.2: Classification of hearing loss
Vertical axis: hearing thresholds in dBHL. Horizontal axis: puretone frequency.
(Image from: Bess & Humes (1995))

The coil-shaped cochlea has 2¾ turns around a central bony canal, the modiolus. The space within each coil is divided into three fluid-filled channels – the scala vestibuli, scala media, and scala tympani (Figure 2.3). The scala media and scala vestibuli are separated by Reissner's membrane, with the basilar membrane separating the scala media and the scala tympani. The hearing organ itself, the organ of Corti, lies on the basilar membrane, serving to convert the mechanical vibrations into neural signals imparted to the brain. The auditory nerve fibres of normally hearing individuals demonstrate stochastic firing properties; that is, their firing patterns are not highly synchronised to the acoustic input. The sensory cells, the hair cells, are mechanoreceptors; they detect the vibrations through the shearing of the stereocilia, and transduce these into receptor potentials for the brain. Up/down movement of the basilar membrane leads to the back/forth shearing displacement of hair cells either via the overlying tectorial membrane, or through the flow of fluid in the space under this membrane (Figure 2.4).

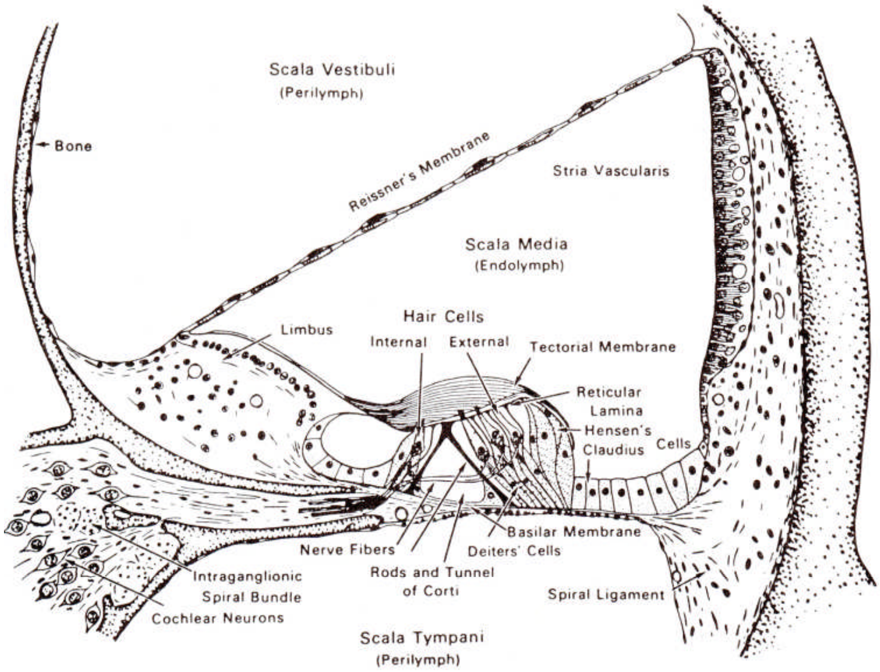


Figure 2.3: Cross-section of the cochlea
(Image from: Gelfand (1998))

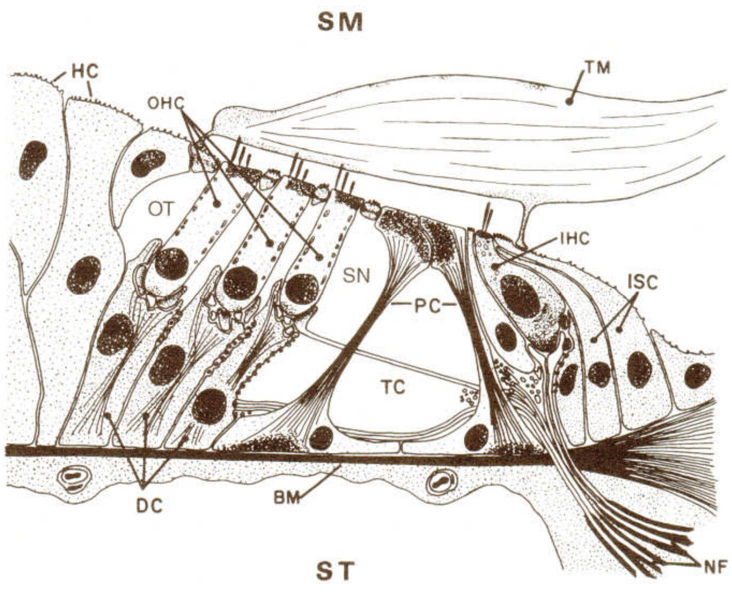


Figure 2.4: The Organ of Corti
BM: basilar membrane; DC: Deiter's cells; HC: Hensen's cells; ISC: inner supporting cells; NF: nerve fibers; SM: scala media; ST: scala tympani; TC: tunnel of corti; TM: tectorial membrane; IHC: inner hair cell; OHC: outer hair cell.
(Image from: Durrant & Lovrinic (1995))

The basilar membrane is stiff, thick, and narrow at the basal end, progressively graduating to be less stiff, thinner, and wider at the apex (Figure 2.5). This variation gives the cochlea its tonotopic organisation (Glossary). In response to an input sound wave, a travelling wave is formed in the cochlear fluid, progressing from the base to the apex of the cochlea. Different frequencies are associated with differing points of maximal displacement along the basilar membrane with low frequencies stimulating apically and higher frequencies exciting more basally. In essence, each frequency creates its own travelling wave which is spectrally analysed by the cochlea – the location of the wave's peak becomes one of the primary bases from which pitch decisions are made by the brain (place pitch - Glossary). The range of human hearing is reported to be approximately 15 Hz to 15 kHz, although we are less sensitive to sounds at extreme frequencies (Moore, 2003b).

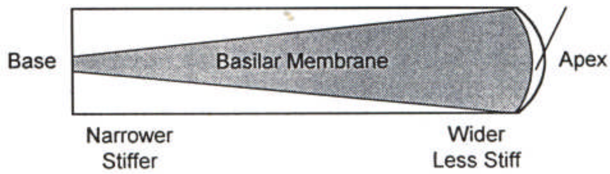


Figure 2.5: Schematic representation of an unwound basilar membrane
(Image adapted from: Gelfand (1998))

2.3 PSYCHOACOUSTICS

The study of psychological responses to a physical stimulus is the primary focus of psychophysics experiments. Psychoacoustics is a branch of psychophysics related to sound and hearing, involving the quantification of relationships between the acoustic stimulus and the psychological response. This field is a widely encompassing area incorporating many topics, each of which have substantial research debate surrounding them. It is beyond the scope of this thesis to cover this field in any depth, and as such, the discussion below serves only to provide an overview of the physical stimulus and psychophysical correlates for the three most pertinent considerations arising in music perception - loudness, pitch, and timbre.

2.3.1 Loudness

Whereas intensity is the physical parameter of a stimulus, loudness is its psychophysical correlate. It is the perceptual response to a sound's amplitude, ordered on a scale from quiet to loud (Clark, 2003). Other than intensity, factors such as duration may also affect loudness, although to a lesser extent (Gelfand, 1998; Lieberman & Blumstein, 1988). The psychophysical scale of loudness is measured by the sone, where one sone refers to the loudness of a 1 kHz puretone at 40 dBSPL; two sones is the loudness of a tone perceived to be twice as loud as this reference. The relationship between sones and intensity is affected by critical bands or auditory filters. The detail of this is beyond the scope of this thesis, with Gelfand (1998) and Moore (2003b) providing more information. Essentially, the relationship between loudness and intensity differs depending upon whether the individual components comprising the complex sound stimulus fall within one critical band, or if the energy is spread over greater than one band. The loudness perceived is greater for the latter scenario, due to greater loudness summation.

Whereas the sone is the psychophysical unit of loudness, loudness level is measured via the phon; equal phons implies equal loudness perception, although not necessarily equal physical magnitude. As 40 phons refers to the loudness level of any sound of equal loudness to a 1000 Hz tone presented at 40 dB, it could also be said that one sone is the loudness that equates to the 40 phon loudness level. To obtain equal loudness at the lower frequencies, more intensity is required than that for the higher frequencies. However, loudness levels at the lower frequencies grow at a faster rate than for the higher frequencies. Humans are most sensitive to sounds between 1 kHz and 5 kHz, with rapid sensitivity deteriorations at the frequency extremes (Moore, 2003b). These levels can be plotted relative to a fixed level for a 1000 Hz tone to form Equal Loudness Level Contours, as shown in Figure 2.6. The contour tends to flatten out for higher sound levels indicating compression in the auditory system's dynamic range, particularly evident at the lower frequency limits of hearing. In sound level meter measurements, the frequently encountered A-weighting, dB(A), is based on the 40 phon equal loudness contour, thereby reducing the low frequency contribution to the final meter reading.

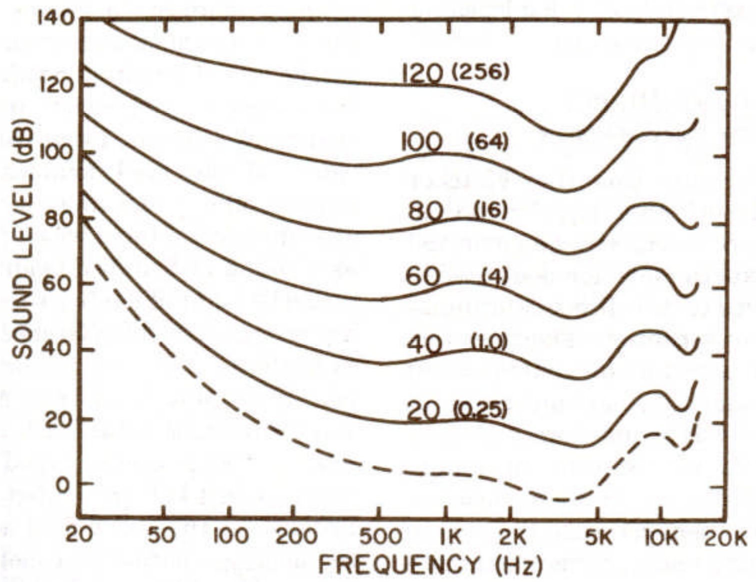


Figure 2.6: Equal Loudness Level Contours
(Image from: Durrant & Lovrinic (1995))

2.3.2 Pitch

"Pitch is related to the repetition rate of the waveform of a sound; for a pure tone this corresponds to the frequency and for a periodic complex tone to the fundamental frequency." (Moore, 2003b, p. 195). In psychoacoustic terms, pitch usually refers to the high to low ordering of sounds on a melodic scale, with systematic variations in pitch providing a sense of melody (Hartmann, 1996; Moore, 2003b). The coding of frequency is the primary, but not sole, determinant of pitch perception with features such as duration and intensity also having an effect.

The unit of pitch is the mel with 1000 mels equating to the pitch of a 1 kHz tone at 40 phons. A sound twice as high would have a pitch of 2000 mels. However, there is a disparity between frequency and pitch – the difference between 100 mels and 200 mels is not the same as the perceived change between 100 Hz and 200 Hz. The pitch of a complex tone is affected by the F0 along with its harmonics, particularly those falling in the range from 500 Hz to 2000 Hz. For some complex tones, energy at the F0 is not required for the pitch of the F0 to be perceived - the pitch of tones with a F0 less than approximately 500 Hz is determined by the harmonics lying between 500 Hz and 2000

Hz (Rasch & Plomp, 1982). The role of harmonics is crucial to pitch perception – if two sounds are played with the same harmonics but the component at the F0 is omitted from one, the pitches will sound the same (i.e. related to the F0), and only their quality (i.e. timbre) will differ (Sekuler & Blake, 1994).

For the normally hearing listener, pitch perception for puretones is associated with both the peak of the travelling wave on the basilar membrane, and the temporal pattern of neural firing (place pitch, and periodicity pitch, respectively - Glossary). Pitch perception is usually best between 20 Hz and 5000 Hz; higher frequencies tend to lack a definite pitch with lower frequencies described as having a rattling character (Rasch & Plomp, 1982). For the acoustic hearing of complex sounds, the perceptual process is rather more involved than that for puretone stimuli. A host of models have been suggested as possible means by which the normal auditory system determines the pitch of complex tones. Although no one particular model can account for all of the phenomena or anomalies associated with pitch perception, the majority of the currently preferred theories can be divided into two classes – pattern-recognition models and temporal-based models. Although it is beyond the scope of this thesis to detail these models, a simplistic explanation of the two classes is provided below to enable comparison of pitch perception processes for normal acoustic hearing to that for electrically stimulated hearing associated with cochlear implant use (as discussed in Chapter 3, section 3.4).

Pattern-recognition models are largely based on the perception of spectral cues whereby the frequencies of the individual components in the signal are used to determine the pitch of the complex tone (Moore, 2003b). These models comprise two main stages – a frequency analysis of the input signal followed by the recognition of the resulting pattern via a central processor. The basilar membrane acts as a bank of bandpass filters, applying a mechanical Fourier-like transform to break down the complex sound wave into its frequency components (Fletcher, 1940; Helmholtz, 1863; Moore, 2003b; Rasch & Plomp, 1982). With the filter widths being wider for the higher frequencies than the lower frequencies, the lower harmonics of a sound are more likely to be fully resolved than the higher ones. It is these resolved components that are then used by the central processor to ascertain the pitch of the complex tone. This processor acts as a harmonic template to match the intervals of the resolved harmonics with the central template (Arehart, 1994; Goldstein, 1973; Moore, 2003b). Exactly how this template undertakes this matching process and how these matched responses are subsequently used to determine pitch is still a matter of contention, with a range of theories having been propounded (Goldstein, 1973; Terhardt, 1974; Thurlow, 1963). Nonetheless, as this model depends on the resolved components of a complex sound, proponents of this theory propose that the lower harmonics of a complex sound contribute most to the pitch percept.

Whereas the pattern-recognition models stem from the spectral resolution of the individual frequency components of a complex sound, the alternative class of pitch perception models are temporal-based. Their underlying principle is that subsequent to a complex sound being spectrally analysed into frequency-related channels, the pitch is determined by analysing the interactions between resulting firing patterns, or more specifically, the intervals between successive neural spikes (Moore, 2003b). Similar to the pattern-recognition models, the exact way that these inter-spike intervals are calculated and then combined to determine pitch is still subject to debate. For example, Licklider (1951), Meddis & Hewitt (1991), and Meddis & O'Mard (1997) favour theories based on an autocorrelation function where the intervals between each and every other pulse are analysed and then summed across all channels. On the other hand, Carlyon et al. (1998) are more inclined towards the derivation of pitch from the first-order intervals. Carlyon et al. (2002) proffered a model where a weighted sum of the first-order intervals between successive spikes was calculated, with longer intervals thought to dominate the pitch percept. Regardless of which theory may hold true, temporal-based models of pitch perception involve examining the firing patterns of the auditory nerve fibres. Although the precise method in which one derives the pitch of a complex tone is still unclear, it would appear that neither model is used in isolation; rather, aspects from both models are used in combination to enable effective pitch perception.

As the F0 does not have to be present in the stimuli for a distinct pitch to be perceived, as is commonly mentioned in relation to the ‘missing fundamental’ phenomenon, this suggests that pitch information is conveyed by the higher harmonics of a complex sound. Although there is some contention as to exactly which harmonics may dominate the pitch percept, there is general agreement that for most complex sounds, the dominant harmonics are usually between the first and fifth, with a trend for the dominant region to decrease as the F0 increases. Plomp (1967) reported that for frequencies above 1400 Hz, the F0 itself dominates the pitch percept, however for frequencies below 1400 Hz, it is the second and higher harmonics that are the most important. For the lower frequencies up to 700 Hz, the third and higher harmonics dominate the percept, with the fourth and higher harmonics being the most important for complex tones with F0 up to 350 Hz.

It is also pertinent to note that pitch can be perceived from stimuli with no distinct spectral peaks, such as the unresolved higher harmonics of complex sounds, noise bursts, or gated white noise. These types of stimuli preclude the place mechanism from contributing to a pitch percept. For example, for many common everyday sounds, the lower harmonics of the sound are ‘resolved’; that is, they excite distinct sites on the basilar membrane. The perception of these resolved components would involve the use of available spectral cues. The higher harmonics of the sound, though, tend to have a flatter shape, with less distinction between the peak and trough. As the auditory filters in the cochlea become broader with increased frequency, the higher harmonics of a complex sound tend to remain unresolved, with several harmonics simultaneously interacting along the basilar membrane. That is, specific sites along the basilar membrane will be responding to several of these higher harmonics, resulting in a complex pattern of vibration. As this vibration pattern repeats at a rate equal to the F0, the pitch of the unresolved components can subsequently be determined from the repetition rate of the interacting harmonics. Although the salience of the pitches perceived from this kind of stimuli is likely to be diminished when compared to signals with a spectral component, research by Kaernbach & Bering (2001) as well as Moore & Rosen (1979), indicates that pitch arising from temporal envelope information can still provide reliable musical information to enable most listeners to discriminate musical intervals and melodies.

Further evidence of pitch being perceived from purely temporal information is demonstrated with psychoacoustic research involving noise stimuli. For example, for repeated busts of noise, the entire basilar membrane would be periodically stimulated corresponding to the rate of the noise bursts, as opposed to having a well-defined peak of stimulation occurring at a certain point along the membrane. Research by Harris (1963), Miller & Taylor (1948), and Pollack (1969) demonstrated that pitch-like sensations could be elicited using interrupted noise stimuli for a limited range of low interruption rates. Burns & Viemeister (1976, 1981) reported that their subjects experienced a pitch sensation corresponding to the modulation rate when listening to sinusoidally amplitude-modulated noise stimuli, up to a rate of between 800 Hz and 1000 Hz. These pitch sensations conveyed musical pitch information that enabled subjects to recognise melodies and musical intervals. These findings are further indication that temporal regularities for certain waveforms can provide a pitch percept to the listener.

However, the pitch cues derived from modulated noise stimuli appear only to be available at low modulation frequencies (Burns & Viemeister, 1976, 1981; Harris, 1963; Miller & Taylor, 1948; Pollack, 1969). Similar to the pitch perceived from the unresolved harmonics of a complex sound, the pitch arising from amplitude-modulated noise is also less salient than pitch associated with puretone stimuli, or other waveforms providing an explicit spectral structure. That is, a spectral basis to stimuli provides a more salient pitch percept than percepts derived from purely temporal information (Burns & Viemeister, 1976, 1981; Moore & Rosen, 1979).

Pitch perception is also affected by variations in frequency and intensity, with the nature of this relationship being dependent on the complexity of the sound and the individual themselves (Sekuler & Blake, 1994). Morgan et al. (1951) published research into the interactions between these three attributes. They found that in general, the degree of pitch change with intensity was small, with most subjects experiencing little or no change of pitch as the intensity varied. However, there was a high level of individual variation and a few atypical cases experienced large pitch shifts. These pitch changes displayed some degree of frequency dependency. For the low frequencies, increased intensity was more associated with decreased pitch percepts, whereas at the high frequencies, increased intensity tended to result in higher pitch percepts. Cohen (1961) corroborated these findings further reporting that for the middle frequencies, there was no pitch change, and any changes at the frequency extremes were generally less than 2.5% of the initial presentation frequency.

It is worth clarifying the difference between two commonly used psychoacoustic pitch assessments – tasks of frequency selectivity and frequency discrimination. *Frequency selectivity* refers to "the ability to resolve the sinusoidal components in a complex sound." (Moore, 2003b, p. 65). *Frequency discrimination*, however, implies "the ability to distinguish a change or difference in frequency when the two tones are played one after the other." (Clark, 2003, p. 299).

2.3.3 Timbre

Unlike pitch and loudness, timbre is a subjective, multi-dimensional attribute related to differences in sound spectra. According to the Acoustical Society of America (1960) (in Gfeller et al., 2002b, p. 349), timbre is "that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar." It essentially incorporates the features of a sound that do not directly relate to pitch or loudness. These include the spectral power distribution, temporal envelope, rate and depth of any amplitude or frequency modulations, inharmonicity of the partials, and the transients present in the signal (Kohlrausch & Houtsma, 1989). Grey (1977) identified three different spectral dimensions to timbre: i) rise time (onset or attack time); ii) spectral centroid (the frequency band with the most energy, contributing to the perception of brightness or dullness); and iii) spectral flux (the number of components in the spectrum, and the spread of these). The temporal envelope also affects the perceived timbre; for example, modifying or substituting the attack phase can render a previously recognised timbre unrecognisable (Kohlrausch & Houtsma, 1989).

Each instrument has a unique timbre partially emanating from its harmonic structure, particularly the number and spacing of the higher harmonics. For example, a clarinet has only odd-numbered harmonics and consequently generates widely spaced harmonics. If the number of the harmonics was increased, and/or the spacing between these harmonics narrowed, the tone may be rated as having a harsher quality (Rasch &

Plomp, 1982). Accurate timbral perception is important for instrument recognition, the aural organisation or structuring of a composition, and auditory scene analysis when listening to musical ensembles. Further, it is also a highly contributory factor to the quality of the listening experience (Gfeller et al., 1998, 2002b). The shape, structure and method of sound production for each instrument directly impacts upon the resulting timbre, with the relevant waveform providing cues for instrument identification. For example, the attack time may indicate the mode of sound production, such as being blown versus being plucked (Clark, 2003). Absolute identification of instruments is dependent upon a host of other factors, including the waveform's envelope, periodicity, sound spectrum, temporal characteristics, transients, and the preceding and proceeding sounds. For example, the flute's simple harmonic structure along with the small noise burst preceding each tone characterises its unique 'breathy' quality (Moore, 2003b). Further, tones from differing pitch ranges often have different relative spectra. For example, whereas low-pitched piano sounds have little energy at the F0 with stronger higher harmonics, the higher-pitched piano notes have a stronger F0 and weaker high harmonic components (Kohlrausch & Houtsma, 1989).

The perception of timbre is also influenced by pitch, and vice versa. Beal (1985), Crowder (1989), and Pitt & Crowder (1992) found that normally hearing listeners were faster, and more accurate, at discriminating whether two sequentially presented pitches were the same or different when the timbre of the two sounds was the same. A difference in the timbre of the two sounds impeded their perception of pitch. This effect was exacerbated in those with little or no music experience.

2.3.4 The Perceptual Consequences of a Cochlear Hearing Loss

The above discussion on loudness, pitch, and timbre perception largely pertained to normally hearing adults. However, a host of researchers have reported that a cochlear hearing loss can alter these perceptions. With respect to loudness, recruitment is commonly reported in those with cochlear hearing loss, where the rate of loudness growth with increased suprathreshold intensity levels is steeper than normal. This serves to decrease the listener's dynamic range as well as exaggerate the perceived fluctuations in the dynamic variation of the input sound; that is, acoustic stimuli will seem to fluctuate more in loudness (Moore, 1995, 1996).

A host of researchers have reported decreased pitch discrimination ability for both pure tones and complex tones in subjects with a cochlear hearing loss, along with perceptual anomalies such as inconsistencies in pitch-scaling tasks (Moore, 1995; Moore & Carlyon, 2005; Moore & Glasberg, 1988a; Moore & Peters, 1992). A cochlear loss impacts upon both the temporal- and place-based cues used in pitch perception (Moore & Carlyon, 2005; Moore & Skrodzka, 2002). The degree of this effect is unpredictable though, and not strongly correlated with absolute hearing thresholds. Reduced frequency selectivity resulting from increased auditory filter bandwidths is commonly reported in this type of hearing loss, with the ensuing psychophysical tuning curves tending to be broader and less defined (Arehart, 1994; Moore, 1995, 1996; Moore & Peters, 1992; Summers & Leek, 1994). Moore (1996) reported that hearing thresholds greater than approximately 40-50 dBHL result in asymmetrically shaped auditory filters with bandwidths approximately two times wider than those for normal hearing. These broader bandwidths would decrease the resolvability of low-order harmonics, impeding the perception of fundamental frequencies (Moore & Moore, 2003). This would not only have a deleterious effect on pitch perception, but timbre perception would also be affected as the spectral shape perceived would be altered, having less detail in the excitation pattern (Moore, 1995). Summers & Leek (1994) reported that the spectral smearing resulting from the broader and more asymmetric auditory filters diminishes the perceived amplitude difference between the signal's peaks and valleys. This spectral flattening could also make it more difficult for the listener to accurately perceive the pitch and timbre of a complex sound.

According to spectrotemporal theories, abnormal pitch perception associated with cochlear hearing loss is also resultant from impaired temporal-processing skills, as the quantity and quality of the periodicity cues and temporally-encoded spectral cues would be diminished (Arehart, 1994; Florentine & Buus, 1988; Glasberg et al., 1987; Moore & Skrodzka, 2002; Nelson & Freyman, 1987). As a consequence of poorer frequency selectivity, receiving relatively less spectral information, and a reduced ability to use temporal fine-structure information, hearing-impaired listeners tend to be more reliant on the temporal envelope cues for pitch information than listeners with normal hearing (Moore, 1995; Moore & Carlyon, 2005; Moore & Moore, 2003). For example, they may place more weighting towards the temporal information from the unresolved harmonics of a complex tone, and less on the spectral information of the resolved harmonics, as compared to listeners with normal hearing for whom spectral information is usually more salient. However, as previously discussed in section 2.3.2, the use of temporal-based envelope cues for pitch perception may not be as effective as the spectrally-derived cues (Burns & Viemeister, 1976, 1981; Kaernbach & Bering, 2001; Moore & Carlyon, 2005).

2.4 MUSIC

Krumhansl & Iverson (1992) identified four basic psychological attributes to musical sounds – pitch, duration, loudness, and timbre. Whilst pitch is predominantly, but not exclusively, a derivative of frequency, and loudness of intensity, timbre involves the perception of a larger number of factors. Music perception primarily involves pattern perception, be it rhythmic, pitch, loudness, or timbral variations (Gfeller et al., 1997). Whereas the sequencing or patterning of pitches forms the musical correlates of melody and harmony, the sequencing of durations or temporal patterns forms the foundation of rhythm. Variations in loudness are often referred to as dynamics, and the perception of different timbres underlies instrument differentiation (Krumhansl & Iverson, 1992). However, although these attributes are separate entities, the combinations of, and interactions between the different attributes largely contribute to music as we commonly know it. For example, listening to a melody entails both pitch and duration (rhythm) perception. Further, variables pertaining to the individual listener, such as one's music experience, prior training, listening preferences, age, culture, or demographics, may also affect music listening.

Melody is comprised of a series of pitches derived from a musical scale, successively organised in a meaningful manner. Individual notes in melodies gain their meaning from the interactions between, and relationships to, other notes in the sequence. For example, a melody can be transposed, thus using an entirely different set of absolute pitches, without the subjective quality or meaning changing for the listener. However, if the original absolute pitches were used and the order changed, the result would be an entirely different melody with a different sound and meaning. The recognition of a melody is dependent on both familiarity with, and the ability to perceive its various structural features, such as the overall contour, relative pitch changes from one note to the next, and rhythm patterns.

2.5 MUSIC STIMULI IN COMPARISON TO SPEECH STIMULI

Regardless of the type of sound, the stages of sound production outlined earlier - energy to initiate a vibration, the vibration itself, its transfer to a sound body, and the propagation of the vibration through a medium, are fundamentally the same. For a musical instrument, the initiating energy may constitute blowing, bowing, hitting, or plucking to stimulate vibrations of materials such as a string, drum skin, or air column. The sound body is often a wooden or metal box, tube, or board. For voice, the instigating energy is through air from the lungs creating either vibrations of the air column from the vocal cords (voiced sounds), or air turbulence (unvoiced sounds), with the oral and nasal cavity acting as the sound body. Both speech and music are largely quasi-periodic signals comprising complex sound waves with the frequency, temporal, intensity, and timbral components presented in an organised manner. Both speech and music signals have a spectra and envelope that varies in time. The spectral envelope of both signals are often characterised by the presence of formants (Glossary) produced by the resonance of the vocal tract or the acoustic properties of the instrument itself. Other differentiating features of the spectral envelope include periods of silence in the signal, a changing pattern of variation in the relative amplitudes of the harmonic components, and varying onset and offset transients (Krumhansl & Iverson, 1992; Wolfe, 2002). These acoustic features change for different sources such as different speakers or instruments.

Although speech and music share numerous similarities, they also differ in many ways. The fundamental frequencies for speech are predominantly between 80 Hz and 500 Hz, with the range of fundamental frequencies in music being significantly greater. For example, the piano alone has a F0 range from 27.5 Hz to 4186 Hz (Sekuler & Blake, 1994). Further, the loudest sounds of speech, even when shouted, tend to be less than 85 dBSPL, and are more often in the vicinity of 75 dBSPL; general conversational speech approximates 60 dBSPL. Irrespective of the style, the louder components of music are often in the range of 100 dBSPL to 110 dBSPL, potentially reaching peaks of up to 118 dBSPL (Chasin, 2003).

Categorical perception of speech phonemes, essential for accurate speech recognition, is dependent on the signal's overall spectral shape. The perception of the acoustic components of the individual phonetic segments comprising speech is subsidiary to the perception of its overall spectral shape, with the acoustic patterns of individual phonemes varying depending on the preceding and proceeding sounds (Moore, 2003b). That is, a phoneme produced in isolation will have a different acoustic pattern to the same phoneme produced in a speech context. Perception of the first and second formants of the harmonic spectrum, as well as the transients between the formants is critical to speech recognition. Accurate perception of the F0 itself is not imperative to speech recognition for non-tonal languages such as English (Wolfe, 2002).

On the other hand, accurate perception of the F0 of individual notes is more important for music recognition. The spectral shape, formants, and the onset or offset transients between formants affect the resulting timbre, but they are not fundamental to the identification of the melody. In speech, these same elements provide vital cues to the listener as to what is actually being said. For accurate melody recognition, it is the information related to pitch, its duration, and the temporal regularities of the pitches that are critical (Wolfe, 2002).

It is also worth noting that apart from differences in the physical parameters of the signals themselves, speech and music also differ in their broader functional roles, as well as in the processing skills required for their interpretation by the listener. Speech is a discursive form of communication with individual words largely having a predefined meaning. Music, on the other hand, is often more abstract and non-discursive in its role, not necessarily having a clearly defined semantic function as does speech. Music perception frequently entails the simultaneous processing of multiple input sources such as concurrent instruments, multiple rhythms, or counter-melodies. Even if the sound is being produced by a single instrument, the resolution of simultaneous notes may be required for polyphonic instruments, such as the piano or guitar. In speech, the listener may be more focused on a single sound source (i.e. the speaker), with speech recognition often being aided by the presence of additional information such as lip reading, body language, situational cues, language conventions, grammatical rules, and prosodic cues.

In summary, the contrasts in the nature and parameters of music and speech signals, along with the different perceptual skills that listening to these sounds entail, may partially contribute to the disparity between current CI users' ability to accurately perceive speech versus music stimuli. This may be further confounded by the limitations of current implant technology in processing complex sounds. This will be discussed further in the next chapter.

CHAPTER 3: COCHLEAR IMPLANTS

With the previous chapter discussing the perception of sound via acoustic stimulation, this chapter focuses on the electrical stimulation of hearing via a CI. Section 3.1 provides an overview of the CI in general, and sections 3.2 and 3.3 respectively detail the speech processors and speech-processing strategies integral to the device. With the continually evolving nature of CI technology, the focus of these sections is primarily on current-day devices, processors, and strategies pertinent to this research. However, as older devices and processors have led us to today's implants, a brief overview of key historical devices and strategies is incorporated. Some of these older devices and strategies have been used in the music perception research studies reviewed in the next chapter.

Within each of these sections describing the major components of an implant, a specialised subsection is included focusing specifically on the Nucleus devices and products relevant to this research. These implants, manufactured by Cochlear Limited, were the only implant types used in this study. The CI22 and CI24 implant systems are outlined in section 3.1.1, with the associated speech processors, both body-level and ear-level options, described in section 3.2.1. The speech-processing strategies implemented in these processors are covered in the subsequent section (section 3.3), commencing with a comparison of the major types of stimulation – analogue stimulation, feature-extraction-based pulsatile stimulation, and filterbank-based pulsatile stimulation (sections 3.3.1 – 3.3.3). Whilst the former two types of strategies were not used by the subjects in this study, they are important considerations with regard to the development of both current-day, as well as future speech-processing strategies. The filterbank-based strategies are of most relevance to this study, and sections 3.3.3.1 – 3.3.3.4 review the three most widely implemented strategies at the present time – the Spectral Peak strategy (SPEAK), Continuous Interleaved Sampling (CIS), and the Advanced Combination Encoder (ACE) strategy. A brief outline of some of the newer hybrid strategies is provided in section 3.3.4. Although not utilised in this current study, these hybrid strategies have been included to provide an indication as to prototype strategies currently under development or trial; they may become relevant for future work in this area.

As the combination of the implant, the speech processor, and the speech-processing strategies impact upon the sound perceived by the user, the effect of the CI, as a complete system, on the perception of musical stimuli is discussed in section 3.4. The ensuing section (section 3.5) compares the CI to the other major audiological habilitative device for those with a moderately-severe to profound hearing loss – the hearing aid (HA). Whereas the CI directly stimulates surviving auditory neurons using electrical impulses, the HA amplifies sound to provide acoustic stimulation via the normal auditory pathway. The devices differ in regards to their sound-processing parameters and thus, in the resulting output presented to the user. That is, the same auditory stimulus may result in a different sound percept for a CI user than for a HA user.

3.1 OVERVIEW OF THE COCHLEAR IMPLANT

In most cases of sensorineural hearing loss, the loss of auditory neurons is often subsidiary to the loss of hair cells. Based on this premise, the cochlear implant (CI) operates by bypassing the outer ear, middle ear, and the hair cells to electrically stimulate the surviving auditory neurons directly (Loizou, 1998). The current criteria for cochlear implantation varies from clinic to clinic, but broadly constitutes a moderately-severe to profound bilateral sensorineural hearing loss in the speech frequency range (Figure 3.1). Generally most clinics require potential adult implantees to meet certain speech perception limitations following trials with appropriately fitted HAs, along with a range of other medical, psychosocial, and audiological evaluations. Both children and adults can receive CIs. The criteria advocated by Dowell et al. (2004) were speech perception scores using sentence stimuli of less than 70% in the best-aided listening condition, and less than 40% in the ear to be implanted. These criteria have been adopted by many CI clinics in Australia, including those involved in this study.

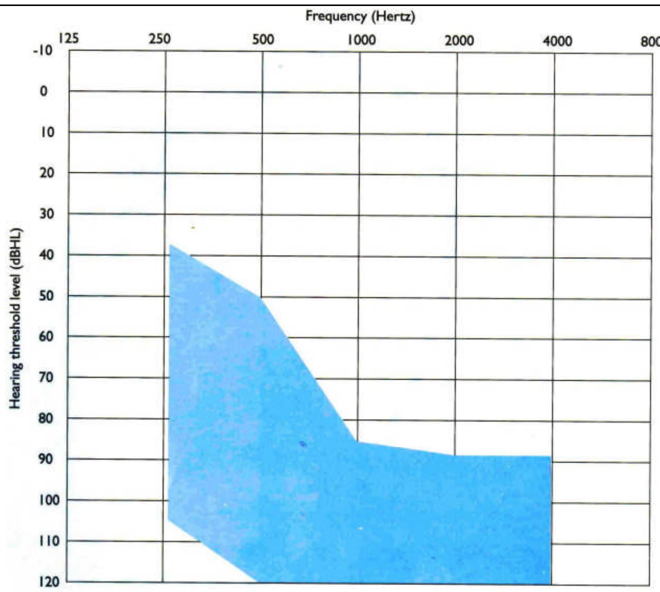


Figure 3.1: Audiogram representing hearing thresholds currently considered for a cochlear implant

The blue section broadly represents hearing thresholds of potential patients currently qualifying for an implant. These thresholds, though, are considered secondarily to the prospective patient's speech perception ability. Many CI clinics in Australia, including those involved in this study, currently utlilise a speech perception criterion for potential implantees of <70% in the best-aided listening condition, and <40% in the ear to receive the CI, using sentence stimuli.
(Image courtesy of: Cochlear Ltd.)

There have been a multitude of CIs commercially available to prospective patients since a pioneering single-electrode device was investigated by Djourno & Eyries (1957). An outline of the early history of CIs is provided by Luxford & Brackmann (1985). Since these initial experimental devices, numerous companies, and devices have found their way on to the market, with differing levels of success and longevity. Technology has progressed from the early single-channel implant systems (Glossary), such as the *House 3M* implant, and percutaneous links (Glossary) (such as the *Ineraid Symbion* implant) through to the current-day multi-channel implants (Glossary) with their transcutaneous links (Glossary). An outline of some of the previous, now-obsolete, CI manufacturers is provided in Appendix 1, along with references that provide more detailed information. The evolution and expansion of cochlear implantation has been rapid, particularly in the last decade. At the 1995 National Institutes of Health's Consensus Development Conference into CIs, there were approximately 12000 implantees world-wide. It was one recommendation of this conference that the criteria for implantation be expanded to include patients whose hearing thresholds are classified at a 'severe' level, with speech perception scores using sentence stimuli less than 30% in the best-aided condition

(National Institutes of Health Consensus Development Conference, 1995). Since then, the expeditious progress of cochlear implantation has seen its growth to over 50000 implantees globally (Moore & Carlyon, 2005), and ever-broadening clinical implantation criteria.

Currently there are three companies manufacturing devices approved by the American Food and Drug Administration (FDA) – Med-El (Combi 40+), Advanced Bionics Corporation (HiRes), and Cochlear Limited (Nucleus 24). A summary of these devices is provided in Table 3.1. It should be noted that both Cochlear Limited and Med-El have recently released new commercial devices - the *Freedom* system (Cochlear Limited), and the *Pulsar CI*^{100} and *Sonatati*^{100} systems (Med-El). However, as these devices were not available at the time of this study, and have yet to be extensively described in the research literature, they have not been considered in this thesis. All current implants are essentially based on the same underlying principles with very similar components (Figure 3.2). Each CI comprises both a surgically implanted internal package and externally worn components. The internal components consist of a receiver-stimulator package containing a magnet and antenna (or receiving coil) housed in a ceramic or silastic-coated titanium case, connected to an electrode array. The receiving coil with magnet is implanted into the mastoid, and communicates with the external transmitting coil worn in a magnetically aligned headset. This receiver-stimulator package decodes the radio frequency signal transmitted from the speech processor and converts it into an electrical current used to stimulate the cochlear nerve fibres via the implanted electrodes.

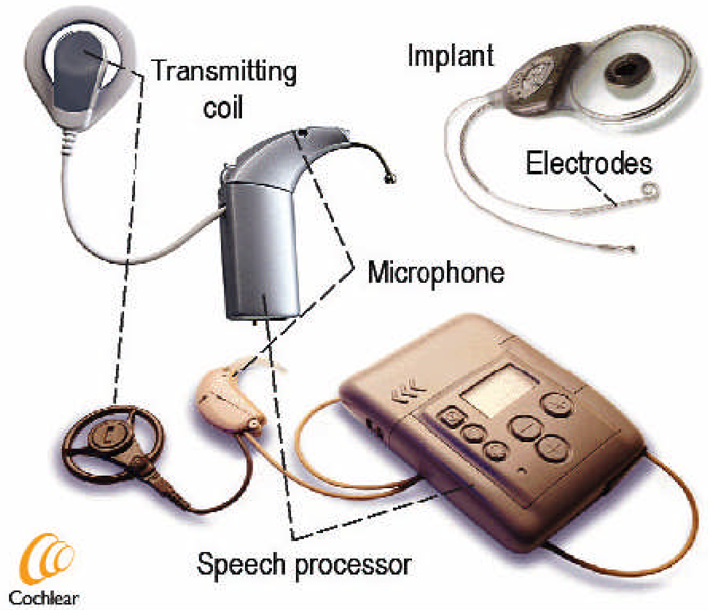


Figure 3.2: Major components of a cochlear implant
(Image courtesy of: Cochlear Ltd.)

Many currently-implanted devices utilise both intracochlear and extracochlear electrodes. The extracochlear electrode(s) act as a ground source for monopolar stimulation (Glossary) and are placed either under the temporalis muscle and/or on a plate attached to the receiver-stimulator package. The intracochlear electrodes, of which there are currently between 12 and 24 depending on the device, are arranged on a carrier inserted into the scala tympani of the cochlea to an optimal depth of between 25 mm and 31 mm from the round window. This corresponds to approximately the first 1.5 turns of the cochlea. The scala tympani is a surgically accessible site in close proximity to the tonotopically arranged spiral ganglion cells. Both flexible and pre-curved arrays are currently utilised; whereas the former tends to sit against the lateral wall of the scala tympani, the latter is designed to curve inwards and lie near the inner wall (Figure 3.3 & Figure 3.4). It is thought that placement close to the inner wall would improve spatial specificity and reduce thresholds which would theoretically reduce power consumption as less stimulation current is required.

Table 3.1: Comparison of Current Commercial Devices

	Cochlear Limited – Nucleus CI24	Clarion – Hi Res	Med-El – Combi 40+
Year FDA approval	2001	2002	2001
Material of casing	Titanium case with silicone envelope	Titanium case with silicone envelope	Ceramic
Electrode arrays available	Straight, Contour Advance (CA), Double	Flexible (*HiFocus IJ*), Perimodiolar (*HiFocus Helix*)	Standard, Compressed, Split-compressed
# Electrodes	22 + 2 extracochlear	16 + 1 extracochlear	12 pairs + 1 extracochlear Split compressed: 7 + 5
# Channels	22 channels	16 channels	12 paired sites
Span of active electrodes	CA: 15 mm Straight: 17 mm	IJ: 17 mm Helix: 13 mm	Standard: 26.4 mm Compressed: 13.1 mm
Electrode spacing	CA: 0.4 mm apically, to 0.8 mm basally Straight: 0.75 mm	IJ: 1.1 mm Helix: 0.85 mm	Standard: 2.4 mm Compressed: 1.1 mm Split-compressed: 1.1 mm
Optimal Insertion Depth	25 – 28 mm	IJ: 25 mm Helix: 18 – 21 mm	Standard: 31.3 mm Compressed: 13 – 15 mm
Stimulation type	Sequential Pulsatile	Sequential and Simultaneous Pulsatile	Sequential Pulsatile
Maximum stimulation rate	14 400 pps	83 000 pps	18 180 pps
Stimulation Modes	Monopolar, bipolar, common-ground	Monopolar	Monopolar
Speech Processors	Sprint; Esprit 3G	Platinum, Auria	Tempo+
Speech-processing Strategies	ACE, SPEAK, CIS	HiRes-S, Hi-Res-P	CIS+
Neural Response Telemetry	Yes	Yes	No
Reliability	Adults: 99.6% (1yr)[#] Children: 98.7% (1yr)[#]	Adults and Children: 99.2% (12 months)[*]	Adults and Children: 99.73% (36 months)[o]
MRI compatible	Yes, surgery required	Yes, up to 0.3 Tesla	Yes, no surgery required; up to 1.5 Tesla

Sanderson (2004) * Advanced Bionics (2004) o Med-El (2005)

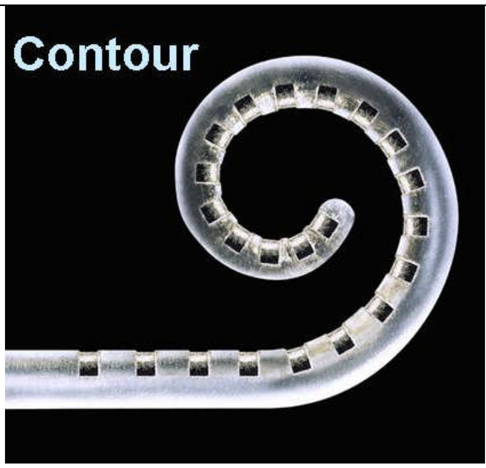


Figure 3.3: Contour Electrode Array

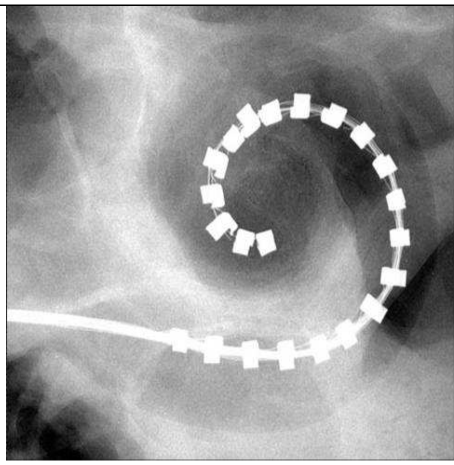


Figure 3.4: Positioning of the Contour Array

(Images courtesy of: Cochlear Ltd.)

A transcutaneous link connects these internal components to the external parts of the system via a radio frequency link across intact skin. There have been previous CIs utilising percutaneous links, however the associated biologic and medical risks arising from a direct connection through the skin has led to them being superseded by transcutaneous transmission in commercially-available clinical devices. The external components comprise a microphone, speech processor, transmitter coil, and associated connecting cables. The microphone, usually worn behind the ear or less often as an optional lapel microphone, is an input transducer, converting the acoustic sounds into electrical signals. These electric signals are then transmitted to the speech processor via the connecting cable. Most microphones have a broad frequency response with reduced sensitivity to low frequency vibrations resulting from movement. Both directional and omni-directional microphones (Glossary) are used in current devices, depending upon the manufacturer. The speech processor converts these input signals into patterns of electrical stimulation produced by the electrodes, with the parameters for stimulation being programmed into the processor via a patient's 'MAP' (Glossary). This conversion process is achieved through various stages, the specifics of which vary depending upon the type of strategy in use. For example, for the most commonly implemented strategy type, filterbank strategies (Glossary), the electrical signals from the microphone are firstly amplified to levels suitable for further processing before being split into a number of frequency bands, each band corresponding to one stimulating channel of the implant. The output of the filterbank (Glossary) is then sampled to determine the sequence of stimulation in accordance with the speech-processing strategy implemented. The electrical amplitudes derived from the selected filter outputs are compressed into levels appropriate for electrical hearing, corresponding to the patient's electrical dynamic range. This information is then encoded into a radio frequency signal to be transmitted to the internal receiver-stimulator package to activate the implant. The integrity of the implant system, as well as the parameters for stimulation, are usually established and verified via a manufacturer-specific programming interface.

There are three common modes of stimulation – monopolar, bipolar, and common-ground (Glossary) (Figure 3.5). Monopolar stimulation can only be implemented on devices with an extracochlear electrode as current flows between the extracochlear electrode and an intracochlear electrode. In bipolar stimulation, current flows between two intracochlear electrodes, and in common-ground stimulation, the current flows from one electrode to all the others connected together. As the mode of stimulation determines which electrodes are activated and the degree of current spread between them, this may affect 'T' and 'C' levels (Glossary), spatial specificity, pitch perception, and power consumption (Battmer et al., 1993; Cohen et al., 2001; Pfingst et al., 1995b; Rebscher et al., 2001). For example, monopolar stimulation through the use of an extracochlear electrode enables a wider current spread than the bipolar mode, potentially exciting a larger number of auditory neurons (Pfingst et al., 1995a; Rebscher et al., 2001; Shepherd et al., 1993). This then enables lower stimulation thresholds and reduced power consumption (Battmer et al., 1993; Busby et al., 1994; Cohen et al., 2001; Teig et al., 1997; Zwolan et al., 1996). Busby et al. (1994) demonstrated that bipolar stimulation provided more localised current distribution than either monopolar or common-ground stimulation; this enabled a more localised set of auditory neurons to be stimulated and thereby potentially provided some implant users with better spatial specificity. However, the same study also showed that bipolar stimulation required greater current levels for subjects to achieve their respective 'T' and 'C' levels than the use of the other two stimulation modes. There is no definitive proof though, that better spatial specificity would in turn improve speech perception ability (Hughes & Abbas, 2003; von Wallenberg et al., 1995).

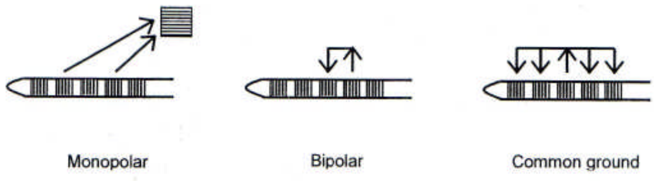


Figure 3.5: Stimulation modes
In monopolar stimulation, a reference or ground electrode is situated remote to the active electrode. Current flows between this remote electrode and the active intracochlear electrode. In bipolar stimulation, current passes between two active intracochlear electrodes. For common-ground stimulation, one intracochlear electrode acts as the active electrode, with current passing from this electrode through all of the other intracochlear electrodes.
(Image from: McDermott (2004))

The main type of stimulation used in current CIs involves the use of constant-current biphasic pulses (Glossary), presented sequentially to active electrodes (Figure 3.6, Figure 3.7, & Figure 3.8). The use of these discrete pulses avoids the creation of a direct current which may damage surrounding tissue, with the non-simultaneous presentation helping to control potential channel interactions. The amplitude of each pulse is extracted from the envelope of the filtered waveform. One consequence of currently-utilised pulsatile stimulation strategies is the limited representation of fine-temporal details from the original input signal. As the output of the filterbank is smoothed for each stimulation cycle, only the temporal envelope cues are preserved. The rate used to sample this temporal envelope would also affect the amount of temporal information available to the CI user. Pulsatile stimulation is applicable to both the current filterbank strategies (CIS, SPEAK, and ACE), as well as the older feature-extraction strategies (Glossary) (F0F2, F0F1F2, and MPEAK) that will be discussed later in this chapter.

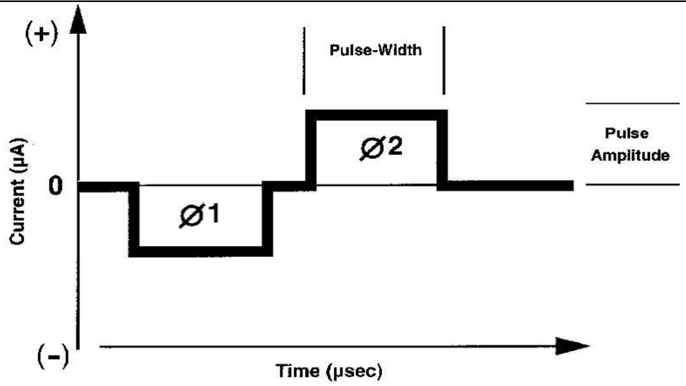


Figure 3.6: Biphasic pulse

(Figure courtesy of P. Seligman)

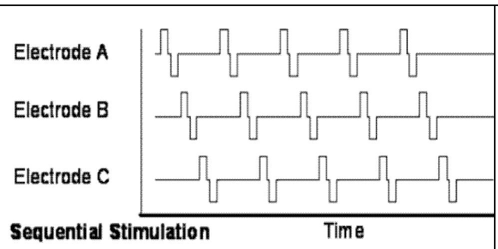


Figure 3.7: Sequential pulsatile stimulation

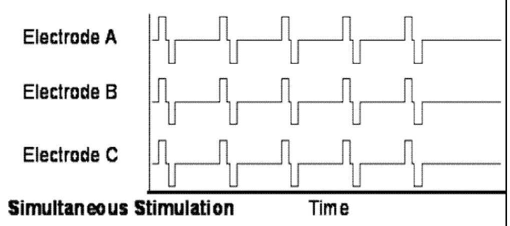


Figure 3.8: Simultaneous pulsatile stimulation

(Figures courtesy of P. Seligman)

The alternative to pulsatile stimulation is analogue stimulation (Glossary) where an electrical analogue of the input waveform is presented to the active electrodes simultaneously (Figure 3.9). These analogue-based strategies use a continuously-varying current to present the acoustic information to all electrodes as a continuous waveform. Analogue stimulation maintains the fine-temporal information of the original waveform in the resulting electrical waveform as the output of the filters are not smoothed or sampled, as occurs with pulsatile stimulation. However it is unclear as to whether CI users can perceive and effectively utilise the extra fine-temporal information (McKay, 2004; Wilson et al., 2004b; Zeng, 2002).

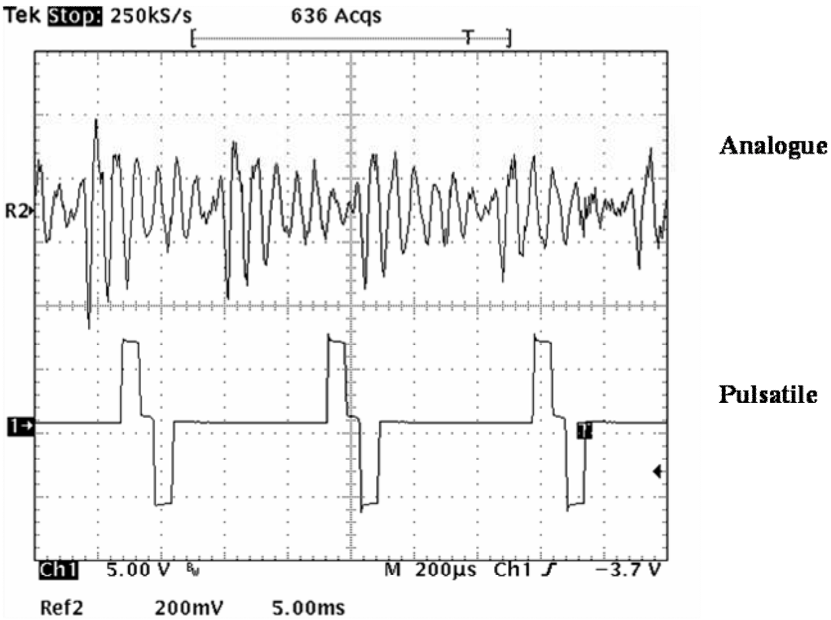


Figure 3.9: Analogue and pulsatile stimulation waveforms
(Figure courtesy of P. Seligman)

3.1.1 Nucleus Cochlear Implants

The Nucleus series of CIs are manufactured by Cochlear Limited. Three different FDA-approved Nucleus implants were used in this study - the Nucleus 22 (CI22M), the Nucleus 24 (CI24M), and an updated Nucleus 24. This latter updated CI24 has the option of either a straight electrode array (CI24K), or the perimodiolar *Contour* array (CI24R) (Glossary). A summary of these devices is provided in Table 3.2. The initial Contour array was modified to include a '*Softip'*, and renamed the '*Contour Advance*'. However for simplicity, both versions of this perimodiolar array will be collectively referred to as the Contour array from here on. It should be noted that in 2005, the *ESPrit Freedom* was commercially released by Cochlear Ltd. This most-contemporary implant system incorporates a new range of components, devices, accessories, and features, including an updated implant, a body-worn speech processor, an ear-level speech processor, and new programming software. However, with this product range not being commercially available at the time this research was conducted, it will not be discussed in any more detail.

The CI22M implant consists of a titanium-encased receiver-stimulator package attached to 22 intracochlear electrodes on a silicone carrier. With no extracochlear electrodes, only bipolar or common-ground stimulation can be utilised in up to 20 channels with a maximum stimulation rate (Glossary) of approximately 4000 pps. The CI24M has 22 intracochlear and two extracochlear electrodes with a smaller receiver-stimulator package than the CI22M. The first of the extracochlear electrodes is a ball electrode placed under the temporalis muscle, with the second being a plate electrode situated on the receiver-stimulator package itself. These extracochlear electrodes enable monopolar stimulation to be used, in addition to the bipolar and common-ground modes, in up to 22 channels. The CI24M also increases the maximum stimulation rate to around 14400 pps, and includes Neural Response Telemetry to externally provide information relating to the implant's integrity and function. The receiver-stimulator package of the CI24R is an updated version of the CI24M, designed to be more rugged and reliable. It is slightly smaller and more symmetric than its predecessor, although the internal electronics are identical to the CI24M. The major modification associated with the CI24R was the introduction of the pre-curved perimodiolar electrode array, the Contour array. Consisting of 22 half-banded electrodes designed to be placed adjacent to the modiolar wall, the array aims to improve spatial specificity and reduce power consumption. The CI24K implant is also available using the same updated receiver-stimulator package as the CI24R, but with a straight array, for patients with whom the Contour array is contraindicated. Both the straight and Contour arrays are designed to be inserted around 25 mm to 28 mm into the cochlea, approximating to the first 1.5 turns of the cochlea. All of the CI24 devices use essentially the same programming software, external equipment, and accessories (Co-operative Research Centre for Cochlear Implant and Hearing Aid Innovation, 2003a, 2003b; Seligman, 2003b). Although the various CI24 devices vary in their physical structure, the internal electronics of the devices are identical and hence, they will be collectively grouped together and referred to as CI24 implants from this point forward.

Table 3.2: Comparison of Current Nucleus Cochlear Implants

	CI22M	CI24M	CI24R	CI24K
Features	• 22 intracochlear electrodes • Maximum stimulation rate: 4000 pulses/sec • Maximum 22 channels • No telemetry	• 22 intracochlear electrodes, and 2 extracochlear electrodes • Maximum 22 channels • Thinner electronics package • Wider current range supported • Telemetry • Allows higher stimulation rates (up to 14400 Hz) • Smaller in size • Allows compliance, impedance, and neural response telemetry	• Designed for use with the *Contour* electrode array - a perimodiolar array • 22 half-banded electrodes and 2 extracochlear electrodes • Thin, tapered, 15 mm array, inserted with the aid of a stylet • More rugged and reliable electronics packaging than CI24M • Smaller size, more symmetric, with vertical pedestal on sides • Internal electronics identical to CI24M	• Straight-array version of the CI24R • This straight array has 22 full-banded electrodes, along with the 2 extracochlear electrodes • This array (17 mm) is longer than the perimodiolar array to allow for placement along the lateral wall of the cochlea • Internal receiver-stimulator package itself is identical to the CI24M device
Volume bone excavation	1400 mm^3	380 mm^3	380 mm^3	380 mm^3
Stimulation Modes	Common-ground, Bipolar	Common-ground, Bipolar, Monopolar	Common-ground, Bipolar, Monopolar	Common-ground, Bipolar, Monopolar
Year of FDA approval	1985: Adults 1990: Children	1998	2000	2000

Seligman (2003a, 2003b)

3.2 SPEECH PROCESSORS

Most currently implanted CIs have both body-worn, along with ear-level, speech processor options. All processors enable a range of speech-coding strategies to be implemented, and allow the user to select from multiple programs. Each has a range of controls, alarms, and accessories, along with an associated programming suite for mapping (Glossary) the processor. A comparison of the most commonly utilised, commercially-available speech processors is provided in Table 3.3. As a general rule, previous ear-level processors, being smaller in size and thus more restricted in their power supply, tended to have less programming flexibility than their body-worn counterparts. However, recent technological improvements have enabled the implementation of most of the commonly utilised strategies into the manufacturers' respective ear-level processors.

Table 3.3: Comparison of Current Commercial Speech Processors

	Nucleus Body-level	Ear-Level	Advanced Bionics Body-level	Ear-Level	Med-El Ear-Level
Name	Sprint	Esprit 3G	Platinum	Hi-Res Auria	Tempo+
Programs	4	2	3	3	3-9
Size	103 mm x 67 mm x 23 mm	51 mm x 19 mm x 14 mm	40 mm x 69 mm x 22 mm	27 mm x 21 mm x 12 mm	25 mm x 15 mm x 8 mm
Microphone	Directional	Directional	Omni-directional	Omni-directional: In-the-ear 'T-mic', or standard earhook	Omni-directional
Inbuilt Telecoil	No	Yes	No	No	No
Strategies	ACE, CIS, SPEAK	ACE, CIS, SPEAK	HiRes-S, HiRes-P	HiRes-S, HiRes-P	CIS+
Input Dynamic Range	32 dB	30 dB	Up to 80 dB (adjustable)	Up to 80 dB (adjustable)	75 dB
Batteries	1 AA	3 hearing-aid batteries (675)	1 rechargeable lithium ion	4 rechargeable, or 2 AA (via a 'powerpak' attachment)	Rechargeable, or 3 hearing-aid batteries
Other signal-processing features	ADRO	Whisper	Allows for current steering – "virtual channels"		

3.2.1 Nucleus Speech Processors

For all current Nucleus implants, the implantee has the option of either a body-level or an ear-level speech processor. The body-level processor for the CI24 implants is the *SPrint* which was commercially released in 1997. The SPrint processor allows implementation of the *Continuous Interleaved Sampling* (CIS), *Spectral Peak* (SPEAK), or *Advanced Combination Encoder* (ACE) strategies to a maximum total stimulation rate of 14400 Hz. The processor has four programs, a volume and/or microphone sensitivity control (Glossary), an automatic sensitivity control option (Glossary), and various other alarms and options for the user's convenience. Unlike its ear-level equivalent, the SPrint also offers the option of *Adaptive Dynamic Range Optimisation* (ADRO) (Glossary) – an adaptive signal processing algorithm designed to optimise the dynamic range within individual frequency bands (James et al., 2002). A microphone, usually worn behind the ear (although a lapel-worn microphone is available), is connected to the processor, with the option of attaching alternative external input sources such as FM systems or telecoil-based loops (Glossary) instead. In the older

CI22 implants, the body-level processor was the *Spectra22* which enabled the SPEAK strategy to be implemented. An automatic sensitivity control program, along with a microphone sensitivity control adjustment were the only features available. For the CI22, an ear-level speech processor, the *ESPrit22* was commercially released in 2001. Again, SPEAK is the only strategy implementable in this processor with its two programs and a single volume or microphone sensitivity control.

For the CI24, the first ear-level device was the *ESPrit*, introduced in 1998. This processor, however, has been largely superseded with the release of an updated ear-level processor for the CI24, the *ESPrit3G*, in 2002. The ESPrit3G allows both high-rate ACE and CIS to be implemented, along with the low-rate SPEAK. Like its predecessor, this processor has two programs, a single volume or sensitivity control, and a range of alarms and options. Further, the ESPrit3G has an inbuilt telecoil, along with various sound-optimisation and noise-suppression options such as '*Whisper*' (Glossary) (Cochlear Ltd, 2000; McDermott et al., 2002; Seligman, 2003a). Apart from the ESPrit, the other speech processors were used by subjects in this current study. A summary of the currently available speech processors is provided in Table 3.4.

Table 3.4: Comparison of currently-utilised Nucleus speech processors

	Sprint	Esprit	Esprit22	Esprit3G
Description	• Body-level processor for all current Nucleus implants • 4 programs • Supports high-rate strategies • Volume and/or sensitivity control • Auto Sensitivity Control option • LCD screen • Single or double battery pack • Stimulation up to 14400 Hz • Two specialised integrated circuits: (i) Analogue (ii) Digital signal processing • Allows input from FM or infrared systems, TV, or telecoil	• Ear-level processor for CI24M • 2 programs • Auto Sensitivity Control option • Volume or sensitivity control • 22 filters • Frequency analysis from 62 Hz to 10000 Hz • Power-saving features: reduces power usage when silent, decreases the stimulation rate in noisy environments, and can also adjust power according to skin flap thickness. • 2 Zinc-Air batteries	• Ear-level processor for CI22M • 2 programs • Volume or sensitivity control • 2 Zinc-Air batteries	• Ear-level processor for CI24R and CI24K • 2 programs • Inbuilt telecoil • Allows direct audio input, and input from induction loops or FM systems • Option of Auto Sensitivity Control (infinite compression), or Whisper noise-suppression (2:1 compression ratio) • Volume or sensitivity control • Stimulation up to 14400 Hz • 3 Zinc-Air batteries used
Current Strategies Supported	• SPEAK • ACE • CIS	• SPEAK • Low-rate ACE	• SPEAK	• SPEAK • ACE • CIS
Input Dynamic Range	• 32 dB	• 30 dB	• 30 dB	• 30 dB
Year Released	• 1997 	• 1998 	• 2001 	• 2002 

(Seligman, 2003a, 2003b) (Images courtesy of Cochlear Ltd.)

3.3 SPEECH-PROCESSING STRATEGIES

The speech-processing strategy plays a primary role in determining the resulting sound perceived by the implant user. Speech-processing strategies analyse and convert the acoustic input signal picked up by the microphone into electrical stimulation patterns. These electrical patterns are subsequently transmitted to the electrodes of the CI. Spectral information is represented as the variations in waveform amplitude across electrodes, with the temporal information being represented via the temporal fluctuations of the stimulating waveform presented at each electrode. The processing strategy enables an electrical representation of the input sound by defining various parameters of stimulation. These parameters include the rate of stimulation, the specific electrodes to be activated, the order of activation for these electrodes, the type of waveform or pulse to be used, the current amplitude, and the sampling rate (Glossary). These details form part of the patient's MAP, and are programmed into the speech processor by the clinician. As different strategies convey different features of the input sounds to the wearer, each with their own unique parameters of stimulation, the resulting sound perceived by the wearer could potentially differ from one strategy to the next. A brief outline of the principles underlying significant strategies follows. Although some of the strategies mentioned are now obsolete, these strategies provided the fundamentals from which both current strategies were derived, and future strategies may evolve. Further, some of these strategies were used by subjects in research studies discussed in Chapter 4. A more comprehensive review of speech-processing strategies along with sample speech perception results is detailed by Loizou (1998). For each of the major strategies below, references are also given which provide more information.

3.3.1 Analogue Stimulation-Based Strategies

3.3.1.1 Compressed Analogue (CA) strategy

The *Compressed Analogue* (CA) scheme was used in the now-obsolete Ineraid-Symbion and UCSF/Storz devices. It was the precursor to the more-contemporary Simultaneous Analogue Stimulation strategy described in the next section. In the CA strategy, the input signal is compressed with fast-acting automatic gain control

(Glossary) and usually filtered into four contiguous frequency bands covering the speech frequency range. Spectral and temporal patterns are represented by amplitude and temporal fluctuations, with the outputs of the filterbands being amplified before being simultaneously presented to each channel. However, there were several issues that impeded the effectiveness of this strategy. Firstly, although theoretically the presentation of maximal amounts of the original signal seems appealing, only a small proportion is actually used by the wearer due to physiological limitations of the human auditory processing system. For example, frequency modulations are only perceivable up to approximately 200 Hz to 400 Hz, above which the CI user is limited in their use of the temporal information to perceive pitch (Townshend et al., 1987; Vandali et al., 2005; Zeng, 2002). Secondly and more restrictedly, simultaneous stimulation has been shown to result in channel interactions which in turn decrease electrode independence, reduce the salience of channel cues, and cause uncontrolled and unpredictable loudness variations, all of which can impair speech perception and sound quality (Loizou, 1998; McKay, 2004; Wilson et al., 1991, 1993).

3.3.1.2 Simultaneous Analogue Stimulation strategy (SAS)

Based on the above CA scheme, the *Simultaneous Analogue Stimulation (SAS)* strategy was developed and is commercially available to Clarion implant users. Although no longer available with their current Hi-Res implants, SAS is still clinically utilised by some patients with the preceding CI models. The design and development of this strategy aimed to overcome the difficulties that arose with its predecessor by substituting the original input compression system (Glossary) of the CA strategy with a logarithmic channel-mapping function designed to convert the acoustic signal into appropriate levels for electrical stimulation. This function is applied to the output of each bandpass channel in order to try and provide a more-normal loudness growth function within each individual channel, and to maximise the use of the electrode's dynamic range. The back-end compression is combined with modiolar electrode placement and an updated electrode configuration, designed to limit the lateral spread of current along the cochlea. Initial implementations of SAS were not able to consistently achieve sufficient loudness levels due to the stimulation being too localised with the original bipolar stimulation mode. Hence the stimulation mode was modified to increase the spatial distance between electrodes to approximately 1.7 mm in order to provide sufficient loudness. SAS has seven channels available for selection, as opposed to four in the original CA strategy.

In its most common implementation, SAS divides the digitsed input signal into seven frequency bands via a filterbank, with the output of each filter being compressed into the patient's electrical stimulation range. These processed signals are then re-converted into a continuous analogue waveform for simultaneous activation of the electrodes. The default setting samples the signal at 13000 samples per second per channel, with a cumulative rate of 91000 samples per second for the seven channels. Compared with the CIS, ACE, and SPEAK strategies, SAS aims to provide more of the fine-temporal detail from the acoustic signal in the resulting electrical waveform (Kessler, 1999; Loizou et al., 2003; Wilson, 2004). Kessler (1999) provides more information on SAS.

3.3.2 Pulsatile Stimulation-Based Strategies - Feature-Extraction Strategies

Original attempts at speech-processing strategies with the Nucleus CI systems were based on physiological principles whereby fixed filters were used to model the auditory nerve's firing patterns. However, these resulted in loudness fluctuations from interacting electrical fields and concomitantly, poor speech perception. As a result, feature-extraction strategies were developed, based on the principle of extracting and transmitting components of the input signal deemed important for speech recognition. Most speech sounds primarily comprise a fundamental frequency (F0), and a series of harmonics (Glossary) above this. For example, voiced sounds, produced when air is forced through vibrating vocal cords, are periodic complex sounds having a low F0 and an acoustic spectrum containing harmonics over a wide range of frequencies. The resonant cavities of the vocal tract create formants – that is, frequency regions of high intensity in the speech spectrum. These formants, and the relationship between the formants, play important roles in aiding speech perception. For example, vowel recognition is primarily reliant on the perception of the first two formants along with durational cues. Consonant recognition is more complex and often requires the perception of both spectral and temporal components of the speech spectrum.

Derived from these underlying principles, the overall parameters of stimulation for feature-extraction strategies were formulated around the premise that loudness could be controlled by the amplitude of electrical stimulation, subjective pitch by the pulse rate of stimulation, and timbre, on a scale from dull to sharp, by electrode position. Accordingly, the amplitude of the acoustic signal was coded as electrical amplitude, voice pitch (i.e. the F0) as the pulse rate, and the speech formant frequencies as electrode positions. The pilot F0F2 strategy utilised formant-extraction algorithms, and was designed primarily to supplement lip reading, endeavouring to provide speech cues not available from lip reading. It presented three features of the speech signal – i) the amplitude of the incoming signal, via the amount of current or charge; ii) the F0, via the rate of biphasic pulse stimulation; and iii) the frequency of the second formant (F2), represented by varying the stimulation site along the electrode array. Later psychophysical research showed that amplitude envelope information was an important cue for consonant perception, and that the addition of amplitude and frequency estimations from the region corresponding to the first formant (F1) provided additional cues to aid detecting amplitude envelope changes, as well as extra voicing information (Blamey et al., 1984, 1987a, 1987b). Blamey et al. (1987a) also reported that F1 frequency information was better preserved in the presence of background noise than F2 frequency information. The F0F1F2 strategy, developed in the mid-1980's, presented frequency and amplitude information for the first two formants.

The culmination of these feature-extraction strategies was the *Multipeak* (MPEAK) scheme adopted in the late 1980's where amplitude information from three high frequency bands (2 kHz to 2.8 kHz, 2.8 kHz to 4 kHz, and 4 kHz to 6 kHz) was presented onto three fixed electrodes to provide additional information for consonant perception (Figure 3.10). For voiced sounds, the lower two bands' outputs stimulated the two more-apical of these three electrodes. For unvoiced sounds, the output amplitudes of all three high frequency filters were used to activate the three fixed electrodes, whilst the output amplitude of F2 activated a fourth electrode at a location corresponding to the frequency of this formant. Whereas the F0F2 and F0F1F2 strategies were implemented in the *Wearable Speech Processor* (WSP), an updated speech processor, the *Miniature Speech Processor* (MSP) was introduced in association with the MPEAK strategy. Whilst each subsequent strategy led to improved speech perception performance in quiet listening environments, all strategies proved largely unsatisfactory and ineffective for listening in background noise or with simultaneous speakers (Clark, 1995, 1997; Hollow et al., 1995; Swanson, 2003). These strategies were in common use with Nucleus CIs up until 1994, at which time they were largely superseded by the filterbank strategies.

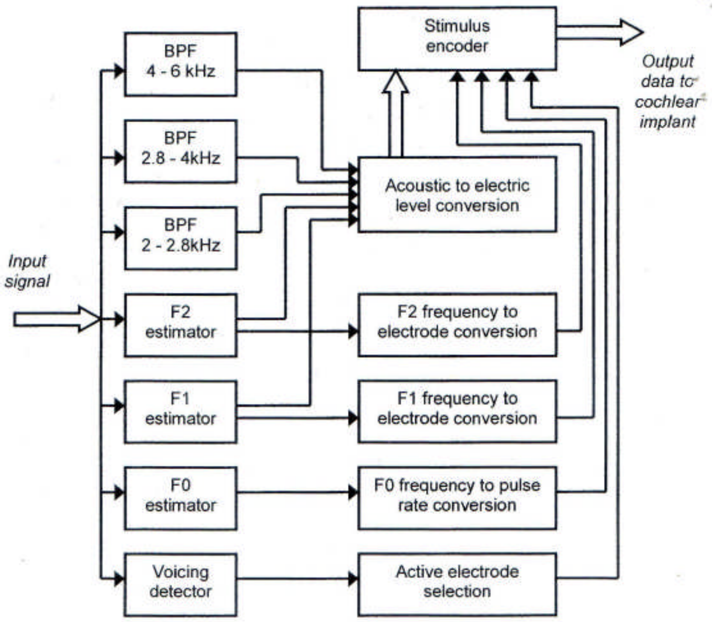


Figure 3.10: Schematic diagram of the MPEAK strategy
(Image from: McDermott (2004))

3.3.3 Pulsatile Stimulation-Based Strategies - Filterbank Strategies

Following from the above feature-extraction strategies, a new strategy was created for the Nucleus CI – the *Spectral Maxima Sound Processor* (*SMSP*) (McDermott et al., 1992; McKay et al., 1992). Along with SPEAK and ACE, these strategies can be broadly described as variants of an '*n-of-m*' strategy (Glossary) where during each cycle of stimulation, the 'n' largest signal envelopes from 'm' bandpass channels are presented to selected electrodes. To minimise redundancy, these strategies aim to present only the parts of the signals with the greatest amounts of spectral information as the parts of a signal with amplitudes significantly below the peak provide minimal, if any, information to assist speech perception (Wilson, 2000). In all filterbank strategies, the incoming sound is split into frequency bands, with each band allocated to one channel of the implant. The number and width of these filters not only varies for each strategy, but may differ from one implantee to the next depending on the type of implant, stimulation mode, electrode insertion depth, and the presence of electrode anomalies. In Nucleus implants, the bands are logarithmically spaced for frequencies above approximately 1000 Hz, with linear spacing below this to mimic the tonotopic arrangement of a normally hearing ear (Seligman & McDermott, 1995; Vandali et al., 2000; Wilson, 2000). The outputs of the filterbank are analysed to determine the channels with the greatest amplitude for each stimulation cycle, with these channels often being referred to as ‘maxima’ (Glossary). The number of channels used (m) and the number selected in each cycle (n), along with the rate of stimulation, varies for the different strategies. It should be pointed out that the above-mentioned filter spacing was based on research by Zwicker (1961) related to the critical band function. However, more-recent research by Glasberg & Moore (1990) indicates that this filter spacing may not be entirely correct, particularly for auditory filters with very low or very high centre frequencies.

For the SMSP strategy, the input signal was sent to 16 bandpass filters with centre frequencies from 250 Hz to 5400 Hz. In each stimulation cycle, the relative amplitude of these filters’ outputs was compared to obtain the six channels with the greatest amplitude. Each of these maxima was then assigned to one electrode with stimulation occurring in decreasing order of amplitude. The rate of stimulation was kept at a constant 250 pps per channel using non-simultaneous biphasic pulses. A detailed description of the SMSP strategy is provided in McDermott et al. (1992).

The three most commonly utilised filterbank strategies in Nucleus systems are currently SPEAK, CIS, and ACE. The current study involves users of either the SPEAK or ACE strategies. The three strategies vary in both the number of channels used, as well as the rate of stimulation. The CIS strategy uses fewer channels and a high rate, thereby relying more on temporal information, whereas SPEAK presents more spectral information at lower stimulation rates. ACE aims to combine the benefits of both high-rate stimulation and a greater number of channels to provide more temporal and spectral information. Figure 3.11 demonstrates the differences in the stimulation patterns provided by these three strategies, in response to the word “choice”.

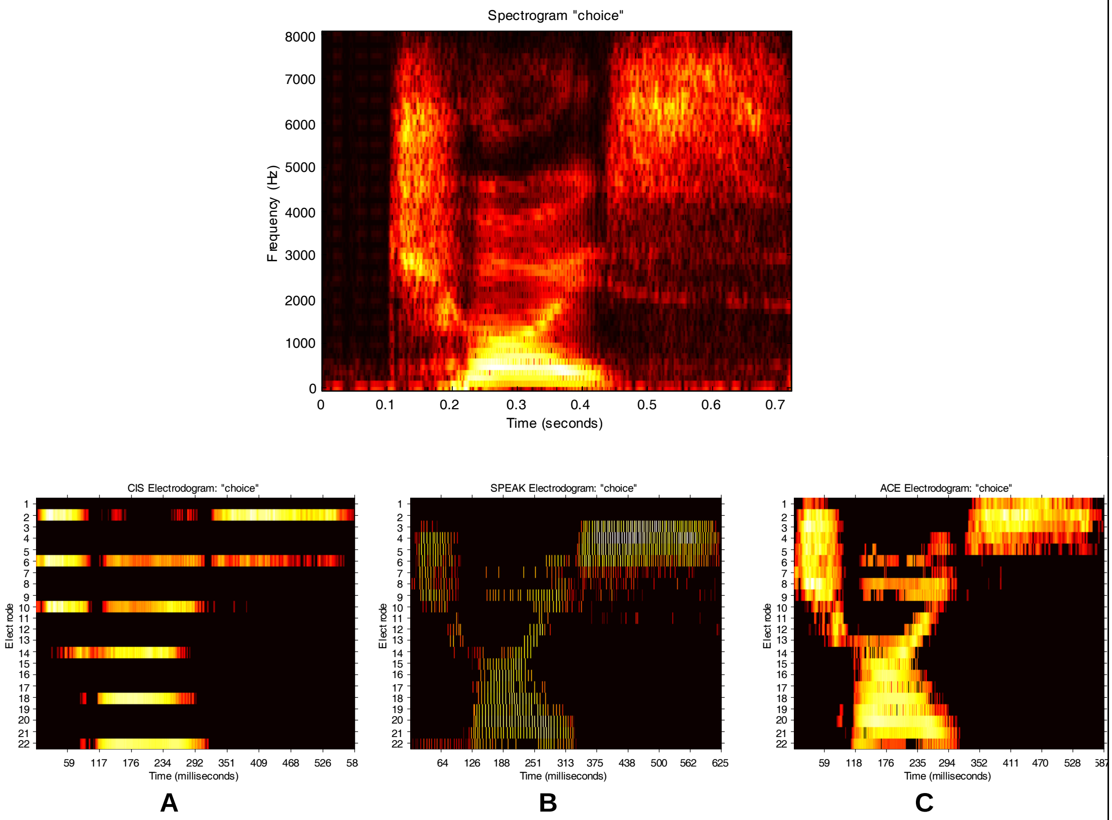


Figure 3.11: Comparison of the information provided by different speech-processing strategies

The larger top diagram is a spectrogram for the word “choice”. This plots the amount of energy in a frequency band as a function of time. Time is represented on the X-axis, frequency on the Y-axis, with the lighter colours of yellow and white representing areas with greater energy present.

The three diagrams below it are the resulting electrodograms for: A) the CIS strategy; B) the SPEAK strategy; C) the ACE strategy. The electrodogram presents similar information to the spectrogram, with the X-axis representing time, the Y-axis representing the individual electrodes of the CI, and the colour reflecting the rate of stimulation for each electrode.

The CIS representation has only 6 activated electrodes, with high rates of stimulation for each. SPEAK has a greater number of activated electrodes, but each is stimulated at a lower rate. ACE has both a large number of activated channels, along with high stimulation rates on each electrode.
(Figures courtesy of B. Swanson)

3.3.3.1 Spectral Peak Strategy (SPEAK)

The SMSP was the prelude to the currently utilised SPEAK strategy which increased the number of filters to 20, and allowed for a variable number of maxima to be selected, usually between six and nine. Used in the Nucleus speech processors, SPEAK is a ‘roving’ stimulation strategy dependent on having many stimulation sites (usually 20 channels) in conjunction with a low stimulation rate around 250 Hz. Compared to the CIS strategy, SPEAK focuses on providing more spectral, as opposed to temporal, cues for the listener. Each channel of the implant is allocated to a separate frequency band, encompassing a frequency range from approximately 116 Hz to 8000 Hz, with stimulation occurring for the 1 to 10 channels deemed to have the most energy for that stimulation cycle (Figure 3.12). The filterbank is programmable (i.e., the filter scaling can be varied), however in its default setting, the 20 bandpass filters are linearly spaced below 1850 Hz, and logarithmically spaced above this. At the output of each filter is an amplitude detector programmed to detect the output's peak – filter outputs with the greatest amplitude are flagged as one of the maxima (Glossary). The filterbank is then scanned for the preset number of channels with maximal output; these are then used to generate the stimulating pulse train (Seligman & McDermott, 1995). In SPEAK, the order of stimulation is tonotopic from base to apex, whereas in the SMSP strategy, stimuli were presented to the selected electrodes in decreasing order of amplitude. The number of maxima selected per stimulation cycle (Glossary) is dependent upon the spectrum of the input sound; a minimum amplitude level at the filter's output must be reached for stimulation to occur. Thus battery power is conserved by not stimulating channels with little energy as these are unlikely to aid speech perception. SPEAK adopts an adaptive stimulation rate that varies according to the number of maxima selected in each cycle and the parameters set in the processing program - increasing the number of maxima decreases the rate. Similarly, if less maxima are selected, less electrodes are activated in that cycle, and the rate increases to compensate. In effect, this enables the strategy to run at its maximum possible speed (Seligman & McDermott, 1995). The most common implementation of SPEAK in the SPrint device consists of six maxima selected in each cycle out of 20 channels, at a cycle rate averaging 250 Hz . For the ear-level ESPrit devices, eight maxima may be used (Co-operative Research Centre for Cochlear Implant and Hearing Aid Innovation, 2003a; Seligman, 2003a). The stimulus rate 'jitters' (Glossary) around a 250 Hz average to help eliminate a low-frequency pitch percept that may be present when using constant low-rate stimulation (Loizou, 1998; Seligman & McDermott, 1995; Skinner et al., 1994; Wilson, 2000). Further detail on the SPEAK strategy is provided by Seligman & McDermott (1995), and Skinner et al. (1994).

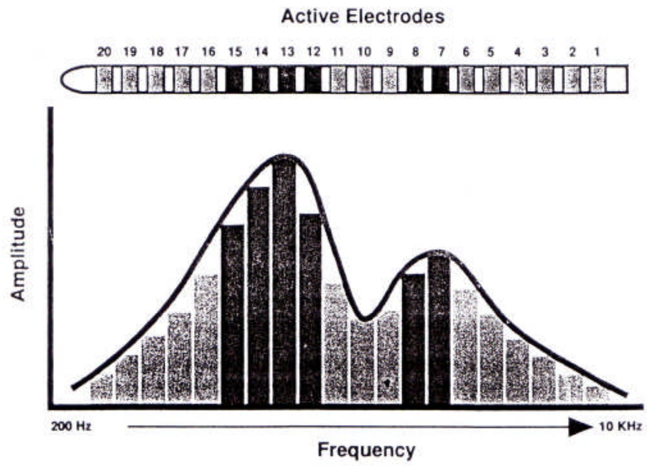


Figure 3.12: Diagram representing maxima detection in the SPEAK strategy
(Image from: Seligman & McDermott (1995))

3.3.3.2 Continuous Interleaved Sampling (CIS)

Unlike the spectral focus of the SPEAK strategy, the CIS strategy prioritises temporal information by employing a high stimulation rate across a small number of channels. Research with this strategy has indicated that the use of more than eight channels is unlikely to significantly improve speech perception scores in quiet listening environments (Dorman et al., 1997; Loizou et al., 1999). A schematic diagram of the CIS strategy is provided in Figure 3.13. Similar to the SPEAK and ACE strategies, the input signal is pre-emphasised with frequencies below 1.2 kHz being attenuated at a rate of 6 dB per octave (Glossary) to allow the softer, higher-frequency speech components to be more audible. For the CIS strategy, the output of this pre-emphasis filter is then divided into a smaller number of frequency bands (commonly six or eight wider bands, as opposed to the 20 narrower bands of SPEAK), with fixed-rate sampling of the filterbank's output usually between 833 pps and 1111 pps per channel. The filter outputs are compressed to fit the dynamic range of electrical hearing before being converted into biphasic pulse trains. The energy within each band (i.e. the filter's output) determines the amplitude of the pulse train delivered to selected electrodes with variations in the pulse amplitude representing the level variations present in the input signal. Stimulation occurs sequentially from base to apex using interleaving biphasic pulses at a high rate, usually above 900 Hz, to adequately represent waveform modulations (Loizou, 1998; Wilson, 2000). Unlike the SPEAK strategy where the activated electrodes vary from one time-window to the next, the CIS strategy stimulates fixed electrode sites at a constant rate; the exact rate used depends upon the pulse duration and interpulse interval. CIS, ACE, and SPEAK do not extract or represent specific features of speech such as the F0, nor do they separate voiced from unvoiced sounds. CIS uses the temporal envelope information at the output of each bandpass filter to vary the amplitude of the stimulating pulse train, and employs high stimulation rates to ensure adequate representation of these waveform modulations (Wilson, 2000). The development of CIS is described by Wilson et al. (1993).

CIS+, currently implemented in the Med-El speech processor, is a modified version of the standard CIS strategy utilising a Hilbert transform and a wider frequency range when compared to the preceding models (Med-El, n.d.). According to the manufacturer, the Hilbert transform replaces the Fast Fourier transform, and the wave rectification techniques of other CI systems, to provide a more accurate representation of the temporal information from the original signal (Med-El, 2006).

The *HiRes* strategy is the default option in the current Clarion *HiRes* implant. A derivative of the CIS strategy using higher stimulation rates, HiRes divides up the incoming sound into a maximum of 16 channels. The envelope information from each of these channels is used to modulate a biphasic pulse train, delivered to the corresponding intracochlear electrodes (Spahr et al., 2005). Either sequential (HiRes-S) or simultaneous paired (HiRes-P) stimulation is available. Further information on this strategy can be obtained from Frijns et al. (2003) and Koch et al. (2004). The CIS strategy, including its derivatives such as CIS+ and HiRes, use all available channels in each cycle of stimulation, as opposed to ‘roving’ selection strategies such as SPEAK or ACE, where a predetermined number of channels are used for each stimulation cycle. That is, these latter strategies may not use all of the available channels in each cycle of stimulation.

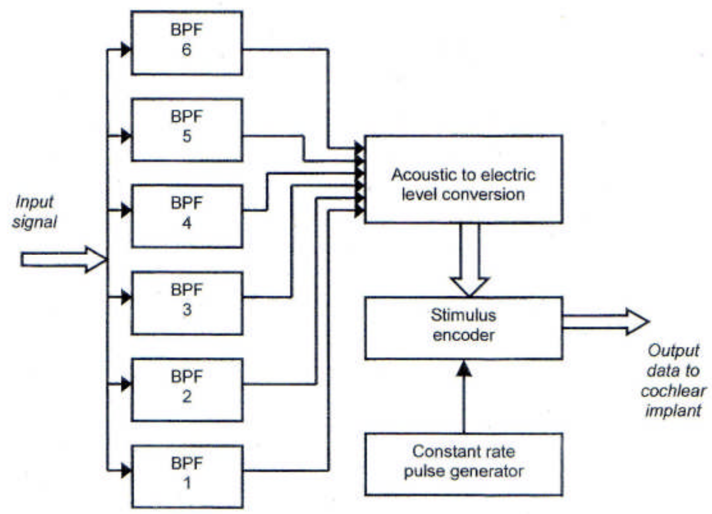


Figure 3.13: Schematic diagram of the CIS strategy
(Image from: McDermott (2004))

3.3.3.3 High Stimulation Rates

Both the CIS strategy and the ACE strategy (to be discussed in the next section) use higher stimulation rates than the SPEAK strategy. CIS employs this high rate across a small number of channels, with ACE incorporating the high stimulation rate in conjunction with a greater number of channels. With manufacturers continually promoting the high total stimulation capabilities of their device, it is worth considering the effects of higher stimulation rates in providing temporal cues for the wearer. There is usually a trade-off between stimulation rate and the number of channels that can be used per stimulation cycle; fewer channels potentially decrease spectral detail with lower rates providing less temporal information. It should also be noted that as higher stimulation rates require greater power consumption, it is not as straight-forward as utilising the highest possible stimulation rate.

In sequential pulsatile stimulation strategies, the bandpass filters' outputs are smoothed, thereby retaining only the temporal envelope information from the input acoustic stimulus. Higher rates sample this envelope information more often to provide a better representation of the temporal cues, however the extent to which this may benefit speech recognition is unclear. Loizou et al. (2000) reported that higher pulse rates improved open-set speech recognition for the CIS strategy, with most of their subjects obtaining their best performance when using the highest rate tested, 2100 pps. On the other hand, Vandali et al. (2000) found that stimulation rates greater than 250 pps did not result in any significant improvement in speech perception scores. These subjects were using the ACE strategy. However, in that same study, most of the subjects indicated a subjective presence for the highest rate condition, 1615 pps, for listening to music. It was speculated by the authors that a higher rate may improve sound quality without having a direct effect on speech intelligibility per se. High levels of intersubject variability have been reported in numerous studies, though, with no one particular rate providing all subjects with optimal performance (Vandali et al., 2000; Wilson et al., 1995, 1997).

Research by Rubinstein et al. (1999) and Wilson et al. (1997) also suggested that the use of very high pulse rates may enable neural firing patterns to more closely resemble that which occurs for acoustic hearing by allowing stochastic resonance to be incorporated into the neural discharge patterns for electrical stimulation. This would theoretically improve temporal resolution. Unlike acoustic stimulation of hearing, electrical stimulation results in highly deterministic neural firing patterns, where the firing patterns are tightly phase-locked to the stimulus at low pulse rates (Wilson et al., 1997). That is, the stochastic properties associated with acoustic hearing are not observed for electrically stimulated hearing at low pulse rates. However, Rubinstein et al. (1999) and Wilson et al. (1997) showed that high-rate stimulation, above approximately 4000 pps per electrode resulted in desynchronisation between the neural response and the stimulation rate, allowing shorter temporal intervals to be coded. This should thereby enable neural responses to more accurately follow finer variations in the temporal envelope.

Another research finding of relevance to this discussion are reports by Busby et al. (1993) and McKay et al. (1994) that carrier frequency rates less than approximately four times the modulation frequency of the stimulus diminish the accuracy for perceiving the modulation frequency. Low carrier frequency rates do not allow the modulation pattern to be sufficiently sampled in order to provide reliable pitch cues. Thus for a stimulus with a F0 of 200 Hz, stimulation rates greater than 800 pps would be required (McKay et al., 1994).

3.3.3.4 Advanced Combination Encoder (ACE)

This ‘n-of-m’-type strategy is often described as combining the features of the CIS and SPEAK strategies. It uses both a large number of channels, most commonly 22, in conjunction with a high stimulation rate to provide the combined benefits of additional temporal and spectral cues to the wearer, as opposed to prioritising one of these cues over the other. The balance between stimulation rate and the number of maxima selected can be optimised for the individual. The flexibility of the ACE strategy enables its implementation with either an electrode selection technique as used in SPEAK, or by using all available channels in each stimulation cycle similar to CIS. The precise parameters are stipulated by the clinician when mapping the processor. A schematic diagram of this strategy is provided in Figure 3.14.

The input signal is analysed to create ‘m’ filterbands (usually 22) with these bands being linearly spaced from 188 Hz to 1312 Hz, and logarithmically spaced above this to 7938 Hz. Adjacent filterband outputs may be combined if required. ‘n’ of these ‘m’ bands are then selected for each time-window, based on the output amplitude of each filterband. The corresponding electrodes are then tonotopically activated with an amplitude conversion function (Glossary) being applied to ensure that ensuing stimulation levels are appropriate for, and within the dynamic range of the wearer (Vandali et al., 2000). In summary, ACE provides the following options programmed into the speech processor: up to 22 filterbands to cover the input frequency range, as opposed to 8 to 12 for CIS; up to 20 maxima, in contrast to the 6 to 9 maxima for SPEAK; and a stimulation rate between 250 pps to 2400 pps per channel, as opposed to 250 pps for SPEAK. More information pertaining to the development of this strategy is available from Vandali et al. (2000).

In implementing the ACE strategy with the CI24 system, there is a maximum total stimulation rate across all channels of 14400 pps. The ACE strategy is only available in the SPrint or ESPrit3G devices, most commonly with 8 to 10 maxima being selected across the 22 channels. Stimulation rates vary widely from 250 Hz to 2400 Hz, although rates above 1800 Hz are less common due to the higher power consumption requirement (Co-operative Research Centre for Cochlear Implant and Hearing Aid Innovation, 2003a; Seligman, 2003a). Moderate-rate ACE (900 Hz) is currently the default choice for the Nucleus 24 devices in view of its flexibility and reported higher speech perception scores than obtained with previous strategies such as SPEAK (Holden et al., 2002; Skinner et al., 2002).

The Med-El Combi40 implant implemented a version of this strategy called the "*N-of-M*" strategy. It was essentially very similar to the ACE strategy based on the same high-rate approach. However, "N-of-M" used 8 or 12 channels, rather than the 22 channels available with ACE, with overall stimulation rates up to 12500 pps (Loizou, 1998).

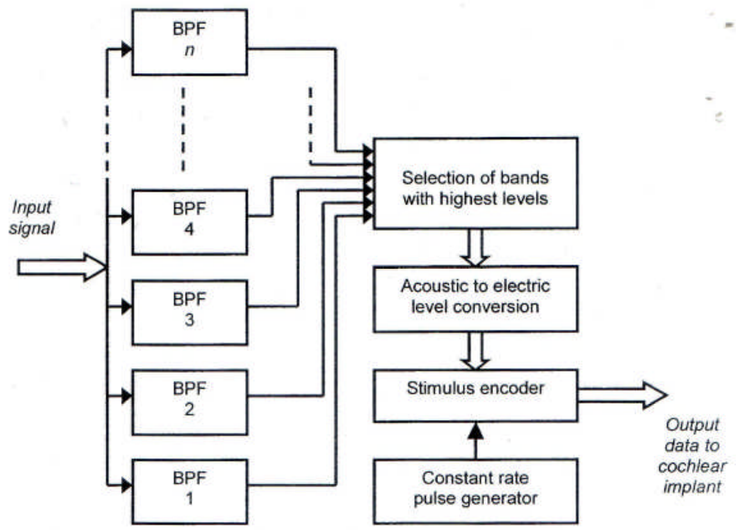


Figure 3.14: Schematic diagram of the ACE strategy
(Image from: McDermott (2004)).

3.3.4 Other Hybrid Or Combination Strategies

There are a range of other experimental high-rate, hybrid, and combination strategies currently utilised or being trialled with the Clarion CI. These include the *PPS* (Paired Pulsatile Sampler), *MPS* (Multiple Pulsatile Sampler), and *HAP* (Hybrid Analogue Pulsatile) strategies. The first two are derivatives of CIS, but activate two or more channels simultaneously. The third implements both SAS and CIS concurrently with some channels being preselected for CIS stimulation, and others for SAS stimulation. Usually the CIS-allocated channels are associated with processing the lower-frequency components of the sound using monopolar stimulation, whilst the higher-frequency information is processed by the SAS-allocated channels using bipolar stimulation. This strategy is not widely available, and is not implementable in their current commercial speech processor. Kessler (1999), Loizou et al. (2003), and Wilson (2004) provide further descriptions of these essentially prototype strategies.

3.4 THE PERCEPTION OF MUSICAL STIMULI WITH A COCHLEAR IMPLANT

Integrating the information presented in Sections 2.3 and 2.4 along with the preceding section on CIs draws attention to some of the issues and difficulties that may arise when listening to music through a CI. A listener’s ability to accurately perceive music with a CI would also be further affected by variables such as the type of implant used, speech-processing strategy employed, audiological history, and length of hearing loss, to name a few (Limb, 2000). The manner in which the implant codes the different characteristics of the signal will affect perceptual attributes such as the pitch, timbre, and loudness of the sound.

The use of electrical stimulation, as occurs with a CI, results in a different sound percept than that experienced through acoustic hearing. Existing research indicates that those with a cochlear hearing loss, including CI users, have temporal resolution skills equivalent to those of the normally hearing population (Moore & Glasberg, 1988b, 2001; Shannon, 1989, 1992). It is the discrepancy between these populations on frequency- and spectral-based tasks that has the greatest impact on music perception.

In current filterbank speech-coding strategies, only the temporal envelope information is retained, with the fine-frequency information being eliminated. A bank of bandpass filters divides the input signal into a number of frequency bands. The envelope information is then extracted via full wave rectification and low-pass filtering within each band, usually using a filter between 200 Hz and 400 Hz. This results in the elimination of temporal fine-structure detail, with the remaining envelope information being used to modulate the fixed-rate pulse train (Loizou, 1998). Research by Smith et al. (2002) has suggested that the fine-structure information may be more important for pitch perception than the envelope cues.

Accurate perception of western music requires the listener to discriminate frequency modulations as small as 6%, which corresponds to approximately one semitone (Glossary). Pitch perception is fundamental to melody recognition and music appreciation. There are two main ways that users of multi-channel CIs perceive pitch. The first is via the temporal domain where either modulating the amplitude, or changing the rate of the stimulating pulse train can provide pitch cues to the listener. The percepts elicited by these two methods share various similarities and, at low rates or modulation frequencies, both can provide reliable and largely predictable pitch percepts (McKay et al., 1994, 1995). However, variations of the modulation frequency will only provide a reliable pitch change if the carrier rate is sufficiently high to adequately sample the temporal waveform. A carrier rate less than approximately four times the modulation frequency may result in spurious pitch percepts (McKay et al., 1994, 1995). Pitch cues can also be provided via the spectral domain, with changes in the place of stimulation also giving pitch information (Busby et al., 1994; McDermott & McKay, 1994; McKay et al., 1996).

Research has indicated that for CI subjects, increasing the rate of steady pulse trains delivered to single electrode sites from approximately 50 Hz to 300 Hz is associated with a corresponding increase in perceived pitch (Eddington, 1980; Fearn & Wolfe, 2000; McKay & McDermott, 1996; Pijl, 1995, 1997b; Tong & Clark, 1985; Townshend et al., 1987; Vandali et al., 2000). For pulse rates below 50 Hz, a buzz-like sound with no salient pitch is often reported, with rate increases above 300 Hz providing little change in the perceived pitch. Pijl & Schwarz (1995) activated individual electrode pairs with pulse trains where the pulse rate was determined by the F0 of well-known melodies. They reported that, even without rhythm cues, subjects could identify melodies when the pulse rate varied corresponding to the pitch variations in the melody. Results were best with the lowest fundamental rates assessed (75 pps and 100 pps), decreasing to chance levels for the fastest rate assessed (400 pps). Performance was also significantly better with apical electrode stimulation than basal stimulation. In a different study, Pijl (1995) similarly found melody identification scores to be better for pulse trains delivered to apical electrode sites than basal sites, with subjects affirming apical stimulation to sound more musical. A second experiment by Pijl & Schwarz

(1995) demonstrated that for low pulse train modulation rates, variations in the pulse rate could be perceived as musical intervals when organised into pairs adopting the same ratio difference as that used in western music. A subsequent study by Pijl (1997a) showed that at low pulse rates, these intervals could be perceived and labelled by the CI subjects with comparable accuracy to that achieved by musically untrained subjects with normal hearing. Overall, existing research indicates that musical pitch information can be conveyed via the temporal parameters of electrical stimulation over a limited range of low stimulation rates.

Most current speech-processing strategies, however, use pulse trains delivered at a constant, relatively high rate. They do not vary the stimulation rate as a consequence of features in the input signal, as was the case with earlier feature-extraction approaches described in section 3.3.2 where an estimate of the F0 was used to control the stimulation rate (Clark, 2003; McDermott, 2004). Although the pulse rate does not vary, psychophysical studies have demonstrated that pitch information may alternatively be obtained from the variations in the pulses' amplitude. The amplitude modulation depth is derived from the estimated amplitude of the input signal envelope within a frequency band, and confined to within the boundaries of the 'T' and 'C' levels mapped for each electrode. Amplitude-modulated pulse trains delivered at relatively high rates to the implant provide rapid temporal fluctuations in the electric stimuli; these fluctuations can provide a pitch percept which can be used to convey musical information (Geurts & Wouters, 2001; McKay & McDermott, 1996; McKay et al., 1994, 1995; Pijl, 1995, 1997b). Amplitude modulations may be particularly important when spectral cues are degraded, as is often the case when listening to complex acoustic stimuli through a CI (Shannon et al., 1995, 2004). CI users appear able to derive reliable pitch cues from amplitude modulations in the low modulation frequency region only, which is similar to the limitations associated with pitch cues elicited from varying the pulse rate (McKay et al., 1994, 1995). With the upper boundary for temporal pitch coding being around 300 Hz, many CI subjects would have difficulty obtaining reliable pitch cues from temporal variations in stimuli with a F0 above approximately middle-C (McKay, 2004; Pijl, 1997a; Pijl & Schwarz, 1995; Shannon et al., 2004; Tong & Clark, 1985; Zeng, 2002).

The salience of amplitude-based pitch cues is dependent upon a sufficient modulation depth and a high carrier rate (Geurts & Wouters, 2001; McKay et al., 1994, 1995). As mentioned earlier, it has been recommended that the carrier pulse rate of the speech-processing strategy should be at least four times greater than the modulation frequency of the stimuli. At lower carrier rates, envelope periodicity cues are less reliable as they are more sparsely sampled, and hence less detail is present in the modulation waveform (McKay et al., 1994, 1995). Geurts & Wouters (2001) examined the effect of modulation depth on pitch discrimination using two sinusoidally-modulated pulse trains presented to a single channel. Their results indicated that pitch discrimination deteriorated with smaller modulation depths, although the critical value of the modulation depth required for reliable discrimination varied between subjects. There was also a saturation effect noted - a point above which subsequent increases in modulation depth did not result in further improvements in discrimination scores. The authors noted that performance was better for the lower modulation rate 150 Hz condition than the 250 Hz comparison; increased modulation depths were required at 250 Hz to achieve equivalent performance levels to those at 150 Hz.

Another important factor impacting on an implantee's ability to extract pitch information from the waveform relates to the consistency of the alignment of the phase of the amplitude modulations across electrodes. Phase misalignments can inhibit the salience and consistency of pitch cues as the periods of these modulations are no longer in phase across the electrodes (Figure 3.15). That is, the location of the peaks on one electrode may not be consistently aligned with the peak location on another electrode (McKay, 2004; Moore, 2003a; Shannon et al., 2004). The perceptual consequence of these misalignments will depend on whether the CI user extracts pitch information from such amplitude modulations independently for each electrode, in which case pitch perception would not be altered, or if they integrate the information across the activated electrodes. In the latter case, phase differences could negate each other and the perceived modulation pattern may no longer reflect the F0 (McDermott, 2004). Psychophysical research indicates that temporal information is combined when electrodes are closely spaced. Once the electrode separation has exceeded a certain distance, the degree of which varies between individuals, temporal patterns are perceived independently. In other words, the effect of phase shifts may be negligible only for widely spaced electrodes (McKay & McDermott, 1996).

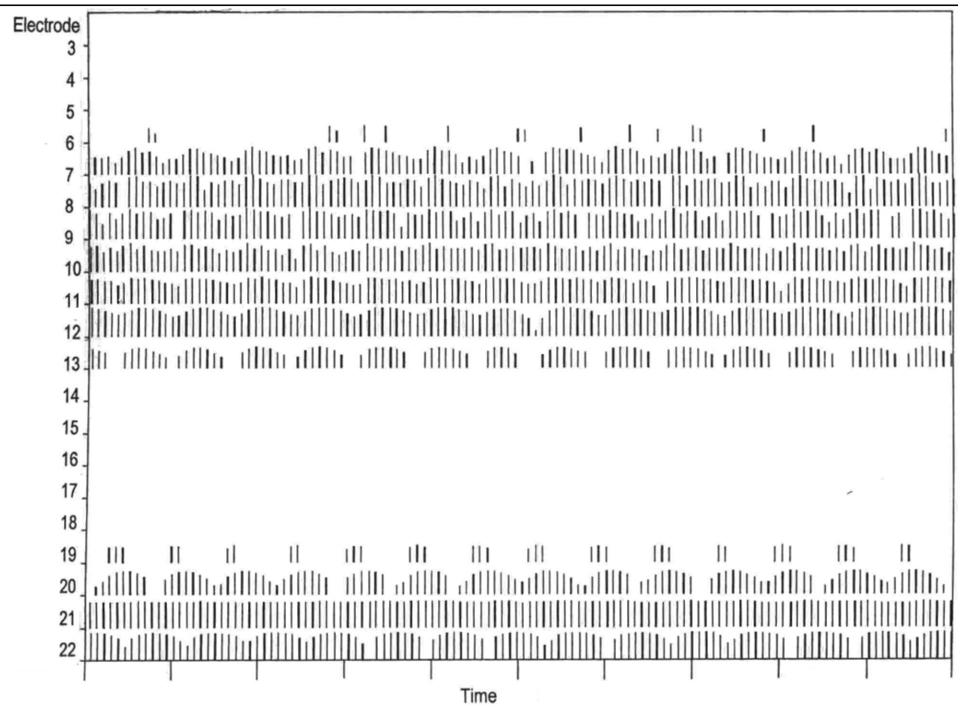


Figure 3.15: Electrodogram showing phase misalignments
Electrodogram for the vowel /i/, sung by a male singer at a F0 of approximately 139 Hz. The numbers on the Y-axis correspond to the implanted electrodes of a Nucleus CI24 implant, implemented with the ACE speech-processing strategy. 22 is the most-apical electrode, and 3 is the most-basal. The X-axis represents time, with a total duration of 100 ms. Amplitude modulations present on each activated electrode can provide a pitch percept. However, these modulations may not necessarily be aligned across the electrodes, as can be observed for the 3 most-apical electrodes in this figure.
(Figure courtesy of D. Tsang)

There is a large difference in the ability, and range over which implantees can extract pitch information from temporal variations in the electric stimuli. For example, some implantees can discriminate small changes in the rate of steady pulse trains over a wider range than others (Shannon et al., 2004; Tong & Clark, 1985; Zeng, 2002). Although the precise reason for this is unclear, factors such as differences in musical ability, memory for melodic pitches, variable effects of auditory deprivation, central processing issues, pitch processing deficits, as well as the distribution of surviving auditory neurons may impact on performance in such psychophysical tasks (McKay, 2004; Pijl, 1995).

In addition to these temporal cues, changes to the location where the electrical stimuli are delivered can also contribute to the perception of pitch for CI users (Figure 3.16). Fearn & Wolfe (2000) propounded that for current pulsatile speech-processing strategies, both rate and place cues will contribute to a subject's perception of pitch when low stimulation rates are used, however the salience of these rate cues will diminish for higher stimulation rates. Therefore pitch percepts for higher frequencies would be largely based on place cues.

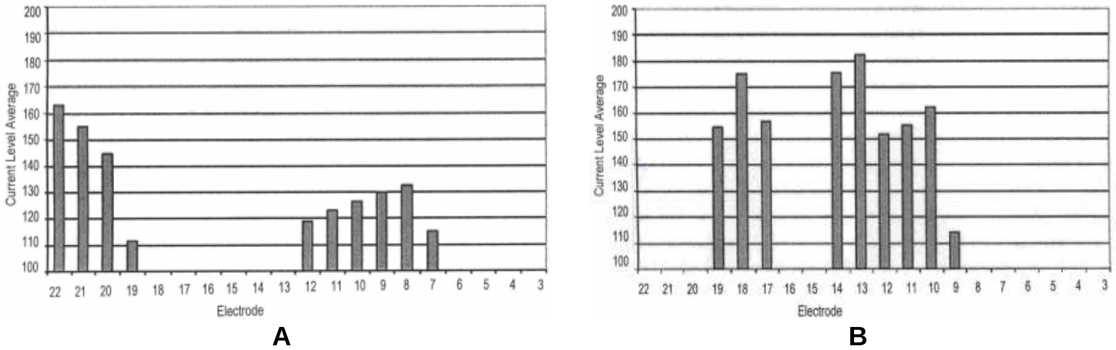


Figure 3.16: Change in the place of stimulation for different F0s

This figure demonstrates the change in the current level distribution across electrode positions for two stimuli with different F0 sung by two different singers, one male and one female. The stimulus for both sounds was the sung vowel /i/. A) was sung by the male singer, at a F0 of approximately 98Hz. B) was sung by the female singer at a F0 of approximately 740 Hz. The X-axis corresponds to the 22 electrodes of the Nucleus CI24 implant, where 22 is the most-apical, and 3 is the most basal electrode. The Y-axis represents the average current level of stimulation at that electrode site. The speech-processing strategy was the ACE strategy.

A basal shift in the place of stimulation can be observed from the lower F0 for A to the higher F0 of B. (Figures courtesy of D. Tsang)

The main frequencies for speech perception from 500 Hz to 3000 Hz correspond to a 14 mm length of the cochlea between 10 mm and 24 mm beyond the round window (Gantz, 1987). In the afore-mentioned studies investigating the role of temporal cues for pitch perception, comparisons were made for differing stimulation rates at single electrode sites. Similarly, the role of place cues can be assessed by using constant-rate pulse trains delivered to differing electrode sites to activate different neural populations. This tends to provide a range of pitch sensations often described on a 'sharp' to 'dull' scale (Clark, 2003). Research has demonstrated that electrical stimulation of different cochlear locations with constant-rate pulse trains can provide differing pitch percepts, largely corresponding to the cochlea's tonotopicity (Busby et al., 1994; McDermott & McKay, 1994; McKay et al., 1996; Townshend et al., 1987; von Wallenberg et al., 1995). Generally, pitch increases when a constant-rate pulse train activates a more-basally positioned electrode, although the ability to distinguish the pitch of one electrode from another may be dependent upon the degree of spatial specificity when activating individual electrodes (Busby et al., 1994; Moore, 2003a; Tong & Clark, 1985; Townshend et al., 1987). Multi-channel CIs are based on the assumption that different electrodes stimulate non-identical neural populations in order to provide a range of pitch percepts. There is limited evidence from a case-study by McDermott & McKay (1997) that changing the stimulation site whilst maintaining the same stimulation parameters can provide some degree of musical pitch information.

There is often a mismatch in the tonotopic place of stimulation, particularly for the low frequencies, with stimulation potentially occurring at a more-basal location than what occurs for normal acoustic hearing. This is due to the implanted electrodes not usually being inserted deep enough to stimulate regions below around 1000 Hz. For example, a 500 Hz signal may stimulate a location on the basilar membrane normally tuned for 1000 Hz, thereby creating a frequency-to-place mismatch in the frequency information presented to the implantee. A study by Oxenham et al. (2004) demonstrated the importance of place-specific information in extracting the F0 from a complex sound. In their study, when temporal information of harmonic stimuli was presented to an incorrect tonotopic location along the basilar membrane, subjects were largely unable to perceive a salient pitch sensation. It was proposed that the usual higher cortical-level processing of pitch does not occur in this situation. Hence it may be that the presentation of temporal information stimulating incorrect basilar membrane locations may impede CI users' perception of pitch. The ability of CI users to perceive temporal information in the range between 50 Hz and 500 Hz is imperative to their pitch perception as it is at these low frequencies that the implant user may be able to derive the most reliable pitch cues from amplitude modulations representing the F0 (Kong et al., 2004; McKay, 2004; Oxenham et al., 2004; Shannon et al., 2004). Studies of pitch ranking using sung vowel stimuli in conjunction with the SPEAK speech-processing strategy have tended to show improved pitch-ranking ability with lower F0 stimuli, particularly when the F0 was below the upper temporal pitch limit for both tones (McDermott, 2004).

The nature of the relationship between the spectral information in the signal and the resulting place of stimulation is impacted on by a multitude of issues. Unlike some of the above-mentioned research where stimulation was evoked on single electrodes only, in processing a complex acoustic signal, there is activity on several electrodes concurrently. This gives rise to the potential of channel interactions, decreased electrode independence, and reduced spatial specificity. As pointed out by McDermott (2004), this may even occur for puretones if the frequency in question crosses over between adjacent filters, causing the excitation of neighbouring electrodes. Similarly, the longitudinal spread of current in the cochlea may result in a large population of auditory neurones being excited, thereby decreasing the specificity of place-pitch cues (Shannon et al., 2004; Townshend et al., 1987).

A host of other variables can impact on a subject’s ability to use place cues to perceive pitch. These include those related to the electrode (e.g. insertion depth, placement, and miscellaneous anomalies), the speech processor (e.g. the processing strategy specifications, stimulation mode, or current path), interaction with other features of the stimuli (such as loudness levels, or pulse duration), and patient factors (e.g. pathological processes, auditory neuron survival, neural density, tissue impedance surrounding the array, and the distribution of target neurons relative to the activated electrode location) (Busby et al., 1994; McDermott, 2004; McKay, 2004).

Spatial and temporal cue-based percepts are independent – that is, stimulation of individual electrodes using pulse trains gives rise to two simultaneous but different percepts, one rate-related, the other place-related (McKay et al., 2000; Moore & Carlyon, 2005). Interpretation of these two percepts utilises two separate processing mechanisms (Tong & Clark, 1985). It has been proposed that place changes may contribute more to changes in perceived timbre than pitch per se (McDermott & McKay, 1997; McKay, 2004; Moore & Carlyon, 2005). Several authors have noted that it is often hard for subjects to differentiate between differences in pitch and timbre (Beal, 1985; Krumhansl & Iverson, 1992; Pitt, 1994; Pitt & Crowder, 1992; Warrier & Zatorre, 2002). Pitt (1994) reported that timbre appeared to be a more-dominant dimension than pitch, particularly for non-musicians. Musicians were more accurate in processing the two dimensions independently than non-musicians. The latter subject group had particular difficulty differentiating timbral changes when the pitch remained the same, frequently reporting that both dimensions had changed. Both Erickson (2003) and Handel & Erickson (2004) reported that musically trained and musically untrained subjects had difficulty ignoring pitch variations in making timbre-based decisions. However, it would be very difficult for researchers in psychophysics to definitively differentiate whether subjects were judging pitch or timbre changes, or a combination of both, when undertaking pitch perception tasks (McDermott, 2004). Thus, should the pitch percept be unclear to the subject and/or should the variation of timbre be a more-dominant percept, it is possible that a CI user may respond to the change of timbre as opposed to pitch per se.

Pitch perception is significantly worse when subjects listen with their speech processor, as opposed to when pulse trains are directly presented to the electrode array (Pijl, 1997a). It appears that whilst processing techniques implemented by these strategies provide most implant users with efficient and reliable cues for speech perception, the same processing techniques also impede the wearer from effectively extracting F0 information from a complex sound, thereby having an adverse affect on pitch perception. This reduced ability to extract the F0 information may stem from a range of factors. In addition to the issues raised in the above discussion on pitch perception with electrical stimulation, current speech-processing strategies usually examine the envelope modulations of the input signal rather than the rapidly varying components of the fine structure (Oxenham et al., 2004; Rubinstein & Hong, 2003). Smith et al. (2002) suggested that the provision of the fine-structure information may be beneficial in improving pitch perception for CI users, with their research amongst normally hearing subjects indicating that this fine-temporal information may be more important in providing pitch cues than the envelope cues. However, the extent to which CI users can perceive this fine-temporal information is unclear (Lobo et al., 2002; Wilson et al., 2004a). Research indicates that 300 Hz is approximately the upper limit for the temporal code with most CI users only being able to reliably perceive stimulation rate variations up to around 300 pulses per second (McDermott & McKay, 1997; Townshend et al., 1987), or detect pitch differences for frequencies up to approximately 300 Hz (Zeng, 2002). This suggests that even if the fine-structure information could be reliably presented to CI users, the listener may only be able to perceive the gross features of the information due to perceptual limitations (Wilson et al., 2004a).

Another factor related to the pitch perception of CI users is that current processing strategies vary both the spatial and temporal properties of the stimuli, depending upon the input signal. The combined perceptual effect of these variations is often unpredictable and highly deviant between individuals. If the cues are not consistent, the listener may get contradicting information regarding the direction or nature of the pitch change.

Although the filterbanks employed in current implant systems were largely designed to mimic the filtering properties of the basilar membrane as best as possible, the two filtering mechanisms also differ in several key ways. Unlike the non-linear, level-dependent auditory filters of a normal cochlea, with their continuous centre frequencies, the filterbanks of the major commercial implant devices involve a pre-determined number of overlapping filterbands (up to 22 for the Nucleus CI24 system). Each of these filterbands have fixed centre frequencies and variable bandwidths. Research has indicated that the lower harmonics of a complex sound may not be fully resolved by the filterbands and/or that that there may be greater interaction between the frequency components within each separate filterband (McKay, 2005). Even if the lower harmonic components were fully resolved by the filterbands, properties of the fixed filters would prohibit the exact frequency of the harmonics from being determined. The implantee would only be able to derive which filter the signal fell into via the electrode that was activated, however they would not be able to determine the exact frequency of the signal component. Further, if the resolved components fell into adjacent filters and activated two or more adjacent electrodes, the CI user would be unlikely to be able to resolve the independent places of stimulation.

There is also evidence that the mode of stimulation can affect the ensuing pitch percept, although the precise extent and nature of this effect is a matter of conjecture (von Wallenberg et al., 1993, 1995). Changing the stimulation mode alters the size of the electrical field and the place of the peak potential, which in turn affects the size of the population of the auditory nerve fibres being stimulated, along with their rate of discharge (Pfingst et al., 1995 a, 1995b). von Wallenberg et al. (1995) reported that bipolar stimulation tended to give a higher percept of pitch than monopolar stimulation, with the latter being subjectively selected by all subjects as providing a more comfortable and clearer sound. An earlier paper found bipolar stimulation provided a larger range of pitch percepts (von Wallenberg et al., 1993). Busby et al. (1994) compared monopolar, bipolar, and common-ground stimulation for nine subjects. The majority of the pitch anomalies occurred with common-ground stimulation, although some irregularities were also observed with a bipolar configuration. Additionally, the authors reported that subjects experienced vastly different pitch percepts when stimulating the same electrode using the different stimulation modes trialled. Contrary to von Wallenberg et al. (1995), Busby et al. (1994) found a monopolar configuration to provide higher pitch estimates than the bipolar or common-ground modes, particularly conspicuous at the apical end of the electrode array.

As mentioned in Chapter 2, for the normally hearing population, pitch shifts related to changes in intensity for puretones are minimal, generally being less than 2% to 3% (Cohen, 1961; Morgan et al., 1951). Conversely, in cochlear implantees, significant degrees of interaction have been found between pitch and electrical stimulation level. For example, Pijl (1997a) reported that higher stimulation levels resulted in lower pitch percepts, although there are often large variations in the degree and direction of these shifts across both subjects and electrodes (Pijl, 1997a; Townshend et al., 1987). This pitch-intensity relationship may be partially attributed to the fact that increasing stimulation levels will result in more auditory neurons being stimulated, thereby encompassing a wider excitation area (McKay & McDermott, 1998; Tong et al., 1983). This in turn could change the pitch perceived as different nerve fibre populations are stimulated. The nature of this pitch-intensity relationship is further affected by pathological processes in the cochlea, and the characteristics of the surviving neurons. This interaction was one of the limitations of the Compressed Analogue speech-processing scheme; the simultaneous stimulation resulted in uncontrolled and unpredictable loudness variations, along with inconsistent pitch percepts.

The perception of pitch through the CI for each individual would be dependent on extraneous factors such as the rate of neural survival, the density of the surviving neurons and their spread throughout the cochlea, the current paths in the cochlea, electrode placement, and any electrode irregularities. Anomalies in these factors, such as a poor rate of neural survival, or a wide spread of electrical current in the cochlea, can result in a host of perceptual anomalies such as inaccurate perceptions of pitch changes, fluctuating percepts, irregular pitch shifts, or pitch reversals, to name a few (Busby et al., 1994; Cohen et al., 1996; Collins et al., 1997; McDermott, 2004; Townshend et al., 1987).

Many of the above-mentioned factors are also pertinent issues in the perception of timbre with a CI, with both pitch and timbral percepts being related to the spectral envelope of the input signal. As previously discussed, changes in the place of stimulation may result in timbral variations more than pitch variations, with less musically-experienced subjects experiencing difficulty in differentiating between these two sensations. Pitch and timbre are not entirely separable attributes of sound (Beal, 1985; Krumhansl & Iverson, 1992; McDermott & McKay, 1997; McKay, 2004; Pitt, 1994; Pitt & Crowder, 1992; Warrier & Zatorre, 2002). Accurate timbre perception requires the perception of both the signal's temporal envelope, and the energy spectrum of its harmonic components. Changing the frequency and/or amplitude of the harmonics, or modifying features of the temporal envelope such as the attack (or rise) time will alter the perceived timbre (Handel, 1989; Kohlrausch & Houtsma, 1989). Existing CI processors using constant-rate stimulation strategies only conduct a crude spectral analysis of the input signal. Although fine-spectral details are not essential for speech recognition in quiet situations, with normally hearing individuals being able to tolerate significant levels of spectral distortion in optimal listening environments, spectral selectivity appears to be considerably more important for listening to music stimuli (Shannon et al., 2004). As discussed with relation to pitch perception, the coding of spectral shape in CIs is limited as a result of insufficient stimulation channels, decreased specificity in mapping frequency bands to electrodes, a lack of precision in conveying temporal and spectral detail, phase misalignments, and individual subject and physiological factors (Moore & Moore, 2003). Accurate timbre perception may also be further affected by the presence of perceptual spectral smearing for many CI users, possibly arising from factors such as current spread around the electrode, and neural survival characteristics (McDermott, 2004). These issues may in part account for CI users' poor performances on tasks involving the identification of musical instruments or other complex sounds.

In summary, the perception of music stimuli through a CI is affected by a host of variables ranging from the actual process of electrically stimulating the cochlea, through to the sound-processing undertaken by the speech processor, as well as the specifics of the individual implantee. These serve to impact upon the listener's perception of the resulting sound, and, as will be seen in the next chapter, the degree and nature of this effect varies immensely across individuals.

3.5 HEARING AIDS IN COMPARISON TO COCHLEAR IMPLANTS

Although the focus of this research is on CIs, a generic overview of HAs is provided in this section so that appropriate comparisons can be made. It is beyond the scope of this thesis to cover the multitude of different HA types, designs, options, earmould acoustics, electronic circuitry, or signal-processing strategies available; Dillon (2001) and Staab (2002) provide more comprehensive information.

Like CIs, HAs aim to present the input acoustic information to a listener in a form accessible and useful for the impaired auditory system. However the manner in which the two devices achieve this differ. Unlike the CI, which uses electrical stimulation to directly excite surviving auditory neurons, HAs are electroacoustic devices that amplify sounds to sufficient levels that enable acoustic stimulation of the impaired ear via the normal hearing mechanism. Although there are a wide variety of HAs available, the basic components of an air-conduction aid are similar (Figure 3.17 & Figure 3.18). An input transducer, in the form of an omni-directional and/or directional microphone(s), detects and converts sound pressure variations into an electrical signal. This signal is then amplified, the means, manner and specifics of which vary for different aids and individuals, before being delivered to a receiver. This receiver converts the modified electrical signal back to an acoustic sound wave delivered to the wearer's ear canal (Staab, 2002). Coarse adjustments to the final sound can also be made by more mechanical means – predominantly via changes to the earmould and connecting tubing.

For example, modifying the sound bore, damping, or venting will vary the final frequency response along with the perceived sound quality, whilst changing the style of the earmould and/or materials utilised will alter attributes such as loudness and timbre (Dillon, 2001). Some HAs also allow for alternative inputs such as telecoils (Glossary), and/or enable an Assistive Listening Device to be connected directly to the aid.

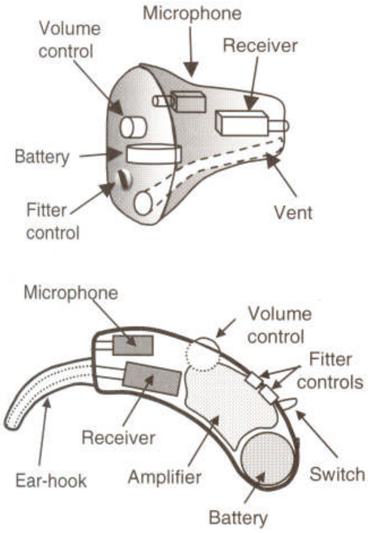


Figure 3.17: Components of a hearing aid
The top diagram represents an in-the-ear model, with the bottom diagram being a behind-the-ear model.
(Image from: Dillon (2001))

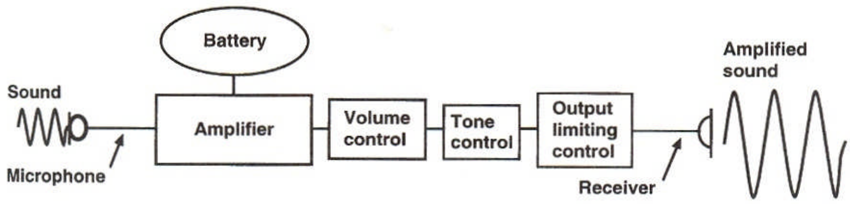


Figure 3.18: Schematic diagram of the components of a hearing aid
The microphone acts as the input transducer, picking up the input sound and converting these sound waves into electrical signals. This signal is then amplified by analogue or digital technology, whereby the electrical waveform is made larger. In most HAs, there are tone controls to adjust the degree of amplification in different frequency ranges, as well as an output limiting control (i.e. output compression) to ensure that loud sounds are at an appropriate level for the wearer. Finally, the receiver converts the electrical waveform back to an acoustic output signal, transmitted to the HA user. In multi-channel HAs, the input sound picked up by the microphone may be divided into a number of separate frequency bands, each with its own gain settings. An input limiting control (i.e. input compression) may also be used to provide the listener with better access to softer acoustic sounds.
(Image from: Bess & Humes (1995))

The gain provided by the HA refers to the degree of amplification; in most current aids, this can be varied across the frequency range to suit an individual's hearing loss, with the measured gain across the whole frequency range being referred to as the 'frequency response' of the HA. Analogue and digital HAs are available, with the latter enabling increased flexibility and control over the sound processing. Promoted benefits of digital aids include the automatic reduction of background noise, feedback cancellation, improved speech perception, more precise and individualised compensation for loudness recruitment, and quicker, automated responses to the listening situation and input stimuli through the use of adaptive algorithms. There are also digitally-programmable HAs which involve analogue circuitry, but enable programming via an external interface, such as a computer, through the inclusion of a digital memory for storing processing parameters. Compared to digital HAs, digitally-programmable aids do not convert input sounds into digital signals, and are less flexible in their programming and sound-processing options (Staab, 2002).

Hearing aids are also classified as single-channel or multi-channel devices (Glossary), although the term 'channel' is not used with the same denotation as for CIs. For HAs, single-channel devices process sound through a single electrical circuit to modify the entire frequency range or one section of the range. Thus, any frequency adjustment in these aids via a filter affects a single broad frequency range. Multi-channel HAs divide the input signal into two or more frequency bands, enabling separate amplification parameters to be set for each band. Subsequent to processing within each individual channel, the outputs of these bands are combined for the receiver. Unlike CIs where it is widely accepted that multiple channels benefit the user, the effect of multiple channels in HAs is less defined with conflicting reports as to their effect (Staab, 2002). It may be of most benefit to those with a sloping hearing loss where more amplification is required for some frequency ranges than others (Dillon, 2001). A sub-optimally programmed multi-channel HA, such as one with inappropriate gain settings for each channel, could distort the overall balance and relationships between the different frequency components. Such imbalance in the amplification may subsequently affect the perceived sound quality.

The normally hearing ear has a 'comfortable' dynamic range of approximately 100 dB, and spans a frequency range from around 20 Hz to 15 kHz, and up to 20 kHz in children (Moore, 2003b; Swanson, 2003). The presence of recruitment, however, in a sensorineural hearing loss results in a steeper than normal loudness growth, and consequently a reduced dynamic range. Often this recruitment is frequency-dependent, being more evident at some frequencies than others (Hansen, 2002). Accordingly, some HAs are set up to compensate for this, incorporating a limiting system to ensure that the resulting sound is kept within the listener's dynamic range (Dillon, 2001). The implementation of compression within a HA system will impact upon the resulting sound quality, although the nature and degree of this effect is dependent upon the parameters chosen, and varies across patients (van Buren et al., 1999). Most current devices allow for the implementation of both input and output compression. Whereas input compression may be used to improve the listener's access to the softer sounds of speech, output compression largely serves to ensure that loud sounds do not become uncomfortably loud, or distorted for the wearer. Several types of limiting systems such as peak clipping, peak rounding, or automatic gain control (Glossary) can be incorporated into the device to both maintain the output within a suitable dynamic range for the user, as well as ensure that the resulting sound is not too loud. A comprehensive description of these limiting systems is available in Dillon (2001).

There is no definitive proof, however, of one particular setup being advantageous over another. For example, to restore normal loudness perception, the use of multi-channel non-linear compression with short time constants is commonly suggested. These shorter time constants allow the temporal level changes of the input signal to be monitored, and are thought to more closely mimic a normal functioning cochlea. However, Hansen (2002) proposed that shorter time constants increased the perception of noisiness as any background noise is unintentionally amplified. The subjects in Hansen's (2002) study demonstrated an overall preference for longer release times with low compression thresholds for listening to speech, and to a lesser extent, non-speech signals such as music. van Buren et al. (1999) conducted research comparing different compression parameters. They reported that the use of wide-band compression in a single processing band with a compression ratio (Glossary) of two did not decrease the perceived sound quality. However, subjective quality ratings for music and speech sounds decreased as the number of independent compression channels was increased, as well as when the compression ratio was increased. Further, linear amplification was always judged to be more pleasant than non-linear amplification.

There is little published research that objectively investigates which HA settings could provide the most benefit for music perception. As HAs are predominantly fitted to optimise the perception of speech, this may have repercussions for listening to music. For example, as reported by Chasin & Russo (2004), speech signals have a long-term speech spectrum that HA prescriptions aim to match. On the contrary, music's long-term spectrum resembles low-pass filtered noise with no specific target that can be used for fitting purposes. Chasin & Russo (2004) suggested that the parameters used in optimising HAs for speech may not necessarily be the optimal setting for listening to music stimuli.

Hearing aids can be fitted to accommodate a wide range of levels and configurations of hearing loss from mild through to profound impairments. In general, they provide amplification between about 100 Hz and 5000 Hz with different devices providing differing levels of gain. However, of significance for CI technology is that the potential benefit provided by HAs is minimal for the two extreme levels of hearing loss - mild and profound. Those with a profound level of loss rarely obtain significant speech recognition benefit from a HA (Bess & Humes, 1995), and hence these patients may consider, and benefit from, a CI.

In order to fulfill candidacy for a CI, most clinics require the patient to undergo assessments with optimally fitted HAs; a CI would then be considered if the patient is unable to obtain sufficient speech perception benefit from HAs. The criteria for this vary across clinics, but are generally derived from the mean scores obtained by their current CI users – that is, whether the CI can potentially provide the prospective patient with higher speech perception scores than they achieve with their HAs, based on the data of existing implantees. Figure 3.19 provides a guide to the probability of a CI user exceeding a particular speech score for three different speech tests. For example, the results of this graph show that adopting the traditional implantation criteria of an open-set speech perception score of 40% in the best-aided listening condition for sentence stimuli presented in quiet, would provide a 67% chance of the potential recipients achieving a better speech perception score post-surgery. However this graph was based upon the results of Skinner et al.'s (1994) study which involved users of the Nucleus CI22 device implemented with the SPEAK speech-processing strategy. Therefore it would be reasonable to expect better outcomes with current devices and strategies. Figure 3.20 presents more-recent speech perception results obtained from a retrospective study of 92 postlingually deafened adults using a range of Nucleus implants and strategies. These scores were for sentence stimuli obtained between three and six months post-surgery. The mean score was 80%, with a median of 91%. Of more interest, though, is the first-quartile score of 68%, indicating that if a potential recipient’s pre-implant sentence perception score was 68% or less, there would be a 75% chance of them obtaining a post-surgery score higher than this. This research gave rise to many clinics expanding their implantation criteria to sentence perception scores less than 70% in the best-aided listening condition, and less than 40% in the ear to be implanted. It is important to bear in mind, though, that decisions with respect to the suitability of a patient for an implant are made on an individual case-by-case basis, after consideration of a range of factors; the benefit they currently obtain with HAs is just one of these factors.

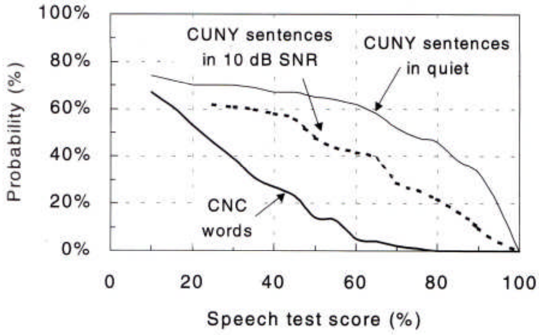


Figure 3.19: Probability of exceeding specific speech scores, post-implant

X-axis: pre-implant speech test scores for the 3 types of speech assessments.
Y-axis: the probability of an implanted adult exceeding that speech score once they have received their CI.
The data for this graph was based on the results of Skinner et al.'s (1994) study involving adult implantees using the Nucleus CI22 system in conjunction with the SPEAK strategy.
(Image from: Dillon, 2001, p. 229)

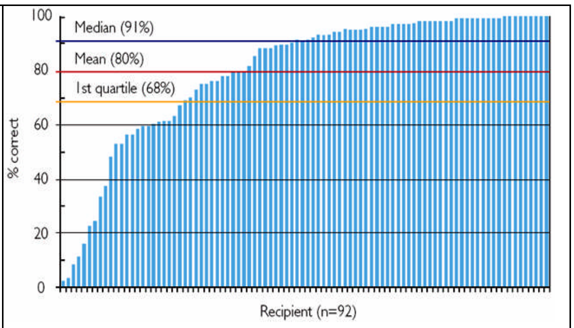


Figure 3.20: Speech perception results obtained 3 – 6 months post-implantation

Each column in this graph represents the speech perception result for sentence stimuli obtained from a CI recipient, arranged from lowest through to highest. The results were collated from postlingually deafened adults tested with their Nucleus implant only, between 3 and 6 months post-surgery. A range of Nucleus implants and speech-processing strategies were incorporated into this data set.
(Image and data from: Dowell et al., 2004, p. 2)

In summary, both the HA and CI aim to enable perception of acoustic information for a person with a hearing impairment. However, the manner in which they achieve this differs; the HA acoustically stimulates hearing via the normal hearing mechanism, whilst the CI uses electrical stimulation of hearing. Accordingly, the two devices vary in their components, operational functioning, fitting options, and methods of processing the input sound. The CI becomes an option for those with significant levels of hearing loss who are unable to obtain adequate speech perception benefit from their HAs. Despite their differences, one of the primary aims of both devices is to improve the wearer's ability to perceive speech; hence the devices are designed and programmed accordingly. This may, however, affect the user's perception of music stimuli, as will now be reviewed in the next chapter.

CHAPTER 4: LITERATURE REVIEW

While Chapter 3 concentrated on psychoacoustic research about the direct effect of electrical stimulation on the perception of the fundamental elements of music, this chapter reviews research involving music perception in more realistic listening situations. Initially, the music perception skills of the general hearing-impaired population, including children, is summarised to provide a broad perspective on the topic (section 4.1). Since there is very little research investigating the music perception abilities of adult HA users, literature comparing children with hearing impairments to those with normal hearing will at least provide some points of comparison. From there the literature review primarily focuses on research conducted with postlingually deafened adult CI recipients (section 4.2).

4.1 MUSIC PERCEPTION OF PEOPLE WITH HEARING IMPAIRMENTS

It is well established that people with hearing impairments, including CI users, perceive rhythm approximately as well as those with normal hearing (NH). Initial studies into the effect of hearing loss on music perception were predominantly undertaken with primary school-age children with prelingual hearing losses utilising HAs or vibrotactile devices. These collectively report that hearing loss has a negative impact upon melody perception, and to a lesser extent, rhythm perception (Darrow, 1979, 1984, 1987; Klajman et al., 1982; Korduba, 1975).

Darrow (1984) studied various elements of rhythmic reproduction, finding that children with hearing impairments scored equivalently to children with NH on all tested measures, except for melodic rhythm duplication. The author attributed this latter lower score to environmental factors – that is, a comparative lack of musical training, as opposed to true perceptual limitations. This supported the earlier work of both Korduba (1975) and Rileigh & Odom (1972), also with children. However, in a later study using Gordon's (1979) Primary Measures of Music Audiation (PMMA), Darrow (1987) found that children with severe to profound hearing impairments performed significantly worse on both the tonal and rhythm subtests, and thus the overall composite score, when compared to the NH age norms. The PMMA is a standardised test of music perception developed by Gordon (1979) for children. Its two subtests, tonal and rhythm, each incorporate 40 stimuli pairs for comparison, to be assessed by the listener as sounding the ‘same’ or ‘different’. In Darrow's (1987) study, the children with hearing impairments performed better on the rhythm than the tonal subtest across all of the age groups tested, in contrast to the NH age norms which demonstrated higher scores for the tonal subtests. Klajman et al. (1982) compared the musical abilities of 130 children with hearing impairments aged 7 to 19, with a range of hearing levels, to 104 similarly aged children with NH. They tested rhythm, pitch, melodic, tonal, and timbral perception, along with musical memory, as a precursor to identifying a subset of the hearing-impaired subjects to receive specialised music training. Their rhythm test, involving the direct reproduction of a rhythm pattern, assessed the child’s sense of rhythm, duration, and rhythmic memory. They found that children with hearing impairments performed significantly better than the NH group on measures testing their sense of rhythm and duration. However, the children with NH scored better than those with a hearing impairment on the pitch, pitch memory, melodic memory, and melody reproduction assessments, with little difference between the groups for the various timbre perception tasks (Klajman et al., 1982).

4.2 MUSIC PERCEPTION WITH COCHLEAR IMPLANTS

As CIs have developed, they have become the preferred habilitative and rehabilitative device for both children and adults with severe or profound hearing losses. With a wealth of research validating the benefits that a CI can offer in terms of speech perception, more-contemporary research studies have begun to compare the music perception skills of paediatric and adult CI users to NH subjects. Stordahl (2002) and Vongpaisal et al. (2004) found that children with NH were significantly better than children with CIs in recognising melodies presented in a closed-set format. The CI subjects in Vongpaisal et al.'s (2004) study performed at a level close to chance in identifying the single-line melodies played on the piano, despite rhythmic cues remaining intact. These findings were replicated by Nakata et al. (2005) in a study involving Japanese children using CIs; the children performed at chance level in identifying familiar melodies presented without their lyrics, either as instrumental or melody-line versions. Stordahl (2002) also observed that the children's behaviours during testing differed. Those with CIs exhibited few external signs suggesting familiarity with a song, and they would listen to the entire item before making a decision or response. On the other hand, the children with NH selected their answers immediately, demonstrating less variability in their results and error patterns. The CI subjects in Stordahl's (2002) study appraised the melody items to be less pleasant than the NH subjects. The inclusion of words with melodies has been reported to result in both higher recognition and appraisal scores in prelingually deafened children (Nakata et al., 2005; Vongpaisal et al., 2004).

However, direct comparisons of findings from paediatric populations to adult implant users may be confounded by subject characteristics; particularly the differences in their physical, cognitive and social development, as well as the prelingual nature of most hearing losses in the paediatric population. The majority of children with CIs would have either been born with a congenital hearing impairment, or would have developed their hearing loss at a very young age. As such, comparisons to postlingually deafened subjects cannot be justified as the latter would have an internal representation of sound through a better-hearing mechanism. Those with prelingual losses have learnt to hear via the implant which obviously does not provide the same representation of sound as normal acoustic hearing. Accordingly, with this study involving adults with hearing losses acquired postlingually, the following literature review will concentrate primarily upon research conducted with postlingually deafened adults. The general music listening habits of these CI users will be discussed before turning to their perception of the specific musical features investigated in this study - rhythm, pitch, timbre, and melody. A variety of articles are reviewed in order to give a general perspective on the existing status of research in this area, to provide an understanding of the current consensus of CI users' abilities on various music tasks, as well as to provide the background information from which the aims, hypotheses and justification for this current study were derived.

4.2.1 General Music Listening Habits

Gfeller et al. (2000b) designed and implemented a questionnaire to assess the musical background, listening, and enjoyment of 65 adults using a variety of multi-channel CIs (29 Clarion, 17 Nucleus, 11 Med-El, and 8 Ineraid implants). Seventy-seven percent of respondents stated that they listened to music and/or were involved in music-related activities prior to their hearing loss. The amount of time spent listening to music post-implantation was significantly lower than pre-implant, with one-third commenting that they avoided music due to its aversive sound. The second part of Gfeller et al.'s (2000b) study classified 67 CI subjects into three groups according to how long ago they had received their implant. All subjects were then asked to record the number of hours spent listening to music in a diary. Despite the high degree of variability within and between the groups, there was a general trend for the more-recently implanted CI users to record higher levels of daily music participation than those who had been implanted earlier. However, many recipients reported music to sound strange and noisy, and in some cases so poor that they deliberately avoided it.

Similar findings have been reported elsewhere. In Leal et al.'s (2003) study, 86% of 29 CI subjects stated that they spend less time listening to music post-implant than pre-implant, with 38% reporting that they did not like listening to music. Mirza et al. (2003) surveyed 35 CI patients in regards to their appreciation of music prior to the acquisition of a hearing loss compared to whilst they were listening with their CI. According to the authors, most of the respondents listened to music before becoming deaf, with 46% listening to music once implanted. No respondents reported that they listened to music in between the time they became deaf and when they received their CI. In rating their enjoyment of music out of 10 (where 0 = not at all, and 10 = very much), mean ratings were 8.7 prior to hearing loss, compared to 2.6 with the implant. For the 16 respondents who reported listening to music routinely post-implantation, the mean enjoyment rating was 5.6, compared to 9.3 pre-hearing loss. Sixty-nine percent of the 35 respondents said that they were disappointed with the sound of music through the CI. In a survey reported by Tyler et al. (2000), 83% of 63 adult CI users recorded lower levels of enjoyment for music post-implantation, with 51% responding that music sounded unpleasant or difficult to follow. Brockmeier et al. (2004) administered a questionnaire to compare self-reported music perception and listening preferences for 104 CIS users compared to 69 SPEAK users. Results were similar for the two strategies, with 60% of CIS users and 57% of SPEAK users stating that they listened to music with the CI. Of the CIS and SPEAK users respectively, 47% and 45% reported the quality of music to be pleasant, with 30% and 35% responding that music sounded natural.

Collectively, these studies suggest that the sound of music through the implant is sub-optimal, and does not allow the user to fully appreciate musical stimuli. Implantees appraise the sound to be noisy, disappointing, and unenjoyable, and consequently, they spend less time listening to music when compared to pre-implant and/or pre-hearing loss levels. There is no indication that one type of CI or speech-processing strategy is preferable for music appreciation, although a wide range of music listening habits and preferences is prevalent across the implant population.

It is worthwhile to note that in many of these subjective, comparative studies, it is ambiguous as to exactly when the implantee was being asked to make their comparative judgments to. Some studies used the term ‘pre-implant’ in their publication (Leal et al., 2003), others stated ‘prior to hearing loss’ or ‘prior to a profound loss’ (Gfeller et al., 2000b; Mirza et al., 2003), whilst some authors appeared to use the terms interchangeably (Gfeller et al., 2000b). The exact point in time that the subject was making their personal comparisons to would have significant implications for the results obtained. For many subjects, there would have been a period of time prior to implantation where they were using HAs. The length of this time period, and the severity of their hearing loss during this time, would have varied among subjects. Hence the question arises as to what previous period of time the respondents were comparing to, with some possibilities being:

i) A period of time where they had completely NH;
ii) A period of time where they had a hearing loss, but were not using HAs;
iii) A period of time where they used HAs; or even
iv) A period of time before implantation when HAs were no longer of any benefit, and they relied mainly on lipreading or visual cues for communication.

Consequently, in interpreting the afore-mentioned results, it is unclear as to whether the comparative responses were based on sound percepts with unaided or aided hearing, and the severity of the hearing loss at that time. Further, if the respondents were making judgements based on when they had normal or near-normal hearing, it must also be asked how long ago this period of time was, and the clarity of their recollection or memory for these musical sounds. These confounding issues make the interpretation of some of these results ambiguous.

4.2.2 Rhythm Tests

Analogous to the results reported with children, adult subjects with NH and CIs also perform similarly on measures of rhythmic or temporal discrimination (Gfeller et al., 2000b; Gfeller & Lansing, 1991, 1992; Gfeller et al., 1997; Schulz & Kerber, 1994). Gfeller & Lansing (1991) administered the PMMA to 18 postlingually deafened CI subjects (10 Nucleus, 8 Ineraid implantees). Mean scores on the rhythm subtest (88%) were higher than on the tonal subtest (78%). The same authors then expanded this initial study by presenting the same tests to 34 postlingually deafened CI subjects (17 Nucleus, 17 Ineraid implants) (Gfeller & Lansing, 1992). Similar results were obtained, with the CI subjects demonstrating significantly greater accuracy on the rhythm than tonal subtest. The uniqueness of this was highlighted by the authors as in prior studies using the PMMA as an assessment tool, other population groups, including subjects with NH, adults with a brain injury, and those older than 65 years, performed better on the tonal than rhythm subtest. Only the CI recipients, and the children with hearing impairments (as cited in relation to Darrow's (1987) study in section 4.1) provided the reverse results. It was speculated that this anomaly may be related to the consideration that for those with significant levels of hearing loss, they may only be able to perceive very limited frequency-based information through their HA, resulting in gross temporal cues (i.e. rhythm) being comparatively more salient. Therefore the HA user may become more attuned to using these gross temporal cues than frequency cues. This consideration would still be applicable for users of a CI, as the majority of implantees would have had a significant hearing loss and/or used HAs prior to implantation. Thus, those with a history of a longstanding significant hearing loss may be more aware of, if not more reliant upon rhythmic information, with the lack of access to frequency-based information potentially having an adverse effect on music perception (Gfeller & Lansing, 1992).

It is worth clarifying at this point, the differentiation between ‘gross temporal cues’ that impart a sense of rhythm in music, as opposed to temporal cues which provide a sense of pitch. Temporal patterns in the frequency range of 0.2 Hz to 20 Hz provide a distinctive rhythm to musical stimuli, whereas higher-frequency components of the acoustic signal provide the pitch information, as discussed in the previous chapter (Gfeller et al., 1997; Looi et al., 2004; McDermott, 2004) .

Schulz & Kerber (1994) assessed the music perception skills of eight single-channel CI users and seven NH subjects. Results indicated that the CI subjects were more accurate at the rhythm than pitch perception tasks, with their scores on the rhythm pattern reproduction task being slightly higher than the NH controls. Leal et al. (2003) conducted two rhythm tasks – one was classified by the authors as a discrimination task (same/different), with the second being classified as an identification task where the subject had to additionally indicate the point at which the rhythm changed. Twenty-nine postlingually deafened adults with at least 3 months experience with their CI24 device undertook the task. The mean score on the discrimination task was 95%, with 24 of the 29 subjects scoring 90% or higher. This indicates that the majority of the CI subjects had little difficulty in discriminating between rhythm patterns. There was no control group comparison in this study.

Overall, these results indicate that CI and NH subjects perform approximately equivalently on rhythm perception tasks. However some of these studies also suggest that the pitch perception of CI users is degraded compared to NH listeners. This will be further illustrated in the following section.

4.2.3 Pitch Tests

The collective findings across a range of studies indicate that CI users perform significantly worse than NH controls on pitch-based tasks. For example, in Gfeller & Lansing 's (1991) study, CI subjects scored 78% on the tonal subtest of the PMMA, compared to 88% for the rhythm subtest. In their 1992 extension of the initial study, scores were 85% and 78% for the rhythm and tonal subtests respectively, with CI subjects unequivocally reporting the tonal subtest to be harder than the rhythm subtest (Gfeller & Lansing, 1992).

Gfeller et al.'s (1997) study comparing the now-obsolete F0F1F2 and MPEAK speech-processing strategies showed that CI subjects, regardless of which strategy they utilised, performed significantly worse than NH subjects on the PMMA tonal subtest. This was particularly apparent for item pairs which had the same melodic contour, but differed in their absolute pitches. It should be noted, though, that these feature-extraction strategies would have predominantly conveyed low-frequency formant information, and are no longer in clinical use. Schulz & Kerber (1994) assessed pitch perception for eight users of a single-channel CI programmed with a now-obsolete analogue processing scheme. In the melodic-direction perception task, subjects were required to assess whether a tonal sequence played on the piano was ascending, descending, or unchanging in pitch. The results of this assessment indicated that, as expected, the NH subjects found the task to be easy, scoring nearly 100% regardless of the distance between notes, or the direction of the change. Subjects with CIs scored between 68% and 84%, suggesting that although they could undertake the task with varying degrees of accuracy and ease, their overall scores were still significantly below that of the NH control group. Schulz & Kerber (1994) also asked subjects to differentiate whether two tones played on a piano were the same or different in pitch. The results indicated that whilst NH subjects could reliably differentiate between notes one semitone (Glossary) apart, which approximately equates to a 6% difference in the F0, the CI group were only able to differentiate intervals larger than a major second or a minor third (i.e. 12% to 19% frequency ratio). The authors reported that these results were fairly consistent across the 11 octave (Glossary) frequency range used in this task. In another task requiring subjects to vocally match the pitch of the musical note presented, CI subjects averaged 22% correct, with the NH subjects scoring 45%. In a separate study, Fujita & Ito (1999) found that five of eight Nucleus CI22 users discriminated the higher of two notes between four and ten semitones apart. However, the remaining three subjects could not discriminate between two notes, one octave (i.e. 12 semitones) apart. Again, it must be kept in mind that seven of the eight subjects in this study used the now-obsolete MPEAK strategy, with the other subject using the SPEAK strategy.

Leal et al. (2003) conducted two pitch perception tasks with 29 CI24 users, 20 using the ACE strategy and nine using the SPEAK strategy. The first task was a same/different comparison of 12 pitch pairs. The second task consisted of eight paired musical excerpts where the listener had to state whether the pitch became higher or lower, and where this change occurred. For the same/different comparison, the mean score was 10.8 out of 12 (90%), with 69% of subjects scoring above 88%. In the higher/lower task, subjects averaged 5.9 out of 8 (74%), with 48% of subjects scoring above 75%. A significant correlation was found between this latter score for the higher/lower task and both post-implant listening habits as well as musical background scores. There was no significant relation between music perception scores and which of the two speech-processing strategies the subjects used.

Gfeller et al. (2002a) also investigated pitch perception by comparing CI users with NH controls in a higher/lower pitch discrimination task. The goal of the task was to determine the smallest interval size where subjects could judge that two notes were different in pitch; this was termed their 'difference limen for pitch'. The stimuli consisted of two one-second tones recorded on a synthesised acoustic piano. The results of this test revealed a significant discrepancy between NH and CI subjects' skills. Whereas the mean difference limen for NH subjects was 1.13 semitones (range: 1 to 12 semitones), the CI population's performance was highly variable with a mean difference limen of 7.56 semitones, and a range of between 1 and 24 semitones (standard deviation: 5.18 semitones).

However, in interpreting the results of the complex tone discrimination task in Gfeller et al.'s (2002a) study, a potentially critical issue in regard to the methodology adopted should be accounted for. An adaptive procedure was implemented to determine this difference limen for pitch. The algorithm employed was based on the staircase method where the interval size decreased after a predetermined number of correct responses, and similarly increased after a predetermined number of incorrect responses. If the subject made at least nine correct responses from 11 presentations, the interval size was reduced. The interval size was increased when three incorrect responses were made at a particular interval level. The process was continued, up to a maximum of 55 presentations, until a certain interval size was deemed to be significantly correct; the interval size one smaller than this was judged by the algorithm as incorrect. The pitches tested spanned three octaves, from 73 Hz to 553 Hz, in a two-alternative forced-choice response format. The potential flaw with this procedure was that it implicitly assumed that as the interval size increased, the subject was more likely to get it correct, and conversely, when the interval size decreased, an incorrect decision would be more likely. However, this assumption does not necessarily hold true as CI subjects are highly variable in their pitch-ranking ability, and their accuracy is affected by the F0 of the stimuli (Gfeller et al., 1997; Looi et al., 2004; McDermott, 2004). For example, even if the interval size was held constant, a subject may vary from chance performance to 100% correct or even 100% incorrect (indicative of pitch reversals), as the F0 of the reference note changes. Hence, larger intervals do not necessarily correspond with increased discrimination accuracy, or vice versa. Even studies with NH subjects tend to use a standard reference frequency in determining frequency difference limens, as opposed to randomly varying the frequencies over a wide range, as was the case in Gfeller et al.'s (2002a) study. Hence it is possible that a significantly different result, or difference limen, could be obtained should the study be replicated using a different frequency range, or even a different set of reference F0s within the same range.

Due to the poor pitch perception scores for implant subjects, researchers have been proposing a host of approaches to endeavour to improve or compensate for this limitation. As some of these suggestions are discussed in Chapter 10, it will suffice at present to state that one such approach currently attracting interest is to combine the use of residual acoustic hearing with the CI's electrical stimulation for suitable patients. This may be achieved through the use of a HA in the contralateral ear (Kong et al., 2005; Tyler et al., 2002), or unilaterally through the use of either a modified surgical technique (Kiefer et al., 1998, 2005; Skarzynski et al., 2003) and/or a shorter electrode array (Gantz & Turner, 2003, 2004; Gantz et al., 2005) to preserve as much low-frequency hearing as possible. Gfeller et al. (2004) investigated whether this combination of electric and acoustic hearing in the same ear, with a short electrode array (as detailed in Gantz et al., 2005), could assist a CI recipient in their ability to perceive pitch. It is worth emphasising at this point, though, that the short electrode array is only viable for a select group of patients who have significant levels of low-frequency acoustic hearing. Furthermore, its development is still at an early phase, and it is not yet commercially available. The results of Gfeller et al.'s (2004) study showed that the six subjects implanted with this short electrode array performed significantly closer to 22 NH subjects than a group of 41 CI subjects utilising a conventional longer electrode array. For example, in ranking puretone intervals presented one semitone apart at a low F0 (around 131 Hz), both the NH and short electrode array subjects scored close to 100%. The CI subjects with the conventional array scored approximately 70% for this same task. Ostensibly, as also mentioned by the authors (Gfeller et al., 2004), these results provide some initial evidence that the use of low-frequency acoustic hearing in conjunction with the implant has the potential to improve the perception of pitch for these bimodal device users. For the subjects implanted with the short array in Gfeller et al.'s (2004) study, their comparatively good pitch perception also served to benefit them on a subsequent melody recognition task (see section 4.2.5). In interpreting the results of this study, it should be noted that no information was provided regarding the subjects' levels of post-surgery residual hearing. Greater levels of post-surgery residual hearing would be likely to assist with music perception tasks. Therefore in comparing the results of CI subjects with a short array to those with a long array, the greater levels of residual hearing for the former group may account for their better performance on the music tests, rather than the type of array they had.

Despite a wide variety of methodologies, testing protocols, task requirements, implant types, speech-processing strategies, and range of subjects participating in the above studies, the existing literature concurs that the pitch perception skills of conventional CI users are significantly poorer than those of NH subjects, leading to an adverse effect on overall music perception. The pitch perception of postlingually deafened HA users has not been specifically investigated, except in relation to psychoacoustic studies researching the perceptual consequences of a cochlear hearing loss on frequency selectivity, or other pitch-based tasks. Some of these studies were mentioned in Chapter 2, Section 2.3.4.

4.2.4 Timbre Tests

Unlike pitch and loudness, timbre is a multi-dimensional attribute related to differences in sound spectra, and its perception is usually assessed in music studies via instrument identification tests. In a musical sense, timbre differences would allow a listener to differentiate between two instruments playing at the same pitch and loudness. Timbral perception assists with both instrument recognition and auditory scene analysis when listening to musical ensembles, and its perception is a contributory factor to the quality of the listening experience. The sound quality of timbral features may contribute more to the implant users' satisfaction when listening to music than instrument identification skills, particularly for non-musicians where aesthetic enjoyment is a primary function of music (Gfeller et al., 2002b). Timbre can also assist music listening by systematically organising musical information into structural sections, such as theme and variations, and by conveying the underlying emotion of the music. The perception of timbre in a musical context is further confounded by variables such as the number of music lines or parts (i.e. the harmony), style or genre, complexity, rhythm, tempi, pitch, articulation and phrasing (Gfeller et al., 1998, 2003).

Timbre perception is related to the acoustic signal's frequency spectrum and amplitude envelope, along with the changes in these two attributes over time. For CI users, the perception of the signal's spectral shape is important, with variations in the spectral characteristics of the acoustic signal changing the perceived timbre (McDermott, 2004). Hence, the manner in which CIs code these spectra will affect the listener's ability to differentiate between timbres. The way in which current CIs represent spectral information is largely inadequate for accurate recognition of many musical instruments with tasks involving instrument identification with implants usually reporting very poor overall results compounded by high levels of individual variability (Gfeller et al., 1998, 2002c; Looi et al., 2004).

It is also worth considering that given that timbre comprises all of the perceptual attributes once pitch and loudness are controlled for, it is possible that the perceptual quality and distinctiveness of timbre may be diminished or ambigious in the absense of pitch. As mentioned in Chapter 2, there can be significant perceptual interactions between pitch and timbre, particularly for those with little or no formal musical training (Beal, 1985; Crowder, 1989; Pitt, 1994; Pitt & Crowder, 1992). That is, for non-musicians, perceptual variations in one dimension may impact upon percepts in the other dimension. In view of the consideration that few CI users would have had high levels of musical training, the ability of implantees to recognise tuned musical instruments (i.e. instruments which can convey a melody-line) may also be affected by their reduced pitch perception skills.

Gfeller et al. (1998) investigated the recognition and appraisal of the trumpet, clarinet, violin, and piano in 28 Clarion CI users programmed with the CIS speech-processing strategy, compared to 41 NH listeners. Appraisal results were made via a rating scale on a continuum from 0 to 100, where 0 equated to “dislike very much” and 100 corresponded to “like very much”. Overall appraisals from the CI population were below those of the NH population, with significant differences found between specific instruments. NH subjects were not only significantly more accurate than implantees in recognising the four different musical instruments, but they were also more confident with their responses and tended to make consistent, ‘justifiable’ errors, such as confusions within the same instrument family. CI users demonstrated diffuse, non-systematic error patterns, and often sought reassurance during testing as to the accuracy of their responses. The ratings provided by the CI subjects were significantly correlated with the amount of reported post-implant music listening ($r = 0.49$; $p = 0.008$), and with musical background scores ($r = 0.41$; $p = 0.028$), as recorded in the study’s musical background questionnaire. Speech perception results were not predictive of instrument recognition ability or appraisal, and only a weak correlation was found between recognition scores and overall musical experience scores, or the length of hearing loss. No significant correlation was found between accuracy and appraisal scores.

In a larger study, Gfeller et al. (2002c) found a significant difference between CI and NH subjects’ ability to recognise eight different musical instruments playing the same seven-note melodic sequence formed with equal-duration notes. Each instrument was presented three times in the soundfield, with subjects making their decisions from a larger set of 16 different instruments. Subjects were also asked to appraise the overall pleasantness of the eight instruments using the same scale as in their earlier study described above (Gfeller et al., 1998). NH subjects in the larger study scored 91% correct whereas the CI recipients only scored 47%, with correspondingly lower general appraisal scores (Gfeller et al., 2002c). Furthermore, higher-frequency instruments such as the flute, violin, and piano played in its upper registers, were perceived by CI subjects to have a noisier and duller quality than the corresponding appraisals provided by the NH subjects. Again, no significant correlations were found between speech perception scores and the general appraisal or recognition scores (Gfeller et al., 2002c).

Leal et al.'s (2003) smaller-scale instrument identification task utilised short melodies played on the piano, trombone, and violin. Each instrument was presented once only in a closed-set recognition task, with a score out of three being obtained. Sixty-nine percent of the subjects identified all three instruments correctly, however it should be noted that subjects had a one-in-three chance of selecting the correct instrument (i.e. chance score of 33%).

Schulz & Kerber's (1994) research with single-channel implants found that NH subjects scored significantly higher than CI subjects (90% and 36% respectively) in a closed-set recognition task incorporating five instruments. The authors also conducted a subjective rating task where subjects were asked to rate, on a scale from 1 to 5, the level of appeal for rhythms played by: i) a solo drum, and ii) a drum accompanied by other instruments. For the CI subjects, the solo drum presentations were rated to be more appealing than the accompanied-drum versions, with the latter stimuli reported to sound hazy and less distinct. These findings led the authors to postulate that increased aural complexity led to decreased subjective ratings by CI users.

Stainsby (2001) conducted research to assess the amount of frequency-spectrum information present in a complex sound that is available to a listener. A correlation between the internal and physical spectra of a sound was calculated (i.e. the relationship between the actual frequency spectrum and the amount of this spectrum perceived by the listener), enabling the frequency selectivity of 4 NH, 5 CI, and 3 HA subjects to be compared. Steady-state stimuli comprising 10 synthesised musical sounds (5 musical instruments and 5 sung vowels) with a F0 of 494 Hz were used, where the temporal envelope variations had been removed. The results of the study were that the correlations between the physical and internal spectra for the subjects with hearing impairments were weaker, but still significant, than those for the subjects with NH. This may have been due in part to the broader auditory filters associated with cochlear hearing loss, along with the broader filters of the speech processor’s (Spectra 22) filterbank. The CI subjects obtained a stronger spectral correlation than the HA subjects. However there were only three HA subjects, all of whom had severe to profound hearing losses, involved in the study. For all subject groups, the ability to discriminate between the stimuli was related to the strength of the correlation between the internal and physical spectra. The better this correlation, the better their discrimination score. However the subjects’ identification scores were significantly poorer than their discrimination scores, and not related to frequency selectivity. This may have been due to the consideration that cues from variations in the temporal envelope, absent in the steady-state stimuli used in the study, were necessary for the identification of musical stimuli. Overall, the results of Stainsby's (2001) research suggest that although some degree of timbre perception is possible through the CI to enable discrimination of musical sounds, the level of spectral information conveyed is not the same as for listeners with NH, nor is it sufficient for instrument identification.

To summarise, the perception of timbre through implants is currently inadequate to enable a CI user to accurately and reliably identify musical instruments and/or appreciate their diverse and unique qualities. Existing literature collectively affirms that CI users are not only significantly worse than NH subjects on instrument identification tasks, irrespective of the instruments assessed, or the subject’s musical background, but that they also rate musical instruments to be less pleasant sounding. There is no published literature investigating the instrument identification abilities of postlingually deafened adult HA users, either in comparison to NH subjects, or to CI recipients.

4.2.5 Melody Tests

The other major component incorporated into many music perception test batteries is the recognition of melodies. In many ways, this task is an extension of the pitch perception task; it could be considered that whereas pitch discrimination is an analytic-type task, melody recognition is a synthetic extension, a skill requiring more than just pitch perception, with direct relevance and functionality in real life. Although pitch perception is an integral part of melody recognition, it is by no means the only element to consider; for example, the perception of rhythm, lyrics, timbre, genre or musical style must also be considered. Gfeller et al. (2002a) stated that melodies comprise specific pitch patterns sequentially organised into cohesive melodic units. The recognition of melodies is affected by one's familiarity with, and ability to sufficiently perceive, the various structural elements of the melody including melodic contour, the absolute or relative intervals, and the rhythm. Additional features such as the lyrics, or a specific singer or instrument, may also aid recognition. Many NH people are able to recognise varying degrees of modification of the melody, as well as the addition of harmony.

Fujita & Ito's (1999) study of eight implantees (7 using the MPEAK strategy, and 1 using SPEAK) incorporated two melody recognition tasks. In the first task, nursery rhymes were presented in two formats: i) sung with verbal cues, and ii) keyboard only with no verbal cues. In the former mode, subjects identified an average of 3.9 out of 10 tunes open-set, and 5.3 out of 10 tunes closed-set. However, without the vocal cues, scores dropped to 1.7 and 2.1 out of 10, open- and closed-set respectively. For the second task, four nursery rhymes with identical rhythms played at the same speed were tested, thereby precluding the use of rhythmic cues to aid identification. In this task, the CI subjects were unable to identify the melodies, performing at chance level only. Schulz & Kerber (1994) also compared the melody recognition of CI subjects to that of NH subjects using melodies classified by the authors as rhythmically structured or unstructured. Results indicated that the CI recipients (40% to 55%) scored significantly lower than the NH subjects (96% to 98%), with the scores from the former group decreasing as tonal and rhythmical complexity increased. That is, CI subjects found rhythmically structured tunes easier to identify than rhythmically unstructured tunes. Given the pitch-perception anomalies commonly associated with the use of CIs, it has been speculated that CI recipients are potentially far more reliant on rhythm cues for melody recognition than those with NH (Gfeller et al., 2000a, 2002a). This is in keeping with the findings reported earlier that CI users are better on rhythm- than pitch-based tasks.

Leal et al.'s (2003) task involved identifying eight familiar songs presented closed-set in three contexts: i) orchestra only with no verbal cues; ii) melody-line, played by a solo piano with no verbal cues; and iii) orchestra with verbal cues. Only 1 of the 29 subjects identified more than 50% of the melodies in the first condition, increasing to 14 (48%) when the task was simplified to the solo piano presentation. Significantly, when verbal cues were added to the stimuli, 96% of the subjects identified more than 50% of the melodies correctly, and scores for this condition were correlated with speech perception scores. As reported earlier, this corresponded with both Fujita & Ito's (1999) finding that the inclusion of verbal cues provided significant assistance with melody recognition for CI users, along with Schulz & Kerber's (1994) proposition that increased musical complexity had the opposing effect.

In one of the larger studies in this area, Gfeller et al. (2002a) compared 49 CI users utilising either the CIS, SPEAK, or ACE speech-processing strategies, to 18 NH adults on their ability to recognise familiar melodies. These piano-played melodies were presented without vocal cues, both melody-line only, and melody with harmony, with the scores from both presentation formats being combined for the analysis of the results. The results showed that the NH subjects were able to identify significantly more melodies than the CI subjects. CI subjects scored 0% to 44% correct overall (mean=13%) with two-thirds of the correctly identified melodies having being classified by the authors as rhythmic in nature. NH subjects scored between 13% and 69%, with a mean of 55%; just over half of these correctly identified melodies were in the rhythmic category. A moderate negative correlation was found between melody recognition scores and age for the CI population, but only weak correlations were observed between the melody recognition scores and the length of profound hearing loss, or length of implant use. Moderate correlations were also found between some of the speech perception measures and melody recognition scores. As was the case in the instrument identification task, CI subjects were far less confident in identifying melodies than NH subjects, often waiting for the entire melody to be played and/or seeking reassurance as to the accuracy of their responses. These reports are in keeping with the observations mentioned earlier by Stordahl (2002) in her research with children. Gfeller et al. (2002a) surmised that melody recognition was a complex process for CI users, requiring not only good pitch discrimination, but also a multitude of other perceptual and cognitive skills.

In a more-recent study, Gfeller et al. (2005) investigated the recognition of familiar melodies adopting a slightly different approach. ‘Real world’ music stimuli were used, subdivided into three musical styles or genres – pop, country and western, and classical. The excerpts for the two former styles also included lyrics whereas the excerpts for the classical genre were entirely instrumental. The authors hypothesised that the presence of lyrics in the extract would enhance recognition scores in this open-set task. In another variation from the traditional melody recognition tests, the responses from each participant in the study were analysed in two formats or levels – ‘accurate recognition’ and ‘attribute recognition’. The former involved the subject correctly identifying the melody being played, whereas the latter constituted the recognition of an attribute within the melody, without being able to identify the melody itself. Examples of ‘attribute recognition’ were being able to name the artist or composer, repeating the lyrics of the extract, or singing back the melody line. This two-tier analysis was conducted to allow closer investigation of which melodic traits were best perceived by the listener. Up to 36 familiar melodies were incorporated in the open-set task. For the 59 postlingually deafened adults included in the study, the mean recognition score of 15.6% was significantly lower than the score of 54.7% for the 30 NH control subjects. As hypothesised, the CI group was significantly less able to identify the purely instrumental extracts of the classical genre compared to the other two styles where lyrics were present in the extract. Further, the authors reported that the most commonly recognised attributes for the CI subjects were lyric-based, such as being able to repeat lyrics in the extract, or recognising the singer’s voice. In comparison, the NH group were significantly more accurate at identifying the pop or classical extracts than the country and western excerpts with the presence of lyrics in the extract having less of an impact. This led Gfeller et al. (2005) to speculate that with musical features such as pitch, harmony, or timbre not being effectively transmitted to CI recipients, these subjects become more reliant upon traits such as vocal cues that are more salient through the implant.

Gfeller et al. (2003) investigated the appraisal of melodies, both in regards to personal preference, and perceived complexity, also across the same three musical styles of pop, country and western, and classical. The CI subjects provided similar ratings across the three genres with a strong preference for stimuli perceived to be ‘simple’. This was in contrast to the NH group who demonstrated definite stylistic preferences, along with a preference for stimuli perceived to be more complex. The authors hypothesised that it was possible that the CI subjects could not differentiate between the three styles, hence the uniformity across their ratings. In fitting with the identification findings reported in the previous paragraph, the CI recipients gave significantly lower appraisal ratings than the NH group for stimuli in the classical genre. Whereas both the country and western, and pop styles tended to have strong, easy-to-follow beats in addition to the vocal cues, the excerpts from the classical style were void of any lyrics or vocal cues, and were subjectively reported to sound more complex (Gfeller et al., 2003).

The use of combined acoustic and electric hearing by CI recipients for melody tests has also been investigated. Kong et al. (2005) compared the melody recognition skills for five CI subjects utilising a HA in the non-implanted ear, across three listening modalities – CI-alone, HA-alone, and both devices simultaneously. Three sets of 12 familiar melodies devoid of rhythm cues (i.e. having pitch cues only) were generated; one set for the low-frequency range, one for the mid-frequency range, and the other covering the high frequencies. For each frequency range, the melodies were presented three times for each of the three listening modalities. A practice session was provided prior to testing, with feedback on accuracy being given during the testing process. Results reflected the wide individual variability common to similar studies. For the HA-alone condition, scores ranged from 19% to 90%; for the CI-alone condition, scores ranged from 8% to 81%; and with both devices, scores spanned from 21% to 92%. The HA-alone score (mean = 45%) was, on average, 17 percentage points better than the average CI-alone performance, with little difference between the HA-alone and bimodal conditions.

In another study incorporating CI recipients using combined acoustic and electric hearing, Gfeller et al. (2004) investigated the melody recognition performance of five subjects implanted with a short electrode array which enabled them to use both modalities for hearing unilaterally. These subjects attained open-set melody recognition scores close to those of the NH subject group (84% and 85% respectively), with a large discrepancy between these scores and the mean score of 31% obtained by 27 implantees using a standard long electrode array. Hence, for subjects using the short electrode array in the study, the simultaneous use of residual acoustic hearing in combination with the implant resulted in melody recognition scores more similar to those of the NH subjects than the conventional CI device users. It should be reiterated that the short electrode array is only suitable for a limited group of patients. These potential recipients tend to have steeply sloping hearing losses, and could therefore have significantly greater levels of post-surgery residual hearing than conventional CI recipients. It is feasible that a CI recipient with a conventional long array may perform equitably to a recipient with a short array, should they have similar levels of post-surgery residual hearing. That is, as mentioned in section 4.2.3, it is the level of residual hearing, rather than the type of electrode array, that is the important variable in studies of this kind.

From the results of these studies involving the simultaneous use of acoustic and electric hearing, it may be surmised that the improved pitch perception provided by acoustic stimulation, as reported in section 4.2.3, translates to improved melody recognition skills for these bimodal device users. In the same way, the cumulative findings from studies involving CI subjects using a traditional CI device indicate that their diminished pitch perception ability does have an adverse affect on their ability to recognise melodies with which they were familiar. The effect of the unique pitch and timbre percepts obtained from electric, as opposed to acoustic, stimulation can also be indirectly observed in the disparity between appraisal ratings and listening preferences provided by CI and NH subjects for melodies across the various genres or styles.

4.3 REVIEW

This chapter has reviewed some of the existing literature pertaining to the performance of subjects with hearing impairments on various music perception tasks. Due to the lack of published research studies investigating the music perception skills of postlingually deafened adult HA users, the review concentrated primarily on studies undertaken with CI recipients. A wide range of procedures, methodologies, subjects, and hypotheses have been adopted and investigated by different researchers. Despite this, it is largely accepted that those with significant levels of hearing impairments, regardless of their age or the type of listening device they use, are not disadvantaged on rhythm-based tasks. Their performance on tasks requiring spectral analysis, however, is significantly poorer than that of their NH counterparts, particularly evident in studies involving CI users. Examples of these tasks include pitch discrimination, instrument identification, and melody recognition. These pitch- and timbre-based tasks involve the perception and integration of frequency-related information from the input signal, such as its spectral shape, temporal modulations, and harmonic structure. It may be that the subject with a hearing impairment is unable to obtain sufficient cues in order to successfully perceive pitch or timbre, and for CI users, this may be further confounded by the available cues providing inconsistent information. There is a high level of both inter- and intra-study variability depending upon the stimuli and methodology adopted in the research. One consistent finding, though, is the large variability between individual implantees' abilities to accurately perceive music, with no single variable or explanation being able to account for this.

Nevertheless, irrespective of the variability and the host of potential causes underlying this variability, the literature reviewed unequivocally demonstrates that the majority of CI users are significantly poorer than NH subjects at pitch discrimination, instrument identification, and melody recognition tasks. As would then be expected, this impaired perception of pitch and timbre is also reflected in studies of a more qualitative or subjective nature, where CI subjects have a tendency to rate music to sound different and/or less pleasant, appreciate different types and styles of music, and commonly spend significantly less time listening to music when compared to adults with NH.

CHAPTER 5: OVERVIEW OF THE CURRENT STUDY

5.1 RATIONALE

With the rapid growth of implant technology and the ever-increasing number of implantees, a substantial body of research is now available comparing the music perception skills of CI users to NH subjects as overviewed in the previous chapter. These findings collectively indicate that whilst CI users perceive musical rhythm approximately as well as those with NH, the performance of CI users on other music perception tasks such as instrument identification, melody recognition, and pitch discrimination is far less adequate (Fujita & Ito, 1999; Gfeller & Lansing, 1991, 1992; Gfeller et al., 1997, 1998, 2002a, 2002c; Leal et al., 2003; Looi et al., 2004; McDermott, 2004; Schulz & Kerber, 1994). This compromised performance may be associated with a range of factors including those arising from the processing of the acoustic signal through the CI to enable electrical stimulation of hearing, the design and function of the CI device, along with the characteristics of the listener's impaired auditory system. Electrical stimulation of hearing via an implant results in a unique percept of sound different from that arising from acoustic stimulation. Further, the effect of electrical stimulation on the sound perceived can be unpredictable, and highly variable from one implant user to the next.

Published studies have compared CI users to the NH population, but there are very few studies making comparisons to HA users with similar levels of hearing loss. Several authors have speculated that HA users may perform differently from CI users, but there has been little research to verify this. As mentioned in the previous chapter, Kong et al. (2005) conducted research involving bimodal device users; that is, adult CI subjects who also used a HA in the contralateral ear. However, this relatively small study involving five subjects only investigated melody recognition and speech perception. The current study compared the music perception skills of subjects who used a CI (when tested implant-alone), to subjects who only used a HA (i.e. no CI) across a broad range of musical tests. Hence the results for the HA condition in this study were obtained from subjects who relied primarily on HA(s) to amplify sounds in order to stimulate their hearing via normal auditory pathways, as opposed to those who used a HA to supplement the information from a CI. A preliminary report briefly mentioned within Gfeller's (2000) article suggested that children with prelingual hearing losses utilising CIs were more accurate on a range of rhythmic, pitch, and melodic perceptual tests than children with hearing impairments utilising traditional amplification devices such as HAs. However, little detail was provided as to the basis of this speculation.

A further consideration is that psychoacoustic research has shown that subjects with a cochlear hearing impairment also experience diminished pitch and timbre perception when compared to NH adults (Arehart, 1994; Moore, 1995, 1996; Moore & Peters, 1992; Summers & Leek, 1994). The extent to which an individual may be affected is highly variable; however, it appears that adults with a greater degree of hearing loss tend to perform worse on pitch- and timbre-based tests than those with a lesser degree of loss (Moore, 1995, 1996). Therefore the question arises as to whether some of the discrepancy between the results of CI users and NH subjects could be attributable to differences in hearing thresholds, and the physiological changes associated with a cochlear hearing loss, irrespective of the mode of stimulation used to elicit hearing sensations (i.e. electric or acoustic). Accordingly, this study aimed to compare the music perception of adults with hearing impairments using electrically stimulated hearing, via a CI, to that of adults who met the audiological criteria for a CI, but used acoustically stimulated hearing, assisted with HAs. This enabled the two subject groups to be better equated on the variable of overall hearing levels, and presumably more comparable in regard to the number of surviving auditory neurons in the cochlea than for comparisons made to NH subjects. Studies by Nadol et al. (1989) and Otte et al. (1978) have reported that the number of spiral ganglion cells in the cochlea generally decreases with increased levels and length of hearing loss. Further, these same studies also reported a decreased number of spiral ganglion cells with increased age. With many existing music perception studies comparing CI users to NH subjects reporting a significant difference between the groups for the variable of age, it would also follow for there to be a difference in the number of spiral ganglion cells between the groups solely on the basis of age.

In this study, the comparison between subjects using a CI and those using a HA was achieved by: i) comparing experienced CI users utilising current speech-processing strategies to HA users who met the audiological criteria for a CI, and ii) comparing a group of subjects on the waiting list for an implant, tested pre-implantation with their HA, and subsequently post-surgery with a CI. There has been no research into any aspect of music perception comparing the same subjects pre- to post-implant surgery, and few studies involving only current, commercially-available speech-processing strategies.

5.2 IMPLICATIONS AND SIGNIFICANCE

Over the last two to three decades, the CI has progressed from initially being an experimental device, to an accepted alternative when conventional amplification was of no benefit, through to now being the preferred habilitative or rehabilitative treatment for those with a severe to profound bilateral sensorineural hearing loss. The number of implantees is continually increasing as CI technology becomes more widely available, furthered by improvements in speech perception outcomes, increased publicity, and an ageing population. The broadening of the implantation criteria has enabled a larger number of patients to qualify for, and potentially receive, a CI. As current and future CI candidates and recipients tend to have greater levels of residual hearing than those who received their implant at an earlier time, they will inevitably become more discerning and expectant about outcomes they hope to obtain with the implant. Potential CI recipients need to be adequately counselled on the possible effect that implantation may have on music perception, in addition to speech perception.

With most implantees being able to attain excellent speech perception in quiet listening environments, regardless of the device they use, current and potential CI users are hoping to be able to listen to, and enjoy, other acoustic stimuli such as music. Music is often prioritised by many patients as it can serve not only to aid relaxation or reminiscence, but has auxiliary social and affective purposes as well. For example, a CI or HA user may be more inclined to go to church, attend a concert, or participate in a variety of other social functions if they are able to enjoy and appreciate the music being played. Music can also play an interpretive role through setting or reflecting the mood for a film or social event; this would enhance the quality of the overall experience or event for the participant. Therefore, music perception and appreciation goes beyond just being able to identify the specific musical instruments or works. Research of this kind is significant in terms of its potential to improve the quality of life of tens of thousands of people with significant hearing impairments. The results of the experiments in this study may provide a better understanding of both the music perception abilities of those with a moderately-severe to profound hearing loss who use CIs or HAs, as well as the factors that might limit their perception of music. Existing speech processors for CIs along with ‘music listening programs’ for HAs do not code music so that it sounds “normal”. Whilst many HA companies have incorporated music programs into their advertised HA features in an attempt to improve the perceived quality of musical sounds through the device, the manner in which each manufacturer addresses this differs, with no ubiquitous or well-accepted formula, and only limited success thus far.

5.3 AIMS AND HYPOTHESES FOR THIS STUDY

The current study aimed to investigate the music perception skills of postlingually deafened adults using a CI in comparison to postlingually deafened adults with similar levels of hearing impairment using a HA. This was achieved via two methods. Firstly, subjects who had utilised a CI for greater than one year (CI subject group) were compared to those using a HA (HA subject group) on tests assessing rhythm discrimination, pitch ranking, instrument perception, and melody recognition. Secondly, in order to eliminate some of the inter-subject variability that arises when making such between-group comparisons, a group of patients on the waiting list for a CI (WL subject group) was tested prior to implantation whilst utilising their HAs, and then again post-surgery, with their CI. These two methods cumulatively allowed the assessment of: 1) the music perception skills of CI users when compared to HA users with a moderately-severe to profound bilateral hearing loss, 2) the effect of cochlear implantation on an individual’s music perception skills; and 3) other subject factors that may affect a CI or HA user’s perception of music.

In consideration of the existing research findings reviewed in Chapters 3 and 4, the following hypotheses were determined for this research:

1) That the experienced CI users (CI subject group) would score lower than the HA

users (HA subject group) on the pitch, instrument identification, and melody tests, but not the rhythm test;

2) That subjects on the waiting list for a CI (WL subject group) would score higher on the pitch, instrument identification, and melody tests when tested with their HA pre-implantation than post-surgery with their CI; and

3) That subjects utilising a HA (i.e., both the HA subject group and the WL subject group when tested with their HA pre-implantation) would rate music to sound more pleasant than the subjects utilising a CI (i.e., the CI subject group and the WL subject group when tested post-implantation).

The materials, stimuli, and procedures used to investigate these hypotheses are outlined in the next two chapters.

CHAPTER 6: METHODS - MATERIALS

This chapter details the materials developed for this study. Section 6.1 describes the development of the music test battery, the process of obtaining and recording the stimuli for the tests, as well as the task requirements and subjects' response mode for each test. As a search of both research and clinical literature revealed no current widely-accepted published questionnaire assessing musical background, musical training, or listening preferences, two questionnaires were designed for this study: (i) the Music Training and Experience Questionnaire, and (ii) the Music Listening and Enjoyment Questionnaire. These are described in section 6.2.

6.1 THE MUSIC TEST BATTERY

6.1.1 Development of the Music Test Battery

A summary of the test battery appears in Table 6.1. Three fundamental components of music perception - rhythm, pitch, and timbre (in the form of instrument perception) were assessed, along with a melody recognition test where both the pitch and rhythm cues were preserved. It should be noted that in order to ensure that it was identification abilities being assessed, and not musical knowledge for the instrument perception and melody recognition tests, each subject's familiarity with the instruments, ensembles, and melodies was verified prior to testing. All of the subjects were familiar with all of the instruments and melodies utilised in the test battery. The order of the stimuli constituting each individual test or subtest was fully randomised, with the order of the tests within the battery being pseudo-randomised. For example, the presentation order of the vowels within each of the pitch subtests was randomised, however testing for each of the different interval sizes was completed before the next interval size was assessed. Written standardised instructions were developed and provided to subjects detailing the requirements for each of the tests or subtests. An opportunity to ask questions was provided prior to the commencement of each test, with no feedback given to subjects during the testing itself. Subjects were informed that they had to guess if they were unsure of an answer, and responses were entered directly into the computer.

Table 6.1: Summary of the Music Test Battery

Test	Subtests	Stimuli	Response	Scoring
Rhythms Discrimination		38 pairs of rhythm	Same-Different	/38
Pitch Ranking		4 sets of sung vowels: Male /a/, Male /i/, Female /a/, Female /i/	Higher-Lower	
	1) One-octave	24 intervals per set		/96
	2) Half-octave	24 intervals per set		/96
	3) Quarter-octave	32 intervals per set		/128
Instrument Perception	1) Single Instrument	12 instruments or ensembles	A – Closed set: list of 12 instruments or ensembles	1) /48 Rating: /10
	2) Solo Instrument with background accompaniment	4 presentations of each	B – Quality Rating: scale from 1 - 10	2) 2 runs; each /48 Rating: /10
	3) Music Ensemble	48 stimuli per subtest		3) /48 Rating: /10
Melody Recognition		10 melodies, 2 presentations of each	Closed set: list of 10 melody names	/20

6.1.2 Rhythm Test

The stimuli for this test were derived directly from Gordon's (1979, 1986) *Primary Measures of Music Audiation* (PMMA) rhythm subtest with minor modifications made to the presentation format. This is a standardised music aptitude test which has been previously used by other researchers across a variety of fields. As mentioned in Chapter 4, the stimuli in the rhythm subtest of the PMMA consist of short pairs of rhythm patterns, presented one after the other. Each individual rhythm pattern comprises a series of tones generated at the same pitch, formulated to be one measure (or musical ‘bar’) in length. Gordon (1986) provides more detail regarding the generation and recording of the original stimuli, which was then slightly modified for the purpose of this study.

Although Gfeller & Lansing (1992) evaluated the PMMA to be an appropriate measure for assessing rhythm discrimination, their research also highlighted a problem in administering the original test’s format. As the test was designed for children to mark their response on an answer sheet, each rhythm pair was preceded by a verbal prompt.

However, when implementing this test with implantees, Gfeller & Lansing (1992) discovered that subjects found it difficult to discriminate between the verbal prompt and the rhythm proper. In view of this, these prompts were removed for the purpose of the current study. This was achieved by firstly converting the initial audio tape recordings of the PMMA stimuli into a single computer WAV file via the computer software program '*Cool Edit Pro*'. This enabled each verbal prompt to be isolated and deleted, with 38 of the resulting rhythm pairs being saved as separate WAV files. This provided the additional advantage of allowing the item presentation order to be fully randomised through the software program '*MACarena*'. As a result, the order of presentation of rhythm pairs differed from that used in the original PMMA test, and varied randomly from one subject to the next. The use of MACarena also eliminated the need to pause the tape between test items; subjects could have as much time as they required to select 'same' or 'different', improving time efficiency. The 1.5 seconds of silence between the two rhythm patterns was preserved, but the 5-second break between test pairs from the original PMMA became redundant, as a response was required before the program presented the next pair of stimuli.

Subjects were required to provide a verbal response as to whether the two rhythm patterns constituting each pair were the same or different. They were asked to ignore any differences in pitch or loudness that they may perceive. That is, they were required to assess only the rhythmic properties of the stimuli; they did not have to identify how the patterns differed. A score out of 38 was obtained.

6.1.3 Pitch Test

This test comprised three subtests, each essentially identical in format, but with differing interval sizes constituting the pitch stimuli. The first subtest consisted of note pairs one-octave (12 semitones) apart, the second subtest utilised half-octave (6 semitones) intervals, and the third subtest assessed quarter-octave (3 semitones) intervals. It should be noted that the one-octave subtest was generated at a later date than the other two subtests; testing of the half- and quarter-octave intervals with initial subjects revealed numerous scores consistent with chance-level performance, highlighting the need to avoid a 'floor effect'. In view of this, the one-octave subtest was generated for administration with all subsequent subjects. The first three subjects in the CI subject group were not available to undertake the one-octave subtest when that extra condition was introduced into the study; they only undertook this subtest in the second administration of the test battery, approximately 4 months later.

The stimuli development process was identical for all three subtests. Recordings of the sung vowels /i/ (as in ‘heed’) and /a/ (as in ‘hard’) were obtained from trained male and female singers, encompassing a wide pitch range. A selection of these were then converted into WAV files to constitute the signals for this test. Each signal consisted of two different notes, of the same vowel and sung by the same singer, at the designated interval size. Each note was set up with a linear rise/decay ramp of 30 ms, with the two notes being presented sequentially, ascending or descending, separated by 500 ms of silence. A total of eight WAV files for each pair of pitches were created – four where the first note was higher than the second note (i.e. descending), and four in the reverse order (i.e. ascending). Initially, the levels of all the stimuli for each subtest across the four vowels were normalised as a group, using the ‘equal loudness contour’ group normalisation function in the software ‘Cool Edit Pro’. This function applied a manufacturer-derived algorithm to the selected stimuli to enable the presentation levels for both the male and female stimuli to be consistent across the range of pitches tested. Subsequent to this, the levels of the two notes forming each pitch pair were randomised by 6 dB below the initial comfortable loudness presentation level. This was undertaken in order to reduce the potential for loudness cues biasing the average pitch comparisons. A wide range of F0s were incorporated into the test (Table 6.2).

Subjects were required to state which of the two notes in each pair was higher in pitch, ignoring any difference in the loudness of the notes. The one-octave and half-octave subtests provided scores out of 96, whilst the quarter-octave subtest was out of 128. Using a two-alternative forced-choice response format, the chance score was 50%, with scores significantly less than 50% indicating pitch reversals (i.e., the pitch being ranked in the opposite direction to the change of the F0).

Table 6.2: Fundamental frequencies of the pitches included in the pitch test

Interval size		Fundamental frequency of pitches comprising each interval
Subtest 1 One-octave	Female	*C4-C5* (262-523 Hz); *D#4-D#5* (311-622 Hz); *F#4-F#5* (370-740 Hz)
	Male	*G2-G3* (98-196 Hz); *A#2-A#3* (117-233 Hz); *C#3-C#4* (139-277 Hz)
Subtest 2 Half-octave	Female	*C4-F#4* (262-370 Hz); *F#4-C5* (370-523 Hz); *C5-F#5* (523-740 Hz)
	Male	*G2-C#3* (98-139 Hz); *C#3-G3* (139-196 Hz); *G3-C#4* (196-277 Hz)
Subtest 3 Quarter-octave	Female	*C4-D#4* (262-311 Hz); *D#4-F#4* (311-370 Hz); *F#4-A4* (370-440 Hz); *A4-C5* (440-523 Hz)
	Male	*C#3-E3* (139-165 Hz); *E3-G3* (165-196 Hz); *G3-A#3* (196-233 Hz); *A#3-C#4* (233-277 Hz)

6.1.4 Instrument Perception Tests

The instrument perception tests incorporated two separate types of assessment - an identification task, and a quality rating task. Both tasks used the same test stimuli which are described in more detail in sections 6.1.4.1 – 6.1.4.3. Akin to the pitch test, the stimuli for the instrument test were subdivided into three subtests, each involving the same procedural format. The first subtest used single-instrument stimuli, the second used solo instruments with background accompaniment, and the final subtest evaluated music ensemble sounds. For each subtest, four extracts of 12 different instruments or ensembles were included (i.e., 48 stimuli per test). Each of these 5-second extracts was obtained from high quality, commercially-available compact-disc sound recordings of music works representing each instrument or ensemble. The use of four differing extracts as opposed to identical stimuli was viewed as more realistic, enabling a variety of styles, instrumentations, and tempi to be incorporated. As per the pitch test, the overall levels were firstly normalised across all 48 excerpts within the subtest using 'Cool Edit Pro', to minimise the effect of frequency variations on the perceived loudness of the extract. Subsequently, the levels of the four extracts for each instrument or ensemble were randomised within a 6 dB range below the initial comfortable loudness level.

For the identification task, a closed-set procedure was adopted whereby the subject was presented with a list of the twelve instruments or ensembles, each with a matching picture. Two different closed-set lists were compiled; one for the first and second subtests, and another for the third subtest. The subject was required to point to, or name, which stimulus they thought was playing, resulting in a score out of 48 for each of the subtests.

Two identification runs for the second subtest were conducted in order to assess whether the supplementary information pertaining to the background accompaniment impacted upon music perception, and if so, the nature of this interaction. Gfeller et al. (2003) reported anecdotal statements from implantees indicating that prior knowledge of stimuli, such as information regarding what they were hearing, or familiarity with the music, made it easier for them to follow. Therefore, for the first run of the second instrument identification subtest in this study, subjects were only told that the stimuli comprised a solo instrument with background accompaniment, without specifying the nature of this accompaniment. For the second run, subjects were additionally informed that the background ensemble for each extract was an orchestra, before being required to identify the relevant solo instrument from the closed-set list. This allowed the effect of prior knowledge to be assessed.

For the quality rating task, the same stimuli, and thus the same closed-set lists, were used as per the identification task. However, for this rating assessment, as each extract was being played, the researcher identified the instrument or ensemble in question to the subject. The subject was subsequently required to rate how pleasant each of the extracts sounded, on scale from 1 to 10, where 1 referred to the extract being ‘very unpleasant’, and 10 corresponded to ‘very pleasant’. If they could remember how this instrument or group sounded before they lost their hearing, they were asked to compare the excerpt’s sound to this memory. If they could not remember the specific instrumental sound, they were asked to rate the overall pleasantness of the musical excerpt. It should be emphasised that the instructions were not changed for the WL subjects from the pre-implant to post-implant test blocks. That is, the subjects were asked to make their judgements based on their memory of the specific instrumental sound pre-hearing loss. Therefore subjects utilising a CI (both the CI subject group and the WL group post-surgery) did not make their comparisons to the sound of an instrument through a HA, but rather to the instrumental sound prior to having a hearing loss. Subjects were encouraged to use the whole range from ‘1’ through to ‘10’, and an opportunity was provided for them to comment on, or provide extra detail, regarding the sound they heard.

6.1.4.1 Subtest 1

Based on pilot tests (Looi & McDermott, 2002), 12 commonly-heard solo instruments were selected for this subtest. These instruments covered the four instrumental families - strings, woodwind, brass, and percussion, across a wide pitch range. Both tuned and untuned percussion instruments were included, along with vocal stimuli. The following instruments constituted this subtest: male singer, female singer, piano, guitar, bass drum (or timpani), drum kit, xylophone, cello, violin, trumpet, flute, and clarinet. Table 6.3 provides the approximate note ranges for each of the solo instruments included in the first and second instrument perception subtests, with Table 6.4 providing the approximate F0 equivalents for these musical notes.

Table 6.3: Approximate note range for stimuli in the instrument perception subtests 1 and 2 (range across the 4 extracts)

Subtest 1: Single Instrument		Subtest 2: Solo Inst + Accompaniment	
Bass Drum	G2 – Eb3	Bass Drum	G2 - C4
Cello	C2 - C5	Cello	C#2 - E4
Clarinet	D3 – D6	Clarinet	G3 - Bb5
Drum Kit	N/A	Drum Kit	N/A
Flute	A4 – A5	Flute	C4 - D6
Female Singer	Bb3 – C#5	Female Singer	E4 - A5
Guitar	D3 - B4	Guitar	G3 - B4
Male Singer	E3 - C4	Male Singer	G2 - G4
Piano	G1 – Eb6	Piano	A1 - A7
Trumpet	A3 – Bb5	Trumpet	G3 - A5
Violin	A3 – Eb6	Violin	D4 - B6
Xylophone	C5 - E7	Xylophone	Bb3 - E7

Table 6.4: Fundamental frequencies for musical notes

Note	Hz	Note	Hz	Note	Hz	Note	Hz	Note	Hz	Note	Hz	Note	Hz
A1	55	A2	110	A3	220	A4	440	A5	880	A6	1760	A7	3520
A#1	58	A#2	117	A#3	233	A#4	466	A#5	932	A#6	1865	A#7	3729
B1	62	B2	123	B3	247	B4	494	B5	988	B6	1976	B7	3951
C2	65	C3	131	C4	262	C5	523	C6	1047	C7	2093	C8	4186
C#2	69	C#3	139	C#4	277	C#5	554	C#6	1109	C#7	2217	C#8	4435
D2	73	D3	147	D4	294	D5	587	D6	1175	D7	2349	D8	4699
D#2	78	D#3	156	D#4	311	D#5	622	D#6	1245	D#7	2489	D#8	4978
E2	82	E3	165	E4	330	E5	659	E6	1319	E7	2637	E8	5274
F2	87	F3	175	F4	349	F5	698	F6	1397	F7	2734	F8	5588
F#2	92	F#3	185	F#4	370	F#5	740	F#6	1480	F#7	2960	F#8	5920
G2	98	G3	196	G4	392	G5	784	G6	1568	G7	3136	G8	6272
G#2	104	G#3	208	G#4	415	G#5	831	G#6	1661	G#7	3322	G#8	6645

6.1.4.2 Subtest 2

In an extension of the previous subtest, the same 12 instruments were presented to subjects, but in a different context. This subtest utilised the above instruments, still in a solo role, but with the inclusion of background musical accompaniment (i.e., a solo instrumentalist accompanied by a music ensemble). The addition of background accompaniment gave rise to two major considerations. From one perspective, it added to the complexity of the stimuli, which, according to some previous research, may negatively impact upon hearing-impaired subjects' perception or appreciation of stimuli. For example, if the background accompaniment resembled noise for the listener, this could have a deleterious effect on music listening (Schulz & Kerber, 1994). On the other hand, though, the accompaniment may have served to provide extra rhythmic, timbral, and/or pitch cues to aid perception or appreciation. It is well accepted in studies of speech perception that contextual cues, and prior knowledge or information may aid hearing-impaired subjects' speech recognition performance; is there a music equivalent?

No adjustment was made to the signal-to-background level between the solo instrument relative to the accompaniment. Instead, as part of the initial test development process, NH subjects were asked to verify if the excerpt contained an obvious, easily perceived solo instrument with the orchestra playing a background accompaniment role only. It was felt that manipulation of the stimuli parameters to a certain, pre-determined ratio would be unrepresentative and unreflective of true everyday perceptual abilities; different music styles, recordings, and works have varying signal-to-background ratios between the soloist and the accompaniment. The listener would generally be unable to adjust or manipulate this ratio for everyday listening situations.

6.1.4.3 Subtest 3

The stimuli comprising this subtest consisted of 12 different music ensembles, each playing as a cohesive, unified group without a soloist. The ensembles chosen covered a variety of instrumental combinations, genres, styles, and group sizes. The selected ensembles were:

- choir (four-part; a cappella)
- orchestra
- jazz band (instrumental only – no voice)
- rock band (instrumental only – no voice)
- country and western group (instrumental only – no voice)
- string quartet
- percussion ensemble (varying instrumental combinations)
- violin + piano (duet)
- cello + piano (duet)
- male singer + piano (duet)
- female singer + piano (duet)
- 1 male and 1 female singer, with piano accompaniment (trio)

Where practical and appropriate, a range of styles was incorporated, such as a choir singing a four-part chorale or folk songs, and a jazz band playing the blues or swing. Further, the four extracts per ensemble type were not restricted to one group, nor one recording. For example, different rock bands were included in the four extracts for ‘rock band’, and different combinations of instruments constituted the four recordings for the ‘orchestra’ or ‘jazz band.’ It was felt that this would increase the generalisability of the obtained results, ensuring that they did not represent the perception of only one type of recording, or one musical style.

6.1.4.4 Instrumental Picture Verification Procedure

In order to ensure that the pictures accompanying the lists of instrument or ensemble names were clear, unambiguous and identifiable, the pictures were initially verified with adult members of the general public. For each of the instruments and ensembles that comprised the closed-set tests, the researcher selected two pictures from primary school music resource books. These pictures were then photocopied to approximately equal sizes (7 cm x 7 cm for the solo instruments, and 10 cm x 10 cm for the ensembles). Ten members of the general public were asked to view these pictures presented in a random order, and to name the instrument, or ensemble, represented in the picture. Any picture that was not correctly identified by at least 9 of the 10 assessors was replaced. Subsequently, the two pictures corresponding with each instrument or ensemble were paired. These pairs were then shown to a further five different members of the general public, accompanied by the name of the instrument or ensemble. The adults were asked to select which of the two pictures most clearly represented the target instrument; this choice was recorded by the researcher with the most popular picture from each pair being incorporated into the instrument test.

6.1.5 Melody Recognition Test

Ten well-known melodies were recorded two times each, once with the preset clarinet sound and once with the preset oboe sound on a *Yamaha PSR-276* portable keyboard. These melodies had been identified in a separate study conducted by the author and colleagues (Looi et al., 2003) as ten of the most-familiar melodies to an Australian population, inclusive of those with no hearing impairment, as well as both HA and CI users. The ten melodies chosen were (in alphabetical order): Advance Australia Fair, Baa Baa Black Sheep, For He’s a Jolly Good Fellow, Happy Birthday, Jingle Bells, O Come All Ye Faithful, Old Macdonald Had a Farm, Silent Night, Twinkle Twinkle Little Star, and Waltzing Matilda. These melodies included both rhythmically complex tunes (e.g. ‘For He’s a Jolly Good Fellow’), and rhythmically simple tunes (e.g. ‘Twinkle Twinkle’).

Each melody was played in C major, centering around middle-C on the keyboard, at a speed of 100 beats per minute (i.e. crotchet = 100). All of the notes for each melody fell in the range from C3 to C5 (131 Hz to 523 Hz). It is acknowledged that there are often slight variations to the rhythms of such commonly known melodies; therefore, for this study, melodies were recorded in the simplest rhythm appropriate for that tune. Where the song consisted of both a verse and chorus, only the chorus was recorded. Subsequently, the first 15 seconds of each melody was extracted; 15 seconds was deemed to be a sufficient length to enable recognition of known melodies without unnecessarily prolonging the test. Having all melodies the same length was considered preferable to playing the entire melody, as some melodies were relatively short (e.g. ‘Happy Birthday’ ~ 15 seconds), whilst others were comparatively long (e.g. ‘Advance Australia Fair’ and ‘O Come All Ye Faithful’ ~ 45 to 50 seconds). Further, 15 seconds was approximately the length of the shortest melody in the set. The last 5 seconds of each extract was linearly ramped to zero amplitude. Six NH subjects initially verified the recognisability of these melodies.

A closed-set identification format was employed where subjects selected from a list of the 10 melody names. With each melody being presented two times within a run, a score out of 20 was obtained.

6.2 DEVELOPMENT OF QUESTIONNAIRES

6.2.1 Introduction

Two questionnaires were designed for this study: (i) the Music Training and Experience Questionnaire (MTEQ), and (ii) the Music Listening and Enjoyment Questionnaire (MLEQ) (included in Appendices 2 and 3). These questionnaires were based on one developed by Gfeller et al. (2000b), with significant modifications made to ensure its applicability to both the Australian population and this study.

The MTEQ assessed previous music training and participation levels, whilst the MLEQ examined a subject's music preferences and opinions. Collecting both qualitative and quantitative information in either a face-to-face or take-home format, the results of the questionnaires were examined to enable classification of the subjects into various categories, based upon their music experiences and listening habits. These classifications enabled correlations to be calculated between such subject variables and perceptual performance.

6.2.2 Music Training And Experience Questionnaire (MTEQ)

The information obtained from the subjects' responses to this questionnaire formed the primary basis from which the music experience scores were calculated in this study. The questionnaire was given to all subjects, including those with NH, to complete. The first six questions asked respondents to detail the number of years and relevant ages for which they were involved in either instrumental or academic music lessons, or participated in ensembles or other musical activities. The seventh question asked them to subjectively rate their knowledge of music history and theory, their ability to read and play music, as well as their overall music ability. Subjects who used a HA or CI were also asked if they had been involved in music lessons, groups, or activities since being fitted with the relevant device.

Subjects' responses were collectively used to determine a music experience score, as reported in Tables 7.1, 7.2, and 7.3 at the end of the next chapter (p. 125-127). Subjects were ranked into three categories: '0' indicated no formal music training or participation in music activities or music classes; '1' indicated having had instrumental lessons for 2 years or less, and/or participation in music activities or music classes for 5 years or less; and '2' indicated those subjects who had formal instrumental lessons for more than 2 years, and/or participation in informal music activities or music classes for more than 5 years.

6.2.3 Music Listening And Enjoyment Questionnaire (MLEQ)

This questionnaire was only administered with the subjects who had hearing impairments involved in this study (i.e. not the NH subject group). Its main purpose was to gauge the effect of the hearing device on a subject's appreciation of, and time spent, listening to music. For example, subjects were asked to estimate the amount of time that they spent listening to music both prior to their hearing loss, and now whilst utilising their current device. The following categories were provided for making this self-rating: 0 = 'never'; 1 = 'occasionally'; 2 = 'sometimes'; 3 = 'often'; and 4 = 'very often'. These ratings are reported in Tables 7.2, 7.2, and 7.3 (p. 125-127, at the end of Chapter 7) as their 'pre-hearing loss listening score' ('pre-HL listening score') and 'current listening score'. Subjects were also asked to compare the sound of different types of music stimuli from a time when they had better hearing and were not using a HA or CI, to the present time whilst listening with their current device. The questions encompassed areas including general enjoyment of music, and listening habits, as well as asking respondents to provide ratings for specific musical instruments, ensembles, and music genres. The last part of the questionnaire asked subjects to review a list of tactics and traits that may impact upon their listening experience, such as familiarity with the work, the simplicity of the music, or the attributes of the surrounding listening environment. The qualitative nature of the MLEQ also provided respondents with the opportunity to make further comments on their experience of listening to music with a CI or HA.

In a modification to this questionnaire, a variation was developed for use with WL subjects only, for administration after their CI operation – the MLEQ-WL Post Questionnaire (Appendix 4). The questions and format of this modified version were largely identical to the original, except that it aimed to compare the sound of music with a CI versus a HA. That is, subjects were asked to make the relevant comparisons for time spent listening to music as well as sound quality judgements between post-surgery, whilst listening with the CI, to before their operation when listening with HAs. By integrating responses provided on the initial MLEQ, the administration of the MLEQ-WL Post version also allowed a comparison to be made between the amount of time currently spent listening to music with the CI, and the amount of time spent listening to music prior to having a hearing loss. The 'CI listening score' reported in Table 7.3 (p. 127) was based on the subject's rating of the time they spent listening to music post-surgery, whilst using the implant.

Three more general questions were added to this follow-up MLEQ to allow the author to gauge a WL subject's satisfaction with the CI. It would not be unreasonable to consider that their general satisfaction with the new device may impact upon their responses to the music rating comparisons. Subjects were asked to self-rate the amount of difference that the implant had made to their speech perception, compared to using HAs, on the following scale: 1 = made it worse; 2 = no change; 3 = made it a little better; 4 = made it somewhat better; 5 = made it much better. There was also a question asking them to rate their overall satisfaction with the CI on a scale from 1 to 5, where 1 = unsatisfied; 2 = indifferent; 3 = a little satisfied; 4 = somewhat satisfied; 5 = very satisfied. A third question asked subjects whether overall the CI met their expectations – 'no', 'yes', or 'unsure'.

It should be noted that the purpose of both versions of the MLEQ was two-fold. Firstly, the responses allowed the author to gauge if factors such as the time spent listening to music, and for the WL subjects, general satisfaction with the CI, were predictive of perceptual abilities on the music tests. That is, if there were other subject factors that may affect a CI or HA user's perception of music. Secondly, the questionnaire also collected a substantial amount of qualitative and descriptive information that may be of interest for future studies. Only the information having a direct bearing on the results or findings of this study are presented in Chapters 8 and 9.

CHAPTER 7: METHODS – SUBJECTS & PROCEDURES

This chapter firstly details the three different research subject groups who were involved in this study, as well as a comparative group of NH subjects who verified the individual tests (section 7.1). Section 7.2 then describes the procedures used for testing the research subjects.

7.1 SUBJECTS

Ethical approval for this study was obtained from the Royal Victorian Eye and Ear Hospital's Human Research Ethics Committee, and the experimental, subject selection, data collection, and data storage procedures were in full accordance with these ethical requirements. All subjects were informed of their rights, and they signed the appropriate consent forms prior to the commencement of testing. They were free to withdraw from the study at any time, and participants were not paid for their involvement in the research.

7.1.1 Normally Hearing Subjects

Although the aim of this research was to compare the performance of CI users to HA users on a variety of music perception tasks, it was felt that the music tests should be verified on a group of NH subjects in order to confirm the tests' appropriateness and feasibility before administration to the subjects with hearing impairments. Therefore, 10 NH subjects (7 females, 3 males) were recruited to undertake the quarter-octave pitch test, the three closed-set instrument identification tests, and the melody recognition test. The one-octave and half-octave pitch tests were not tested for time-efficiency. As it was expected that the NH subjects would be able to reliably rank pitches a quarter-octave apart, and therefore for a 'ceiling effect' to be observed in the results, it was assumed that testing of larger, and hence easier, intervals was not necessary. All of the NH subjects had bilateral hearing thresholds ≤ 25 dBHL at 250 Hz, 500 Hz, 1000 Hz, 2000 Hz, 4000 Hz, and 8000 Hz. This was verified by an audiologist prior to testing. The subjects ranged in age from 22 to 51 years (mean: 32.9 years).

It should be emphasised that testing with NH subjects was performed for test verification purposes, and as such, their results were not statistically compared to the CI or HA groups' results. The mean scores for the NH group for each test were greater than 95%, indicating that the NH subjects found the tests to be relatively easy.

7.1.2 Cochlear Implant Subject Group (CI Group)

Fifteen postlingually deafened adult users of the Nucleus CI system (7 male, 8 female) were recruited from two CI clinics, one in Melbourne, and one in Brisbane, Australia. Subject details appear in Table 7.1 (p. 125, at the end of this chapter). Potential subjects were sent letters to invite them to be involved in the study. The fifteen subjects ranged in age from 36 to 75 years (mean: 60.4 years). There were eight users of the more-recent CI24 system (both the CI24M and CI24R devices), and seven users of the preceding CI22 system. Six subjects used body-worn speech processors (4 SPrint, 2 Spectra22), and nine used ear-level devices (4 ESPrit22; 5 ESPrit3G). There were eight subjects using the ACE strategy, with stimulation rates ranging from 275 Hz to 1800 Hz per electrode, and seven using the SPEAK strategy. The pre-surgery audiogram of the group's average hearing thresholds at octave frequencies between 250 Hz and 8000 Hz for the implanted ear is included in Figure 7.1.

The speech perception scores reported in Table 7.1 were obtained using CUNY (City University of New York) sentences presented at 65 dBSPL from a loudspeaker in a sound-treated room. For subjects whose clinical files did not have a record of such scores obtained with their current speech-processing strategy, the author conducted this speech test during the course of the study. The 'music experience score', along with the 'pre-hearing loss listening score' (pre-HL listening score) and 'current listening score', were determined from the subject's responses on the MTEQ and MLEQ respectively, as described in the previous chapter.

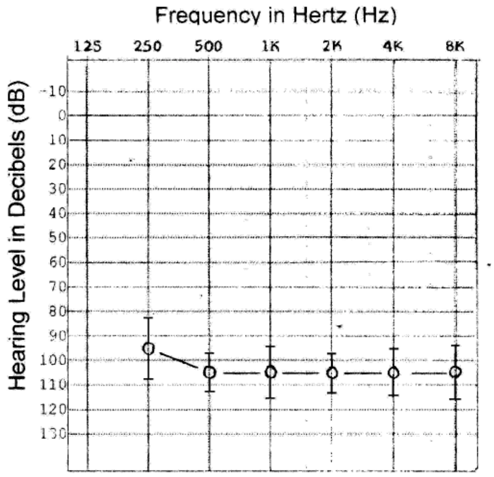


Figure 7.1: Average thresholds of the CI subject group

This audiogram presents the average pre-surgery unaided hearing thresholds of the CI subject group, for the ear implanted with the CI. The circle represents the mean hearing threshold across the 15 subjects for that frequency, with the error bars indicating 1 standard deviation. These results were obtained from the subject's audiology files. For these mean thresholds, a maximum figure of 110 dB was used. That is, where a subject's threshold was equal to, or greater than 110 dB, or noted on the audiogram as being beyond the limits of the audiometer, 110 dB was recorded for that frequency.

7.1.3 Hearing Aid Subject Group (HA Group)

Fifteen postlingually deafened adult HA users were also recruited from a range of sources including CI clinics, as well as general audiology clinics. Details pertaining to these subjects appear in Table 7.2 (p. 126, at the end of this chapter). Recruitment procedures were similar to those for the CI subjects, except that HA subjects were required to meet the audiological criteria to qualify for a CI in terms of level of hearing loss, and speech perception scores. These criteria included having a bilateral moderately-severe to profound sensorineural hearing loss between 1 kHz and 4 kHz, and speech perception scores for sentence stimuli presented auditory-alone in quiet at 65 dBSPL of less than 70% in the best-aided condition, and less than 40% for the ear recording the poorest scores. In order to ensure that these criteria were met, potential subjects' aided speech perception abilities were initially tested by the researcher using CUNY sentences presented at 65 dBSPL from a loudspeaker in a sound-treated booth. These sentence perception scores were obtained for each ear individually, as well as binaurally. Those adults who fulfilled the criteria detailed above were then invited to proceed with the music testing. Recruited subjects ranged in age from 49 to 80 years (mean: 64.7 years). The audiogram of the group's average hearing thresholds at octave frequencies between 250 Hz and 8000 Hz for the ear tested in this study is included in

Figure 7.2. Speech perception scores are listed in Table 7.2. As subjects utilised their personal HA(s), a range of HA brands were involved; all were digitally-programmable or digital behind-the-ear models, as listed in Table 7.2. None of the subjects used a special music listening program or novel device setting for the tests. As per the CI subject group, the HA subject's responses on the MTEQ and MLEQ determined their 'music experience score', 'pre-HL listening score', and 'current listening score', listed in Table 7.2.

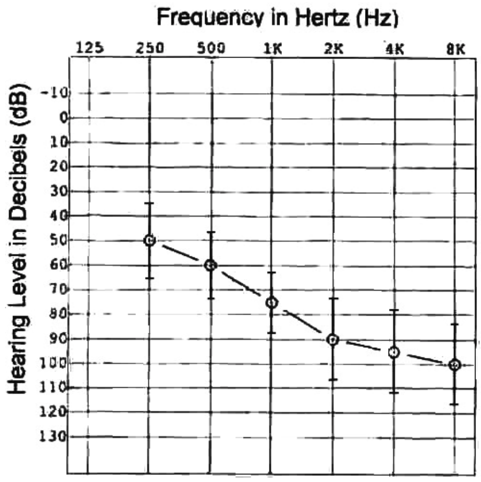


Figure 7.2: Average thresholds of the HA subject group
This audiogram presents the average unaided hearing thresholds of the HA subject group. The circle represents the mean hearing threshold across the 15 subjects for that frequency, for the ear used by the subject to undertake the tests in this study. The error bars indicate 1 standard deviation. These results were obtained from the subject's audiology files. For these mean thresholds, a maximum figure of 110 dB was used. That is, where a subject's threshold was equal to, or greater than 110 dB, or noted on the audiogram as being beyond the limits of the audiometer, 110 dB was recorded for that frequency.

7.1.4 Waiting List Subject Group (WL Group)

Nine subjects (7 male, 2 female) on the waiting list for an implant, who then subsequently received their implant, also participated in the study. Details of these subjects are outlined in Table 7.3 (p. 127, at the end of this chapter). Subjects were recruited from the same two clinics as the CI subject group, with all of the required pre- and post-implant audiological and medical assessments being conducted by the relevant clinic. All of the subjects were postlingually deafened adults ranging in age from 41 to 71 years (mean: 54.3 years). They were all implanted with the Nucleus CI24R implant, and used the ACE strategy, with rates ranging from 250 Hz to 1200 Hz per electrode.

Both pre- and post-implantation speech perception scores for CUNY sentences were recorded, as reported in Table 7.3. The pre-implant speech scores were obtained as part of the CI assessment process, with the post-implant scores being obtained during the CI evaluation testing applicable to each clinic. The most recent post-implant CUNY sentence score that preceded the commencement of the post-implant music testing block is reported. The subject group's average pre-CI unaided hearing thresholds for the octave frequencies between 250 Hz and 8000 Hz are included in Figure 7.3. The 'music experience score', 'pre-HL listening score', and 'HA listening score' in Table 7.3 were determined from responses on the MTEQ and MLEQ respectively. The 'CI listening score' was ascertained from the MLEQ-Post version, as described in Chapter 6, section 6.2.3.

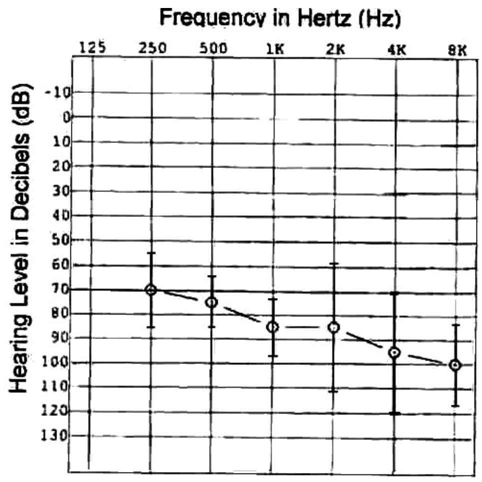


Figure 7.3: Average thresholds of the WL subject group
This audiogram presents the average pre-surgery unaided hearing thresholds of the WL subject group. The circle represents the mean hearing threshold across the 9 subjects for that frequency, for the ear used by the subject to undertake the pre-CI surgery test block of this study. The error bars indicate 1 standard deviation. These results were obtained from the subject's pre-CI assessment audiogram. For these mean thresholds, a maximum figure of 110 dB was used. That is, where a subject's threshold was equal to, or greater than 110 dB, or noted on the audiogram as being beyond the limits of the audiometer, 110 dB was recorded for that frequency.

7.2 PROCEDURES

7.2.1 Music Test Battery Administration

The music test battery detailed in Chapter 6 was administered on two occasions with all hearing-impaired subjects. For the WL group, one run of the battery was conducted pre-implant (test block 1), with the second run conducted around 3 months post switch-on of the implant (test block 2), to coincide with the 3 months speech-perception assessments often conducted by the audiology clinic. In order to assess the potential of a learning effect biasing the within-group comparison for the WL subjects, the HA and CI groups were also tested twice, approximately 4 months apart. The tests were generally administered over two to three sessions, each of around 1 to 1½ hours duration, although this was flexible, and organised to accommodate individual subjects.

Prior to the commencement of testing, the MTEQ and MLEQ were administered. In order to ensure that questions were understood and correctly interpreted, the author went through both questionnaires with each subject. The amount of time required for this varied between subjects. Those who then expressed confidence in independently completing the questionnaires were given them to take home. If the subject appeared hesitant in interpreting the questionnaire, or did not wish to complete it at home, both questionnaires were administered face-to-face in the initial session.

7.2.2 Stimuli Presentation

The individual WAV files comprising each test or subtest of the battery were saved into the software program '*MACarena*' which automatically randomised the presentation order of stimuli, and allowed the subject's response to be recorded for later analysis. All stimuli were presented from a *Dell 'Latitude'* laptop computer, connected to an external *Creative* '*Soundblaster Extigy'* soundbox. Both the CI subjects, and the WL subjects in their post-implant testing phase, used their own speech processor, processing strategy, and preferred processor settings for the study, although optional specialised compression algorithms such as '*Whisper*' were disabled for the tests. Similarly, the HA subjects, including the WL subjects pre-implant, used their own HAs, listening program, and usual device settings for the testing. That is, subjects did not use a separate music listening program, or a novel device setting with unique gain, frequency response and/or compression parameters targeted to music listening situations. All of the testing was performed in sound-treated audiology rooms or booths.

Where possible, direct audio input (DAI) was used to present the test stimuli to the subjects. DAI enabled the presented sound to bypass the microphone system on the device, and concurrently eliminated some of the variations in frequency response that may have otherwise arisen should the stimuli have been presented via a loudspeaker in the soundfield. With testing occurring at a few different clinics, the use of DAI avoided the variations in the frequency response resulting from having different loudspeakers, as well as different acoustic and physical characteristics for the numerous test rooms. For subjects using a CI, the DAI setup involved connecting the speech processor directly to the soundbox of the computer via an audio cable. For those using a HA, DAI was configured via an audio shoe, where available, attached to the base of the HA, with a cable connecting this audio shoe to the computer's soundbox. These connection cables, for both the CI and HA, were manufactured to provide the same frequency response as would occur if the sound had been picked up by the microphone on the respective device. For situations where DAI was not possible, as not all HAs enabled DAI or had appropriate audio shoes, a neck-loop system was utilised, via the telecoil on the HA. The neck loop was plugged directly into the soundbox; when the subject switched to the 'T' mode on their HA, this disabled the microphone and allowed them to pick up the stimuli through the loop system.

For NH subjects, *Etymotic ER4B* flat-frequency response earphones were used, with the earphones being plugged directly into the soundbox. The manufacturer describes the ER4B as being referenced to a flat diffuse sound field, and appropriate for use in perceptual research where the goal is to emulate the same frequency response at the eardrum as would occur for sounds presented in a live situation (i.e. without any equalisation for loudspeakers). The earplugs of the Etymotic earphones are designed to be inserted into the ear canal, and therefore also passively act to attenuate external background noise.

For the HA subjects, the ear with which the subject obtained better speech perception scores was used for testing, or in cases with similar or fluctuating losses, the ear which the subject identified as their preferred ear. This was done in order to match the procedures for the WL subjects. In the majority of cases, a CI is implanted into the poorer-performing ear, often one with minimal or no residual hearing. In view of the consideration that some potential CI recipients only utilise a monaurally-fitted HA in their better-hearing ear, whilst some others may have insufficient residual hearing in the ear to be implanted to feasibly undertake the tests, it was necessary to conduct pre-implant testing with WL subjects utilising their better-hearing ear. For 8 of the 9 WL subjects in this study, this was the non-implanted ear. The exception to this was subject 1 who had a fluctuating hearing loss and recorded similar speech perception scores for both ears. The clinical decision regarding which ear was to receive the implant was not made until just prior to his surgery, at which time the pre-surgery music testing had been completed. As he ultimately received the CI in his very marginally 'better-hearing' ear, pre- and post-surgery music test results were obtained from the same ear. Post-surgery, all of the WL subjects were tested using only their CI. That is, subjects were not assessed bimodally using a CI in conjunction with a HA.

In order to ensure that stimuli were presented at a comfortable level for each subject, subjects were individually asked to verify that a calibration noise was of a comfortable loudness, prior to testing. The calibration stimulus consisted of continually repeating 1-second bursts of ICRA (International Collegium of Rehabilitative Audiology) noise, separated by 1 second of silence. Developed by the International Collegium of Rehabilitative Audiology (ICRA, 2005), the ICRA noise is a collection of sounds designed for use in clinical HA testing. Signals were designed to have well-defined spectral and temporal characteristics, reflecting real life speech signals and babble noise. Dreschler et al. (2001) provides more specific information pertaining to the development and acoustic characteristics of the ICRA noise. This noise stimulus was set up as a WAV file and substituted the original calibration WAV file in MACarena. For each subject, the calibration noise was initially presented at a low level before being gradually increased until the subject judged the sound to be at a comfortable level. The process was repeated, with the average of the two levels being used. This procedure was performed prior to the commencement of testing for both test blocks. The level of the calibration noise was equal to the 'initial comfortable loudness presentation level' mentioned earlier in relation to the pitch, instrument, and melody tests, around which level randomisation was performed.

Table 7.1: Cochlear Implant Subject Group's Details

Aetiology: C/P=Congenital/Progressive
Experience with the device: Subjects marked * had been reimplanted. Total number of months with both devices is reported
Speech perception: Word score (%) for CUNY sentence test, tested CI only, presented in quiet
Music Experience Score: Ascertained via the MTEQ (range: 0 – 2; see Chapter 6, section 6.2.2)
Pre-HL listening score (i.e. pre-hearing loss listening score), and *Current listening score*: Ascertained via the MLEQ (range: 0 – 4; see Chapter 6, section 6.2.3)
Mode of stimulation: MP=monopolar; BP=bipolar; CG=common ground

Subject (M/F)	Age	Aetiology	Device experience (months)	Speech perception score	Music experience score	Pre-HL listening score	Current listening score	Type of CI	Ear	Processor	Strategy	Stimula-tion mode	No. chan-nels	HA in other ear
1 (M)	47	C/P	16	95	2	3	3	24R	R	Sprint	Ace 1200Hz	MP	22	N
2 (M)	67	Otosclerosis	60	75	2	3	2	24M	L	Sprint	Ace 1200Hz	MP	22	N
3 (F)	45	C/P	22	99	2	1	1	24M	R	Sprint	Ace 1800Hz	MP	22	N
4 (F)	36	Rubella	108	61	0	3	4	22M	R	Esprit22	Speak	BP	20	N
5 (F)	72	C/P	24	96	1	3	1	24M	L	Esprit3G	Ace 720Hz	MP	20	Y
6 (F)	56	C/P	17	100	1	3	3	24R	L	Esprit3G	Ace 900Hz	MP	20	Y
7 (M)	71	Trauma	300*	97	2	4	1	24M	L	Sprint	Ace 275Hz	MP	20	N
8 (F)	75	C/P	38	95	2	3	0	24R	R	Esprit3G	Ace 900Hz	MP	20	Y
9 (F)	70	C/P	180*	84	0	4	0	22M	L	Esprit22	Speak	CG	20	N
10 (F)	61	C/P	18	78	0	4	1	24R	L	Esprit3G	Ace 500Hz	MP	20	Y
11 (M)	48	C/P	135	72	0	3	1	22M	R	Esprit22	Speak	BP	18	N
12 (F)	66	Meningitis	211	90	2	3	0	22M	R	Esprit3G	Speak	Varied	16	N
13 (M)	69	C/P	138	37	0	2	0	22M	L	Esprit22	Speak	BP	20	N
14 (M)	64	C/P	184	79	0	2	2	22M	L	Spectra22	Speak	BP	16	N
15 (M)	59	Trauma	185	94	0	2	0	22M	L	Spectra22	Speak	CG	16	N

Table 7.2: Hearing Aid Subject Group's Details

Aetiology: C/P=Congenital/Progressive
Speech perception: Word score (%) for CUNY sentence test, obtained in the best aided condition, presented in quiet
Music experience score: Ascertained via the MTEQ (range: 0 – 2; see Chapter 6, section 6.2.2)
Pre-HL listening score (i.e. pre-hearing loss listening score), and *Current listening score*: Ascertained via the MLEQ (range: 0 – 4; see Chapter 6, section 6.2.3)
Mode of presentation: DAI=direct audio input; Loop=neck loop system in conjunction with the telecoil

Subject (M/F)	Age	Aetiology	Device experience (months)	Speech perception	Music experience score	Pre-HL listening score	Current listening score	Type of HA	Ear tested	Mode presen-tation
1 (F)	62	Viral	96	48	2	2	0	Phonak Supero	R	DAI
2 (F)	56	Otosclerosis	240	65	1	2	2	Phonak Perseo 311dAZ	L	Loop
3 (F)	56	C/P	276	51	2	4	2	GN Resound Canta7	L	Loop
4 (F)	61	C/P	384	38	1	4	2	Bernafon PB675	L	DAI
5 (F)	74	Unknown	180	0	0	2	1	Phonak Supero	L	DAI
6 (M)	67	C/P	492	23	0	1	3	Bernafon PB675	L	Loop
7 (M)	76	Infection	240	7	1	2	2	Phonak Supero	R	DAI
8 (M)	70	Otosclerosis	264	48	1	2	4	Phonak Supero	R	DAI
9 (F)	60	Otosclerosis	408	67	2	3	2	Phonak Supero	R	DAI
10 (F)	80	Meniere's	360	17	2	2	1	Phonak Supero	R	DAI
11 (M)	70	Noise Exp	120	50	0	1	1	Phonak Supero	R	DAI
12 (M)	70	Unknown	120	27	0	2	0	Oticon Digifocus II	R	Loop
13 (F)	49	C/P	156	56	2	4	2	Phonak Supero	R	DAI
14 (F)	62	Meniere's	180	63	2	4	1	Siemens Music Pro	R	Loop
15 (M)	57	Unknown	96	63	0	3	1	Phonak Sonoforte2	R	DAI

Table 7.3: Waiting List Subject Group's Details

Aetiology: C/P=Congenital/Progressive
Speech pre-CI: Pre-implant word score (%) for CUNY sentence test, obtained in the best aided condition, presented in quiet
Speech post-CI: Best post-implant word score (%) up to 3 months post switch-on, for CUNY sentence test, tested CI only, presented in quiet
Music experience score: Ascertained via the MTEQ (range: 0 – 2; see Chapter 6, section 6.2.2)
Pre-HL listening score (i.e. pre-hearing loss listening score), and *HA listening score*: Ascertained via the initial MLEQ (range: 0 – 4; see Chapter 6, section 6.2.3)
CI listening score: Ascertained via the follow-up MLEQ, administered post-implant (range: 0 – 4; see Chapter 6, section 6.2.3)
Mode of presentation: DAI=direct audio input; Loop=neck loop system in conjunction with the telecoil
Mode of stimulation: MP=monopolar
HA in other ear: Whether the subject wore a HA in their contralateral ear, post-CI

Sbjt (M/F)	Age	Aetiology	HA Type	HA exper-inece (mths)	Speech pre-CI	Music exper-ience score	Pre-HL listening score	HA listening score	Mode presen-tation	Type of CI	Ear CI	Speech post-CI	CI listening score	Processor	Strategy	Stimu-lation mode	No. chan-nels	HA in other ear
1 (M)	45	Meniere's	Siemens Prisma 2	60	23	2	4	0	DAI	CI24R	L	100	1	Esprit3G	Ace 900Hz	MP	20	N
2 (M)	70	Otoscle-rosis	Oticon Ergo	96	64	1	4	0	DAI	CI24R	L	99	1	Esprit3G	Ace 900Hz	MP	20	Y
3 (M)	51	Meniere's	Phonak Claro	18	61	1	2	0	Loop	CI24R	L	96	1	Esprit3G	Ace 900Hz	MP	20	Y
4 (M)	71	C/P	Bernafon LS16D	300	20	0	2	0	DAI	CI24R	L	94	1	Sprint	Ace 1200Hz	MP	22	Y
5 (F)	60	Familial	Siemens Prisma 2	600	3	0	2	2	Loop	CI24R	L	100	1	Esprit3G	Ace 900Hz	MP	20	Y
6 (F)	50	German Measles	Phonak Piconet	336	67	0	3	1	DAI	CI24R	R	99	2	Esprit3G	Ace 250Hz	MP	20	Y
7 (M)	46	Familial	Siemens MusicD SP	456	57	1	0	2	Loop	CI24R	L	99	3	Esprit3G	Ace 900Hz	MP	20	Y
8 (M)	41	Familial	Widex Senso	420	40	1	3	1	Loop	CI24R	L	98	2	Esprit3G	Ace 900Hz	MP	20	Y
9 (M)	55	C/P	BE-15 (from UK)	636	21	1	2	0	Loop	CI24R	R	88	1	Esprit3G	Ace 900Hz	MP	20	N

CHAPTER 8: RESULTS

Due to the quantity of data, and the multiple cross-group analyses required for this research, the results of the study are presented over two chapters. This chapter includes only the raw data scores for the three hearing-impaired subject groups for the music test battery, along with general comments on the overall observations and broad trends noted. The data is graphically presented in the next chapter where statistical analyses are undertaken, and comparisons are made both within and between the groups in order to address the research aims and hypotheses. As mentioned in Chapter 7, the NH subjects who verified the test battery averaged greater than 95% on each test, indicating that they found the tests to be relatively easy.

8.1 HEARING AID SUBJECT GROUP

This section details the results from the 15 HA subjects, tested on two occasions whilst using their HA, approximately 4 months apart. For each test comprising the music test battery, a table is presented detailing the mean score and standard deviation (SD) for both test blocks individually (i.e. 'test block 1' and 'test block 2'), as well as the average when the scores from the two test blocks were combined. It should also be noted that these HA subjects in part act as the control group for the WL subject group, who were tested pre-implant with HAs, and then approximately 4 months later with the CI.

8.1.1 Rhythm Test

As can be seen in Table 8.1, there was a high level of inter-subject consistency, with small SD figures for each test block. There was also little difference between the mean scores from the two test blocks.

Table 8.1: HA group's rhythm test results from each test block, and combined test blocks

	Test block 1		**Test block 2**		**Combined test blocks**	
(%)	Mean	*SD*	Mean	*SD*	Mean	*SD*
Rhythm	94.21	*4.79*	94.04	*3.65*	94.12	*4.19*

8.1.2 Pitch Test

Table 8.2 presents the overall results for the two test blocks of each subtest. As would be expected, subjects were more accurate when ranking the larger interval sizes, with the largest SD being observed for the quarter-octave subtest. Individual mean scores for the combined test blocks ranged from 76% to 98% for the one-octave subtest, 67% to 98% for the half-octave subtest, and 59% to 99% for the quarter-octave subtest.

Table 8.2: HA group's pitch test results from each test block, and combined test blocks

Subtest (%)	Test block 1		Test block 2		Combined test blocks	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Pitch 1: One-octave	86.61	*11.79*	93.54	*6.18*	90.17	*9.86*
Pitch 2: Half-octave	80.21	*10.36*	87.22	*6.98*	83.72	*9.38*
Pitch 3: Quarter-octave	71.67	*12.24*	77.76	*10.13*	74.71	*11.47*

8.1.2.1 Individual Sung Vowel Results

Table 8.3 presents the mean scores for each vowel, averaged across the two test blocks. For each subtest, mean scores were better for the female-sung vowels than the male-sung vowels, with the only exception to this being for the /a/ vowel in the one-octave subtest.

Table 8.3: HA group's pitch test results for the individual vowels, and singer's sex

Combined test blocks (%)	Female sung stimuli			Male sung stimuli		
	/a/	/i/	Female#	/a/	/i/	Male##
Pitch 1: One-octave	86.95	*98.20*	92.57	92.78	*82.78*	87.78
Pitch 2: Half-octave	88.20	*91.81*	90.00	78.33	*76.53*	77.43
Pitch 3: Quarter-octave	73.65	*81.77*	77.71	70.63	*72.81*	71.72

Mean of the female-sung /a/ and /i/ results ## Mean of the male-sung /a/ and /i/ results

8.1.3 Instrument Identification Test

Table 8.4 presents the mean identification scores from the three subtests. There was wide variability between individual subjects' scores, with mean scores for the combined test blocks ranging from 53% to 97% for the first subtest, 31% to 89% for the second subtest, and 19% to 86% for the third subtest. As explained in Chapter 6, section 6.1.4, two runs of the second subtest were conducted within each test block; these are referred to as '1st run' or '2nd run' in the table. The HA subjects appeared to find the instrument identification test harder than any of the other music perception tests, with lower mean percentages being recorded.

Table 8.4: HA group's instrument identification test results from each test block, and combined test blocks

Subtest (%)	Test block 1 Mean	*SD*	Test block 2 Mean	*SD*	Combined test blocks Mean	*SD*
Subtest 1: Single Instrument	65.00	*14.67*	72.08	*12.12*	68.54	*13.70*
Subtest 2: Instrument with Accompaniment (*1st run*)	49.03	*17.25*	52.92	*18.71*	50.97	*17.79*
Subtest 2: *(2nd run)*	50.56	*14.31*	54.72	*17.30*	52.64	*15.74*
Subtest 2: *(mean of 2 runs)*	49.79	*15.59*	53.82	*17.73*	51.81	*16.68*
Subtest 3: Ensembles	44.72	*19.45*	48.75	*19.76*	46.74	*19.37*

8.1.3.1 Error Analysis For The Instrument Identification Test

Table 8.5, Table 8.6, and Table 8.7 are confusion matrices detailing the subjects' responses, combined across both test blocks, for each of the three subtests. For the matrix corresponding to the second subtest (Table 8.6), the responses from the two runs conducted within each test block were also combined. The list of instruments in the first column corresponds to the stimuli presented to the subject, with the horizontal listing in the first row reflecting the response provided by subjects. The numbers in each cell correspond to the subjects' responses, reported as a percentage of the total number of presentations for that instrument or ensemble. The highlighted cells extending diagonally across the table reflect the percentage of correct identifications for each item. For example, for subtest 1 (Table 8.5), the cello was correctly identified 59% of the time, but was identified as a clarinet 6% of the time. These matrices enable closer investigations of trends, error patterns, or common confusions in the subjects' responses for the identification tests, and are further discussed in Chapter 9.

Table 8.5: HA group's confusion matrix for subtest 1

Response Given - % Correct

Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	59.2	5.8	0.8	0.0	2.5	5.8	0.8	4.2	15.0	0.0	0.0	5.8
Clarinet	2.5	65.0	0.0	17.5	0.0	5.0	0.0	0.8	5.8	2.5	0.0	0.8
Drum Kit	0.8	0.0	69.2	0.0	0.8	0.0	28.3	0.0	0.0	0.8	0.0	0.0
Flute	0.8	18.3	0.0	46.7	0.0	0.0	0.0	13.3	11.7	6.7	2.5	0.0
Guitar	4.2	1.7	5.8	0.0	46.7	31.7	1.7	5.8	0.8	1.7	0.0	0.0
Piano	0.8	0.0	0.0	0.8	0.0	95.8	0.0	0.0	0.0	2.5	0.0	0.0
Timpani	0.0	0.0	36.7	0.0	0.0	0.0	63.3	0.0	0.0	0.0	0.0	0.0
Trumpet	1.7	9.2	0.0	3.3	1.7	0.0	0.0	71.7	11.7	0.0	0.8	0.0
Violin	17.5	11.7	0.0	4.2	0.0	2.5	0.0	0.0	63.3	0.8	0.0	0.0
Xylo	0.0	0.0	0.0	4.2	0.0	12.5	0.0	0.0	0.0	83.3	0.0	0.0
Fem Sing	1.7	4.2	0.0	0.8	0.8	0.0	0.0	2.5	4.2	0.0	85.0	0.8
Male Sing	9.2	1.7	0.0	0.0	0.0	0.0	0.0	0.0	4.2	0.0	11.7	73.3

Single instrument stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.6: HA group's confusion matrix for subtest 2 (combined runs)

Response Given - % Correct

Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	60.4	3.8	1.3	1.3	5.8	4.6	0.0	1.7	18.3	0.4	1.3	1.3
Clarinet	10.0	40.0	0.0	20.8	0.0	4.2	0.0	0.8	19.6	1.7	2.5	0.4
Drum Kit	1.3	0.4	69.6	0.4	0.8	0.4	23.8	1.7	0.4	1.3	0.0	0.0
Flute	6.3	18.8	0.4	32.5	0.4	1.3	0.0	7.5	26.7	0.8	5.4	0.0
Guitar	0.8	4.6	1.3	0.0	30.0	50.4	0.0	0.4	1.3	11.3	0.0	0.0
Piano	4.2	1.7	0.4	2.5	0.8	82.1	0.4	0.4	4.2	3.3	0.0	0.0
Timpani	0.8	0.0	34.2	0.0	0.0	0.0	64.2	0.4	0.0	0.4	0.0	0.0
Trumpet	9.6	16.7	0.4	11.3	1.3	2.1	0.0	20.0	30.0	2.1	4.2	2.5
Violin	8.3	12.5	0.0	14.2	1.3	0.4	0.0	3.3	40.4	0.4	12.9	6.3
Xylo	8.3	2.9	5.8	8.8	7.5	20.4	2.1	4.2	15.8	18.8	2.5	2.9
Fem Sing	0.8	2.1	0.0	1.3	0.0	0.4	0.0	2.5	4.2	0.4	84.6	3.8
Male Sing	2.9	1.3	0.0	1.7	0.4	0.0	0.0	2.5	2.9	0.4	8.8	79.2

Instrument with accompaniment stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.7: HA group's confusion matrix for subtest 3

Response Given - % Correct

Stimuli	C&W	Choir	Jazz	Orch	Percus	Rock	Str Qt	Vln+Pno	Cel+Pno	Male+Pno	Fem+Pno	M+F+Pno
C&W	24.2	0.0	10.8	2.5	6.7	0.8	6.7	25.0	19.2	0.0	4.2	0.0
Choir	3.3	60.0	2.5	2.5	0.8	0.8	3.3	0.8	0.8	4.2	8.3	12.5
Jazz	15.8	0.0	35.0	3.3	2.5	1.7	10.8	12.5	11.7	1.7	2.5	2.5
Orch	1.7	15.0	1.7	49.2	0.0	3.3	14.2	1.7	5.0	1.7	6.7	0.0
Percus	3.3	0.0	1.7	0.8	63.3	5.0	6.7	7.5	10.0	0.8	0.8	0.0
Rock	4.2	0.8	8.3	6.7	25.8	41.7	5.8	0.8	5.0	0.0	0.0	0.8
Str Qt	1.7	8.3	4.2	31.7	3.3	1.7	26.7	5.0	7.5	1.7	5.8	2.5
Vln+Pno	2.5	0.0	1.7	7.5	0.0	0.0	28.3	40.0	13.3	0.8	5.8	0.0
Cel+Pno	0.0	2.5	1.7	10.0	0.0	0.8	7.5	31.7	40.0	2.5	2.5	0.8
Male+Pno	3.3	0.8	1.7	6.7	0.8	3.3	2.5	1.7	3.3	69.2	0.8	5.8
Fem+Pno	3.3	1.7	3.3	1.7	0.8	0.8	1.7	2.5	0.8	5.0	71.7	6.7
M+F+Pno	0.8	7.5	0.0	3.3	0.0	0.8	1.7	3.3	1.7	8.3	32.5	40.0

Ensemble stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

8.1.4 Instrument Quality Rating Test

As can be seen in Table 8.8, the HA subjects provided higher overall ratings for the single-instrument subtest than for the accompanied instrument or ensemble subtests. Table 8.9 provides the mean ratings for each instrument (or ensemble) in each of the subtests.

Table 8.8: HA group's mean rating out of 10 from each test block, and combined test blocks

Subtest (/10)	Test block 1		Test block 2		Combined test blocks	
	Mean	*SD*	Mean	*SD*	Mean	*SD*
Subtest 1: Single Instrument	6.48	*1.61*	7.02	*1.45*	6.75	*1.53*
Subtest 2: Instrument with Accompaniment	6.02	*1.59*	6.57	*1.34*	6.30	*1.47*
Subtest 3: Ensembles	5.60	*1.64*	6.30	*1.65*	5.95	*1.66*

Table 8.9: HA group's mean rating out of 10 for each instrument in each subtest

Subtest 1		Subtest 2		Subtest 3	
Instrument	Rating	Instrument	Rating	Ensemble	Rating
Cello	6.28	**Cello**	6.57	**C&W**	5.48
Clarinet	7.34	**Clarinet**	6.60	**Choir**	6.44
Drum Kit	6.48	**Drum Kit**	6.11	**Jazz**	5.63
Flute	6.36	**Flute**	6.23	**Orch**	6.03
Guitar	5.74	**Guitar**	5.97	**Percus**	5.93
Piano	7.78	**Piano**	7.53	**Rock**	4.86
Timpani	6.81	**Timpani**	6.53	**Str Qt**	6.27
Trumpet	6.99	**Trumpet**	5.53	**Vln+Pno**	6.40
Violin	7.18	**Violin**	6.59	**Cel+Pno**	6.06
Xylophone	6.96	**Xylophone**	4.98	**Male+Pno**	5.99
Fem Sing	6.44	**Fem Sing**	6.42	**Fem+Pno**	6.56
Male Sing	6.59	**Male Sing**	6.49	**M+F+Pno**	5.75

Subtest 1 and 2 stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Subtest 3 stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

8.1.5 Melody Test

As a group, the HA subjects performed well on this test; the mean score was 91% (Table 8.10), with 10 of the 15 subjects scoring between 95% and 100%. There was a large degree of variability between subjects, though, with the mean of the combined test blocks' scores ranging from 35% to 100%.

Table 8.10: HA group's melody test results from each test block, and combined test blocks

Test (%)	Test block 1 Mean	Test block 1 *SD*	Test block 2 Mean	Test block 2 *SD*	Combined test blocks Mean	Combined test blocks *SD*
Melody	89.33	*16.35*	91.67	*16.11*	90.50	*15.99*

8.1.5.1 Error Analysis For The Melody Test

Using the same format as the confusion matrices for the instrument identification tests, Table 8.11 details the error patterns for the melody test, combined for the two test blocks. The melodies presented are listed in the first column, with the corresponding response provided listed horizontally. The numbers in each cell represent the percentage of times each response was made. For example, Waltzing Matilda was identified correctly 100% of the time, with For He's A Jolly Good Fellow identified on all but two presentations (98%).

Table 8.11: HA group's confusion matrix for the melody test

Response Given - % Correct

Stimuli	Adv A F	Baa Baa	For He's	Happy BD	Jingle	Old McD	O Come	Silent	Twinkle	Waltzing
Adv A F	90.0	0.0	1.7	0.0	0.0	1.7	5.0	0.0	1.7	0.0
Baa Baa	0.0	86.7	0.0	0.0	0.0	1.7	0.0	0.0	11.7	0.0
For He's	0.0	0.0	98.3	1.7	0.0	0.0	0.0	0.0	0.0	0.0
Happy BD	0.0	0.0	3.3	91.7	1.7	3.3	0.0	0.0	0.0	0.0
Jingle	0.0	0.0	0.0	0.0	88.3	3.3	3.3	1.7	3.3	0.0
Old McD	1.7	1.7	1.7	0.0	0.0	86.7	5.0	0.0	3.3	0.0
O Come	0.0	0.0	3.3	0.0	0.0	0.0	93.3	1.7	1.7	0.0
Silent	1.7	0.0	1.7	0.0	0.0	0.0	5.0	91.7	0.0	0.0
Twinkle	0.0	16.7	0.0	0.0	0.0	5.0	0.0	0.0	78.3	0.0
Waltzing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

Melodies presented: Advance Australia Fair (Adv A F); Baa Baa Black Sheep (Baa Baa); For He's A Jolly Good Fellow (For He's); Happy Birthday (Happy BD); Jingle Bells (Jingle); Old McDonald (Old McD); O Come All Ye Faithful (O Come); Silent Night (Silent); Twinkle Twinkle Little Star (Twinkle); Waltzing Matilda (Waltzing)

8.2 Cochlear Implant Subject Group

This section details the results from the 15 experienced CI users, again tested on two occasions approximately 4 months apart. The format of the tables is the same as for the HA subject group, in the previous section. Comparisons of these results to those obtained from the other subject groups are made in Chapter 9.

8.2.1 Rhythm Test

Similar to the HA subjects, there was a high level of inter-subject consistency for this test, with low SD values (Table 8.12).

Table 8.12: CI group's rhythm test results from each test block, and combined test blocks

	Test block 1		Test block 2		Combined test blocks	
(%)	Mean	*SD*	Mean	*SD*	Mean	*SD*
Rhythm	92.81	*3.06*	92.81	*3.65*	92.81	*3.31*

8.2.2 Pitch Test

As a group, the performance of the CI users was less accurate and more variable than that of the HA subjects on the pitch-ranking test across all of the interval sizes assessed. Individual mean scores for the combined test blocks ranged from 24% to 91% for the one-octave subtest, 49% to 80% for the half-octave subtest, and 41% to 65% for the quarter-octave subtest. It should be remembered that the chance score for this test was 50%.

Table 8.13: CI group's pitch test results from each test block, and combined test blocks

Subtest (%)	Test block 1 Mean	Test block 1 SD	Test block 2 Mean	Test block 2 SD	Combined test blocks Mean	Combined test blocks SD
Pitch 1: One-octave	69.44* (n=12)	16.57	69.97 (n=12)* 68.19 (n=15)*	16.01 17.57	67.98* (n=15)	16.82
Pitch 2: Half-octave	62.50	*8.18*	66.04	*9.66*	64.27	*8.98*
Pitch 3: Quarter-octave	51.09	*6.26*	52.40	*7.36*	51.75	*6.75*

* Only 12 CI subjects undertook test block 1, whilst all 15 CI subjects undertook test block 2. As explained in Chapter 6, the one-octave subtest was incorporated into the test battery after the half-octave and quarter-octave subtests. Three of the CI subjects were not available to undertake this one-octave subtest during the first test block, and hence only undertook this subtest during the second test block, four months later. Therefore, for the statistical tests reported in Chapter 9, comparisons made on the combined test block scores (such as comparisons between subject groups or music tests) were based on all 15 subjects (overall mean). However, for analyses specifically comparing test block 1 to test block 2 of the one-octave subtest, these calculations were made based only on the results of the 12 subjects who undertook both test blocks. For the test block 2 results in this table, the mean and SD have been provided for both scenarios: i) only the 12 subjects who had results for this subtest from both test blocks, and ii) across all 15 subjects.

8.2.2.1 Individual Vowel Results

Contrary to the HA subjects, the CI subjects tended to be more accurate with the male-sung than female-sung vowels. As can be seen from Table 8.14, this was particularly notable for the one-octave subtest.

Table 8.14: CI group's pitch test results for the individual vowels, and singer's sex

Combined test blocks (%) n=15	Female-sung stimuli /a/	Female-sung stimuli /i/	Female#	Male-sung stimuli /a/	Male-sung stimuli /i/	Male##
Pitch 1: One-octave	53.92	*65.87*	59.90	82.75	*72.75*	77.75
Pitch 2: Half-octave	60.00	*68.75*	64.38	61.95	*66.39*	64.17
Pitch 3: Quarter-octave	40.01	*55.42*	47.71	57.61	*53.96*	55.78

\# Mean of the female-sung /a/ and /i/ results ## Mean of the male-sung /a/ and /i/ results

8.2.3 Instrument Identification Test

The CI group's mean scores were lower than the HA group for each subtest (Table 8.15). Mean of the two test blocks' scores for the CI subjects ranged from 48% to 80% for the first subtest, 28% to 64% for the second subtest, and 23% to 65% for the third subtest. As was the case with the HA subjects, the CI subjects scored higher for recognising the single instrument stimuli than the accompanied instrument, or ensemble stimuli.

Table 8.15: CI group's instrument identification test results from each test block, and combined test blocks

Subtest (%)	Test block 1 Mean	*SD*	Test block 2 Mean	*SD*	Combined test blocks Mean	*SD*
Subtest 1: Single Instrument	58.89	*11.93*	62.78	*11.54*	60.83	*11.70*
Subtest 2: Instrument with Accompaniment (*1st run*)	42.36	*12.22*	45.14	*13.22*	43.75	*12.58*
Subtest 2: *(2nd run)*	45.42	*11.79*	48.47	*13.69*	46.94	*12.65*
Subtest 2: *(mean of 2 runs)*	43.89	*11.90*	46.81	*13.33*	45.34	*12.61*
Subtest 3: Ensembles	41.25	*14.61*	43.89	*15.09*	42.57	*14.66*

8.2.3.1 Error Analysis For The Instrument Identification Test

Table 8.16, Table 8.17, and Table 8.18 detail the CI subjects' responses for each of the three identification subtests. The formatting structure of these tables is as per the explanation provided in section 8.1.3.1 for the HA subjects, with the numbers in each cell corresponding to the percentage of times a response was given, combined across both test blocks. Additionally, the matrix corresponding to the second subtest (Table 8.17) combined the responses from the two runs conducted within each test block.

Table 8.16: CI group's confusion matrix for subtest 1

	Response Given - % Correct											
Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	31.7	7.5	5.0	0.8	7.5	5.0	9.2	2.5	17.5	0.8	3.3	9.2
Clarinet	5.8	40.8	0.0	25.0	0.0	4.2	0.0	8.3	14.2	0.0	1.7	0.0
Drum Kit	0.0	0.0	71.7	0.0	0.0	0.0	28.3	0.0	0.0	0.0	0.0	0.0
Flute	10.0	29.2	0.8	20.8	0.0	0.8	1.7	10.0	14.2	1.7	8.3	2.5
Guitar	3.3	0.0	20.8	0.0	35.0	17.5	7.5	0.0	0.0	15.8	0.0	0.0
Piano	1.7	0.0	2.5	0.0	2.5	87.5	0.0	0.0	0.8	5.0	0.0	0.0
Timpani	0.8	0.8	6.7	0.0	2.5	9.2	76.7	0.0	0.0	3.3	0.0	0.0
Trumpet	3.3	12.5	1.7	4.2	0.0	0.0	0.8	70.8	2.5	2.5	0.0	1.7
Violin	16.7	7.5	0.0	9.2	4.2	2.5	0.0	4.2	55.0	0.0	0.8	0.0
Xylo	0.0	0.0	0.0	2.5	0.0	5.0	0.0	0.0	0.0	92.5	0.0	0.0
Fem Sing	2.5	2.5	0.0	2.5	0.0	0.0	0.0	4.2	2.5	0.0	70.0	15.8
Male Sing	1.7	0.8	0.0	0.8	0.0	0.0	0.0	0.0	0.0	0.0	19.2	77.5

Single instrument stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.17: CI group's confusion matrix for subtest 2 (combined runs)

Response Given - % Correct

Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	40.4	7.9	1.3	2.5	5.4	12.9	2.5	2.1	13.8	1.3	5.8	4.2
Clarinet	12.1	14.6	0.0	21.7	0.0	4.2	0.0	6.7	30.0	2.9	2.1	5.8
Drum Kit	0.4	0.0	58.8	0.0	3.8	0.0	27.5	0.8	1.3	7.5	0.0	0.0
Flute	6.7	12.1	0.0	21.3	0.0	2.9	0.8	10.0	33.8	2.5	5.4	4.6
Guitar	2.9	0.4	2.9	0.0	22.9	57.9	2.9	0.0	1.3	8.3	0.4	0.0
Piano	7.9	1.7	3.8	0.8	3.3	72.9	3.3	0.4	1.7	2.9	0.0	1.3
Timpani	2.9	0.8	18.3	0.0	1.7	7.5	66.3	0.4	0.4	1.7	0.0	0.0
Trumpet	8.3	21.7	0.4	7.1	0.0	0.8	0.8	23.3	26.3	0.4	5.4	5.4
Violin	14.6	7.9	0.0	9.6	0.0	0.8	0.0	1.7	43.8	2.5	10.8	8.3
Xylo	2.5	1.3	8.3	2.1	5.4	17.5	4.6	1.3	7.1	47.5	1.7	0.8
Fem Sing	1.7	4.6	0.0	2.9	0.0	0.0	0.0	1.3	5.8	0.0	49.2	34.6
Male Sing	0.0	0.4	0.0	0.4	0.0	0.0	0.0	0.4	0.0	0.0	14.6	84.2

Instrument with accompaniment stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.18: CI group's confusion matrix for subtest 3

Response Given - % Correct

Stimuli	C&W	Choir	Jazz	Orch	Percus	Rock	Str Qt	Vln+Pno	Cel+Pno	Male+Pno	Fem+Pno	M+F+Pno
C&W	8.3	0.0	8.3	9.2	20.8	10.8	8.3	20.0	13.3	0.0	0.8	0.0
Choir	1.7	63.3	0.0	3.3	0.0	0.8	0.8	0.8	0.0	15.8	4.2	9.2
Jazz	6.7	0.8	40.8	13.3	2.5	5.8	10.0	10.0	6.7	1.7	0.0	1.7
Orch	0.8	6.7	3.3	56.7	1.7	3.3	10.0	4.2	11.7	0.8	0.0	0.8
Percus	5.0	0.0	9.2	0.8	56.7	12.5	10.0	0.0	5.0	0.8	0.0	0.0
Rock	2.5	0.8	5.8	9.2	37.5	41.7	0.8	0.0	0.8	0.0	0.0	0.8
Str Qt	6.7	3.3	6.7	23.3	2.5	0.8	28.3	13.3	10.8	4.2	0.0	0.0
Vln+Pno	3.3	1.7	5.8	8.3	1.7	0.8	36.7	25.0	9.2	5.0	0.0	2.5
Cel+Pno	0.8	1.7	0.0	15.0	3.3	1.7	6.7	25.8	36.7	5.0	0.0	3.3
Male+Pno	7.5	6.7	3.3	1.7	0.0	0.8	1.7	0.8	1.7	66.7	1.7	7.5
Fem+Pno	10.8	0.8	1.7	0.8	0.0	0.8	4.2	2.5	0.0	5.0	59.2	14.2
M+F+Pno	0.0	10.0	0.0	0.8	0.8	0.0	0.8	0.8	0.0	37.5	21.7	27.5

Ensemble stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

8.2.4 Instrument Quality Rating Test

Similar to the trend observed with the HA subjects, the highest mean ratings from the CI subjects were for the single-instrument stimuli, with the lowest mean ratings for the music ensemble stimuli (Table 8.19). However, despite obtaining lower identification scores than the HA subjects, the CI subjects provided higher mean ratings than the HA subjects for all three subtests. Table 8.20 provides the mean ratings for each instrument in each of the subtests.

Table 8.19: CI group's mean rating out of 10 from each test block, and combined test blocks

Subtest (/10)	Test block 1 Mean	*SD*	Test block 2 Mean	*SD*	Combined test blocks Mean	*SD*
Subtest 1: Single Instrument	6.98	*2.06*	7.37	*1.63*	7.18	*1.83*
Subtest 2: Instrument with Accompaniment	6.69	*2.10*	6.97	*1.97*	6.83	*2.00*
Subtest 3: Ensembles	6.54	*2.27*	6.63	*2.03*	6.59	*2.11*

Table 8.20: CI group's mean rating out of 10 for each instrument in each subtest

Subtest 1		Subtest 2		Subtest 3	
Instrument	Rating	Instrument	Rating	Ensemble	Rating
Cello	5.78	**Cello**	6.04	**C&W**	5.93
Clarinet	6.38	**Clarinet**	6.15	**Choir**	7.19
Drum Kit	8.29	**Drum Kit**	7.7	**Jazz**	6.49
Flute	5.93	**Flute**	5.98	**Orch**	6.96
Guitar	6.83	**Guitar**	6.53	**Percus**	7.01
Piano	8.06	**Piano**	7.56	**Rock**	6.25
Timpani	7.87	**Timpani**	7.51	**Str Qt**	6.78
Trumpet	7.5	**Trumpet**	6.43	**Vln+Pno**	6.69
Violin	7.23	**Violin**	7.03	**Cel+Pno**	6.16
Xylophone	8.67	**Xylophone**	7.16	**Male+Pno**	6.89
Fem Sing	6.28	**Fem Sing**	6.44	**Fem+Pno**	6.63
Male Sing	7.31	**Male Sing**	7.43	**M+F+Pno**	5.99

Subtest 1 and 2 stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Subtest 3 stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

8.2.5 Melody Test

The CI subjects appeared to find this test more difficult than the HA subject group, averaging only 52% (Table 8.21). Their combined test blocks' mean scores ranged from 25% to 98%.

Table 8.21: CI group's melody test results from each test block, and combined test blocks

	Test block 1		Test block 2		Combined test blocks	
(%)	Mean	*SD*	Mean	*SD*	Mean	*SD*
Melody	49.00	*23.45*	54.67	*24.89*	51.83	*23.94*

8.2.5.1 Error Analysis For The Melody Test

Table 8.22 provides the confusion matrix for the CI subjects' responses on the melody test, combined across the two test blocks.

Table 8.22: CI group's confusion matrix for the melody test

	Response Given - % Correct									
Stimuli	Adv A F	Baa Baa	For He's	Happy BD	Jingle	Old McD	O Come	Silent	Twinkle	Waltzing
Adv A F	48.3	0.0	3.3	0.0	0.0	6.7	16.7	10.0	10.0	5.0
Baa Baa	5.0	68.3	0.0	3.3	0.0	10.0	1.7	0.0	10.0	1.7
For He's	1.7	1.7	48.3	16.7	3.3	16.7	0.0	3.3	3.3	5.0
Happy BD	3.3	3.3	18.3	51.7	1.7	6.7	6.7	3.3	5.0	0.0
Jingle	0.0	11.7	1.7	1.7	43.3	6.7	5.0	3.3	23.3	3.3
Old McD	1.7	16.7	5.0	3.3	10.0	38.3	8.3	1.7	8.3	6.7
O Come	13.3	1.7	6.7	1.7	1.7	3.3	60.0	5.0	3.3	3.3
Silent	8.3	1.7	0.0	5.0	3.3	5.0	13.3	53.3	3.3	6.7
Twinkle	3.3	28.3	3.3	0.0	6.7	5.0	5.0	0.0	46.7	1.7
Waltzing	5.0	3.3	8.3	5.0	5.0	6.7	1.7	1.7	3.3	60.0

Melodies presented: Advance Australia Fair (Adv A F); Baa Baa Black Sheep (Baa Baa); For He's A Jolly Good Fellow (For He's); Happy Birthday (Happy BD); Jingle Bells (Jingle); Old McDonald (Old McD); O Come All Ye Faithful (O Come); Silent Night (Silent); Twinkle Twinkle Little Star (Twinkle); Waltzing Matilda (Waltzing)

8.3 WAITING LIST SUBJECT GROUP

This section presents the results of the nine subjects who were tested pre-implant with their HA, and then approximately 3 months after the switch-on of their implant. The tables below provide both the mean results and the SDs for the pre-implant and post-implant test blocks for each of the tests. Graphical representations and statistical comparisons between the pre- and post-implantation results are presented in the next chapter.

8.3.1 Rhythm Test

The post-implant mean was less than one percentage point higher than the pre-implant mean for this test (Table 8.23).

Table 8.23: WL group's rhythm test results, pre- and post-implantation

(%)	Pre-Implant		Post-Implant	
	Mean	*SD*	Mean	*SD*
Rhythm	94.74	*3.48*	95.61	*2.63*

8.3.2 Pitch Test

For the pitch test, the subjects' scores were lower post-implantation than prior to surgery (Table 8.24). The pre-to-post implant discrepancy was particularly evident for the one-octave and quarter-octave intervals where scores decreased by 10 and 11 percentage points, respectively.

Table 8.24: WL group's pitch test results, pre- and post-implantation

Subtest (%)	Pre-Implant		Post-Implant	
	Mean	*SD*	Mean	*SD*
Pitch 1: One-octave	83.80	*11.16*	73.84	*14.23*
Pitch 2: Half-octave	72.22	*12.04*	72.11	*12.13*
Pitch 3: Quarter-octave	66.23	*10.05*	54.77	*10.73*

8.3.2.1 Individual Vowel Results

Unlike the HA subject group, the WL group's pre-implant pitch test scores did not show a strong bias according to the sex of the singer (Table 8.25). However, once implanted, substantially higher scores for the combined male-sung than combined female-sung vowels were generally noted (Table 8.26). This was similar to the trend for the experienced CI subject group.

Table 8.25: WL group's pre-implant pitch test results for individual vowels, and singer's sex

Pre-Implant (%)	Female-sung stimuli /a/	/i/	Female#	Male-sung stimuli /a/	/i/	Male##
Pitch 1: One-octave	76.39	89.35	82.87	89.35	80.09	84.72
Pitch 2: Half-octave	68.98	75.93	72.45	73.15	70.83	71.99
Pitch 3: Quarter-octave	57.64	75.69	66.67	66.67	64.93	65.80

Mean of the female-sung /a/ and /i/ results ## Mean of the male-sung /a/ and /i/ results

Table 8.26: WL group's post-implant pitch test results for individual vowels, and singer's sex

Post-Implant (%)	Female-sung stimuli /a/	/i/	Female#	Male-sung stimuli /a/	/i/	Male##
Pitch 1: One-octave	60.19	64.35	62.27	87.96	82.87	85.42
Pitch 2: Half-octave	64.81	75.93	70.37	74.54	73.15	73.84
Pitch 3: Quarter-octave	39.58	60.07	49.83	62.15	57.29	59.72

Mean of the female-sung /a/ and /i/ results ## Mean of the male-sung /a/ and /i/ results

8.3.3 Instrument Identification Test

The WL group were more accurate at identifying musical instruments with their implant than with their hearing aids (Table 8.27). There was a post-implant improvement of 11, 4, and 10 percentage points for the three respective subtests. Consistent with the observations of the CI and HA subject groups, though, was the diminished accuracy in identifying the instruments from subtest 1 to 2, with a further decrease in scores from subtest 2 to 3. This was noted both pre- and post-implantation.

Table 8.27: WL group's instrument identification test results, pre- and post-implantation

Subtest (%)	Pre-Implant Mean	Pre-Implant SD	Post-Implant Mean	Post-Implant SD
Subtest 1: Single Instrument	53.47	14.69	64.81	11.62
Subtest 2: Instrument with Accompaniment (*1st run*)	42.78	10.21	45.33	9.42
Subtest 2: *(2nd run)*	43.11	16.19	48.11	7.85
Subtest 2: *(mean of 2 runs)*	42.94	13.17	46.72	8.58
Subtest 3: Ensembles	35.42	10.52	45.83	8.53

8.3.3.1 Error Analysis For The Instrument Identification Test

Table 8.28, Table 8.29, and Table 8.30 are the confusion matrices for the WL group's pre-implant responses, with Table 8.31, Table 8.32, and Table 8.33 being the post-implant responses. All figures are reported as a percentage of the total number of presentations for each item, with the matrices corresponding to the second subtest combining the responses from the two runs. The overall format of these matrices is as per the description in section 8.1.3.1.

Table 8.28: WL group's pre-implant confusion matrix for subtest 1

Response Given - % Correct

Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	44.4	5.6	5.6	0.0	5.6	2.8	5.6	5.6	16.7	2.8	2.8	2.8
Clarinet	5.6	50.0	0.0	8.3	0.0	8.3	0.0	13.9	13.9	0.0	0.0	0.0
Drum Kit	0.0	0.0	72.2	0.0	0.0	2.8	25.0	0.0	0.0	0.0	0.0	0.0
Flute	0.0	11.1	0.0	33.3	0.0	11.1	0.0	0.0	27.8	8.3	5.6	2.8
Guitar	5.6	2.8	8.3	0.0	27.8	55.6	0.0	0.0	0.0	0.0	0.0	0.0
Piano	0.0	5.6	8.3	2.8	0.0	72.2	2.8	0.0	0.0	8.3	0.0	0.0
Timpani	2.8	0.0	36.1	0.0	0.0	0.0	58.3	0.0	0.0	2.8	0.0	0.0
Trumpet	5.6	13.9	0.0	13.9	0.0	0.0	0.0	50.0	11.1	2.8	2.8	0.0
Violin	11.1	11.1	0.0	8.3	0.0	16.7	0.0	8.3	41.7	2.8	0.0	0.0
Xylo	0.0	0.0	0.0	8.3	2.8	16.7	0.0	0.0	0.0	72.2	0.0	0.0
Fem Sing	2.8	8.3	0.0	0.0	0.0	0.0	0.0	5.6	16.7	0.0	31.1	5.6
Male Sing	11.1	0.0	0.0	0.0	0.0	0.0	0.0	5.6	0.0	2.8	22.2	58.3

Single instrument stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.29: WL group's pre-implant confusion matrix for subtest 2 (combined runs)

	Response Given - % Correct											
Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	44.4	11.1	1.4	2.8	4.2	5.6	1.4	2.8	22.2	0.0	0.0	4.2
Clarinet	5.6	25.0	0.0	16.7	0.0	13.9	0.0	6.9	27.8	2.8	0.0	1.4
Drum Kit	5.6	0.0	66.7	0.0	1.4	1.4	23.6	0.0	0.0	1.4	0.0	0.0
Flute	2.8	6.9	0.0	30.6	0.0	1.4	0.0	6.9	44.4	0.0	6.9	0.0
Guitar	4.2	0.0	2.8	0.0	2.8	63.9	4.2	1.4	0.0	20.8	0.0	0.0
Piano	4.2	2.8	2.8	4.2	2.8	62.5	4.2	1.4	4.2	8.3	2.8	0.0
Timpani	0.0	0.0	23.6	0.0	0.0	1.4	73.6	1.4	0.0	0.0	0.0	0.0
Trumpet	9.7	11.1	0.0	9.7	1.4	1.4	0.0	11.1	41.7	0.0	11.1	2.8
Violin	12.5	5.6	1.4	13.9	0.0	1.4	0.0	5.6	41.7	0.0	8.3	9.7
Xylo	11.1	5.6	11.1	8.3	4.2	8.3	12.5	2.8	18.1	16.7	1.4	0.0
Fem Sing	1.4	6.9	0.0	2.8	0.0	1.4	0.0	5.6	12.5	1.4	66.7	1.4
Male Sing	1.4	1.4	0.0	1.4	0.0	1.4	0.0	2.8	4.2	0.0	13.9	73.6

Instrument with accompaniment stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.30: WL group's pre-implant confusion matrix for subtest 3

	Response Given - % Correct											
Stimuli	C&W	Choir	Jazz	Orch	Percus	Rock	Str Qt	Vln+Pno	Cel+Pno	Male+Pno	Fem+Pno	M+F+Pno
C&W	19.4	2.8	0.0	2.8	8.3	2.8	2.8	30.6	22.2	0.0	5.6	2.8
Choir	0.0	63.9	2.8	11.1	5.6	2.8	0.0	0.0	0.0	2.8	8.3	2.8
Jazz	19.4	0.0	36.1	8.3	2.8	2.8	11.1	11.1	0.0	2.8	2.8	2.8
Orch	5.6	0.0	0.0	44.4	0.0	0.0	25.0	2.8	5.6	0.0	13.9	2.8
Percus	5.6	0.0	16.7	8.3	36.1	2.8	2.8	13.9	11.1	2.8	0.0	0.0
Rock	5.6	0.0	19.4	11.1	25.0	30.6	5.6	0.0	2.8	0.0	0.0	0.0
Str Qt	19.4	5.6	0.0	30.6	2.8	2.8	13.9	2.8	8.3	5.6	5.6	2.8
Vln+Pno	0.0	0.0	0.0	11.1	0.0	0.0	27.8	33.3	16.7	2.8	8.3	0.0
Cel+Pno	0.0	0.0	2.8	8.3	2.8	0.0	19.4	13.9	38.9	11.1	2.8	0.0
Male+Pno	11.1	8.3	0.0	13.9	5.6	0.0	11.1	0.0	0.0	33.3	11.1	5.6
Fem+Pno	8.3	0.0	5.6	0.0	0.0	2.8	2.8	2.8	2.8	11.1	52.8	11.1
M+F+Pno	0.0	16.7	2.8	8.3	0.0	0.0	8.3	0.0	5.6	19.4	16.7	22.2

Ensemble stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

Table 8.31: WL group's post-implant confusion matrix for subtest 1

	Response Given - % Correct											
Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	38.9	2.8	8.3	0.0	2.8	2.8	8.3	2.8	11.1	0.0	5.6	16.7
Clarinet	11.1	38.9	0.0	11.1	0.0	2.8	0.0	8.3	22.2	0.0	2.8	2.8
Drum Kit	0.0	0.0	86.1	0.0	0.0	0.0	13.9	0.0	0.0	0.0	0.0	0.0
Flute	5.6	25.0	0.0	25.0	0.0	0.0	2.8	8.3	11.1	2.8	16.7	2.8
Guitar	8.3	0.0	11.1	0.0	47.2	25.0	0.0	0.0	0.0	8.3	0.0	0.0
Piano	2.8	0.0	2.8	0.0	0.0	88.9	2.8	0.0	0.0	2.8	0.0	0.0
Timpani	0.0	0.0	19.4	0.0	0.0	0.0	80.6	0.0	0.0	0.0	0.0	0.0
Trumpet	2.8	11.1	0.0	2.8	0.0	0.0	2.8	66.7	8.3	0.0	5.6	0.0
Violin	5.6	11.1	0.0	16.7	2.8	5.6	0.0	0.0	55.6	0.0	2.8	0.0
Xylo	0.0	0.0	8.3	2.8	0.0	0.0	5.6	0.0	0.0	83.3	0.0	0.0
Fem Sing	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	77.8	16.7
Male Sing	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	11.1	88.9

Single instrument stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.32: WL group's post-implant confusion matrix for subtest 2 (combined runs)

Response Given - % Correct

Stimuli	Cello	Clarinet	Drum Kit	Flute	Guitar	Piano	Timpani	Trumpet	Violin	Xylo	Fem Sing	Male Sing
Cello	56.9	9.7	0.0	1.4	2.8	8.3	4.2	4.2	9.7	0.0	1.4	1.4
Clarinet	8.3	11.1	1.4	20.8	1.4	6.9	0.0	2.8	27.8	1.4	11.1	6.9
Drum Kit	0.0	0.0	75.0	0.0	4.2	0.0	16.7	0.0	0.0	4.2	0.0	0.0
Flute	5.6	11.1	0.0	13.9	0.0	2.8	0.0	1.4	36.1	0.0	18.1	11.1
Guitar	4.2	0.0	2.8	0.0	30.6	48.6	6.9	0.0	1.4	5.6	0.0	0.0
Piano	6.9	1.4	4.2	2.8	8.3	65.3	2.8	1.4	0.0	6.9	0.0	0.0
Timpani	0.0	0.0	15.3	0.0	0.0	0.0	84.7	0.0	0.0	0.0	0.0	0.0
Trumpet	6.9	8.3	2.8	4.2	1.4	1.4	1.4	31.9	18.1	0.0	16.7	6.9
Violin	9.7	6.9	0.0	6.9	0.0	0.0	0.0	1.4	41.7	0.0	19.4	13.9
Xylo	2.8	4.2	26.4	4.2	8.3	11.1	4.2	0.0	4.2	33.3	1.4	0.0
Fem Sing	4.2	2.8	0.0	1.4	0.0	2.8	0.0	0.0	6.9	0.0	36.1	45.8
Male Sing	2.8	2.8	0.0	2.8	0.0	0.0	0.0	0.0	0.0	0.0	11.1	80.6

Instrument with accompaniment stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Table 8.33: WL group's post-implant confusion matrix for subtest 3

Response Given - % Correct

Stimuli	C&W	Choir	Jazz	Orch	Percus	Rock	Str Qt	Vln+Pno	Cel+Pno	Male+Pno	Fem+Pno	M+F+Pno
C&W	22.2	0.0	16.7	2.8	16.7	8.3	5.6	11.1	16.7	0.0	0.0	0.0
Choir	5.6	50.0	0.0	5.6	0.0	0.0	0.0	2.8	0.0	13.9	5.6	16.7
Jazz	13.9	0.0	33.3	5.6	11.1	2.8	19.4	5.6	0.0	2.8	0.0	5.6
Orch	0.0	11.1	0.0	38.9	2.8	11.1	22.2	5.6	2.8	0.0	2.8	2.8
Percus	5.6	0.0	11.1	0.0	61.1	5.6	5.6	5.6	5.6	0.0	0.0	0.0
Rock	0.0	0.0	8.3	0.0	19.4	69.4	2.8	0.0	0.0	0.0	0.0	0.0
Str Qt	5.6	19.4	2.8	22.2	0.0	0.0	27.8	8.3	8.3	0.0	2.8	2.8
Vln+Pno	8.3	0.0	8.3	16.7	0.0	0.0	22.2	22.2	11.1	5.6	5.6	0.0
Cel+Pno	2.8	0.0	5.6	13.9	8.3	0.0	16.7	16.7	30.6	5.6	0.0	0.0
Male+Pno	2.8	2.8	0.0	0.0	0.0	0.0	5.6	0.0	0.0	83.3	0.0	5.6
Fem+Pno	2.8	0.0	0.0	0.0	0.0	0.0	0.0	2.8	0.0	11.1	72.2	11.1
M+F+Pno	2.8	11.1	0.0	0.0	0.0	0.0	0.0	2.8	0.0	41.7	2.8	38.9

Ensemble stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

8.3.4 Instrument Quality Rating Test

The WL group provided higher ratings for the musical extracts when tested with their implant than when tested with their hearing aids (Table 8.34). This was a similar pattern to the comparison between the CI and HA subject groups. Appraisal scores for the more-complex instrumentations of subtests 2 and 3 were lower than for the single instruments of subtest 1, also in accordance with the trend observed with the experienced CI and HA subject groups. The higher appraisal ratings post-implantation were observed for every subtest in seven of the nine subjects, as well as for the averaged ratings of every instrument or ensemble within every subtest.

Table 8.34: WL group's mean rating out of 10, pre- and post-implantation

Subtest (/10)	Pre-Implant Mean	*SD*	Post-Implant Mean	*SD*
Subtest 1: Single Instrument	5.46	*1.51*	6.90	*1.45*
Subtest 2: Instrument with Accompaniment	4.88	*1.56*	6.86	*1.96*
Subtest 3: Ensembles	4.57	*1.64*	6.66	*2.05*

Table 8.35: WL group's pre-implant mean rating out of 10 for each instrument in each subtest

Subtest 1		Subtest 2		Subtest 3	
Instrument	Rating	Instrument	Rating	Ensemble	Rating
Cello	5.56	**Cello**	5.44	**C&W**	4.28
Clarinet	6.22	**Clarinet**	5.31	**Choir**	4.78
Drum Kit	6.11	**Drum Kit**	5.22	**Jazz**	4.25
Flute	4.91	**Flute**	5.11	**Orch**	4.86
Guitar	4.44	**Guitar**	3.94	**Percus**	5.17
Piano	5.44	**Piano**	5.08	**Rock**	4.03
Timpani	6.39	**Timpani**	6.11	**Str Qt**	5.11
Trumpet	5.11	**Trumpet**	4.39	**Vln+Pno**	5.06
Violin	5.28	**Violin**	4.97	**Cel+Pno**	4.67
Xylophone	5.92	**Xylophone**	4.03	**Male+Pno**	4.11
Fem Sing	4.69	**Fem Sing**	4.36	**Fem+Pno**	4.56
Male Sing	5.39	**Male Sing**	4.61	**M+F+Pno**	4.03

Subtest 1 and 2 stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Subtest 3 stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

Table 8.36: WL group's post-implant mean rating out of 10 for each instrument in each subtest

Subtest 1		Subtest 2		Subtest 3	
Instrument	Rating	Instrument	Rating	Ensemble	Rating
Cello	6.72	**Cello**	7.37	**C&W**	6.53
Clarinet	6.78	**Clarinet**	6.75	**Choir**	6.75
Drum Kit	7.69	**Drum Kit**	7.56	**Jazz**	6.42
Flute	6.22	**Flute**	6.01	**Orch**	6.50
Guitar	6.75	**Guitar**	6.67	**Percus**	6.81
Piano	7.19	**Piano**	6.64	**Rock**	6.50
Timpani	8.00	**Timpani**	8.01	**Str Qt**	7.19
Trumpet	6.42	**Trumpet**	5.89	**Vln+Pno**	6.81
Violin	6.86	**Violin**	6.94	**Cel+Pno**	6.81
Xylophone	7.33	**Xylophone**	6.58	**Male+Pno**	6.83
Fem Sing	6.34	**Fem Sing**	6.61	**Fem+Pno**	6.50
Male Sing	6.72	**Male Sing**	7.14	**M+F+Pno**	6.31

Subtest 1 and 2 stimuli: cello; clarinet; drum kit; flute; guitar; piano; timpani (or bass drum); trumpet; violin; xylophone (xylo); female singer (fem sing); male singer

Subtest 3 stimuli: country & western band (C&W); choir; jazz band; orchestra (orch); percussion band (percus) rock band; string quartet (str qt); violin & piano duet (vln+pno); cello & piano duet (cel+pno); male singer & piano duet (male+pno); female singer & piano duet (fem+pno); trio of a male & female singer with piano (M+F+pno)

8.3.5 Melody Test

The post-implantation score was slightly higher than the pre-implantation score (Table 8.37), contrary to the results of the HA and CI subject groups, where the former group's melody recognition score was substantially better than the latter group's. However, the WL subjects were wide-ranging in their ability to perform this task, with large SDs noted both pre- and post-implant.

Table 8.37: WL group's melody test results, pre- and post-implantation

	Pre-Implant		Post-Implant	
(%)	Mean	*SD*	Mean	*SD*
Melody	75	*25.74*	80	*23.98*

8.3.5.1 Error Analysis For The Melody Test

The confusion matrices from the melody test, pre- and post-implantation, are provided in Table 8.38 and Table 8.39 respectively.

Table 8.38: WL group's pre-implant confusion matrix for the melody test

	Response Given - % Correct									
Stimuli	Adv A F	Baa Baa	For He's	Happy BD	Jingle	Old McD	O Come	Silent	Twinkle	Waltzing
Adv A F	77.8	0.0	5.6	0.0	5.6	0.0	11.1	0.0	0.0	0.0
Baa Baa	0.0	83.3	0.0	5.6	0.0	0.0	0.0	0.0	11.1	0.0
For He's	0.0	5.6	72.2	16.7	0.0	0.0	0.0	0.0	0.0	5.6
Happy BD	0.0	0.0	0.0	72.2	5.6	16.7	0.0	0.0	5.6	0.0
Jingle	5.6	5.6	0.0	0.0	72.2	0.0	11.1	5.6	0.0	0.0
Old McD	5.6	5.6	0.0	0.0	5.6	61.1	11.1	0.0	5.6	5.6
O Come	11.1	0.0	5.6	0.0	0.0	0.0	77.8	0.0	5.6	0.0
Silent	5.6	5.6	5.6	0.0	0.0	5.6	0.0	77.8	0.0	0.0
Twinkle	0.0	11.1	0.0	0.0	0.0	5.6	11.1	0.0	72.2	0.0
Waltzing	0.0	5.6	5.6	0.0	0.0	5.6	0.0	0.0	0.0	83.3

Melodies presented: Advance Australia Fair (Adv A F); Baa Baa Black Sheep (Baa Baa); For He's A Jolly Good Fellow (For He's); Happy Birthday (Happy BD); Jingle Bells (Jingle); Old McDonald (Old McD); O Come All Ye Faithful (O Come); Silent Night (Silent); Twinkle Twinkle Little Star (Twinkle); Waltzing Matilda (Waltzing)

Table 8.39: WL group's post-implant confusion matrix for the melody test

	Response Given - % Correct									
Stimuli	Adv A F	Baa Baa	For He's	Happy BD	Jingle	Old McD	O Come	Silent	Twinkle	Waltzing
Adv A F	72.2	0.0	5.6	5.6	0.0	5.6	11.1	0.0	0.0	0.0
Baa Baa	0.0	83.3	0.0	0.0	0.0	0.0	5.6	0.0	11.1	0.0
For He's	0.0	0.0	88.9	5.6	0.0	0.0	0.0	0.0	0.0	5.6
Happy BD	0.0	0.0	0.0	88.9	0.0	5.6	0.0	0.0	5.6	0.0
Jingle	5.6	5.6	0.0	0.0	72.2	11.1	0.0	0.0	5.6	0.0
Old McD	0.0	5.6	5.6	0.0	5.6	72.2	0.0	0.0	11.1	0.0
O Come	5.6	0.0	5.6	0.0	0.0	0.0	83.3	5.6	0.0	0.0
Silent	0.0	5.6	0.0	0.0	5.6	0.0	0.0	83.3	5.6	0.0
Twinkle	5.6	22.2	0.0	0.0	0.0	0.0	0.0	0.0	72.2	0.0
Waltzing	5.6	0.0	0.0	0.0	5.6	0.0	5.6	0.0	0.0	83.3

Melodies presented: Advance Australia Fair (Adv A F); Baa Baa Black Sheep (Baa Baa); For He's A Jolly Good Fellow (For He's); Happy Birthday (Happy BD); Jingle Bells (Jingle); Old McDonald (Old McD); O Come All Ye Faithful (O Come); Silent Night (Silent); Twinkle Twinkle Little Star (Twinkle); Waltzing Matilda (Waltzing)

CHAPTER 9: DATA ANALYSIS

This chapter provides more detailed analyses and comparisons of the results presented in the previous chapter. Section 9.1 compares the HA subject group to the CI subject group with section 9.2 comparing the pre-to-post surgery results of the WL group. Within each of these sections, a comparison is firstly made between the group's subject factors such as age, experience with the device, and musical experience, as any significant differences between these variables may be important considerations for the subsequent group comparisons of the test scores. Following this, comparisons of the results on the music test battery are made, subdivided into the different tests. Sections 9.3 and 9.4 make comparisons of the music test scores between subject groups using the same type of device (i.e. HA or CI). Section 9.3 compares the HA subject group to the WL group pre-implant, and section 9.4 compares the CI subject group to the WL group post-implant. Finally in section 9.5, correlations were calculated between some of the different music test scores for each group, along with correlations between the test results and key subject variables. The reader is referred back to Chapter 8 for the relevant mean and standard deviation values for each subject group, across the various music tests and questionnaires. Two-tailed statistical tests with a significance value of 0.05 were adopted.

It should be noted that as the purpose of testing the NH subject group was primarily to verify that the music tests were of an appropriate difficulty level, statistical analyses between the music test scores for the hearing-impaired subject groups and the NH subjects will not be made. A non-parametric Kruskal-Wallis test showed no significant difference between all of the subject groups (i.e., NH, CI, HA, and WL subjects) for the levels of music experience ($p = 0.151$), as determined from their responses on the MTEQ.

As mentioned in Chapter 6, two runs of the second instrument identification subtest were conducted for each subject within each test block in order to assess whether supplementary information on the background accompaniment impacted upon music perception. To assess whether there was any significant difference between the scores from these two runs for each subject group, paired t-tests were performed. This showed that there was no significant difference between the two runs of this subtest for any of the three subject groups, within each administration of the music test battery (Table 9.1). In view of this, each subject's identification results for the two runs of the second instrument identification subtest were averaged for the analyses and discussion to follow.

Table 9.1: Comparison of the difference between scores (%) from the two runs of the second instrument identification subtest, for each subject group.

%	HA Group		CI Group		WL Group	
Test block 1	run 1 mean: 49.03 run 2 mean: 50.56	p = 0.457	run 1 mean: 42.36 run 2 mean: 45.42	p = 0.237	Run 1 mean: 42.78 Run 2 mean: 43.11	p = 0.954
Test block 2	run 1 mean: 52.92 run 2 mean: 54.72	p = 0.282	run 1 mean: 45.14 run 2 mean: 48.47	p = 0.175	Run 1 mean: 45.33 Run 2 mean: 48.11	p = 0.520

9.1 HEARING AID SUBJECTS COMPARED TO COCHLEAR IMPLANT SUBJECTS

A graphical comparison of the scores from the CI and HA groups for each of the tests or subtests (excluding the quality rating appraisals) is provided in Figure 9.1. For the proceeding analyses, each subject's scores from the two test blocks conducted 4 months apart were averaged. The raw test results were presented in the previous chapter, in sections 8.1 and 8.2 (p. 129 & 135).

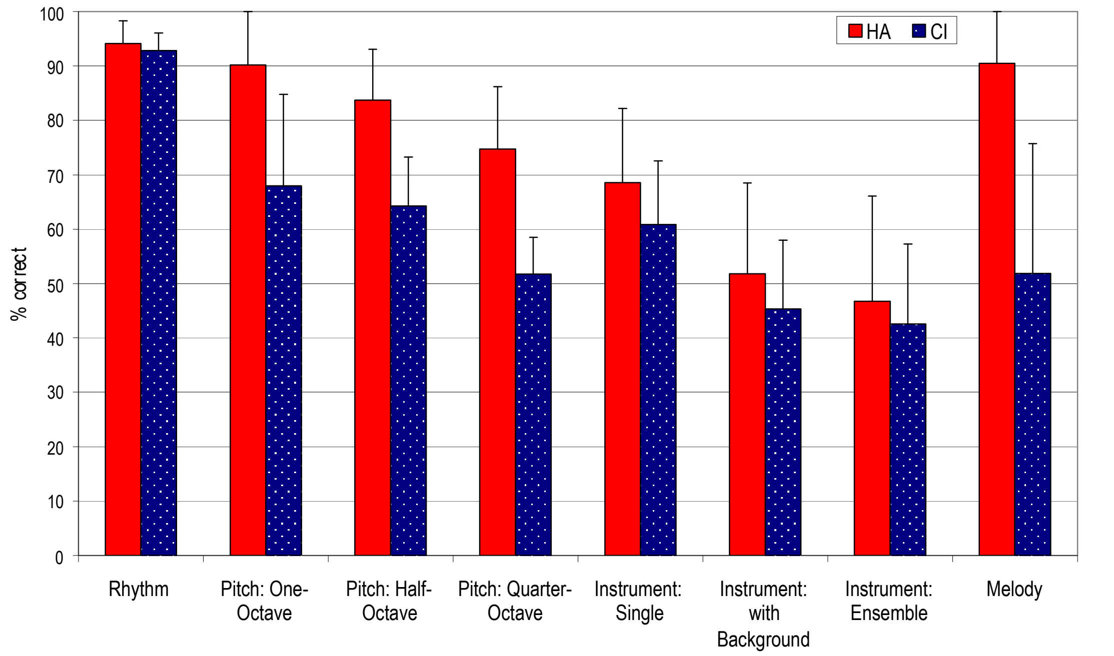


Figure 9.1: Comparison of scores on the music test battery for the HA and CI groups
The X-axis lists the individual tests and subtests, with the Y-axis being the mean % correct scored by each group, averaged across the two administrations of the test battery. The error bars indicate 1 standard deviation.

9.1.1 Subject Factors

Independent-samples t-tests were conducted to compare the HA and CI groups for the subject factors of age, experience with the device, and speech perception, as reported in Tables 7.1 and 7.2 (p. 125 & 126, Chapter 7). This showed a significant difference between the groups for the variables of device experience ($p = 0.002$) and speech perception ($p < 0.001$). The HA group had used their device for a longer period of time than the CI subjects (HA group's mean: 241 months; CI group's mean: 109 months), with the CI group having significantly better speech perception scores (HA group: 42%; CI group: 84%). In view of the inclusion requirement that HA subjects had to meet the implantation criteria for speech perception, the difference in speech perception scores was largely expected. It would also be a reasonable expectation that the HA group would have had more experience with their device than the CI group, although it should be kept in mind that most of the CI subjects would have used HA(s) for varying lengths of time before they received their implant. There was no significant difference for the variable of age (CI group's mean: 60.4 years; HA group's mean: 64.7 years; $p = 0.264$).

Based on responses from the MLEQ and MTEQ, subjects were also assigned a rating for the self-assessed factors of: i) music experience, on a scale from 0 – 2; ii) the amount of time spent listening to music pre-hearing loss, on a scale from 0 – 4; and iii) the amount of time they currently spent listening to music (i.e. around the time of testing), also on a scale from 0 – 4. These scales were outlined in Chapter 6, sections 6.2.2 and 6.2.3. Each subject's ratings appear in Tables 7.1 and 7.2. For the CI subjects, the mean level of music experience was 0.93, the mean rating for the amount of time spent listening to music pre-hearing loss was 2.87, with the mean rating for the level of music listening at the time of this study when using their respective device being 1.27. For the HA subjects, the mean respective ratings were 1.07, 2.53, and 1.60. Mann-Whitney U tests were conducted to assess whether there was any significant difference between the two groups for these three subjective factors. Results indicated that there was no significant difference between the two groups on any of these three factors (music experience: $p = 0.713$; music listening pre-hearing loss: $p = 0.325$; and music listening at the time of this study: $p = 0.345$).

In order to assess whether there was a significant difference in the amount of self-rated time spent listening to music pre-hearing loss compared to the time when the music testing commenced for the CI and HA groups separately, a non-parametric Wilcoxon Signed Ranks test was performed for the two groups. For the CI group, music listening at the time of testing was significantly less than pre-hearing loss estimations ($p = 0.005$). For the HA group, this difference was almost significant ($p = 0.051$).

It should be pointed out that for the CI subjects who used a HA in the contralateral ear, as reported in Table 7.1, responses on listening preferences from the MLEQ may have been provided for a CI-only listening condition, or whilst listening with both the CI and HA. Subjects were not asked to differentiate between these two listening conditions. Similarly, for the HA subject group, subjects were not asked to indicate whether responses were made whilst listening with either one or two HAs.

In summary, there were no significant differences between the CI and HA subject groups for the factors of age, music experience, amount of time spent listening to music pre-hearing loss, or the amount of music listening at the time the questionnaire was completed.

9.1.2 Rhythm Test

An independent-samples t-test showed no significant difference between the HA group's mean of 94% and the CI group's mean of 93% ($p = 0.255$).

9.1.3 Pitch Test

For all three subtests, the HA group's means were higher than the CI group's means (see Table 8.2 and Table 8.13 in Chapter 8). Large SDs were observed for both subject groups in all three subtests, along with differences between scores obtained for female-sung compared to male-sung vowels (Tables 8.3 and 8.14).

A 3-way repeated-measures Analysis of Variance (ANOVA) was conducted using the between-subject factor of group (i.e. CI versus HA subjects), and within-subject factors of interval size (i.e. one-octave, half-octave, and quarter-octave), and singer's sex (i.e. male-sung versus female-sung vowels). This showed highly significant main effects of group ($p < 0.001$) and interval size ($p < 0.001$). As would be expected, average scores decreased with the smaller interval sizes. Although there was no main effect of singer sex ($p = 0.96$), there was a significant 2-way interaction between group and singer sex ($p < 0.001$).

In view of this significant interaction, separate 2-way repeated-measures ANOVAs were conducted for the CI and HA groups to compare the within-subject factors of interval size and singer's sex. This showed that the CI group performed significantly better with the male-sung than female-sung stimuli ($p = 0.035$), whereas the HA group were significantly better with the female-sung stimuli ($p = 0.001$).

One-sample t-tests were also calculated to compare each subject group's performance to the chance score of 50%. This revealed that the CI group's average for the quarter-octave interval (51.75%) was not significantly different to chance-level performance ($p = 0.238$). That is, as a group, the CI subjects were unable to discriminate the pitch of two notes a quarter-octave apart. Their half-octave (64.27%) and one-octave (67.98%) subtest scores were significantly better than the chance score ($p < 0.001$, and $p = 0.001$, respectively). For the HA group, performance for all three interval sizes was significantly above the chance level score ($p < 0.001$ for all comparisons).

9.1.4 Instrument Identification Test

The mean scores for the HA group for subtest 1, 2, and 3 were 69%, 52%, and 47% respectively (Table 8.4), and 61%, 45%, and 43% for the CI group (Table 8.15). To compare the groups' performances across the three subtests, a 2-way repeated-measures ANOVA was conducted. This showed no significant difference for the between-subject factor of group ($p = 0.222$), but a significant difference for the within-subject factor of subtest ($p < 0.001$). There was no significant 2-way interaction between these two terms ($p = 0.529$). For the effect of subtest, post-hoc pairwise comparisons with Bonferroni corrections showed the differences between subtests 1 and 2, as well as between subtests 1 and 3 to be statistically significant ($p < 0.001$ in both cases). The difference between subtests 2 and 3 was approaching significance ($p = 0.062$). That is, the HA and CI groups performed significantly better at identifying instruments in the single-instrument subtest than for the other two subtests incorporating multiple instrumentation.

From the confusion matrices for the HA subjects presented in Chapter 8 (Tables 8.5, 8.6, and 8.7), some observations and trends can be noted. The best recognised single instrument was the piano, recognised 96% of the time, followed by the female singer. In the second subtest, the best recognised accompanied instrument was the female singer, followed by the piano. The duet of a female singer and piano was the best recognised ensemble. Common confusions for subtest 1 were the cello and violin, the flute and clarinet, and the drum kit and timpani, which were all errors within the same instrument family. For the purposes of this thesis, the musical instrument-family classifications were derived from Rimsky-Korsakov's (1891) text on orchestration. In subtest two, the error patterns were more diffuse in nature. For example, whilst the cello was most commonly mistaken to be a violin, the violin was confused to be a clarinet, flute, or female singer (instruments in the same pitch range) more often than it was confused as a cello. The trumpet was more likely to be selected as a violin than as any other wind instrument, having a far broader error pattern than that observed for the first subtest. In the music ensemble subtest, the error patterns were again quite widespread in nature, particularly for the larger ensemble groups such as the jazz, or country and western bands. For all three subtests, the stimuli incorporating the female singer were more often identified than the stimuli incorporating the male singer. Furthermore, in the third subtest, the most common error for the trio of a male singer, female singer, and piano was its selection as a female singer with piano duet, indicating that the male voice in the extract was not adequately perceived.

For the CI subjects (Tables 8.16, 8.17, and 8.18), the most commonly recognised single instruments were the xylophone and the piano, both percussive instruments. For the second subtest, the best recognised sound was the male singer, with the male singer and piano duet being the best recognised ensemble. This is in contrast to the HA subjects where the most recognised accompanied instrument and music ensemble was the female singer, and the female singer with piano duet, respectively.

With regards to the overall error patterns for the CI subjects, it can be observed that in subtest 1, errors tended to be within the same instrument family. For example, although the drum kit was not the most accurately recognised instrument, the only error made by the CI subjects was confusing it with the timpani. The most common errors for the timpani were its selection as a piano or drum kit – two other percussive instruments. Although the CI subjects may not have accurately recognised the sex of the solo singers, they rarely confused the singers with the instrumental stimuli. Other common confusions for the first subtest were the clarinet and flute, as well as the violin and cello, again instruments within the same family. In the second subtest, these within-instrument-family confusions were still apparent, with the percussive instruments tending to be confused with other percussive instruments, or the vocal stimuli being mistaken to have been sung by the opposite sex. A common across-instrument-family error was the incorrect selection of the violin when the clarinet or flute was presented. In the third subtest, the error patterns were more diffuse. For example, the country and western band was confused with most of the other stimuli. Interestingly, the trio of male singer, female singer, and piano stimuli in this subtest was more likely to be identified as a male singer with piano duet by the CI subjects, indicating that the female voice was not perceived in these cases. This is opposite to the results of the HA subject group for the same stimuli.

9.1.5 Instrument Quality Rating Test

Figure 9.2 provides a comparison of the HA and CI groups' mean ratings for the three subtests. A 2-way repeated-measures ANOVA was conducted to compare these ratings across the three subtests. This showed no significant difference in ratings for the between-subject factor of group ($p = 0.386$), but a significant difference for the within-subject factor of subtests ($p < 0.001$), with no significant interactions between group and subtest ($p = 0.686$). Post-hoc pairwise comparisons with Bonferroni corrections showed a significant difference between all three subtests (subtest 1 and 2: $p = 0.001$; subtest 1 and 3: $p < 0.001$; subtest 2 and 3: $p = 0.042$), with appraisal scores being highest for subtest 1 and lowest for subtest 3.

Tables 8.9 and 8.20 in the previous chapter listed the mean ratings provided for each instrument for the HA and CI groups respectively. For the HA group, the piano received the highest appraisal ratings in the first and second subtests, with the female singer and piano duet being rated as the most pleasant-sounding ensemble. The lowest rated instruments or ensembles by the HA subjects for the three respective subtests were the guitar, xylophone, and the rock band. It is worthwhile noting the similarities in these preferences to the results of the instrument identification test. The piano was the single-instrument stimuli recognised most often by the HA subjects, and the second most recognised accompanied instrument. The duet comprising a female singer with piano was the most recognised, and highest appraised, ensemble. Further, at the other end of the scale were the guitar, being both the lowest appraised and the least accurately recognised single instrument, as well as the xylophone, which was the lowest appraised and least recognised accompanied instrument for the HA group.

For the CI subject group, the xylophone was rated to be the most pleasant sounding instrument in the first subtest; it was also the most accurately identified single instrument. Similarly, the cello and flute were both the lowest appraised, and the least accurately identified, single instruments. For the second subtest, three percussion instruments were the highest appraised stimuli – the drum kit, piano, and timpani. The flute received the lowest appraisal ratings. The choir was deemed to be the most pleasant sounding ensemble by the CI subjects, with the country and western band being the least pleasant sounding, along with the least recognised group.

It is interesting to note that CI and HA subjects had relatively different preferences in rating the variety of instruments and ensembles. For the CI subjects, instruments from the percussion family dominated the stimuli receiving higher preferential ratings. For example, the three highest appraised instruments in both subtest 1 and 2 were all from the percussion family, with the percussion ensemble being rated the second most pleasant sounding music ensemble in subtest 3. For the HA subjects, instrumental preferences were more spread out across the different instrumental families.

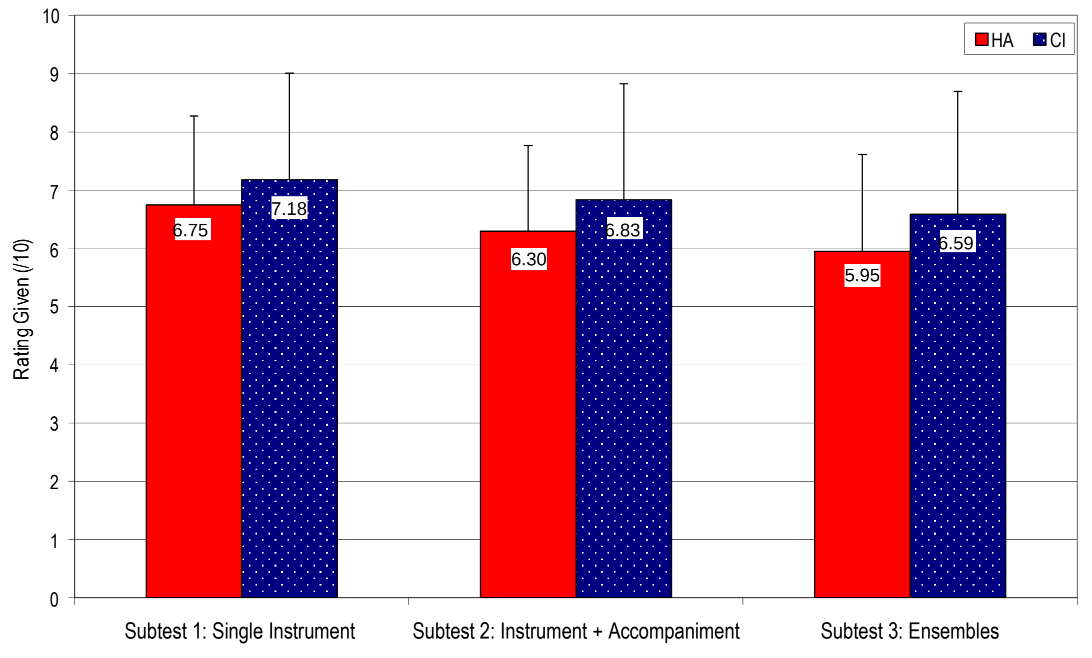


Figure 9.2: Quality rating scores for the three subtests, for the CI and HA groups
The X-axis lists the three subtests. The Y-axis represents the mean rating out of 10, for each group, where ‘1’ = very unpleasant, and ‘10’ = very pleasant. The error bars indicate 1 standard deviation.

9.1.6 Melody Test

The discrepancy between the performance of the CI and HA subjects for each melody presented in this test, as well as the overall mean, is evident from Figure 9.3 which compares the two groups' scores for correctly identifying each of the ten melodies. An independent-samples t-test showed a significant difference between the HA group's mean of 91% and the CI group's mean of 52% ($p < 0.001$) on the melody recognition test.

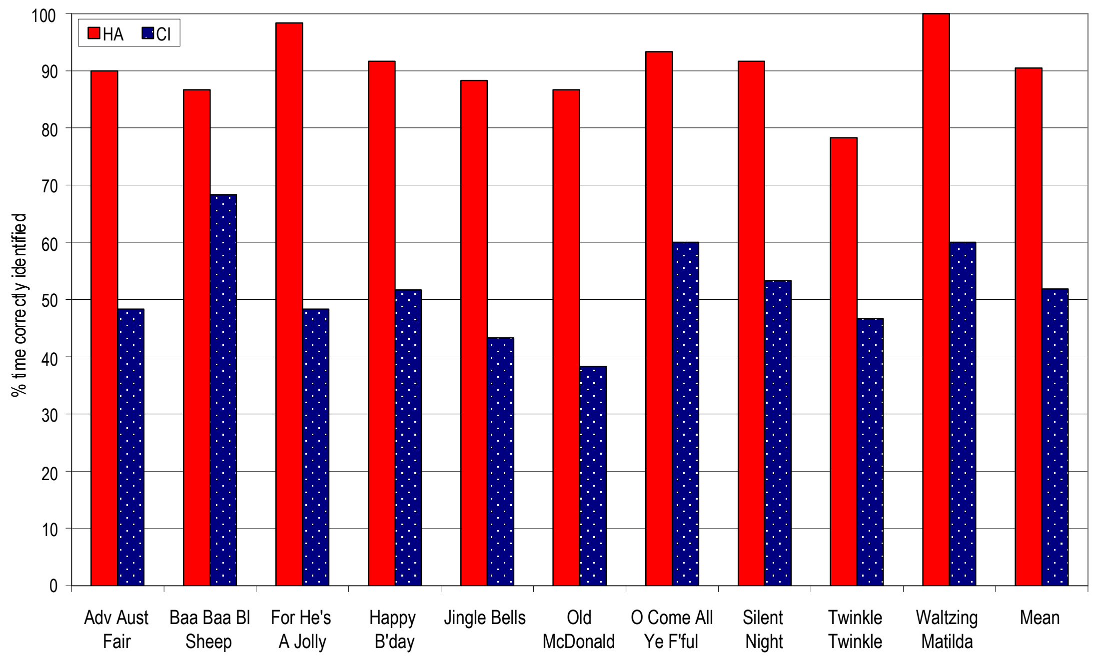


Figure 9.3: Melody test scores for the CI and HA groups
The X-axis lists each of the melodies presented, along with the overall mean. The Y-axis represents the % of times each melody was correctly identified.

For the HA group, Waltzing Matilda was identified correctly 100% of the time, with For He's A Jolly Good Fellow being identified on all but two presentations. The least recognised melody was Twinkle Twinkle, being recognised 78% of the time. The error analysis table in Chapter 8 (Table 8.11) showed that Twinkle Twinkle was most commonly mistaken to be Baa Baa Black Sheep, and similarly, the most common confusion for Baa Baa Black Sheep was Twinkle Twinkle. For the CI group, Baa Baa Black Sheep was the most correctly identified melody, recognised 68% of the time, with Old McDonald being the least correctly identified.

9.1.7 Summary

There was no significant difference between the CI and HA groups for the rhythm test, instrument identification test, or the instrument quality rating test. The CI group scored significantly lower than the HA group on the pitch test, scoring only at chance level for the quarter-octave subtest. The CI group were also significantly poorer at recognising familiar melodies than the HA group.

9.2 WAITING LIST SUBJECTS COMPARISONS – PRE-TO-POST IMPLANT

Figure 9.4 provides a graphical comparison of the pre-to-post implant score differences for the WL group on the identification and discrimination tests in the music test battery. The raw test results were presented in section 8.3 of Chapter 8 (p. 140).

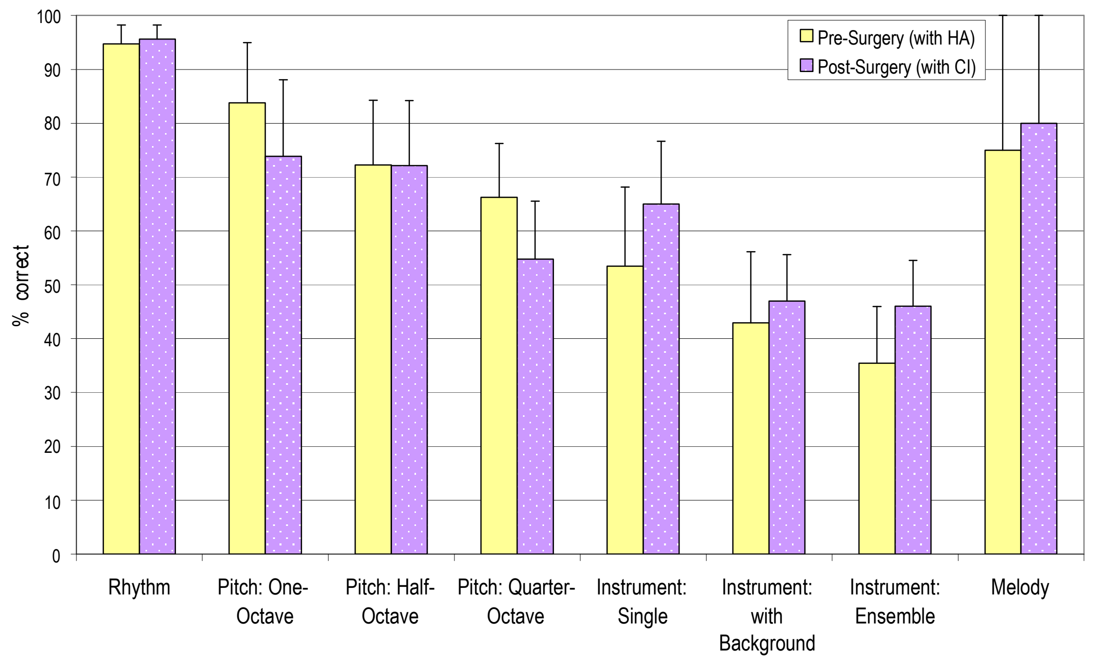


Figure 9.4: Comparison of scores on the music test battery for the WL group pre-surgery and post-surgery

The X-axis lists the individual tests and subtests, with the Y-axis being the mean % correct scored. The error bars indicate 1 standard deviation.

9.2.1 Subject Factors

All of the WL subjects obtained higher speech perception scores when tested with their CI than pre-surgery when using HA(s) (Figure 9.5). These figures were reported in

Table 7.3 at the end of Chapter 7 (p. 127). A paired t-test showed that post-implant speech scores were significantly better than their pre-implant scores ($p < 0.001$).

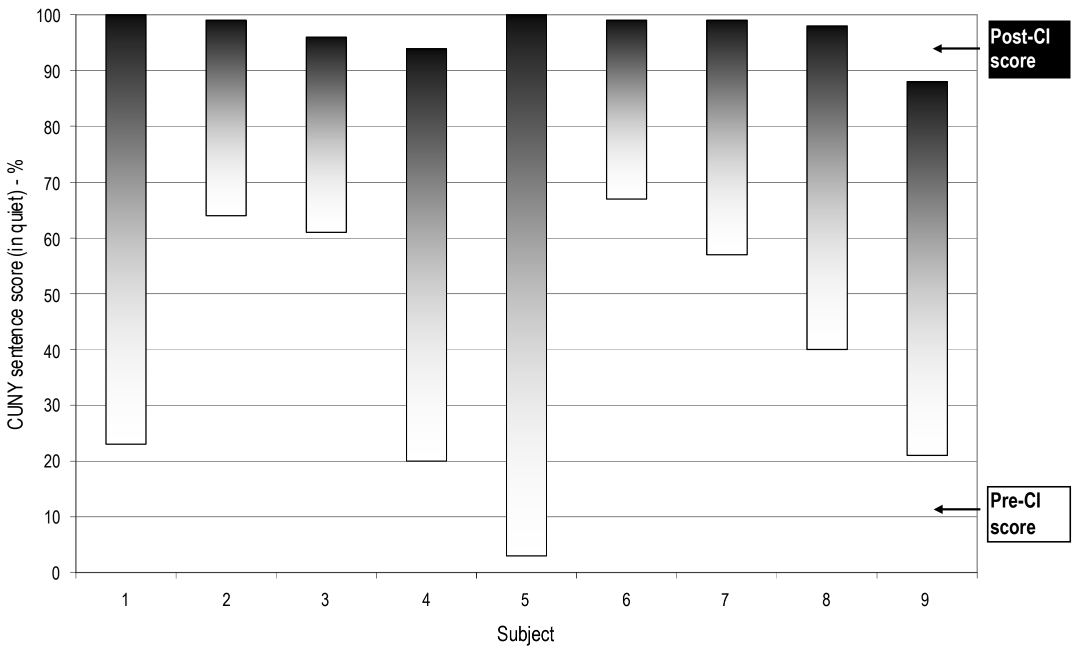


Figure 9.5: Improvement in the WL group's pre-to-post implantation speech perception scores

Each vertical bar corresponds to one of the WL subjects. The bottom line of each vertical bar (white section) indicates the subject's pre-implant speech perception score (%) for their best-aided listening condition. The top line of the bar (black section) represents their post-implant speech perception scores (%) when tested using a CI-only, at approximately 3 months post switch-on. The stimuli for all of these tests were CUNY sentences presented in a quiet listening environment. Hence the length of the bar is a pictorial representation of the improvement obtained by each subject in their speech perception ability from using HAs to using the CI.

Based on their responses from the two MLEQ questionnaires administered with this subject group (i.e., one pre-implant, with a follow-up version post-implant as described in Chapter 6, section 6.2.3), a comparison of the amount of time spent listening to music both pre-hearing loss, and pre-implantation when using HAs, could be made to their current listening habits with the CI. For the WL subjects, the mean rating for time spent listening to music pre-hearing loss was 2.44, as compared to 0.67 for listening with HAs, and 1.44 for listening with the CI. A Wilcoxon Signed Ranks test showed the increased amount of listening from using HAs to the CI to be significant ($p = 0.02$), with no significant difference between the amount of time spent listening to music pre-hearing loss and with the CI ($p = 0.075$; Wilcoxon Signed Ranks test).

The MLEQ-Post version also asked the WL subjects three additional questions pertaining to their satisfaction with the CI. Seven of the nine subjects indicated that "music sounds better with the CI than it did with HAs; it now sounds more pleasant." Only one subject indicated that "music sounded worse with the CI than it did with HAs; I didn't mind the sound of music through HAs, but I don't like the sound through the CI". The other subject conveyed that there was no difference between the two devices for listening to music. When asked to self-rate the amount of difference that the implant had made to their speech perception, compared to using HAs, on the following scale: 1 = made it worse; 2 = no change; 3 = made it a little better; 4 = made it somewhat better; 5 = made it much better; six subjects gave a rating of 5, two subjects gave a rating of 4, and one subject gave a rating of 3. In rating their overall satisfaction with the CI on a scale from 1 to 5, where 1 indicated 'unsatisfied', and 5 indicated 'very satisfied', seven subjects provided the maximum rating of 5, one subject gave a rating of 4, and one subject gave a rating of 3. Of the nine subjects, seven responded that the CI had met their expectations, with the other two circling the 'unsure' response.

As was mentioned earlier whilst discussing the CI and HA groups' MLEQ responses, post-surgery listening preferences for the WL subjects who continued to use a HA in the contralateral ear, (as reported in Table 7.3), may have been provided for a CI-only listening condition, or whilst listening with both the CI and HA. Subjects were not asked to differentiate between the two conditions. It should be reiterated though, that in this study, the post-surgery testing for the WL subjects was conducted solely in a CI-only listening condition, irrespective of whether or not the subject used a HA contralaterally. As was also the case with the HA subject group, pre-surgery responses on the original MLEQ from the WL subjects did not differentiate between listening with one or two HAs.

9.2.2 Accounting for the Learning Effect

Before providing the comparisons of the WL group's pre-to-post surgery music test scores, it is important to explain how the potential for a learning effect in the test results was accounted for in these comparisons. As explained in Chapter 7, the CI and HA subject groups were tested on two occasions, approximately 4 months apart in order for them to act as controls for the pre-to-post surgery comparisons of the WL group. This was an important consideration for interpreting the results of the WL subject group – were the recorded changes in the pre-to-post surgery test scores solely attributable to a learning effect, or did the variable of obtaining an implant have an additional impact on the scores? In order to accurately assess whether the CI contributed to the change in the music test scores, score changes resulting from a learning effect needed to be accounted for.

To achieve this, statistical tests were undertaken to investigate if the change between the pre-surgery and post-surgery test scores for the WL subject group was significantly different to the change in scores between the two test blocks recorded by the CI and HA subject groups (i.e. the second test block score minus the first test block score - "score-difference mean"). In order to see if there was any difference between the groups' score-difference means, a 2-way repeated-measures ANOVA was conducted with a between-subject factor of group (i.e. CI group, HA group, and WL group), and a within-subject factor of subtest. There was a significant difference for the factor of subtest ($p < 0.001$), with no significant main effect of group ($p = 0.529$), and a highly significant interaction between the two factors ($p < 0.001$). This indicates that the degree of change in the scores between the two test blocks for each group was not consistent across the different tests and subtests, which can be observed from the graphical representation of the score-difference mean provided below in Figure 9.6. The score-difference mean results for the quality rating assessment are presented separately in Figure 9.7.

Upon examination of the data, and also visible in the two graphs, there appeared to be little difference between the CI and HA subject groups for the degree of change in scores between the two test blocks. This was confirmed with independent-samples t-tests showing no significant difference in the score-difference mean between the two groups for any of the tests. That is, the extent of the learning effect for the music tests was similar for the CI and HA groups. In view of this, the score-difference means for the CI and HA groups were combined together for subsequent comparisons to the WL subject group (i.e. CI+HA groups' score-difference mean). Independent-samples t-tests between the CI+HA groups' score-difference mean, and the WL group's score-difference mean were then performed in order to assess whether the changes in the pre-to-post surgery test scores for the WL subjects were attributable to more than a learning effect. As the degree of learning effect observed for the CI and HA subjects was similar, it would be reasonable to consider that the WL group would also exhibit a similar learning effect. Hence, a significant ‘p’ value for the independent-samples t-test of the score-difference means would suggest that there was more than just a learning effect contributing to the change in scores for the WL subjects. The results of this test and other relevant analyses for the different music tests are described in sections 9.2.3 – 9.2.7.

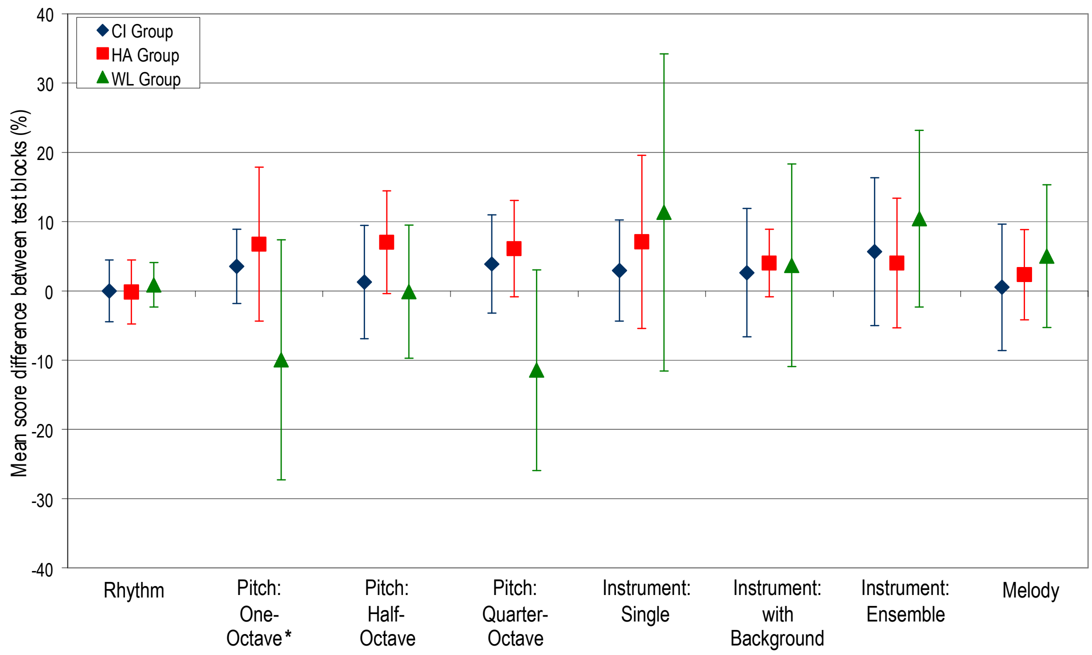


Figure 9.6: Difference between the mean scores of the two test blocks for the three subject groups across the various tests and subtests of the music test battery
Tests and subtests are presented along the X-axis, with the Y-axis representing the mean % point difference between the scores obtained on the two test blocks (i.e. mean of the scores obtained from the second test block minus the mean of the scores obtained on the first test block). The error bars indicate 1 standard deviation. The results of each group are presented separately, with the blue diamond representing the mean difference obtained by the CI subject group, the red square being that for the HA subject group, and the green triangle representing the pre-to-post surgery difference for the WL group. A negative value indicates that the mean score on the first test block was higher than that obtained on the second test block, conducted approximately 4 months later.
** For the one-octave pitch subtest for the CI group, n = 12, as explained in Chapter 7.*

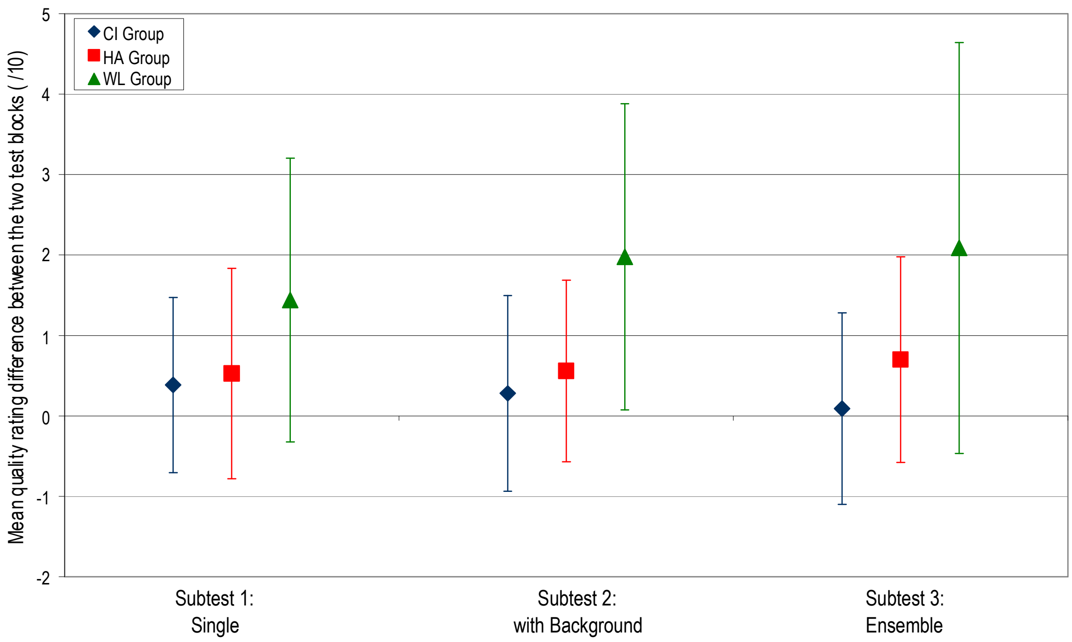


Figure 9.7: Difference between the mean ratings of the two test blocks for the three subject groups for the quality rating assessment

The three subtests are presented along the X-axis, with the Y-axis representing the mean difference between the ratings (out of 10) obtained on the two test blocks (i.e. mean of the ratings obtained on the second test block minus the mean of the ratings obtained on the first test block). The error bars indicate 1 standard deviation. The results of each group are presented separately, with the blue diamond representing the mean difference obtained by the CI subject group, the red square being that for the HA subject group, and the green triangle representing the pre-to-post surgery difference for the WL group.

9.2.3 Rhythm Test

An independent-samples t-test showed no significant difference between the CI+HA groups' score-difference mean and the WL group's score-difference mean ($p = 0.551$). That is, the change in scores for the WL group from 95% ($SD = 3.48$) to 96% ($SD = 2.63$) was not significantly different to the change in scores collectively recorded by the CI and HA groups.

9.2.4 Pitch Test

The WL group's results for the pitch test were presented in Tables 8.24 – 8.26 (Chapter 8). Pre-implant, mean ranking scores for the one-octave, half-octave, and quarter-octave subtests were 84%, 72%, and 66%, respectively. Post-implant, mean scores decreased to 74%, 72%, and 55% for the three respective subtests. An independent-samples t-test comparing the CI+HA groups' score-difference mean to the WL group's score-difference mean was significant for the one-octave and quarter-octave pitch subtests ($p = 0.007$, and $p < 0.001$ respectively). The difference for the half-octave subtest approached significance ($p = 0.061$). It is worthwhile pointing out that the change in the pitch test scores for the WL subjects was in the opposite direction to the other two groups (Figure 9.6). Whereas the CI and HA groups improved from their first to second test run, the WL group's post-surgery scores were lower than their pre-surgery scores.

In order to see if there were any significant differences between the WL group's scores on the three subtests, as well as between the scores for the male-sung and female-sung vowels, 2-way repeated-measures ANOVAs were conducted for the pre-implant and post-implant blocks. Pre-implant, there was a significant difference for the factor of subtests ($p < 0.001$), with no significant difference for the factor of singer's sex ($p = 0.973$). Post-hoc pairwise comparisons with Bonferroni corrections showed the significant effect for the factor of subtest to arise from differences between the scores for the one-octave and half-octave subtest ($p = 0.014$), as well as the one-octave and quarter-octave subtests ($p < 0.001$). The difference between the half-octave and quarter-octave subtests' scores was approaching significance ($p = 0.068$). Mean scores were highest for the one-octave subtest and lowest for the quarter-octave subtest. Post-implant, increased interval size also resulted in higher mean scores. The post-implant 2-way repeated-measures ANOVA showed that both the factors of subtest and singer's sex were significant ($p < 0.001$ and $p = 0.018$, respectively), with the male-sung vowels scoring higher than the female-sung vowels within each of the subtests. Post-hoc pairwise comparisons with Bonferroni corrections showed the significant effect of subtest to arise from the difference between the half-octave and quarter-octave interval scores, as well as between the one-octave and quarter-octave interval scores ($p < 0.001$ for both comparisons). There was no significant difference between the scores from the one-octave and half-octave subtests ($p = 1.00$) post-surgery.

A 1-sample t-test was conducted to assess if there was any difference between the mean scores for each of the subtests, and the chance score of 50%. All pre-implant scores were significantly better than the chance score. However, the post-implant quarter-octave mean of 55% was not significantly different to chance level performance ($p =$

0.219), showing that, on average, subjects were not able to rank pitches one-quarter of an octave apart when using a CI.

9.2.5 Instrument Identification Test

The subjects scored 53%, 43%, and 35% pre-implantation, and 65%, 47%, and 46% post-implantation for the single instrument, instrument with background accompaniment, and music ensembles subtests, respectively (Table 8.27). An independent-samples t-test showed no significant difference between the CI+HA groups' score-difference mean and the WL group's score-difference mean (subtest 1: $p = 0.275$; subtest 2: $p = 0.945$; subtest 3: $p = 0.072$).

To investigate if there was any significant difference between performance across the three subtests, separate 1-way repeated-measures ANOVAs were conducted for the pre-implant, and post-implant scores. Results of these analyses showed that there was a significant difference for the factor of subtest both pre- and post-implant ($p = 0.002$, and $p < 0.001$, respectively). Tests of the within-subjects contrasts showed that both pre- and post-implant, there were significant differences between scores on subtests 1 and 2 (pre: $p = 0.004$; post: $p = 0.001$), as well as subtests 1 and 3 (pre: $p = 0.009$; post: $p = 0.002$). There was no significant difference between the scores of subtests 2 and 3 (pre: $p = 0.107$; post: $p = 0.692$).

Based on the confusion matrices presented in Chapter 8 (Tables 8.28 – 8.33), it was noted that pre-implant, the most recognised single instruments were the drum kit, piano, and xylophone, all recognised 72% of the time. Post-implant, it was the piano, and the male singer that were the most accurately recognised instruments (89%). For the second subtest, the timpani, and male singer were the most recognised stimuli, both pre- and post-implantation. The guitar was the least accurately recognised instrument for both of these subtests pre-implantation. Post-implantation, it was the flute, and clarinet that were the hardest for the subjects to recognise in the first and second subtests, respectively. For the music ensemble stimuli, the choir was the most recognised group pre-implantation, with the male singer and piano duet being the most recognised group post-implantation. The least recognised ensembles were the string quartet when tested with their HAs, and a tie between the country and western band, and the violin and piano duet when using the implant.

Pre-implantation, for both subtests 1 and 2, several of the instruments such as the clarinet, cello, trumpet, flute, and even the female singer were often mistaken to be a violin. With the exception of the cello which represented a confusion within the same instrumental family, the other instruments, including the female singer, had a similar pitch range to the violin. Other common errors for these two subtests were the confusion between the timpani and drum kit, as well as the guitar being misrecognised as a piano. For the third subtest, common errors included confusions between the orchestra and string quartet, the rock band being mistaken to be the percussion group, or the string quartet being incorrectly chosen as a violin with piano duet, or an orchestra.

When performing the same task with the implant, the overall error patterns of these newly implanted WL subjects were in many ways similar to those of the CI subject group discussed earlier. For example, the male and female singers in the first and second subtests were rarely selected to be instrumental stimuli, with the excerpts incorporating the male singer more accurately identified than those incorporating the female singer in all three subtests. Similarly, in the last subtest, the most common error for the trio of a male singer, female singer, and piano was its selection as a duet between a male singer and piano, indicating that the female voice in the extract was not perceived by some subjects.

9.2.6 Instrument Quality Rating Test

A comparison of the ratings given pre- and post-implantation is provided in Figure 9.8. An independent-samples t-test showed that the higher post-implant ratings (Table 8.34) were significantly different to the CI+HA groups' score-difference mean for the second and third subtests ($p = 0.005$, and $p = 0.009$ respectively). The difference for the first subtest was nearly significant ($p = 0.059$).

In order to see if there was any difference between the ratings provided by the WL subjects for the three subtests, separate 1-way repeated-measures ANOVAs were performed for the pre-implant and post-implant ratings. Pre-implant, there was a significant difference for the factor of subtests (p = 0.011), with the within-subject contrasts showing this to arise from the difference between the ratings for subtest 1 and 3 (p = 0.029). Post-implant, there was no significant difference between the subtests' ratings (p = 0.614).

Based on the responses listed in Tables 8.35 and 8.36, it was observed that there was a large degree of similarity in the subject's instrumental preferences pre-to-post implantation. Pre-implantation, for both the first and second subtests, the timpani was the highest rated instrument, with the guitar being the lowest rated. For the third subtest, the percussion ensemble and string quartet were rated to be the most pleasant sounding ensembles, with a tie between the rock band, and the trio of male singer, female singer, and piano, for the least pleasant ensembles. Post-implantation, the timpani remained the highest rated instrument in subtest 1 and 2, with the string quartet receiving the highest appraisal ratings for subtest 3. The lowest appraised instrument or ensemble post-implantation for the three respective subtests were the flute, the trumpet, and the trio of a male singer, female singer, and piano.

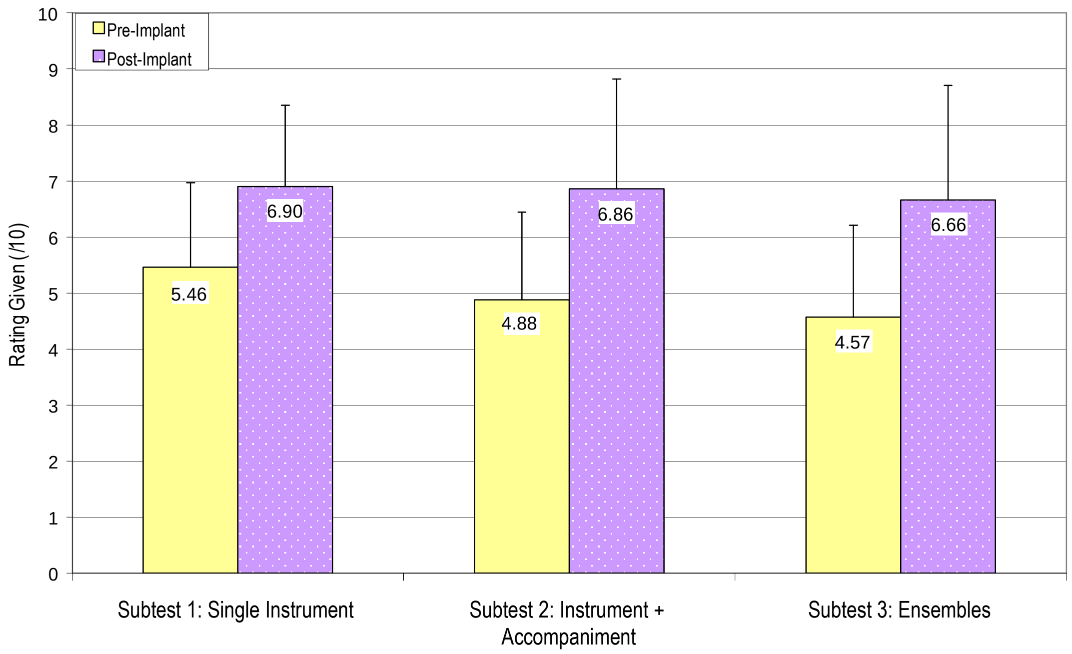


Figure 9.8: Quality rating scores for the three subtests, for the WL group
The X-axis lists the three subtests. The Y-axis represents the mean rating out of 10, where '1' = very unpleasant, and '10' = very pleasant. The error bars indicate 1 standard deviation.

9.2.7 Melody Test

The pre-implant mean was 75% (SD = 25.74), and the post-implant mean was 80% (SD = 23.98). An independent-samples t-test showed no significant difference between the CI+HA groups' score-difference mean and the pre-to-post surgery improvement recorded by the WL group (p = 0.776). When tested with their HAs, Waltzing Matilda and Baa Baa Black Sheep were the best recognised melodies by the WL subjects, whilst with the CI, For He's A Jolly Good Fellow, and Happy Birthday were the best recognised melodies (Tables 8.38 and 8.39).

9.2.8 Summary

There were significant differences between the two sets of score-difference means (i.e. the difference between the scores on the two test blocks for the combined CI and HA groups, and the difference between the pre- and post-surgery test scores for the WL group) for the one-octave and quarter-octave pitch subtests and for the second and third subtests of the quality rating assessment. That is, the degree of change between the WL group's post-surgery and pre-surgery test scores was significantly different to the degree of change between the scores from the two test blocks for the CI and HA subject groups for these subtests only. The differences between the other pre-to-post surgery test scores for the WL group were not significantly different to the improvement in scores associated with the learning effect in the CI and HA subject groups.

9.3 HEARING AID SUBJECT GROUP COMPARED TO WAITING LIST SUBJECTS (PRE-IMPLANT)

Comparing the HA group's scores to the WL group's pre-implant scores allows assessment as to the equivalence of these two groups, prior to the WL group obtaining their CI. Although both groups were experienced HA users, the WL group were in the process of obtaining a different type of device (i.e. a CI) whilst the HA group were continuing to use their HA. A graphical comparison of these two sets of scores for the music test battery is provided in Figure 9.9. The data for this figure and the subsequent analyses was presented in Chapter 8, sections 8.1 and 8.3 (p. 129 & 140). It should be noted that for the analyses and comparisons in this section, only the scores and means from the HA subject group's first test block were used. This was felt to be the most appropriate comparison as the WL subject group only had one run of the test battery with their HA. This also helped to minimise any bias from a learning effect that was present in the results of the HA subject group's second run. To provide a general comparison of the two groups, an independent-samples t-test was conducted comparing the first test block scores of the subjects in the HA group to those of the WL group pre-surgery, across all of the music test and subtest percentages. This showed that overall, the HA group were significantly better across the music test battery than the WL group ($p = 0.003$).

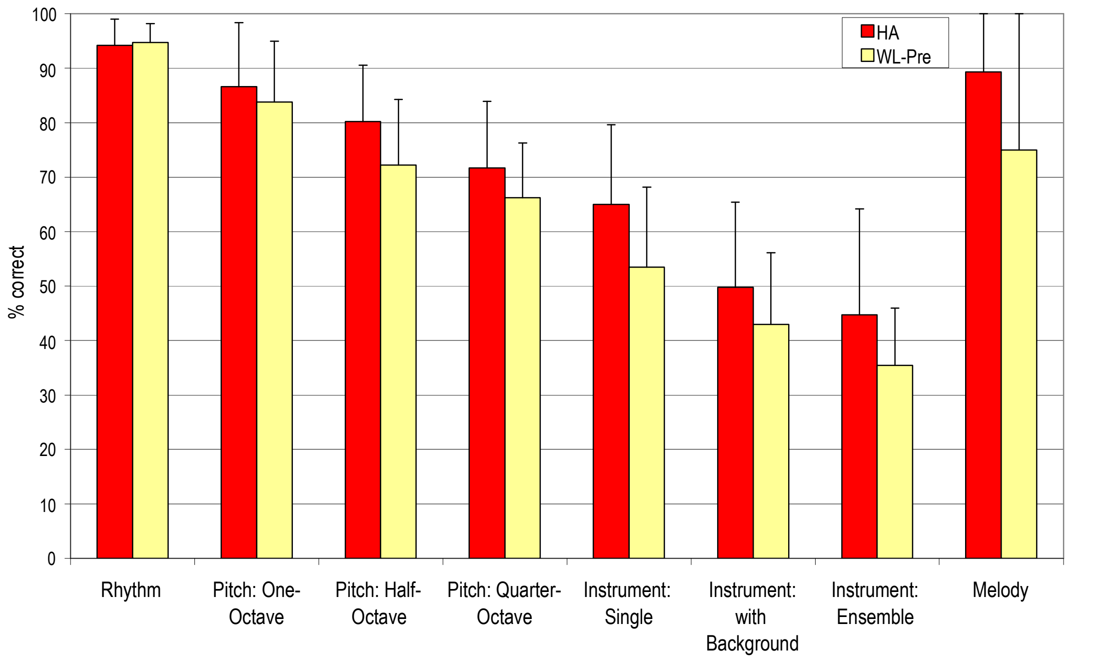


Figure 9.9: Comparison of scores on the music test battery for the HA group and WL group's pre-surgery scores

The X-axis lists the individual tests and subtests, with the Y-axis being the mean % correct scored by each group. The error bars indicate 1 standard deviation.

9.3.1 Subject Factors

A Mann-Whitney U test was conducted to see whether the two groups differed in their self-reported levels of listening to music whilst utilising a HA, based upon the current music listening scores (or HA music listening scores) as listed in Tables 7.1 and 7.3 at the end of Chapter 7. This showed a just-significant difference ($p = 0.048$), with the HA subject group reporting that they listened to music more than the WL subjects pre-surgery. There was no significant difference between the groups for their music experience levels, or their pre-hearing loss listening scores ($p = 0.446$, and $p = 1.00$ respectively).

9.3.2 Rhythm Test

An independent-samples t-test showed no significant difference between the HA group's first test block mean and the WL group's pre-implant mean ($p = 0.778$).

9.3.3 Pitch Test

A 2-way repeated-measures ANOVA was conducted to compare the two groups' performance across the three subtests. This showed no significant main effect for the between-subject factor of group ($p = 0.224$), but a significant main effect for the within-subject factor of subtest ($p < 0.001$), with no significant interaction between group and subtest ($p = 0.339$). That is, both groups performed similarly in their ability to rank pitches across the three subtests (Tables 8.2 and 8.24).

9.3.4 Instrument Identification Test

The results from a 2-way repeated-measures ANOVA showed no significant difference for the factor of group ($p = 0.120$). A significant main effect for the within-subject factor of subtest ($p < 0.001$) was noted, with significantly lower scores from subtests 1 to 2, and 2 to 3, respectively (Tables 8.4 and 8.27). There was no significant interaction between group and subtest ($p = 0.354$). That is, there was no statistically significant difference between the HA group and the WL group when tested with their HAs, in their ability to identify musical instruments or ensembles.

9.3.5 Instrument Quality Rating Test

A 2-way repeated-measures ANOVA showed no significant main effect for the between-subject factor of group ($p = 0.116$), with a significant within-subject factor of subtest ($p = 0.001$). There was no significant interaction between group and subtest ($p = 0.925$). The data were presented in Tables 8.8 and 8.34 in the previous chapter.

9.3.6 Melody Test

An independent-samples t-test showed no significant difference between the HA group's first test block mean (89%) and the WL group's mean when tested with their HAs (75%) ($p = 0.108$) (Tables 8.10 and 8.37).

9.3.7 Summary

There was no significant difference between any of the individual music test scores of the HA group and the WL group pre-implant. However, their overall performance on the test battery (i.e. when combined across all of the tests and subtests) was significantly different ($p = 0.003$), with the HA group scoring higher.

9.4 COCHLEAR IMPLANT SUBJECT GROUP COMPARED TO WAITING LIST SUBJECTS (POST-IMPLANT)

As the subjects in the CI group had at least one year's experience with their device whereas the WL subjects were tested at 3 months post switch-on of the implant, comparing these two data sets enables assessment of whether the newly implanted users involved in this study performed significantly differently to a group of CI users who had greater than one year's experience with the device. A graphical representation of this comparison for the music test battery is provided in Figure 9.10. The data for this table and subsequent analyses was presented in the previous chapter, in sections 8.2 and 8.3 (p. 135 & 140). For the analyses and comparisons in this section, only the means from the CI group's second test block were used. This was to account for the presence of a learning effect in the results of the WL group's post-surgery test results. An independent-samples t-test comparing the scores of the subjects in these two groups across all of the tests and subtests showed that there was no significant difference between the CI group and WL group post-surgery in their overall performance on the music test battery ($p = 0.145$).

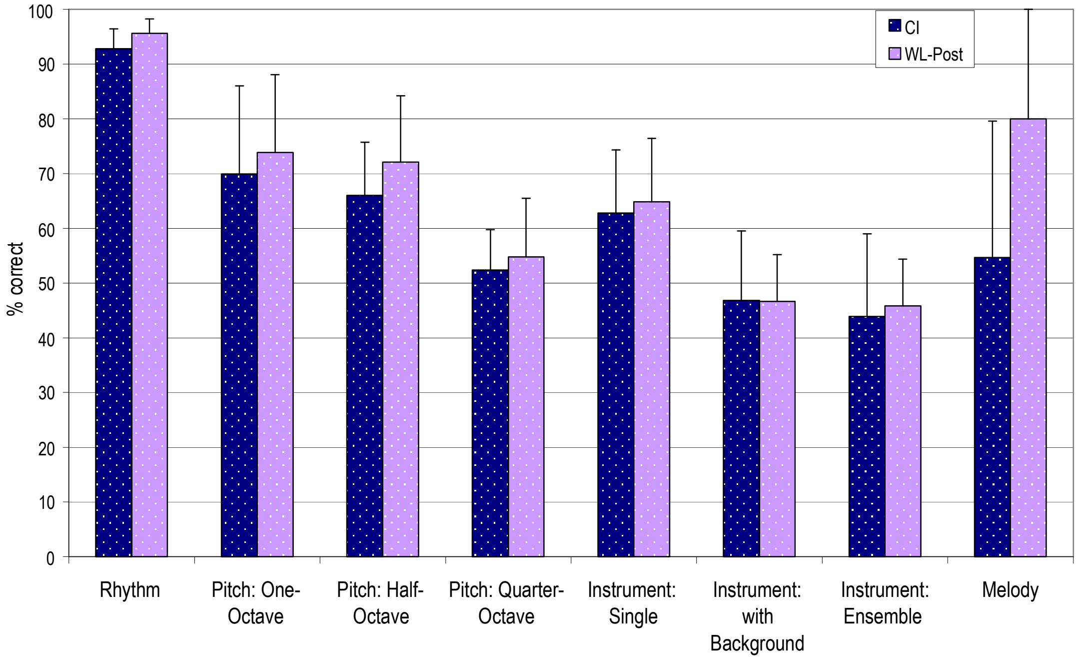


Figure 9.10: Comparison of scores on the music test battery for the CI group and WL group's post-surgery scores
The X-axis lists the individual tests and subtests, with the Y-axis being the mean % correct scored by each group. The error bars indicate 1 standard deviation.

9.4.1 Subject Factors

A Mann-Whitney U test was conducted to see whether the two groups differed in their reported levels of music listening at the time the study was conducted, whilst utilising the CI. This showed no significant difference between the ratings provided by the CI group on the MLEQ and the WL group on the follow-up version of the MLEQ ($p = 0.482$), as listed in Tables 7.1 and 7.3 (end of Chapter 7). That is, the CI group, who had been implanted for a greater amount of time than the WL group, did not spend a significantly different amount of time listening to music when compared to the newly implanted subjects. There was no significant difference between the two groups on their self-reported music experience, or their pre-hearing loss music listening levels (Mann-Whitney U test: $p = 0.770$, and $p = 0.411$, respectively).

9.4.2 Rhythm Test

An independent-samples t-test showed no significant difference between the WL group's post-implant mean of 96% (Table 8.23) and the CI group's second test block mean of 93% (Table 8.12) ($p = 0.057$).

9.4.3 Pitch Test

A 2-way ANOVA was conducted to compare the CI group's performance across the three subtests to those of the newly implanted WL group. This data was presented in Tables 8.13 and 8.24 in Chapter 8. Although there was a significant main effect for subtest ($p < 0.001$), there was no significant difference between the groups ($p = 0.320$), with no significant 2-way interaction ($p = 0.635$). This indicates that the newly implanted subjects of the WL group performed similarly to the experienced implant users (i.e. the CI group) on the pitch-ranking task.

9.4.4 Instrument Identification Test

Using the data reported in Tables 8.15 and 8.27, a 2-way ANOVA showed no significant difference between the groups in their ability to identify musical instruments or ensembles ($p = 0.623$), with no significant 2-way interaction ($p = 0.862$). There was a significant main effect for subtest ($p < 0.001$), with significantly higher scores for the single-instrument subtest than the subtests involving multiple instruments playing simultaneously.

9.4.5 Instrument Quality Rating Test

The results of a 2-way ANOVA based on the data in Tables 8.19 and 8.34 showed no significant difference in the ratings provided by the CI group and the WL group post-implant for the instrumental stimuli ($p = 0.814$). There was a significant main effect for subtest ($p = 0.010$) with the stimuli in the single-instrument subtest being rated as more pleasant sounding than the multi-instrument stimuli of the other two subtests. There was no significant interaction between the two factors of group and subtest ($p = 0.722$). These results indicate that there was no significant difference between appraisal ratings provided by the newly implanted WL subjects and the more-experienced CI subject group.

9.4.6 Melody Test

An independent-samples t-test showed that the newly implanted WL subjects' mean of 80% (Table 8.37) was significantly higher than the CI group's second test block mean of 55% (Table 8.21) ($p = 0.023$) on the melody recognition test.

9.4.7 Summary

With the exception of the melody test, there were no significant differences between the newly implanted WL subject group and the more-experienced CI subject group in their scores on the music tests.

9.5 CORRELATIONS

In order to investigate for potential relationships between the scores on different tests, values for non-parametric Spearman's rho correlation coefficient were calculated for specific comparisons. No adjustments for multiple comparisons were made. Calculations were made to assess: i) if the ability to recognise melodies was associated with pitch ranking or rhythm discrimination skills, and ii) if the ability to correctly identify a musical instrument or ensemble was related to the quality rating assigned to that stimulus.

Correlations were also calculated between performance on the music tests and five subject variables that have been occasionally shown to have some relationship with music perception – age, speech perception, music experience levels, music listening levels prior to having a hearing loss, and music listening levels at the time this study was conducted. As for the between-test correlations, Spearman's rho calculations were performed, with no adjustments being made for multiple comparisons.

9.5.1 Correlations Between The Music Tests

Non-parametric Spearman's rho calculations were performed for the CI, HA, and WL subjects to assess correlations between the melody test scores and scores on the pitch and rhythm tests. For the CI subjects, there was a significant moderate correlation between their melody test score and the mean score across all three pitch subtests ($rho = 0.679$, $p = 0.003$) (Figure 9.11). For the HA group and the WL group (both pre- and post-implantation scores), there were no significant correlations between the melody and pitch test mean scores. There were no significant correlations between the rhythm and melody test scores for any group.

Non-parametric correlations were also conducted to assess whether there was a relationship between the instrument identification scores and their corresponding quality rating appraisals for each subject group. These analyses showed a significant moderate correlation for the HA group (rho = 0.491, p = 0.001), with a slightly weaker significant correlation for the CI group (rho = 0.325, p = 0.029) (Figure 9.12 and Figure 9.13). The corresponding calculations for the WL group were not statistically significant.

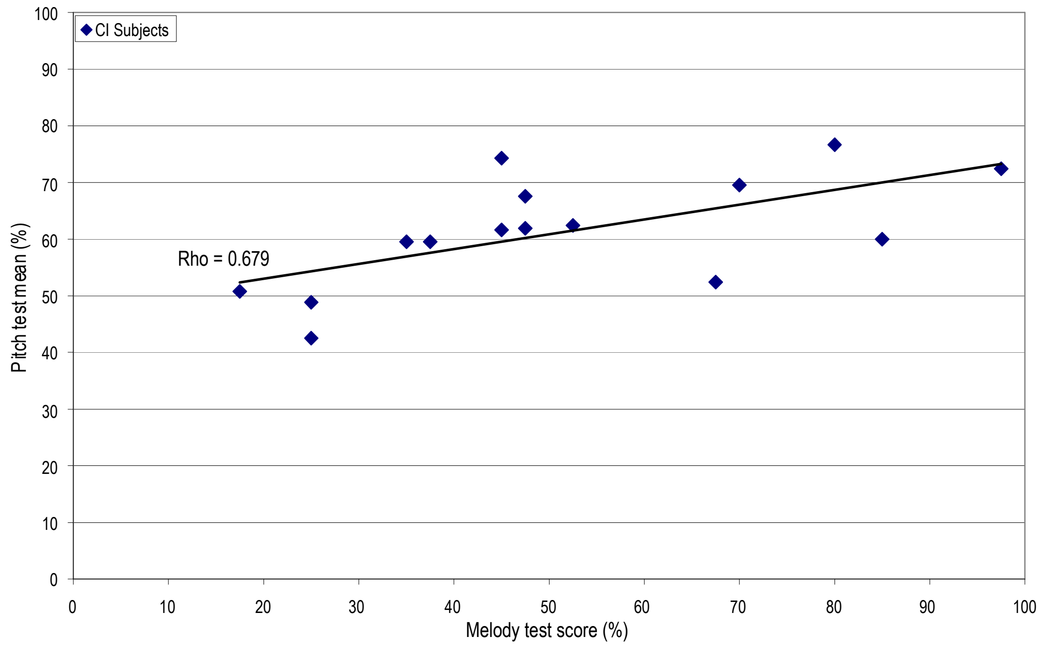


Figure 9.11: Correlation between the scores on the melody and pitch test for the CI subject group

Subject's mean melody recognition score from the combined test blocks are plotted on the X-axis. For the pitch test mean % as plotted on the Y-axis, each subject's mean score across each of the three subtests was averaged. The non-parametric Spearman's rho value was significant (p = 0.003).

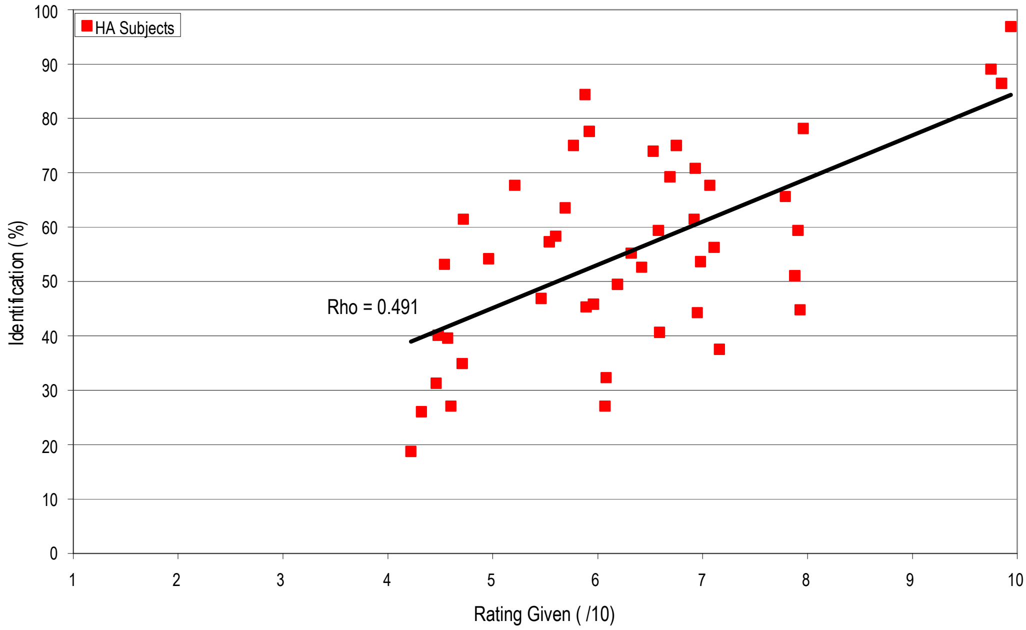


Figure 9.12: Correlation between quality rating and instrument identification scores across the 3 subtests, for the HA group

Each data point represents the correspondence between the mean appraisal rating out of 10 (X-axis) given by a subject for one of the subtests, and the mean identification score (%) for that same subtest (Y-axis). The non-parametric Spearman's rho value was significant (p = 0.001).

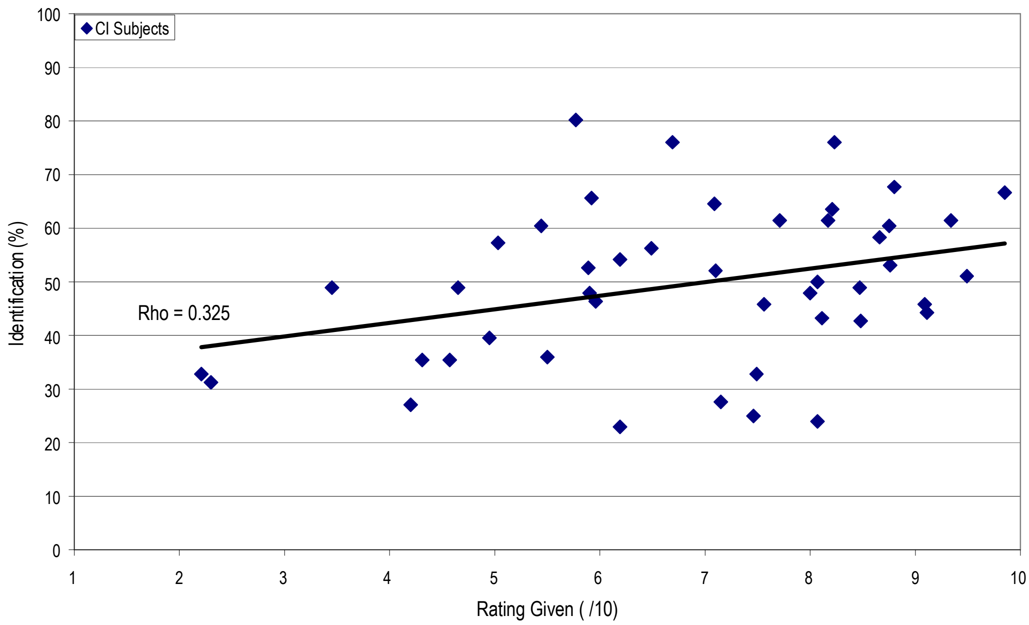


Figure 9.13: Correlation between quality rating and instrument identification scores across the 3 subtests, for the CI group

Each data point represents the correspondence between the mean appraisal rating out of 10 (X-axis) given by a subject for one of the subtests, and the mean identification score (%) for that same subtest (Y-axis). The non-parametric Spearman's rho value was significant (p = 0.029).

9.5.2 Correlations Between Test Scores And Subject Variables

To assess whether the subject variables of: i) age; ii) speech perception; iii) music experience; iv) amount of self-rated music listening prior to hearing loss; and v) amount of self-rated music listening whilst utilising their current device (as recorded in Tables 7.1, 7.2, and 7.3 at the end of Chapter 7), were correlated with the scores or ratings on the tests of: i) melody; ii) pitch; iii) instrument identification; and iv) quality rating, non-parametric Spearman's rho correlations were calculated. Speech perception scores were obtained using CUNY sentence stimuli presented in quiet at conversational levels. The music experience score, along with the two self-rated music listening scores, were obtained from the questionnaires described in Chapter 6, the MLEQ and the MTEQ. For the pitch, instrument identification, and quality rating assessments, each subject's mean across the three subtests was averaged; this figure was then used to calculate the relevant correlation.

For the CI subject group, there was a just-significant moderate negative correlation between age and instrument identification scores (rho = -0.525, p = 0.044); older subjects tended to be less accurate at identifying the musical instruments. A significant moderate relationship was also found between the instrument quality rating scores and the degree of current self-reported music listening (rho = 0.681; p = 0.03). That is, subjects who reported spending more time listening to music were more likely to appraise the instrumental stimuli as sounding more pleasant. The only other significant correlation was between the WL group's pre-implant speech perception scores and their instrument identification scores obtained when tested with their HAs (rho = 0.867; p = 0.012). There were no significant correlations between the above-mentioned subject variables and any of the test scores for either the experienced HA subject group, or the WL group's post-implant assessment.

CHAPTER 10: DISCUSSION

The results of this research partially verified the first hypothesis that experienced CI users would score lower than the HA subject group on the pitch, instrument identification, and melody tests, but not the rhythm test. Consistent with previous research, the CI subjects in this study found tests involving pitch, instrument, or melody perception significantly more difficult than those involving just rhythm perception (Dorman et al., 1991; Fujita & Ito, 1999; Gfeller & Lansing, 1991, 1992; Gfeller et al., 1997, 1998, 2002a, 2002c; Leal et al., 2003; McDermott, 2004; Schulz & Kerber, 1994). As hypothesised, there was no significant difference between the CI and HA subjects on the rhythm test. However, the CI group scored significantly lower on the pitch and melody tests ($p < 0.001$ for both comparisons), although equivalent to the HA group on the instrument recognition tests. The melody recognition results showed a similar wide range in individual abilities as that reported in previous studies (Gfeller et al., 2002a, 2005; Kong et al., 2005), with significant disparity between the results for the HA and CI groups.

The second hypothesis, that the subjects on the waiting list for an implant would score higher on the pitch, instrument identification, and melody tests when tested with their HA pre-implantation than post-surgery with their CI, only held true for the pitch test. Pitch-ranking scores post-surgery with the CI were significantly worse than pre-surgery with the HA for the one-octave and quarter-octave subtests ($p = 0.007$, and $p < 0.001$, respectively). There were no significant differences between the scores obtained with the two devices for the rhythm, instrument identification, or melody tests.

The third hypothesis, stating that subjects utilising a HA (i.e., both the HA subject group and the WL subject group when tested with their HA pre-implantation) would rate music to sound more pleasant than the subjects utilising a CI (i.e., the CI subject group and the WL subject group when tested post-implantation), was not supported by the findings of this research. For the WL group, appraisal ratings for the pleasantness of music stimuli were significantly higher post-surgery (i.e. with the CI) than pre-surgery with the HA (subtest 2: $p = 0.005$; subtest 3: $p = 0.009$). For the comparison between the CI and HA subject groups, the CI group also rated the sounds to be more pleasant than the HA group. Although statistical testing did not reveal the difference between these two groups to be significant, probably due to the large degree of intersubject variability in the ratings given, the trend was consistent across all three subtests.

The performance of the subject groups on each of the tests in the music test battery is now discussed further.

10.1 MUSIC TEST BATTERY DISCUSSION

10.1.1 Rhythm

As expected, performance on the rhythm test was fairly uniform across all subjects and subject groups, with the ceiling effect being observed for all three groups. These findings are consistent with the literature indicating that subjects with hearing impairments can generally discriminate rhythms as well as those with normal hearing (Gfeller & Lansing, 1991, 1992; Gfeller et al., 1997, 200b; Schulz & Kerber, 1994). The perception of rhythm requires the perception of the time-varying envelope fluctuations that occur in the frequency range of approximately 0.2 Hz to 20 Hz (McDermott, 2004). These low rates provide amplitude envelope information, which for music, corresponds to the gross rhythm and tempo of the stimuli. Rates above approximately 50 Hz provide waveform periodicity information, which can aid in the perception of the F0 (Rosen, 1992). Research indicates that most CI user's can perceive rate changes up to around 300 Hz (Grant et al., 1998; Moore, 1995; Pijl, 1997a, 1997b; Pijl & Schwarz, 1995). This approximate cut-off is very important in relation to pitch perception, as discussed in the next section.

10.1.2 Pitch

The CI subject group obtained lower pitch-ranking scores than the HA group for all interval sizes tested, only scoring at chance levels for the quarter-octave stimuli (frequency ratio ~19%). That is, as a group, the CI subjects were unable to discriminate between notes three semitones apart. This result is considerably poorer than that obtained by the subjects in Gfeller et al.'s (2004) study who received a conventional CI (as opposed to a separate group who were implanted with a CI having a modified electrode array). The subjects in that study scored between 70% - 78% correct, depending on the frequency range, for discriminating intervals one semitone apart. This score disparity may be largely attributable to the use of puretone stimuli by Gfeller et al. (2004) as opposed to the more complex, and more realistic, sung-vowel stimuli for this current study. The stimulation pattern arising from the complex nature of the sung-vowel stimuli may have provided conflicting temporal and spatial cues to the listener (discussed later in this section), making pitch perception challenging.

For the WL subject group, scores post-surgery with the CI were worse than pre-surgery with the HA for all three pitch subtests. Statistical analyses showed a significant difference between the degree of change from the pre-to-post surgery scores for the WL group and the degree of change between the two test blocks for the combined CI and HA groups for the one-octave ($p = 0.007$) and quarter-octave ($p < 0.001$) pitch subtests. The change between the WL's scores for the half-octave subtest scores was approaching significance ($p = 0.061$). It is worthwhile keeping in mind that the changes for both the HA and CI groups were in the opposite direction; that is, they improved in their pitch-ranking scores from the first test block to the second test block (refer to Figure 9.7 in the previous chapter). Therefore it could be postulated that if the WL subject group had similarly been tested on two occasions with their HAs (i.e. had they not received the CI), their pitch scores would also have been likely to improve. Contrarily however, their results were poorer on the second test block, when they were tested with the implant. Further, the mean post-implant score for the quarter-octave subtest was not significantly different from the chance score of 50%. That is, as was the case with the CI subject group, the newly implanted WL subjects were also unable to reliably select the higher of two notes, three semitones apart.

Despite being higher than the scores obtained from the CI group, the pitch results from the HA users (both the HA subject group and the WL subjects pre-surgery) were not as good as what one may expect from NH listeners. Schulz & Kerber (1994) and Gfeller et al. (2002a) found that the majority of the NH controls in their studies could reliably rank pitches one semitone apart. The group of NH subjects in this study, whose role it was to verify the music test battery, scored greater than 95% when ranking pitches three semitones apart. The HA group only scored 75%. A number of researchers have reported reduced frequency selectivity arising from increased auditory filter bandwidths in listeners with cochlear hearing losses (Arehart, 1994; Moore, 1995, 1996; Moore & Peters, 1992; Summers & Leek, 1994). Moore (1996) reported that hearing thresholds worse than approximately 40 dB HL to 50 dB HL result in auditory filters with bandwidths approximately two times wider than those for NH. The reduced frequency selectivity due to the wider filter bandwidths would have a deleterious effect on pitch-based tests as the listener would be less able to resolve the lower-order harmonics, affecting their perception of the F0. Diminished pitch perception associated with cochlear hearing loss can be further affected by impaired temporal processing skills which would impact on the listener's ability to derive pitch information from temporal-based (phase-locking) cues (Arehart, 1994; Moore, 1996; Moore & Carlyon, 2005; Moore & Peters, 1992; Moore & Skrodzka, 2002; Oxenham et al., 2004). These factors are relevant considerations for the HA subjects in this study, all of whom had a moderately severe to profound cochlear hearing loss.

Perceiving the pitch of a complex sound involves the listener having to extract F0 information from the complex acoustic signal. In order to extract this information, two different mechanisms are required – i) resolving the individual frequency components present in the signal, and ii) extracting the temporal pitch information from the signal. For an implant user, both of these are mechanisms are affected by a multiplicity of factors discussed below. For the NH listener, both spectral and temporal cues would contribute to the perception of pitch, but the relative degree of contribution from each would vary depending on the parameters of the signal. The contribution of both mechanisms would provide some degree of redundancy to assist pitch perception in more-challenging acoustic listening situations. As discussed in Chapter 2, various models have been proposed by researchers to explain pitch perception for complex tones. Each model varies in the roles and relative importance played by spectral- and temporal-coding mechanisms in deriving pitch. These models can be broadly classified as either place-based, or temporal-based models. The former models rely predominantly on pattern recognition where the resolved partials of complex tones are matched to a central template. The latter models are based on the analysis and integration of neural firing patterns to provide a pitch percept. Common to both classes of theories, though, is the initial spectral analysis; the basilar membrane mechanics of a normal functioning cochlea enables the cochlea to act as a bank of bandpass filters, analysing the input signal and dividing it into its frequency components. As the auditory filter bandwidths are wider at higher frequencies, the lower harmonics of a complex tone are more likely to be fully resolved (i.e., passing through separate filters), whilst the higher-order harmonics may remain unresolved with several harmonics passing through a single filter. The lower harmonics of a complex tone are considered to be the most important in providing the pitch percept (Houtsma & Goldstein, 1972; Moore, 2003a; Plomp, 1967; Ritsma, 1967).

However, F0 information can also be extracted from temporal pitch information when the harmonics remain unresolved. Unresolved harmonics would preclude place information about the harmonic frequencies from being obtained; instead, several harmonics would interact along the membrane, with specific sites being excited by several harmonics simultaneously. Although this results in a complex vibration pattern, the pattern repeats at a rate equal to the F0. Therefore the pitch of a complex sound can be determined from the repetition rate of interacting harmonics (Moore & Rosen, 1979).

Cochlear implant users also rely on temporal and place cues to perceive pitch information; however, it is the preservation, coding, and effective use of these cues that are key concerns pertaining to the perception of pitch through the implant. CIs were designed on the premise that activating different electrode sites would result in different pitch percepts in accordance with the tonotopic structure of the cochlea. However, current filterbanks employed by CI systems differ from the auditory filters of a normally hearing cochlea in several ways. Whereas the cochlea's auditory filters are non-linear and level-dependent with continuous centre frequencies, the filterbanks of the CI22 and CI24 implants only allow for a maximum of 22 overlapping filterbands with fixed centre frequencies. The widths of these filters vary depending on the number of filters being used and the centre frequency, ranging from a width comparable to those in a normal cochlea to ones much wider. A wide filter may preclude the lower harmonics of complex sounds from being fully resolved, making it difficult for the listener to precisely derive the harmonic frequencies, and make reliable pitch judgements. Even if the individual harmonic components were resolved, falling into separate filters, the CI user would only be able to determine which filter the component passed to, as the corresponding electrode would be activated. However, the CI user would be unable to determine the precise position of the signal within that filter's bandwidth and accordingly, the exact frequency of the resolved component. Further, if the resolved components were in adjacent filters which subsequently activated two or more adjacent electrodes, it would be unlikely that the CI user could resolve the places of stimulation to accurately determine pitch information.

This inaccuracy in pitch perception resulting from poor frequency resolution might be further confounded by a mismatch between the frequency of the CI's filter and the corresponding characteristic frequency in the cochlea. With typical electrode insertion depths extending to the first 1.5 turns of the cochlea only, and more-apical sites not being directly stimulated by electrodes, filterbands assigned to active electrodes tend to be lower in frequency than the characteristic frequency normally associated with that stimulation site. Psychophysical experiments have shown that correct tonotopic place of stimulation may be required for accurate pitch perception. For example, Oxenham et al. (2004) reported that when temporal information from the higher harmonics of a stimulus was presented to an incorrect tonotopic location along the basilar membrane, subjects with NH were largely unable to combine the information centrally to give a reliable F0 percept.

For many implantees, the stimulation of individual electrodes does not necessarily result in different pitch percepts, or the spectral definition may be diminished if a wide neural population is stimulated by an active electrode. Should this be the case, their ability to resolve individual frequency components would be affected by channel interactions and spectral smearing. Factors such as the pattern of neural survival, pathological processes in the cochlea, the impedance surrounding the active electrode, and the electrode's proximity to the target neurons would all impact upon the perceptual discreteness of individual pitch sensations, along with the number and order of the different pitch percepts detected. For example, pitch percepts related to electrode position may not necessarily be monotonically ranked on a a scale from low to high by a CI subject, as the electrode position moves from apical to basal (McDermott, 2004; McKay, 2004, 2005).

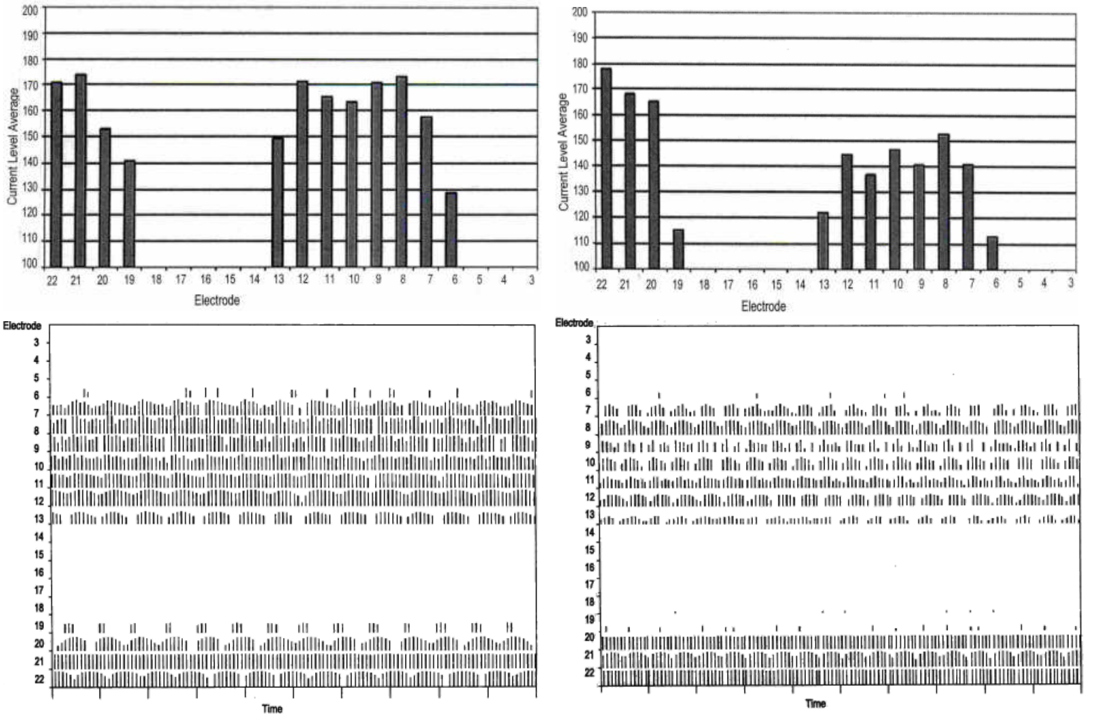
The availability and use of temporal cues also contribute to an implant user's perception of pitch. Pitch information can be conveyed by changes in the rate of the stimulating pulse train. However, the constant-rate stimulation of the ACE and SPEAK strategies applicable to this study would have precluded pitch information from being conveyed in this manner. For the Nucleus implants worn by subjects in this study, temporal cues would be available via amplitude modulations present at the filterbank's output, at a rate corresponding to the input signal's F0. As explained in Chapter 3, research has shown that these amplitude modulations in the envelope of the electric stimuli can provide a pitch percept. The properties and underlying mechanisms related to this pitch percept may be comparable to those experienced by NH listeners when listening to amplitude-modulated noise (Burns & Viemeister, 1976, 1981). The availability and clarity of these modulations are contingent upon having a filter wide enough to encompass more than one harmonic, an overall stimulation rate that is sufficiently higher than the modulation frequency to permit accurate sampling, a sufficiently deep modulation depth, and the filter's output not being smoothed by the processing strategy. An inherent feature of the filters utilised in the speech-processing strategies involved in this study is the progressive attenuation of amplitude modulations above approximately 200 Hz to 300 Hz (McDermott, 2004; Vandali et al., 2000).

How salient and reliable these amplitude-modulation cues are is further dependent upon the alignment of the phase of these modulations across electrode positions. Phase misalignments may occur for a host of reasons, including those related to the acoustic environment, and they can occur across a wide frequency range. If these misalignments are present at the filterbank's output, the resulting overall modulation pattern may no longer reflect the F0 if the modulations are perceptually combined across electrode positions. That is, if the implant user perceptually integrates the information across channels, the resulting modulation pattern perceived may no longer reflect the F0. On the other hand, if the implant user perceives modulation information in each channel independently, the relative phase of the modulations across channels would have no consequence on the resulting pitch percept. The amplitude-modulation pattern within each channel would be identical, with a repetition rate equal to the F0. Research by McKay & McDermott (1996) indicates that CI users may integrate temporal patterns across closely spaced active electrodes, resulting in reduced temporal cues. Electrode spacing needs to exceed a particular distance, the exact size of which varies from one CI user to the next, before temporal patterns on different electrodes are perceived independently. In McKay & McDermott's (1996) study, this distance ranged from 2.25 mm to 7 mm.

It also appears that CI users are only able to extract reliable pitch cues from these amplitude modulations at frequencies up to around 300 Hz, implying that the majority of CI users would have difficulty in obtaining reliable pitch cues from temporal variations in stimuli with a F0 above approximately middle-C (McKay, 2004; Pijl, 1997a; Zeng, 2002). Research has revealed large differences between implantees in their ability to extract pitch information from the temporal variations in electric stimuli, and also for the range over which then can extract the pitch information - some implantees can discriminate smaller changes over a wider frequency range than others (McKay, 2004; Pijl, 1997a; Pijl & Schwarz, 1995; Shannon et al., 2004; Tong & Clark, 1985; Zeng, 2002). The upper rate limit of temporal pitch affects an individual’s ability to determine the pitch of stimuli with higher F0s. That is, an implantee with a higher upper limit of temporal resolution may be able to use amplitude-modulation cues to obtain pitch information across a larger frequency range.

The accuracy of CI users’ responses in pitch-ranking tasks may also be dependent upon which cues the subject uses in making their perceptual judgements. Researchers have suggested that place- and temporal-based cues provide two perceptually independent sensations to the listener (McDermott, 2004; McDermott & McKay, 1997; Moore & Carlyon, 2005; Pijl & Schwarz, 1995). In making perceptual decisions, subjects may apply different weightings to these cues depending upon their salience and availability during each task. Consequently, the accuracy of their pitch perception would be related to which cue was more salient, and the consistency and reliability of the information provided by these cues. Should the cues provide conflicting or incorrect information, the implant user’s ability to make accurate pitch decisions would be affected. For example, it is possible that for an ascending pair of notes, the electrode stimulation could shift more apically due to downward changes in formant frequencies. Shifts in formant frequencies are not necessarily reflective of the direction of change in the F0

(McDermott, 2004). If the site of stimulation shifts apically, a pitch reversal will result for a subject attending to mainly place cues. A representation of this is provided in Figure 10.1.



A – 139 Hz **B – 196 Hz**

Figure 10.1: Conflict in place versus temporal cues

This diagram is an example of a situation where changes in temporal and place cues may provide conflicting pitch information. The two column graphs on the top row represent the distribution of electrical activity (i.e. the current level distribution) across the electrodes. Along the X-axis are the individual electrodes, with 22 being the most-apical and 3 being the most-basal. On the Y-axis is the average current level; the Nucleus CI delivers currents ranging from 10 µA to 1750 µA, divided into a logarithmic scale of 255 levels. Hence these column graphs provide a visual representation of the distribution of electrical activity in the cochlea for a certain stimulus. The two electrodograms on the bottom row depict the stimulation occurring on individual electrodes across a certain period of time. The X-axis represents time, with a total duration of 100 ms. The numbers on the Y-axis correspond to the individual electrodes. The current level graph and electrodogram for the male-sung vowel /i/ is presented. In the left column, 'A', the F0 is 139 Hz. For the right column, 'B', the F0 is 196 Hz.

The current level graphs show a change to the distribution of electrode stimulation from A to B. There is less electrical activity at the basal electrodes for B than A, contrary to the rise in the F0 from A to B. That is, if the subject was attending to place cues, they would have ranked A to be higher in pitch than B. However, the corresponding electrodograms show that amplitude modulations repeating at a rate related to the F0 are also present at most of the active electrodes. If the subject was attending to these temporal cues, and was not affected by the phase misalignments most apparent at the apical electrodes of A, then they would have accurately ranked B as higher than A.

(Figures courtesy of D. Tsang)

Further, the subjects in this study frequently commented that the two notes within a particular pair were the "same" or "very close"; even for intervals of the same distance, one pair of notes may have been judged to have very similar pitches, whereas the notes in the next pair were judged to be obviously different in pitch. It is quite probable that in situations where the pitch sensation was ambiguous or indistinct for the implant user, they could have transferred their attention to alternate cues in their decision-making. For example, more salient sensations such as timbre, or even loudness, may have influenced their judgements. This may have also led to incorrect pitch-ranking decisions should the direction of shifts in electrode selection be inconsistent with the direction of the change in the F0. With researchers suggesting that variations in the place of stimulation affect timbre more than pitch (McDermott, 2004; McDermott & McKay, 1997; Moore & Carlyon, 2005; Pijl & Schwarz, 1995), it is highly possible that perceptual judgements were influenced by variations in timbre. These factors may partially account for some of the inconsistency, unreliability, and variability both within and between individual CI user's pitch perception results.

Other stimulation-related factors, as well as environmental, physiological, and pathological considerations impact upon an individual implantee's perceptual accuracy for pitch. Examples of these variables include their memory for melodic pitches, music knowledge or training; the location, number, and density of surviving neurons in the cochlea; the electrode's placement or insertion depth; the impedance surrounding the electrodes; pathological processes; central processing factors; and the stimulation mode used or electrical current path within the cochlea (McDermott, 2004; McKay, 2004, 2005; McKay & McDermott, 1993; Pijl, 1995). All of these factors contribute to the variability amongst CI users in their ability to perceive pitch. The large variability in the pitch perception skills of individual implantees reported in previous research (Fujita & Ito, 1999; Gfeller et al., 1997, 2002a; McDermott, 2004; Schulz & Kerber, 1994) was also evident in this study. For example, for the one-octave subtest, the CI group's mean scores ranged from 24% to 91% with a standard deviation of 17%. In contrast, the HA group's mean scores for this same subtest ranged from 76% to 98% with a standard deviation of 10%. The wider range of scores for the CI subjects compared to the HA subjects was observed for all three interval sizes assessed.

The issues related to temporal-pitch perception may also account for the greater accuracy demonstrated by the CI subjects for male-sung stimuli than those sung by a female voice. In the pitch test, the CI group were significantly better at ranking the vowels sung by the male vocalist than those sung by the female vocalist ($p = 0.035$). The post-implant pitch test results from the WL subjects demonstrated a similar pattern; they were also significantly more accurate with the male-sung than female-sung vowels ($p < 0.018$), a discrepancy not present in their pre-implant testing scores. As already discussed, it is highly probable that a CI user would use several cues in making pitch judgements. With temporal cues only being available for frequencies lower than approximately 200 Hz to 300 Hz, the lower F0s associated with male voices may have enabled subjects to use temporal modulation cues in making pitch judgements. For pitches in the mid-range, which would have applied to most of the female-sung vowels in this study, two possibilities arise as to why pitch discrimination was less accurate. Firstly, both temporal and place changes may have been perceived by the implant user, but the two types of change provided conflicting information, making ranking unreliable and inconsistent. On the other hand, it may have been possible that whilst the F0 exceeded the subject's upper limit of temporal resolution for using temporal-pitch cues, the difference between the F0 of the pitches in the stimulus pair was not sufficiently large to result in a change in the place of stimulation to provide a pitch cue. Changes in the place of stimulation associated with the F0 are not systematically related to the actual frequency of the fundamental. With place cues at the output of the processor being related to the peaks in the spectral envelope, it is possible that there were no place-cue changes related to the shift in the F0. An additional consideration, as has been stated, is the progressive attenuation of amplitude modulations above approximately 200 Hz to 300 Hz inherent to the speech-processing strategies involved in this study (McDermott, 2004). This may have had a greater adverse effect on the preservation of temporal cues for the female-sung vowels than male-sung vowels, due to the relatively higher F0s for some of the former's stimuli.

In contrast to the CI users, the HA subject group scored significantly better for the female-sung vowels than the male-sung vowels ($p = 0.001$). The lower harmonics of sounds with higher F0s are more likely to be fully resolved, falling into separate auditory filters. Thus the higher F0s for the female-sung vowels would have resulted in a larger number of resolvable harmonics to provide place-pitch cues. For a listener with

NH, the resolved harmonics are important for extracting the pitch percept (Moore & Carlyon, 2005; Ritsma, 1967). As covered in Chapter 2, for acoustic hearing, these place cues have been reported to provide a more salient perception of pitch than temporal cues conveyed by the unresolved harmonics (Arehart, 1994; Moore & Rosen, 1979; Plack & Oxenham, 2005).

The overall results of the pitch test indicate that the CI does not provide wearers with sufficiently reliable pitch cues in order for them always to discriminate between two musical notes. For the largest interval size assessed, the one-octave interval which equated to a frequency ratio of 2:1, subjects only averaged 68% correct. In other words, even when the F0 was halved or doubled, subjects could only accurately rank the pitches around two-thirds of the time. The rate of improvement in pitch-ranking scores with increased interval size from the quarter-octave to the one-octave interval was not uniform. The steepest rate of improvement was from chance level performance for the quarter-octave stimuli to 64% for the half-octave interval. Scores then plateaued, with little difference between performance for the half-octave and one-octave intervals. This suggests that for most of the implant users involved in this study, a frequency separation greater than a quarter of an octave was required to correctly select the higher pitch of two notes, for the sung-vowel stimuli used in this study.

It is important to note when interpreting the results of the pitch-ranking tests that it has been assumed that subjects were making judgements based on a single perceptual dimension – pitch. However, it is possible that subjects were making decisions based on a different dimension, such as timbre, or a combination of perceptual attributes. This may have been particularly relevant when the available pitch cues were weak; in this case the subjects may have shifted their attention to a different perceptual dimension. Some researchers have suggested that variations in the place of stimulation affect timbre more than pitch, however it is not possible to definitively separate these two dimensions in order to be able to state conclusively which dimension subjects were basing their perceptual decisions on (McDermott, 2004; McDermott & McKay, 1997; Moore & Carlyon, 2005; Pijl & Schwarz, 1995). As briefly mentioned in Chapter 2 and 4, studies by Beal (1985), Crowder (1989), and Pitt & Crowder (1992) found that, even for NH listeners, there were significant interactions between the two perceptual dimensions of pitch and timbre. The subjects were less accurate, and took longer, in making their pitch decisions when the timbre of the two sounds varied. This was particularly evident for subjects who had little or no musical training. The results of studies by Pitt (1994) and Pitt & Crowder (1992) demonstrated that if the timbre of two sounds varied whilst the pitch remained the same, non-musically trained subjects were more likely than not to state that the pitch, and not the timbre, had changed. Pitt (1994) postulated that non-musicians did not process the dimensions of pitch or timbre completely independently. He proposed that timbre was a more salient, more heavily weighted attribute than pitch for non-musicians, with this effect being exacerbated when the listener was unsure about the pitch. Considering that many of the hearing-impaired subjects in this study had minimal formal musical training, and also that the CI subjects in particular found the pitch perception task to be difficult, it is plausible that some of the subjects in this study may have made judgements in the pitch task on another perceptual attribute, such as timbre. That is, pitch-ranking tasks such as the one used in this study cannot definitively separate pitch from other percepts, such as timbre.

In summary, the pitch perception results for the subjects using a CI were significantly poorer than for those using a HA. In view that existing research collectively report CI users to be significantly poorer than NH listeners at pitch-based tasks, it may be inferred that the electrical stimulation of hearing impedes the listener’s ability to accurately perceive pitch. This may be attributed to a host of reasons including limitations of the implant device itself, the processing of the input signal, perceptual limitations of the implant user, as well as individual pathological and environmental factors.

10.1.3 Instrument Identification and Appraisal

For the instrument identification subtests, although it was observed that the HA group scored higher than the CI group for all three subtests, statistical analyses did not show these differences to be significant. Seemingly, the better pitch perception skills of the HA subject group did not translate into improved timbre perception. As per the comparison between the CI and HA groups, there was no significant difference in the WL group’s instrument identification scores obtained pre- and post-surgery, with the trend for higher identification scores post-surgery being due to a learning effect.

Given that the NH subjects who verified the subtests averaged above 95%, the comparatively poorer identification results for all three hearing-impaired subject groups may have arisen from a range of factors. Accurate timbre perception requires the perception of both the temporal envelope and the signal's spectral shape. For a NH individual, such spectral selectivity derives from the different frequency components of the acoustic stimulus being separated into different auditory filters, with each frequency component resulting in activity at discrete sites along the basilar membrane. The aim of multichannel CIs is to restore some of this frequency resolution by electrically stimulating sites along the cochlea. However, the degree of discreteness of these individual stimulation sites is not nearly as precise as for NH, and varies from one CI user to the next. Studies have shown that CI users can perceptually resolve complex spectral patterns to different electrode places, to varying extents (Henry & Turner, 2003; Henry et al., 2005; McKay et al., 1996; Tong et al., 1983). Henry et al. (2005) compared subjects with NH, hearing impairments, and CIs on a spectral peak resolution task using rippled-noise stimuli. They found that spectral peak resolution was best for the subjects with NH, intermediate for those with hearing impairments, and worst for those using an implant, with much individual variability within each group.

In CIs, the pattern of electrical stimulation across the electrodes is an electrical representation of the spectral shape of the input acoustic signal. The perceptual representation of this signal spectrum is referred to as the internal spectrum. Experiments adopting forward-masking paradigms allow assessment as to the relationship between the physical or external spectrum of a sound, and the listener's perception of the spectrum (i.e. their internal spectrum) (Laback et al., 2004; Stainsby, 2001; Stainsby et al., 1997). In one such study, Stainsby (2001) reported significant correlations between the internal and physical spectrum for CI and HA users, suggesting that both CI and HA users could achieve some degree of spectral selectivity through their device to enable timbre perception. However the amount of spectral information conveyed was less than for NH listeners, and was largely insufficient to allow the accurate identification of musical instruments. This is consistent with the results of the current study where both CI and HA subjects experienced significant difficulty in tasks of instrument recognition which NH listeners had found to be relatively easy.

As mentioned in the previous section when discussing the discreteness of pitch percepts, perceptual smearing experienced by CI users may be related to individual subject factors such as a wide current spread around the target electrodes, neural interactions, cochlear pathology, or neural survival patterns. The smearing may also be a by-product of the stimulation itself, such as the presence of channel interactions, or it may arise from the use of non-linear amplitude mapping functions in converting acoustic signals to appropriate levels for electrical stimulation (Laback et al., 2004; McDermott, 2004; McKay, 2004). Such spectral smearing, in combination with the coarse spectral analysis of the input signal undertaken by current pulsatile speech-processing strategies such as ACE and SPEAK, may in part account for the difficulty experienced by CI users in tasks related to timbre perception, including the identification of musical instruments. For HA users, perceptual smearing may occur as a consequence of auditory filter anomalies associated with cochlear hearing loss, poor neural survival patterns, and poor frequency selectivity. This may result in diminishing the spectral clarity of the stimuli for the subject (Arehart, 1994; Moore, 1995; Summers & Leek, 1994).

The high levels of inter-subject variability on the instrument perception tasks documented by previous researchers (Gfeller et al., 1998, 2002c; McDermott & Looi, 2004) was also observed in the current study for both the CI and HA subjects. In the single-instrument identification subtest, the CI subjects averaged 61% with a standard deviation of 11%, whilst the HA subjects averaged 69% with a standard deviation of 12%. This subject variability may be attributable to a host of reasons including sociological factors, physiological considerations, musical background or training levels, as well as cognitive, and auditory processing differences (Arehart, 1994; Gfeller et al., 1998 ,2005; Lentz & Leek, 2003; Moore, 1996)

Comparisons of the absolute identification scores obtained in the current study to previous research results are confounded by highly varying methodologies and test requirements between studies. For example, in Gfeller et al.'s (2002c) study, CI subjects averaged 47% in a closed-set identification task involving eight instruments presented three times each via a loudspeaker, whilst in McDermott & Looi's (2004) study, the CI recipients averaged 44% in identifying 16 musical instruments presented four times each via a loudspeaker in a closed-set format. Despite the range of methodologies and protocols, the inequality in performance between the CI and NH subjects is consistently evident across existing studies, with the current research also suggesting that HA subjects with significant levels of hearing impairment may similarly experience difficulty on instrument perception tasks when compared to those with NH.

This study also extended the investigation of timbre perception beyond the single-instrument identification tasks used in most previous studies. Results from the ‘instrument with background accompaniment’ and ‘music ensemble’ subtests reflected the more complex nature of these excerpts. Both the CI and HA groups’ mean scores were significantly lower for the multiple-instrument subtests than for the single-instrument subtest ($p < 0.001$), with greater standard deviations also being observed. For the CI subjects, the mean scores decreased approximately 16 percentage points from the first subtest to the second subtest, and 18 percentage points from the first to the third subtests. For the HA subjects, the decline was on average 17 and 21 percentage points respectively. For the WL subject group, a similar trend was observed with identification scores decreasing from subtest 1 to subtest 2, and then again from subtest 2 to subtest 3, both pre- and post-implantation. For all three subject groups, statistical analyses showed significant differences between the identification scores for the first subtest using single-instrument stimuli and both of the other subtests using multi-instrumental stimuli (CI and HA subject groups - subtest 1 & 2: $p < 0.001$; subtest 1 & 3: $p < 0.001$. WL subject group – subtest 1 & 2: $p < 0.004$; subtest 1 & 3: $p < 0.009$). The additional instruments present in the second and third subtests added to the complexity of the sound which appeared to negatively impact on hearing-impaired subjects’ perception of the stimuli.

In examining the error patterns in the confusion matrices for the instrument identification tests, several observations are worth noting. Gfeller et al. (1998, 2002b, 2002c) reported that whilst their NH subjects tended to make justifiable, consistent errors usually within the same instrument family, the CI subjects demonstrated more diffuse error patterns. The confusion matrix from McDermott & Looi's (2004) study (presented in McDermott, 2004) divided the 16 presented instruments into two categories – percussive and non-percussive instruments. Inspection showed that the CI subjects often made confusions within the same category (e.g., a non-percussive instrument mistaken to be another non-percussive instrument), however the confusions were not necessarily within the same instrument family. For example, the violin and clarinet were often mixed up. In this current study, confusions for the single-instrument subtest tended to be within the same instrument family for both the CI and HA subjects, with substantially more diffuse and across-instrument-family errors apparent in the second and third subtests. There was some similarity between the identification scores and appraisal ratings for both the CI and HA groups. For example, for the CI group, the xylophone was the most recognised and highest appraised single instrument, with the piano being the most recognised and highest appraised single instrument for the HA group. As a generalisation, it was noted that instruments from the percussion family such as the piano, drum kit, and timpani, were more likely to be correctly identified or rated to sound more pleasant. Their distinctive temporal envelopes, typically characterised by faster attack times (i.e. rise times), may have provided more salient durational or rhythmic cues for the subject, in addition to the strong timbre percept. As mentioned in Chapter 2, the temporal envelope of the input signal impacts upon the perceived timbre with Grey (1977) specifying rise (or attack) time as one of three perceptual dimensions to timbre.

The importance of these envelope cues in identifying complex sound stimuli extends beyond just musical sounds. For example, Reed & Delhorne (2005) investigated the recognition of environmental sounds for 11 CI users. The stimuli were 40 sounds, divided into four categories of 10 sounds each – general home, kitchen, office, and outside. In a closed-set identification task, the mean score over the four settings was 79.2%, ranging from 45% to 94% across the subjects. The authors reported that within all four categories, the best-recognised signals had distinctive temporal envelopes or unique durational attributes. Stimulus pairs that were most often confused tended to have similar durations or temporal traits, but differing spectral features. For example, in the kitchen category, the sounds of a cupboard door slam, dishes clanging, and footsteps were accurately identified, whereas the sounds of a dishwasher and running water were commonly confused. The authors concluded that temporal envelope cues were important for the identification of environmental sounds. This is in keeping with the error patterns observed for the instrument identification tasks in this study where instruments with distinctive temporal envelopes, such as percussion instruments, were identified more accurately by the CI subjects than many of the other stimuli.

In many ways, the post-implant error patterns of the WL subjects were comparable to the CI subject group. For example, the percussion instruments along with the stimuli involving the male singer were usually amongst the most correctly identified extracts for both groups. As with the CI subject group, the WL subjects, when tested with the implant, were more likely to accurately identify the excerpts with the male singer than those with the female singer for all three subtests. The most common error for the trio of a male singer, female singer, and piano was its selection as a male singer with piano duet, again suggesting that the female voice was not reliably perceived.

The more-complex instrumentations of subtest 2 and 3 not only resulted in lower identification scores, but it also impacted upon the appraisal ratings for these stimuli. Both the CI and HA subject groups rated the multi-instrumentations of subtests 2 and 3 to be significantly less pleasant than the single instruments in subtest 1 ($p < 0.001$). This is consistent with reports by other researchers suggesting that simpler stimuli are often easier to identify and more pleasant-sounding for CI subjects (Gfeller et al., 2003; Leal et al., 2003; Schulz & Kerber, 1994). However, research has also indicated that this preference for less acoustically complicated music may be opposite to the general preferences of the NH population. Gfeller et al. (2003) reported that the NH subjects in their study provided significantly higher appraisal ratings for the stimuli they perceived to be more complex than for those stimuli perceived to be simple.

Parallels can be drawn with speech perception where it is well established that speech recognition by both CI and HA users in noisy or multi-talker situations is significantly poorer than for a single speaker in a quiet listening environment (Clark, 2003; Dillon, 2001; Hamzavi et al., 2001; Kong et al., 2005; Moore, 1996; Shannon et al., 2004; Zeng, 2004). For CI users, Shannon et al. (1995, 2004) reported that only four to six spectral channels are required for effective speech perception in optimal listening situations. However, for the perception of more-complex signals including speech in noise or musical sounds, substantially more discrete channels are required than are currently available in present-day implants (Kong et al., 2004; McKay, 2005; Zeng, 2004). For HA users, as discussed previously, the nature of their hearing loss may have reduced the saliency of spectral cues (Arehart, 1994; Moore, 1995; Summers & Leek, 1994).

Whilst the differences between the pre-to-post surgery instrument identification scores for the WL subjects were not statistically significant, the higher appraisal ratings post-surgery was statistically significant (subtest 2: $p = 0.005$; subtest 3: $p = 0.009$). The difference for subtest 1 was nearly significant ($p = 0.059$). This preference for the implant was the case in every subtest for seven of the nine subjects. Furthermore, the mean ratings with the CI were higher than those for the HA for every instrument or ensemble within every subtest. This concurred with the WL subjects' responses on the MLEQ (WL-Post version), where 7 of the 9 subjects indicated that music sounded better with the implant than it did with HAs. These higher ratings post-surgery appear to be attributable to more than just a learning effect. There was a significant difference between the change in the pre-to-post surgery appraisal ratings for the WL group compared with the change in appraisal ratings from the two test blocks for the CI and HA groups. It is again worth mentioning that for the combined scores of the two test blocks for CI and HA subject groups, the CI group also rated the sounds to be more pleasant than the HA group. Although statistical testing did not reveal the difference between these two groups to be significant, possibly due to the large degree of intersubject variability in the ratings given, the trend was observed across all three subtests, and consistent with the findings from the WL subject group.

There may be several explanations for this trend of higher appraisal ratings obtained from listeners with an implant than those with a conventional HA. The CI would have provided additional high-frequency information in comparison to the HA. This extra acoustic information would not have been available to subjects when only using their HAs, and could thereby have served to enhance the perceived sound quality and timbre. For example, for some subjects who had very little residual hearing, the HAs may not have provided any more acoustic information than the bass beats. This would have added little to their perception or appreciation of music. The CI may have enabled them to hear a broader spectrum of the original input signal. Although the lack of change in identification scores suggests that the recipients may not have been able to actually identify the source of the higher-frequency sounds (such as a specific instrument), the additional acoustic information could have improved the subjective sound quality.

Another possible explanation for the higher post-implant ratings provided by the WL subject group may pertain to issues of personal bias and subjective views of the device, such as the ‘halo effect’ (Glossary). It would not be unreasonable to assume that the decision to get an implant would have been associated with a substantial personal investment on the subject’s part. Consequently, they would have expected the CI to be a superior, more technologically advanced auditory device that offered greater potential to improve their hearing. This may have biased their perception of the sound quality. Further, the benefits the implant provided them, such as improved speech perception, may also have served to inflate the ratings provided. Post-implant speech reception scores for each of the WL subjects were significantly higher than their pre-implant scores, with the responses from the MLEQ (WL-Post version) confirming that nearly all of the subjects were more than satisfied with the CI. In spite of this, comments volunteered by subjects suggested that the higher post-surgery appraisal ratings reflected more than just subjective expectations. These comments included that they “got more” of the sound when listening with a CI; when they used HAs, they only heard the beat or bass sounds, but with the implant, they could hear more of the higher melody instruments. Some described it as getting a “broader picture” of the music sounds, with more supplementary detail. The responses from the questionnaire also showed that subjects spent more time listening to music post-implantation than pre-implant ($p = 0.02$), which may have contributed to the higher post-implant ratings for music stimuli. Higher levels of post-CI music listening could be expected to result in higher music appraisal ratings, and vice-versa. For the experienced CI subject group, there was a moderate correlation between their instrument appraisal ratings and the amount of time they spent listening to music with the CI (see section 10.4).

One other finding worth mentioning is that the provision of supplementary non-acoustic contextual information in the second instrumental subtest did not assist subjects with identifying the solo instrument in each excerpt. There was no significant difference for any of the subject groups between the initial run of this subtest and the subsequent run where subjects were told more detail about the background accompaniment. However, this was by no means a systematic or comprehensive evaluation as to the role of, or assistance provided by contextual cues for music listening. For example, it may have been the case that a subject had already assumed that the background accompaniment for all of the stimuli was an orchestra, and hence the additional verbal description did not substantially change their decision making. Further, the task instructions asked subjects to ignore the background accompaniment and to identify the solo instrument in the excerpt. Thus information regarding the nature of the background accompaniment may have been irrelevant to the perceptual process used by the subject for the task.

In summary, there was little difference between users of the CI and the HA on tasks of instrument identification in this study, suggesting that neither device enabled the wearer to adequately perceive the signal's spectral shape and/or temporal envelope. However, the subjects using a CI rated music stimuli to sound more pleasant than subjects using a HA; this difference was statistically significant for the WL subject group's pre-to-post surgery comparisons. Finally, all three subject groups found the multi-instrumental stimuli harder to identify, and less pleasant sounding, than the single-instrument stimuli.

10.1.4 Melody

The use of broad temporal cues in the form of rhythm, or structural features such as lyrics assists with melody recognition and the identification of musical styles (Fujita & Ito, 1999; Gfeller et al., 2000a, 2002a, 2003, 2005; Kong et al., 2004; Leal et al., 2003; Vongpaisal et al., 2004). In Kong et al.'s (2004) study involving closed-set recognition of 12 melodies, the CI subjects averaged 63% when the rhythm cues were left intact, dropping to chance level when the rhythm cues were eliminated. Gfeller et al.'s (2002a) study, also involving closed-set recognition of 12 melodies, reported a mean recognition score of 19% for their CI subjects, with 66% of the correctly identified items having been pre-classified by the authors as being of a 'rhythmic' nature. Fujita & Ito's (1999) research found that the inclusion of lyrics greatly improved both the closed- and open-set recognition scores for their eight CI subjects. The melody recognition test in this study's music test battery did not specifically investigate the use of rhythm or vocal cues, with all of the tunes presented as melody-line only with intact rhythm cues. Even with these rhythm cues, though, the CI subjects only averaged 52% correct, significantly poorer than the HA subjects' 91% ($p < 0.001$).

With melody recognition involving the perception of relative pitch distances (i.e. musical intervals), this disparity may be linked to the issues associated with the poorer pitch perception skills of the CI subjects discussed earlier in section 10.1.2. In Kong et al.'s (2005) study of five CI subjects with aided residual hearing in their contralateral ear, melody recognition with a CI was compared to that with a HA using 12 tunes presented as a single melodic line, devoid of rhythm cues. Reflecting the findings of the current study, superior accuracy was obtained with the HA than the CI. However, in Kong et al.'s (2005) study, the average increase of 17 percentage points when tested with the HA was not shown to be a statistically significant improvement due to the large inter-subject differences, and one subject who recorded a reverse pattern of results. Nevertheless, the provision of F0 information by the HA may have assisted with pitch perception, thereby translating to better melody recognition.

For the WL subjects in this study, contrary to the comparisons between the HA and CI subject groups, there was no significant difference between performance pre-surgery with the HA (mean = 75%) and post-surgery with the CI (mean = 80%). The slight increase in the WL subject's scores post-surgery was not significantly different to the degree of the learning effect observed for the CI and HA groups. The lack of difference for the WL subjects between the pre-surgery and post-surgery results for the melody test is somewhat surprising, considering that two different hearing modalities were utilised. It is possible that this lack of difference may have been in part due to the ceiling effect, with two of the subjects scoring 95% or 100% both pre- and post-surgery. Further, it was also noted that there was one subject who found the melody test particularly difficult, scoring 15% pre-implant and 20% post-implant. If the scores of this outlier are eliminated, the mean recognition scores for the remaining eight WL subjects rise to 83% pre-surgery and 88% post-surgery.

The results from the WL group suggest that post-surgery, their performance was more similar to that of the HA subject group than the CI group. The melody recognition score of the WL group post-surgery (80%) was much higher than the experienced CI subject's group mean (52%). The reason for this is not entirely clear, although several suggestions can be propounded.

Firstly, more-recent CI recipients may have had greater exposure to, and a better ability to hear, these melodies pre-implant. The WL group's level of residual hearing pre-surgery was more similar, on average, to the HA group's hearing thresholds than the CI group's pre-surgery thresholds (see Figures 7.1 – 7.3, Chapter 7). Further, the WL group may have had a shorter duration of significant hearing loss pre-implant than the CI group. With the criteria for implantation expanding to include those with lesser degrees of hearing loss, it would be reasonable to expect that more-recent implantees, including those in the WL subject group, to have greater levels of residual hearing pre-surgery and possibly also a shorter duration of significant hearing loss, than those who received their CI some time ago. Based on the better pre-surgery average hearing thresholds of the WL group, their overall neural survival rate may also have been better than for the CI subject group.

Another potential explanation may be that the newly implanted subjects had a better recollection of the melodies whereas the longer-term users may have forgotten the sound of these melodies over the period of time they have had their CI, thus impeding their performance on this recognition task. The CI group had been implanted for an average of 145.2 months (i.e. over 12 years), whereas the WL group had been implanted for 3 months. In other words, the CI group's recall of the specific features for each melody (both rhythm and pitch) may have become 'blurred' or 'faded' as time passed; they once knew the melody (or knew of it), however over the years they have forgotten how it actually sounds.

A third possible reason for the better performance of the WL group than the CI group on the melody recognition test may be related to different levels of motivation. For the WL group, pre-surgery testing may have encouraged them to be more aware of, and more motivated to listen to music post-CI. The subjects may have recalled the names of the melodies incorporated into the test from the first administration of the test battery, and subsequently proceeded to listen to these melodies with the CI, prior to undertaking the post-surgery music test block. The WL subjects reported spending significantly more time listening to music with the implant, as compared to pre-surgery whilst utilising HAs, based on the MLEQ responses ($p = 0.02$). Gfeller et al. (2000b) reported a trend for newly implanted CI users to record higher levels of music listening and participation than longer-term implantees. The higher levels of post-implant music listening by the WL subjects may have also contributed to their melody recognition test results. It could also be possible that this small group of WL subjects were exceptionally good at melody recognition, recording unusually high scores both pre- and post-implant.

In summary, for the melody test, the HA subject group were significantly better than the CI subject group at recognising familiar tunes ($p < 0.001$), however there was no difference between the pre-to-post surgery results of the WL group.

10.1.5 Summary

Overall, comparisons of the results obtained from the CI and HA subject groups indicate that users of the two devices perform similarly on tasks involving rhythm and timbre perception. However, the CI subject group were significantly poorer on the more pitch-specific tasks (i.e., the pitch test and the melody test). The differences between the two modes for stimulating hearing (i.e., acoustic for the HA as opposed to electric for the CI), along with the properties associated with electrical stimulation of the cochlea may account for much of this disparity.

For the WL subjects, as would be expected, the results obtained were in many ways congruent with the comparisons between the CI and HA subject groups. There were no significant differences between the pre-to-post implant results for the rhythm or instrument identification tests, with higher subjective ratings provided with the implant than the HA. Also analogous to the comparison between the CI and HA subject groups, the WL group scored lower on the pitch-ranking task post-surgery, with a similar pattern of greater accuracy for the male-sung than female-sung vowels. The main divergence from the results of the WL group to the comparisons between the CI and HA subject groups was for the melody test where there was no significant difference between the pre- and post-implant scores for the WL group, as compared to the large disparity between the scores of the subjects in the CI and HA groups.

10.2 General Discussion

It is worth considering that for 8 of the 9 WL subjects, the results from the two test runs (pre- and post-surgery) were obtained from contralateral ears. Pre-surgery, the music tests were presented to the ear with which the subjects obtained better speech perception scores. However, with the clinical protocol relevant to the clinics in this study being to implant the ear associated with poorer speech perception scores (and/or hearing thresholds), it was usually the contralateral ear that was tested post-surgery with the CI. Thus the results do not necessarily reflect the true difference between performance with the HA as compared with a CI as a different baseline result may well have been obtained had initial pre-implant testing been conducted utilising the subject's poorer-hearing ear. This was not a viable option, though, as many of the subjects would have had insufficient residual hearing in that ear to enable them to hear the music stimuli at levels adequate to undertake the test battery requirements. Nonetheless, away from the test situation, subjects could have opted to wear a HA in conjunction with the CI for listening to music. Seven of the nine subjects reported that they wore a HA in the contralateral ear, however this study did not record details of how often they used this HA, in which listening situations they used it, or whether they used it for listening to music. Therefore, post-surgery music listening out of the testing environment, such as at home, may vary depending upon which of three possible listening modalities subjects choose to utilise to listen to music (i.e. HA-only, CI-only, or bimodally with the HA and CI simultaneously). The combination of electric and acoustic stimulation is further discussed later in the chapter.

It is also of interest to compare the results of the HA subject group to those of the WL group pre-surgery, and similarly of the CI subject group to the WL group's post-surgery results. Although the subjects for the HA group were required to meet the audiological criteria for an implant, the comparisons between scores from the HA subjects and the WL subjects pre-surgery showed that the perceptual characteristics of these two groups were not identical. The scores and ratings from the HA subject group were higher than those from the WL group for all of the tests and subtests, except for the rhythm test, with statistical analysis showing a significant difference between the overall performance of the two groups across the music test battery ($p = 0.003$). Such disparity between the two groups, particularly for the quality rating task, is not surprising. For example, there may have been differences in the subjects' overall satisfaction with their respective devices. It would be a reasonable assumption that the WL subjects were not fully satisfied with their HA, hence their decision to pursue cochlear implantation. In contrast, the HA subjects were not actively in the process of obtaining a CI, implying that they were probably more satisfied with the HA and/or found it appropriate for their current needs. It could also be speculated that although the HA subjects met the audiological criteria for a CI, they may have generally been better overall performers with their HAs than the WL group. Should this be the case, then it would follow that the WL subjects may have had greater difficulty perceiving complex acoustic stimuli, as well as appraising sounds to be less pleasant than the HA subjects.

Another factor that needs to be considered is that the HA subject group, as a whole, had better levels of low-frequency residual hearing at 250 Hz and 500 Hz than the WL subject group. As shown in Figures 7.2 and 7.3 in Chapter 7, the HA group's average 250 Hz and 500 Hz thresholds were 50 dB and 60 dB respectively, whereas for the WL group, the respective average thresholds were 70 dB and 75 dB. It should be noted that with the current implantation criteria predominantly focusing upon speech perception ability, and to a lesser extent hearing thresholds between 1 kHz and 4 kHz, significant levels of low-frequency hearing would not necessarily have precluded a patient from receiving a CI. The better thresholds at these lower frequencies probably benefited the HA subject group in their perception of pitch, though, by enabling them better access to F0 information.

After the WL subjects had been implanted, their results demonstrated some similarity to those obtained from the more-experienced CI subject group. There was no significant difference between the two groups' overall performance across the music test battery. This similarity implies that there may be little difference between music perception abilities at 3 months post switch-on of the implant compared with having more than one year's experience with the device. This is in keeping with reports by Gfeller (2001) and Gfeller et al. (2000a, 2002b, 2002c, 2005) that unlike speech perception, incidental exposure to music in everyday life does not appear to significantly improve a recipient's music perception skills. The exception to this trend was the result for the melody test in which the newly implanted subjects scored significantly higher than the more-experienced implantees, as discussed in the preceding section.

One over-riding consideration that should be kept in mind when directly comparing the results of the HA subject group to the WL group pre-surgery, and similarly the CI subject group and the WL group post-surgery, is that the groups consist of two separate sets of individuals. Therefore, for example, one would expect some difference between the mean scores of the HA and WL groups, solely based on the fact that there are different individuals involved. In the same way, despite the results suggesting little difference between music perception skills at 3 months post-implantation and with greater than 1 year's experience with the implant, this is based on the results of two different sets of individuals, as opposed to the same set of subjects being assessed longitudinally, over a period of time.

As alluded to at the start of this section, it should also be considered that for the CI subject group, and for the WL subject group post-surgery, the use of a HA in the contralateral ear may have provided additional benefit for music perception. This listening condition was not assessed in this study, but warrants further consideration. In view of the finding that HAs provide more reliable F0 information than CIs to enhance pitch perception, whilst the CI provides additional high-frequency information, the combination of the two devices may be beneficial for subjects with sufficient residual hearing at the low frequencies (Gantz & Turner, 2003, 2004; Gantz et al., 2005; Gfeller et al., 2004; Kiefer et al., 2005; Kong et al., 2005; Tyler et al., 2002). This was also initially discussed in Chapter 4 when reviewing existing pitch and melody perception studies.

Gfeller et al. (2004) presented findings from their research advocating the potential benefit that combined acoustic and electric stimulation may provide for music perception. Six CI users implanted with a short electrode array were compared to 41 implantees with a standard long electrode array, as well as 22 NH subjects. The short-array used in the study had 6 channels on a 10 mm carrier, implanted using a modified surgical technique. Gantz & Turner (2003) outline more detail of this device. The performance of subjects with the short-array in Gfeller et al.'s (2004) study on tasks of pitch perception and melody recognition was more similar to that of the NH subject group than the implantees with the long electrode array. For example, in a pitch-discrimination task utilising intervals across a three-octave range, both the NH and short-array subjects scored close to 100% in determining the direction of the pitch change for puretones one semitone apart in the lowest-frequency octave tested. The long-array subjects scored approximately 70% for the same condition. At the highest frequencies tested, the short-array users scored approximately 94% for the one-semitone interval, whereas the long-array users' performance improved slightly to 78%. Furthermore, in the familiar-melody recognition task, mean results were 87% for the NH subjects, 84% for the short-array users, and 31% for the long-array users. There was greater disparity between the performance of the implant users with the short-array compared to the long-array in the various pitch-based tasks than between the NH subjects and short-array implant subjects in this study. The authors deduced that the preservation of low-frequency acoustic hearing enabled in this study by the short electrode array resulted in improved pitch discrimination ability for those implant users. It should be pointed out that residual hearing can also be preserved by other means, such as modified surgical techniques.

Kong et al.'s (2005) comparison of acoustically- and electrically-stimulated hearing on a melody recognition task involved subjects with a CI in one ear and a HA in the other ear. Each subject was tested in three conditions – CI-only, HA-only, and bimodally with both the CI and HA. The HA-only condition resulted in scores on average 17 percentage points better than for the CI-only condition, with very similar performance for the HA-alone and combined modality conditions. The use of the HA may have enabled some of the lower-frequency fine-structure cues to be preserved, increasing the potential for the subject to extract F0 information from the signal (Kong et al., 2005). As previously mentioned, the signal-processing techniques implemented in most current CIs prevent the lower harmonics from being fully resolved. Only the temporal envelope information is retained, which is insufficient for accurate pitch perception.

As discussed in Chapter 3, section 3.4, amplitude modulations present in the stimulating pulse train can provide pitch information to an implant user. The rate of the amplitude modulations is derived from temporal envelope information extracted from the output of the speech processor's filterbank. The consequence of this is that only the amplitude envelope information is retained, with the fine-structure information being discarded (Kong et al., 2005). It has been noted by many researchers that this fine-structure information, although not required for speech perception in quiet, is more important for music perception and for listening in more difficult acoustic environments (Kong et al., 2004, 2005; McDermott, 2004; McKay, 2005; Shannon et al., 2004; Smith et al., 2002; Wilson et al., 2003b, 2004; Zeng, 2004). This is in part substantiated by research into combining acoustic and electric stimulation to aid music perception. The acoustic mode of stimulation would allow for the partial preservation of the fine-structure information present in the original stimulus. Theoretically, the fine-structure information should aid the perception of complex sounds; however, how to best achieve this, and how much of this additional detail would be perceived by a CI user, is still a matter of conjecture. Several recently trialled approaches for CI speech-processing strategies are discussed in the next chapter.

A secondary finding resulting from the two runs of the music tests was the learning effect observed with the HA and CI subject groups. For all tests and subtests, with the exception of the rhythm test, scores on the second test block were higher than those on the initial test block (Figures 9.7 and 9.8 in Chapter 9). Statistical analyses showed the degree of this learning effect to be similar for the CI and HA groups with the lack of difference for the rhythm test probably being attributable to the ceiling effect.

The learning effect in this study was task-specific, and related to task-familiarity; this study did not include any form of music rehabilitation or a structured training program. It would be interesting, though, to investigate whether specialised training could benefit music perception. Previously-conducted research has suggested that a music training program could benefit some subjects who may want to improve their music perception skills. Zeng (2004) suggested that with there being an appreciable acclimatisation period for new recipients of CIs, a structured rehabilitation program could help in their adaptation and learning process for the new sensory input. Such a program could extend to non-speech stimuli such as music. The potential benefit of a training program is substantiated by the findings of Gfeller (2001) and Gfeller et al. (2000a, 2002b) describing a computer-based program which was developed and administered to attempt to improve various aspects of music perception including song recognition, timbre recognition, song appraisal, and timbre appraisal. The collective findings across these three publications were that this training program could improve a recipient's ability to recognise melodies and instruments within the test situation. However, the generalisability of these findings to a real-world listening situation using naturalistic stimuli was not determined. Moreover, as alluded to by Gfeller et al. (2000a), the improvement in melody recognition scores as a result of training was conceivably the result of subjects learning to develop compensatory and supplementary strategies for recognising and learning melodies, as opposed to improving their ability to perceive pitch cues and extract the F0 information from the input signal. Although the training program was unlikely to remedy the subject's underlying perceptual mechanisms, it may have assisted in familiarising them with the interpretation of the perceived sound. Gfeller et al. (2002b) reported that, post-training, their subjects demonstrated less diffuse error patterns for the instrument identification task. It is likely that subjects had learnt to recognise some of the spectral and/or temporal features of the signal, and subsequently to attribute those features to certain instrumental families. However, it is worthwhile keeping in mind that an improved ability to identify instruments or instrumental families would not necessarily increase a subject's enjoyment or appreciation of music. As will be discussed in the next section, the findings of the current study only showed a weak correlation for the CI subject group between their ability to identify instruments and the corresponding appraisal rating.

Consistent with survey-based research by Gfeller et al. (2000b), Leal et al. (2003), and Mirza et al. (2003), the CI subject group reported spending significantly less time listening to music whilst using their CI compared with pre-hearing loss ($p = 0.005$). For the HA subject group, the lesser amount of time spent listening to music whilst using their HA as compared to pre-hearing loss levels was nearly significant ($p = 0.051$). For the WL group, although the time spent listening to music with the CI was not significantly different to pre-hearing loss estimations ($p = 0.075$), of interest is the finding that the subjects reported spending significantly more time listening to music with the CI than when compared to estimations made when using HAs ($p = 0.02$). Further, seven of the nine WL subjects subjectively rated music to sound more pleasant with the CI than with their HAs. Gfeller et al. (2000b) reported a trend for more-recent implantees to record higher music listening and participation levels than longer-term implantees, although for the current study, there was no significant difference in reported listening levels between the WL and CI subject groups.

One other subsidiary finding worth mentioning was the significant improvement in speech perception scores achieved by the WL subjects with their CI. Speech perception testing with sentence stimuli conducted at each subject's audiology clinic at approximately 3 months post-implantation showed that 8 of the 9 subjects obtained mean scores above 94%, with the ninth subject scoring 88%. The difference between the sentence recognition scores obtained pre-implant in the best-aided condition and those obtained post-implant with only the CI ranged from 32 to 97 percentage points across the nine subjects. Some of the subjects reached a ceiling effect in their post-surgery speech perception scores when tested with sentence stimuli in quiet listening environments. This improvement in speech perception scores obtained in a clinical setting was in keeping with the subject's subjective ratings of the difference the CI had made to their overall speech perception, when compared to HAs. All of the subjects reported that the CI had improved their speech perception to some degree, with six of the nine subjects stating that it had made their speech perception 'much better'. Dowell et al. (2004) reported in their retrospective analysis of speech perception performance of 92 postlingually deafened adults tested between 3 and 6 months post-surgery, that the mean sentence recognition score was 80%, with a median of 91%. The authors surmised that, based on those results, the average postlingually deafened implant user performs at a level approximately equivalent to an adult with a severe hearing loss for recognising speech in optimal listening conditions. Zeng (2004) published a graph collating published speech perception results obtained by users of the Nucleus, Med-El, and Clarion CI devices. This showed that the most-recent sentence recognition scores for implantees using commercial devices from these three manufacturers was between 80% and 90%. The speech perception results for the nine WL subjects in this current research substantiate Dowell et al.'s (2004) recommendation for the implantation criteria to expand to consider patients whose speech perception scores for sentence stimuli is up to 70% using appropriately fitted HAs. For example, subjects 2 and 6, who obtained pre-surgery speech perception scores of 64% and 67% respectively, both improved to 99% when tested at 3 months post switch-on of the implant, using only their CI.

10.3 CORRELATIONS BETWEEN TESTS

In order to assess whether the perception of pitch or rhythm was associated with the ability to recognise melodies for the subjects in this study, non-parametric Spearman's rho calculations were made. For the CI subject group, their difficulty with pitch perception was correlated with their ability to recognise melodies; there was a significant moderate correlation between the pitch and melody test results (Spearman's rho = 0.679). However, this correlation also suggests that there are other factors in addition to pitch perception that account for the variance in melody recognition scores. These factors may include rhythm perception, recall of the melody, cognitive, or auditory processing issues. There were no significant correlations between the scores on the pitch and melody tests for the HA subject group, nor the WL subject group either pre- or post-implant, possibly due to the ceiling effect for the melody test. For the HA subject group, 19 of the 30 separate melody test scores (15 subjects, 2 runs each) were either 100% or 95%, corresponding to a maximum of one error on the test. For the WL group, pre-implant, 4 of the 9 subjects scored 90% or higher, and post-implant, at least 90% was obtained by 5 of the 9 subjects. As mentioned in section 10.1.4, if the melody test scores of one outlying WL subject are removed, the mean melody recognition score for the eight remaining subjects rises to 83% pre-implant and 88% post-implant. The ceiling effect apparent in the rhythm test would also contribute to the lack of any significant correlations between the rhythm and melody tests for any of the subject groups.

For the instrument identification and appraisal assessments, there was a low significant correlation between the CI subjects' ability to identify the instruments and their corresponding appraisal ratings (Spearman's rho = 0.325). For the HA subjects, this correlation was a little stronger at 0.491. The correlations for the WL subject group were not statistically significant, either pre- or post-implant. The ability to identify an instrument or music group does not necessarily lead to a higher appraisal rating. There are many people who appreciate or even love music without having ever received any form of musical training or instruction. After all, in listening to music, one does not necessarily need to know a song, instrument, or artist in order to enjoy the music. In the same way, knowing these details does not categorically ensure that one will like the music. In the broader sense, the appreciation of music is not solely reliant on musical knowledge or musical training. It is also possible for these correlations to arise from a miscellaneous variable, not directly related to the research, impacting on the two test scores.

Existing studies which calculated correlations between a subject's ability to identify a musical sound and the corresponding appraisal rating have provided mixed findings. Gfeller et al. (1998) found no significant association between identification and appraisal scores for their timbre tasks involving the perception of four musical instruments by 28 CI subjects using the Clarion implant implemented with the CIS speech-processing strategy. However, in a study involving 10 adult Nucleus CI22 or CI24 implant users using the SPEAK speech-processing strategy, McDermott & Looi (2004) found a significant moderate correlation between the identification and appraisal scores for their second experiment investigating the perception of 16 musical instrument sounds ($r^2 = 0.67$).

10.4 CORRELATIONS WITH SUBJECT VARIABLES

In existing research with CI recipients, the only relatively consistent correlations between aspects of music perception such as timbre recognition, pitch perception, or melody recognition, and a variety of subject variables, have been for the factors of age (Gfeller & Lansing, 1992; Gfeller et al., 1997, 2002a, 2005), and post-implant music listening habits (Gfeller et al., 1998, 2000b, 2005; Gfeller & Lansing, 1992). Consistent with this, for the CI subjects in this study, a significant moderate correlation was obtained between the current music listening score (as determined from the MLEQ) and the quality rating scores (Spearman's rho = 0.681), along with a significant moderate negative correlation between age and instrument identification scores (Spearman's rho = -0.525). With respect to the first correlation, as one would intuitively expect, the musical excerpts were rated to sound more pleasant by subjects who reported spending more time listening to music with their implant. There may be several explanations for this. For example, increased time spent listening to music may help the recipient to become more acclimatised to the properties and quality of the sound. This exposure would also help to familiarise them with a range of sound percepts to better enable them to attribute certain acoustic or spectral features to specific musical occurrences or instruments. An alternative explanation may be that subjects who rate music to sound more pleasant are more likely to listen to it than those who do not find music to sound pleasant. It is also possible that as a CI user listens to music more, their appreciation of sounds and/or their expectation of the sound quality may change. It may also be that the correlation between rating and music listening time is mitigated by another extraneous variable.

The negative correlation for the subject factor of age to a range of music assessments has also been reported in several previous articles (Gfeller & Lansing, 1992; Gfeller et al., 2005). This diminished music perception in older adults may be a derivative of age-related physiological, cognitive, and central processing changes. In studies involving CI recipients, Gfeller & Lansing (1992) found a moderate negative correlation between age and the scores on both the rhythm and tonal subtests of the PMMA test, whilst Gfeller et al. (2005) obtained a significant negative correlation between age and the ability of subjects to recognise musical excerpts. Studies of speech perception have reported that cognitive deficiencies, physiological changes in the peripheral and central auditory system, and attention issues may impede auditory perceptual performance to varying degrees (Buchman et al., 1999; Flynn et al., 1998; Jerger et al., 1989; Pasanisi et al., 2003). Moore & Peters (1992) reported that some of their elderly NH subjects performed poorer than the younger NH subjects on pitch discrimination tasks. This was ascribed more to a reduced ability to process temporal information rather than diminished cognitive skills consequent to the ageing process. Bruhn (2002) also discussed differences in the musical experiences for the elderly population, arising from physiological, cognitive, and psychological bases.

It is worth elucidating why the variable of length of profound hearing loss was not recorded in this study. Although numerous studies have analysed this in the past, its validity and accuracy is questionable. For many of the subjects involved in these types of studies, their hearing loss tends to be of a progressive nature and it would be difficult to specify an accurate date as to when a subject's hearing loss reached the severe or profound level. Furthermore, the criteria for a CI have expanded to incorporate a larger range of hearing levels, with the current approach being primarily based upon speech perception scores rather than absolute hearing thresholds. Consequently, potential implantees may only have a moderately severe hearing loss at some frequencies, and may never have hearing thresholds that are at a level classified as 'profound'. For the current study, this variable would have been further confounded by the criterion adopted to define the length of profound hearing loss. In existing research, the length of time specified usually represents the time between the onset of a profound hearing loss and receiving the CI; for most implantees, this would include a period when they were using HAs prior to being implanted. As the current study compared the CI users to HA users, that criterion would not have resulted in a valid comparison between the subject groups.

10.5 LIMITATIONS

There are several factors which may need to be accounted for when interpreting the results of this research. As is the case with many studies in this area, relatively small subject numbers, particularly for the WL group, may have implications for the statistical significance and wider interpretation. For example, had more subjects been involved in the study, some of the observed trends in the findings may have reached statistical significance, particularly considering the large standard deviations found on many of the tasks. However, it is also possible that with the relatively small subject numbers of this study, the significant group differences may have been due to individual variability between the subjects. If this were the case, then some of the statistically significant results reported may not be present if the study was replicated with larger subject numbers. Further, the generalisation of this research's findings should be limited to the Nucleus implant system programmed with either the ACE or SPEAK speech-processing strategies. Different implant systems and strategies vary in their processing of the incoming stimuli, potentially resulting in different stimulation parameters and different sound percepts for the listener.

There was a significant difference between the levels of residual hearing for the HA group and the WL group pre-implant. Closer matching of these two groups on the variable of residual hearing may have enabled more direct group comparisons; the better low-frequency thresholds of the HA group may have assisted their music perception. For example, greater low-frequency hearing may have enabled more F0 information to be accessed. The feasibility of such a matching process may be an issue though, as the availability of sufficient subjects in order to undertake the matching process could be an obstacle for many studies.

The inclusion of the WL subjects in this study eliminated some of the inter-subject variability when comparing music perception with a HA to that with an implant. However, intra-subject confounds mentioned earlier (e.g., the “halo effect” or the learning effect) must be accounted for in the interpretation of the results. It would also have been worthwhile to have conducted a second post-surgery run at the 12 month anniversary, although time limitations prohibited this for the current study. By 12 months, the initial uniqueness and novelty factor should have passed, and subjects would be largely accustomed to the device in regards to both its benefits and limitations. They would also have had more opportunity to listen to music, and may therefore be able to report more reliably on their music preferences and experiences. A second retest at this point would also enable one to assess any further changes in music perception skills.

Time and subject limitations made it largely unfeasible to implement the above-mentioned suggestions in this current study. However, they have been outlined here in order to expand the scope for interpretation and generalisability of the obtained results, should this study be replicated in the future.

CHAPTER 11: SUMMARY & CONCLUSIONS

11.1 SUMMARY

This research investigated the music perception of cochlear implant (CI) users in comparison to hearing aid (HA) users who met the audiological criteria for a CI. The music perception skills of 15 experienced Nucleus CI users were compared to those of 15 experienced HA users by means of a music test battery, which was administered to each subject on two occasions approximately 4 months apart. Further, 9 patients on the waiting list (WL) for a CI were tested pre-surgery and then at 3 months post switch-on of their Nucleus implant. All subjects were postlingually deafened adults. Three hypotheses were generated for this research: firstly, that the experienced CI users would score lower than experienced HA users on the pitch, instrument, and melody tests, but not the rhythm test; secondly, that the WL subjects would score higher on the pitch, instrument, and melody tests when tested with their HA pre-implantation than post-surgery with the CI; and thirdly, that subjects utilising a HA (i.e., both the HA subject group and the WL subject group when tested with their HA pre-implantation) would rate music to sound more pleasant than the subjects utilising a CI (i.e., the CI subject group and the WL subject group when tested post-implantation).

The results of the assessments partially supported the first two hypotheses, but not the third. In relation to the first hypothesis, as expected there was no significant difference between the CI and HA subjects on the rhythm test. The CI group scored significantly lower on the pitch and melody tests ($p < 0.001$), but equivalent to the HA subjects on the instrument recognition tests. The second hypothesis only held true for the pitch test (one-octave subtest: $p = 0.007$; quarter-octave subtest: $p < 0.001$), with no significant differences between the pre- and post-surgery test scores for the rhythm, instrument identification, or melody tests. The third hypothesis was not supported by the findings of this research with the subjects utilising a CI rating the music stimuli to sound more pleasant than the subjects utilising a HA (WL subject group: $p = 0.005$ for subtest 2, and $p = 0.009$ for subtest 3).

Considering that previous research has shown listeners with NH to be significantly better than CI users at pitch-based tasks, with this study finding a significant disparity between CI users and HA users with significant levels of hearing loss, it may be inferred that electrical stimulation of hearing with a CI affects the accurate perception of pitch cues. In the current study, the pitch test results of both the experienced CI subjects and the WL subjects post-surgery were poorer than those achieved by subjects utilising a conventional HA. This disparity was particularly evident for the smallest interval size assessed, the quarter-octave, for which implant users were unable to reliably select which of the two notes was higher in pitch. This equates to not being able to discriminate notes having a frequency difference of approximately 19%. Current speech-processing strategies implemented in the majority of commercially-available CIs do not allow the listener to reliably extract F0 information from the input signal for a multitude of reasons. These may include that there are insufficient pitch cues available to the CI user to enable a reliable pitch percept, these pitch cues provide conflicting information, there is poor spectral definition for the available pitch information, and/or the individual is not able to effectively integrate or use the available information.

Based on these poor pitch perception results, it follows that the melody perception scores for CI subjects may also be affected as accurate perception of western music requires the listener to discriminate between intervals one semitone apart, corresponding to a frequency difference of approximately 6%. This was shown to be the case for the CI subject group, whose scores on the melody test were significantly poorer than those obtained by the HA subject group. The CI subject group were only able to recognise just over half (52%) of the melodies presented, even though these melodies included both pitch and rhythm cues. The significant moderate correlation between the melody and pitch test means for this group indicates that whilst pitch perception was associated with the identification of familiar melodies, there were also other factors that played a contributory role. For the newly implanted WL subjects, there was no difference between melody recognition scores obtained with the HA and the CI.

Whilst the pitch perception scores for subjects using a HA were better than those obtained by the CI users, the performance of the former was still significantly poorer than that achieved by NH listeners, possibly attributable to the broader auditory filters and impaired temporal processing skills associated with significant cochlear hearing loss. It is therefore possible that a portion of the disparity reported in existing research comparing CI users and NH subjects may be related to physiological differences in the cochlea. These physiological differences may be a consequence of the cochlear pathology itself, the resulting loss of hearing, as well as the ageing process. Nadol et al. (1989) and Otte et al. (1978) reported that different cochlear pathologies, as well as different levels of hearing loss, result in different rates of spiral ganglion cell loss. Further, the authors also reported an association between the number of spiral ganglion cells in a cochlea and age. Hence, with current perceptual research comparing CI and NH subjects frequently reporting significant differences between the mean age of the two groups, it is also possible that age-related cochlear changes, along with the pathological-related processes, could account for a portion of the disparity between the pitch perception scores of CI and NH subjects.

The higher pitch perception scores of subjects using a HA did not translate into better instrument identification scores. That is, although the subjects who used a HA were better able to extract F0 information from the complex sounds in the pitch-ranking task than those using an implant, their ability to perceive the information from the signal's spectral envelope was more similar to the ability of the CI subjects. The underlying reasons for the difficulty experienced by all three hearing-impaired subject groups in perceiving the signal's spectral envelope would differ for acoustic and electric hearing; however, the resulting effect on the tasks of instrument identification was similar. Compared to the NH subjects who averaged greater than 95% for all three subtests, for the groups using a HA (i.e., the HA subject group and the WL group pre-implant), their means for the three subtests ranged from 35% to 69%. For subjects using a CI (i.e., the CI subject group and the WL group post-implant), the groups' means ranged from 43% to 65% for the subtests.

Consistent across all of the hearing-impaired subject groups was their diminished ability to identify multi-instrumental stimuli compared to single instruments, with significant differences between the mean scores of the three instrument identification subtests. All groups correctly identified more of the single instruments in the first subtest than the more acoustically-complex stimuli of the other two subtests. This disparity was also apparent in the appraisal scores; all subject groups rated the multi-instrumental stimuli to sound significantly less pleasant than the single-instrument excerpts. However, the ability to recognise instruments did not necessarily portend a higher appraisal score, as demonstrated by the weak correlation for the CI group, and the moderate correlation for the HA group, between the results of the identification and appraisal tasks. This suggests that an inability to identify an instrument or ensemble does not necessarily preclude a listener from enjoying its sound.

One surprising finding to arise from the appraisal results, and contrary to the initial hypothesis, was the trend for users of the CI to rate music as more pleasant-sounding than users of the HA. This was unexpected considering that the HA users were better at musical pitch perception tasks than the CI users. Furthermore, in existing research, the electrically stimulated music percepts for CI users have been consistently rated to sound significantly less pleasant than the acoustically stimulated percepts for NH subjects, with CI subjects generally providing largely negative responses on qualitative assessments of sound quality for music. For the WL subjects post-surgery, the appraisal ratings provided were significantly higher than those obtained pre-surgery. For the CI and HA subject groups' comparison, there was a trend for the CI group to provide higher appraisals across all three subtests, although these were not statistically significant. It may be that the CI provided its users with better access to the higher-frequency components of the signal; these additional high-frequency sounds may have resulted in a better sound quality than that obtained with only HA amplification for those with a significant hearing loss.

In addition to the comparisons between the CI and HA group, along with the pre-to-post surgery scores for the WL group, this study also compared groups using the same hearing device (i.e., the WL group pre-surgery compared with the HA subject group, and the WL group post-surgery compared with the CI group). Of interest in these comparisons was the general similarity between the scores from the CI group and the WL group post-surgery on most of the assessments, with the exception of the melody test. With the CI group having used their implant for more than one year, whilst the WL group were tested at 3 months post switch-on of the device, this correspondence between the two groups' scores indicates that incidental exposure to music stimuli in everyday life, after the initial 3 months with the CI, does not necessarily improve music perception abilities.

One overriding feature of this study, and many similar studies, was the large degree of inter-subject variability for all hearing-impaired subject groups. Large score ranges and standard deviations were observed for all of the music tests, except the rhythm test. Individual differences in environmental, physiological, sociological, cognitive, and auditory processing factors would have contributed to this variability.

A supplementary finding to arise from the pre-to-post implant comparisons for the WL subjects in the current study was the significantly better speech perception scores obtained by subjects once implanted. For the nine WL subjects, pre-implant speech perception scores using sentence stimuli ranged from 3% to 67%, with a mean of 40%. Once they had received their CI, the subjects' scores increased to between 88% and 100% (mean = 97%). Such results indicate that, for these patients, the CI is fulfilling its primary function of improving speech perception in quiet listening environments.

In summary, the main findings of this research were that the subjects with a moderately-severe to profound hearing loss using a CI scored significantly lower than those using a HA at perceiving pitch, and also to a large extent, melody recognition. There was no significant difference between the ability of the CI or HA users in this study to discriminate rhythm patterns, or to recognise musical instruments and ensembles. However, there was a trend for those with a CI to rate musical excerpts to sound more pleasant than those with a HA. This difference reached statistical significance for the WL subject group.

The results of this current study are not only consistent with existing research with respect to the overall performance of CI users on music perception tasks, but they also indicate that HA users with similar levels of hearing loss perform at least equal to, if not better than, CI users on these music perception tests. However, despite the differences between scores obtained by subjects using a CI compared to those using a HA, the participants in this study, all of whom had significant levels of hearing loss, were largely unable to achieve satisfactory or effective music perception, regardless of the device they used. These findings support the need for ongoing research into music perception of people with hearing impairments.

11.2 CURRENT DEVELOPMENTS AND FUTURE RESEARCH

In the process of providing an alternative perspective into the music perception abilities of CI users by comparing them to HA users, this research has given rise to a number of questions and issues requiring further research which are briefly outlined in this section. Firstly, additional tests and test procedures are suggested in section 11.2.1 which could provide further information on the ability of subjects with significant hearing impairments to perceive music. Secondly, as the results of this research demonstrate that strategies implemented to process sound in current-day implants negatively affect the perception of music, and particularly the perception of pitch, suggestions for modifications to existing strategies and/or new strategies currently attracting research interest are outlined in section 11.2.2.

11.2.1 Additional Tests and Test Procedures

Whilst the HA users in this study who had significant levels of hearing loss performed better than the CI users on some tasks, their results suggest that they may not achieve optimal music perception either. Music perception for many of these patients would be hindered by physiological changes to the auditory system such as a lack of surviving auditory neurons, and broader cochlear auditory filters; therefore, it is unlikely that perfect music perception would be a realistic goal for the majority of these HA users. However, modifications to the HA's amplification parameters to minimise any distortion, along with ensuring that the frequency response of the aid is appropriate for the individual, may provide some benefit for listening to music. This study did not investigate the effect of different amplification parameters on music perception, and further testing and research is therefore warranted to investigate if any specific modifications to the settings or frequency responses of HAs provide consistently better music perception for HA users. In view of the wide range of acoustic features of different styles and genres of music, it may also be possible that slight variations to the device parameters for different music styles may help to optimise the wearer's music listening experience.

In order to obtain a better understanding of CI recipients' ability to perceive the various elements of music, along with the cues they utilise for such tasks, additional specialised testing may be warranted. For example, to assess the role played by gross temporal or rhythmic cues, two further assessments could be considered. The first would be to undertake the melody test using the same melodies devoid of rhythm cues; in this case, listeners would be solely reliant on the pitch cues to identify familiar tunes. The results of these two conditions (i.e. with and without rhythm cues) could then be compared to assess whether the addition of rhythm cues aids melody recognition, and the extent of any benefit obtained. There are current studies investigating the recognition of melodies without rhythm cues, however few have directly compared performance with and without the presence of rhythm cues. Another more-novel approach may be to test subjects' discrimination of instruments playing in different articulatory styles. Examples of this would be the violin being played with the bow ('*arco*') as opposed to plucking the strings ('*pizzicato*'), the guitar being strummed as opposed to playing single notes, or even a melody being played smoothly ('*legato*') versus with short, crisp notes ('*staccato*') on an instrument. If the subject concludes that the two articulatory styles are two different instruments, it may indicate a strong reliance on the broad rhythmic or slow-modulating temporal cues present in the stimuli, including the attack or decay times, for instrument identification. On the other hand, if a subject can reliably differentiate that the two styles are being played by the one instrument, the use of spectral cues may be playing a larger role for these subjects.

Consistent with existing research, this study found that CI users are largely unable to perceive pitch accurately. The results of the pitch test suggest that CI recipients using the same speech-processing parameters as the subjects in this study may be unable to reliably perceive the higher of two consecutively presented sung vowels, a quarter of an octave apart. Whether this finding would generalise to other musical stimuli, such as notes played by a different instrument, notes in a different F0 range, or sounds sung by a different singer, is a matter for further investigation. There are a host of other tests and investigations that could be conducted to add to the current knowledge-base on the music perception abilities of people with hearing impairments. The suggestions made in this section are only a sample of these, based on the results of this research study.

11.2.2 Implant Technology

As mentioned in Chapter 3, with the fixed-rate pulsatile stimulation strategies usually implemented in current-day CIs, the implant user may need to rely on the amplitude modulations present in the pulse trains to provide pitch information. For these strategies, the rate of the amplitude modulations is determined by temporal envelope information extracted from the output of the filterbank. This results in the fine-structure information being discarded. Recent research suggests, though, that this fine-structure information may be necessary for music perception and for listening in more-complex acoustic environments (Kong et al., 2004, 2005; McDermott, 2004; McKay, 2005; Shannon et al., 2004; Smith et al., 2002; Wilson et al., 2003a, 2004a; Zeng, 2004). Several approaches have been proposed for how best to provide more fine-structure information from the input signal for a CI user. These include using analogue stimulation, developing strategies that more closely replicate the propagation properties of the travelling wave along the basilar membrane, or changing the filtering properties of the envelope detector currently implemented in existing filterbanks (McDermott, 2004; Schatzer et al., 2003; Wilson, 2004; Wilson et al., 2003b, 2004a). Research by Oxenham et al. (2004) sugggests, though, that such temporal fine-structure information would need to be tonotopically mapped to the correct location in the cochlea in order to be of any benefit to the CI user. The use of electro-acoustic stimulation, as discussed in Chapter 10, section 10.2, has also been recommended as a potentially effective way of improving F0 perception for certain CI users who have sufficient levels of residual hearing. The use of acoustic hearing enables some of the lower-frequency fine-structure cues to be perceived, thereby increasing the potential for F0 information to be extracted.

With electrical stimulation resulting in highly synchronous firings of auditory nerve fibres, another suggestion has been to use algorithms that stimulate at very high rates to attempt to restore the stochastic firing properties of neural elements, as occurs for NH listeners. These high rates may then help to desynchronise nerve firings in a population of auditory fibres, potentially enabling shorter time intervals to be coded (Dorman et al., 2002). This was overviewed in Chapter 3, section 3.3.3.3. It should be kept in mind, though, that even if one, or several, of these initiatives prove to be successful, it is still unclear as to whether the listener could actually perceive or use the extra information, and the extent of any benefit they may subsequently derive.

In addition to providing more fine-structure information, a range of other novel sound-processing strategies have been trialled in order to try and improve the perception of pitch and timbre through the CI. Different schools of thought exist as to what might be the most effective approach. In addition to the suggestions just mentioned, a few of the other suggested approaches include representing the F0 by changing the stimulation rate, enhancing the amplitude modulations of the stimulating pulse trains, eliminating the phase shifts that occur when information is combined across electrode positions, increasing the number of discrete stimulation sites in the cochlea, and using higher carrier and/or sampling rates (Geurts & Wouters, 2004; Kessler, 1999; Laneau et al., 2006; Loizou et al., 2003; McDermott, 2004; Nie et al., 2005; Vandali et al., 2005; Wilson, 2004; Wilson et al., 2003a; Zeng, 2004).

Vandali et al. (2005) examined four experimental strategies, in comparison to the more-conventional ACE and CIS strategies, for their performance on a pitch-ranking task using half-octave sung-vowel stimuli; the stimuli were similar to that used in the research discussed in this thesis. Each of the experimental strategies varied in their method of encoding temporal and/or spectral information. One of these strategies coded more of the fine-temporal information whilst the other three adopted deeper F0 modulation depths where the peaks were aligned to occur coincidentally in time across all activated electrodes. Results indicated that the three strategies adopting deeper F0 modulation depths provided significantly better pitch-ranking scores for the male-sung vowel stimuli than the conventional ACE strategy. This was attributed to the three experimental strategies providing more salient cues related to F0 periodicity, as well as minimising interactions between temporal information presented to closely-spaced electrodes.

Laneau et al. (2006) designed and implemented a new sound-processing scheme (*F0mod*) where the envelope information from the output of the filterbank was sinusoidally modulated at the F0 of the input signal. The processor involved two parallel sound-processing blocks – a filterbank, and a F0 estimator. The filterbank used a 512-point Fast Fourier transform (FFT) to analyse the input signal, with the extracted envelope information being split into 22 channels. The bandwidth of each frequency bin was limited to 125 Hz, thus filtering out temporal-pitch cue information. A F0 estimator was used in conjunction with the filterbank; the estimated F0 was limited to between 75 Hz and 593 Hz. Maximal modulation depth was used in all channels with modulations being in-phase across the 22 channels. In comparison to the F0mod strategy, the conventional ACE strategy, as used in this current study, has a 128-point FFT to analyse the input signal, with no F0 estimator as the output of the filterbank is not explicitly modulated at the F0. The 512-point FFT of the F0mod scheme should theoretically have provided better resolution of the lower frequencies. The new processing scheme was compared to the conventional ACE strategy in Laneau et al.'s (2006) study, and results indicated that music perception was better with the new scheme. Fundamental frequency discrimination was reported to be three times better for the F0mod than the ACE strategy when the F0 was less than 250 Hz. Subjects also recognised more melodies with the new scheme. The authors reported that the new scheme was beneficial for low-pitch stimuli, with no clear difference between the two schemes with the highest-F0 stimuli tested, 370 Hz. Overall their results suggested that the explicit modulation of envelope information at the F0, in conjunction with maximal modulation depths and in-phase modulations across all channels as used by the F0mod strategy, enabled subjects to better access and utilise temporal-pitch cues. The perception of place-pitch cues was not changed, however.

Kasturi & Loizou (2005a) hypothesised that alterations to the filterbank of speech processors may be beneficial. They undertook a study involving NH subjects recognising familiar melodies without rhythm cues. The use of logarithmically-spaced filters was compared to four different semitone-spaced filters, each with a different number of channels. The results suggested that the semitone spacing of filters was preferential to logarithmically-spaced filters as fewer channels were required to obtain similar recognition scores. The results with 4 channels based on semitone filter spacing were equivalent to those obtained with 12 channels which had logarithmic filter spacing. Theoretically finer filter spacing could improve pitch perception as it should provide more pitch information; however, this may not actually occur at a perceptual level for CI users. It would be conditional on the increased number of filters providing an increased number of discrete stimulation sites. This presumption may not necessarily be the case, though, dependent upon the spread of electrical stimulation in the cochlea and the characteristics of the surviving auditory neurons.

In another line of investigation, the use of ‘virtual channels’ has also been proposed as a means to improve spectral resolution for CI users by increasing the number of discriminable place-pitch steps perceived by the implantee (Donaldson et al., 2005; Kasturi & Loizou, 2005b; Poroy & Loizou, 2001; Zeng, 2004). Research has shown that one or more intermediate pitch percepts can be produced by the weighted stimulation of two adjacent electrodes using either interleaved or simultaneous pulses (Donaldson et al., 2005; Geurts & Wouters, 2004; McDermott & McKay, 1994; Poroy & Loizou, 2001; Schatzer et al., 2003; Wilson et al., 2003b). In an early study, McDermott & McKay (1994) found that the interleaving of pulses between two closely-spaced electrodes imparted an intermediate pitch percept. More recently, Donaldson et al. (2005) reported that simultaneous dual-electrode stimulation increased the number of place-pitch steps between two- to nine-fold relative to single-electrode stimulation for their CI subjects. Four of the five subjects in Poroy & Loizou's (2001) study were able to detect five different pitches when two fixed electrodes were simultaneously stimulated in five different conditions. Each condition varied the degree of weighting applied between the relative amplitude of pulses sent to each electrode. The subjects stated that the task was easier for apical electrode pairs than basal pairs. Kasturi & Loizou (2005b) reported that the use of virtual channels in the mid-frequency region significantly improved word recognition scores, due to better representation of the second formant. These findings could be exploited to provide finer frequency resolution with existing implant systems and electrode arrays. Research continues into both the theoretical foundation and practical application of dual-electrode stimulation with initial findings suggesting that it may offer some potential for improving pitch perception amongst CI users.

To review, a range of approaches to improve speech-processing strategies are currently attracting research interest. Some of these have focused on trying to provide more temporal information; for example, by providing more fine-structure information, using very high stimulation rates, or enhancing the F0 modulation depths. Others have aimed to improve spectral resolution, such as by modifying the filterbank of the speech processor, or using ‘virtual channels’ to give finer frequency resolution. Due to the complex way that speech-processing strategies impact upon the perception of music and other similarly complex auditory sounds, there are a host of other experimental strategies and approaches that have not been addressed in this thesis. It is not yet clear as how best to improve music perception with electrically stimulated hearing via a CI. With no one particular strategy or approach having yet been shown to be consistently better than others, debate continues as to which direction research should focus upon.

The results and findings of the research described in this thesis may be pertinent as part of the decision-making process for patients considering a CI, as well as for device manufacturers and designers as they continue to develop new technology. With the expeditious growth and continual evolution of CIs, the associated technological innovations, and the success they have had in improving the speech perception abilities of most implant recipients over the last two to three decades, consumers, researchers, and manufacturers alike have been expanding their interest into enhancing perception of other acoustic stimuli, such as music. With the ageing global population, the prevalence rates for hearing impairment will continue to rise. Accordingly, the number of people using a HA and/or CI will also increase. In order to make further improvements to these hearing devices, a better understanding of factors limiting the perception of complex auditory stimuli is required. Such improvements could then benefit overall quality of life for the user and their family. It is hoped that this study may contribute to the ongoing quest to enhance the music listening experience for those with a significant hearing impairment who use a hearing aid and/or a cochlear implant.

GLOSSARY

Adaptive Dynamic Range Optimisation (ADRO)	Available in the SPrint processor of the Nucleus cochlear implant. A digital signal processing algorithm where the gain in each channel is independently adjusted. It aims to improve the audibility of low-level sounds whilst still keeping loud sounds at a comfortable volume. By making the softer components of speech more accessible, this may then improve speech perception, particularly in the presence of background noise.
Amplitude conversion function	A function in cochlear implants to convert acoustic sound levels to levels appropriate for electrical stimulation. Subsequent to spectral analysis, the amplitudes of the selected filter outputs are converted into levels appropriate for electric stimulation suitable for each electrode independently (i.e., between the 'T' and 'C' level set for that electrode). This conversion function is non-linear, with zero attack and release times.
Analogue Stimulation	A continuously varying current is used to present details of the waveform, as opposed to presenting only the signal's envelope information via pulsatile stimulation. It was more commonly used in now-obsolete single-channel devices as opposed to current-day multiple-channel implants.
Automatic Gain Control	A sound-processing technique where the amount of gain provided is dependent upon the level of the input signal. It aims to ensure that variations in the input signal are kept within the user's dynamic range, reducing the amount of gain provided as the input level to the amplifier increases. There are several methods of how this may be achieved; all operate via a feedback loop, examining the peaks in the signal envelope, but may differ in both the range of input levels and speed at which it acts. The amplifier in the device (hearing aid, or cochlear implant) automatically reduces the amount of gain provided as signal levels increase. Unlike automatic sensitivity control, automatic gain control is primarily for short-term level changes occurring in the listening environment.

Automatic Sensitivity Control	A sound-processing technique in cochlear implants to control the sensitivity in relation to background noise. It adjusts the amount of electrical gain applied to sounds received by the microphone, prior to analysis by the speech processor. Aiming to prevent background noise from getting too loud, it examines the troughs in the signal envelope. In contrast to automatic gain control, this sensitivity control is for long-term level changes as the wearer moves from one environment to another.
Biphasic Pulses	Charge-balanced pulses delivering a constant current for electrical stimulation in multiple-channel cochlear implants. The use of these pulses avoids the presentation of a direct current which damages surrounding tissue.
Bipolar Stimulation	Current passes between two active electrodes located proximal to the nerve fibres within the cochlea. Results in more localised stimulation and reduced current spread than monopolar stimulation. In the Nucleus implants, the number of electrodes separating these electrodes determines the mode of bipolar stimulation. Increasing the spatial separation between the two active electrodes (instead of using adjacent electrodes) can reduce the current required for stimulation, but conversely reduces spatial selectivity.
Channel	A frequency region, usually created by the use of filters; the information in these resulting channel(s) can then be subjected to further sound processing. In cochlear implants, the term is also used to define the electrode pair used for stimulation.
C-Level	Comfortable Level. The maximum stimulation level that does not result in an uncomfortable perception of loudness. This level sets the maximum stimulation allowed for each electrode.
Common Ground Stimulation	One intracochlear electrode acts as the active electrode; the selection of the actual electrode varies in time. Current passes from this electrode through all (or most) of the other intracochlear electrodes on its return path.

Compression
A sound-processing technique where sound levels are restricted to a dynamic range appropriate for the user; less amplification is provided for greater sound intensities. It is applied to the input signal, prior to further processing. It should be noted, though, that for cochlear implant users, the conversion of these acoustic inputs into levels appropriate for electrical stimulation of selected electrodes occurs via a non-linear amplitude conversion function subsequent to the sound-processing stage.

Compression Ratio
The degree of compression. For hearing aids, this is the ratio of the change in input sound pressure level to the change in the output sound pressure level, for a certain input level. For cochlear implants the output level of the processor is then further processed to produce the electrical signal used for stimulation.

Conductive Hearing Loss
Hearing loss arising from an impairment to the external or middle ear. This affects the transmission of sound from the external auditory canal to the inner ear; the inner ear is not affected. This type of hearing loss can be potentially treated with medical intervention.

***Contour* Array**
A perimodiolar electrode array designed to curl around and stimulate close to the modiolus. This potentially results in lower stimulus thresholds and more localised stimulation of residual nerve fibres which, in turn, potentially increases battery life and improves spatial specificity. The Contour array is available with the Nucleus CI24R implant. Similar arrays based on this principle are available with other implants. The Nucleus array consists of 22 half-band platinum electrodes designed to be inserted approximately 25 mm into the cochlea. Insertion of the array occurs via a stylet which, once removed, enables the array to curve around the modiolus.

Directional Microphone	Microphones serve as input transducers, converting acoustic into electric signals. Directional microphones are more sensitive to sound from one direction than another. This is achieved by having a forward and a rear port to deliver sounds to those respective parts of the microphone's diaphragm. Sounds from the rear port are delayed or cancelled out using an acoustic damper or resistor. The ratio between the resulting internal time delay this creates to the external time delay (time taken for external sounds to get from one inlet port to the other) determines the directionality of the microphone. Current Nucleus implants use pressure gradient directional microphones. These have cardioid directivity shape; in the freefield, they are most sensitive to sounds from the front, but worn on the ear, they are most sensitive at 45 degrees from the front. The microphones also have an inherent pre-emphasis exploited to boost the weaker higher frequency components of speech.
Feature Extraction Strategy	A class of older, now-obsolete speech-processing strategies predominantly used in Nucleus implants prior to 1994. They were based on the principle of identifying prominent features of speech, such as the fundamental frequency and the lower formants. These were then encoded as patterns of electrical stimulation. The culmination of these strategies was called MPEAK, extracting higher frequency information from three frequency bands (2-2.8 kHz; 2.8-4 kHz; above 4 kHz) in addition to the fundamental frequency and first two formants.
Filterbank	A collection of bandpass filters which divide the incoming signal into a number of frequency bands prior to further processing.
Filterbank Strategy	A class of speech-processing strategies that superseded the feature extraction strategies. These utilise a filterbank (a collection of bandpass filters) to divide up the incoming sound into 'n' frequency bands with each band corresponding to one channel of the implant. The envelope of the waveform is estimated at the output of these filters with the amplitudes of selected bands then being used for electrical stimulation. In the Nucleus strategies, the frequency bands are logarithmically spaced for frequencies above 1000 Hz, and linearly spaced below this.

Formant Areas of high energy in the speech spectrum (spectral peaks). These incorporate a band of harmonics and are critical for vowel perception. The formant frequencies are largely independent of the fundamental frequency.

Fundamental Frequency The repetition rate of the waveform. It is the lowest frequency component of the waveform and also called the first harmonic. It is often, although not necessarily, the primary determinant of the pitch of periodic sound.

Halo Effect An experimental effect where the novelty of a particular intervention or treatment may increase a subject's expectation of the task or its possible outcome, change their level of alertness in performing a task, or modify their perceptions in undertaking the task.

Harmonic Higher frequency components of a periodic sound (i.e., a sound with regular vibration pattern) which are integer multiples of the fundamental frequency. In those with normal hearing, the harmonic structure affects timbre perception with the first 5-6 harmonics contributing most to pitch perception. Also called partials or overtones.

Jitter A speech-processing option available in current Nucleus strategies where small, random variations in the rate of stimulation are made around a central frequency in order that the stimulation is not exactly periodic. It is thought that periodic stimulation at rates below 500 Hz may result in the perception of a tone or buzz at a frequency equal to the stimulation rate due to 'rate-pitch' stimulation. This unwanted stimulation could then conflict with 'place-pitch' information. Thus jitter aims to eliminate the potential of such fixed rate-pitch percepts when stimulating at rates below the pitch saturation limit. In the Nucleus clinical software, jitter is expressed as a percentage of the stimulation period.

Knee point The threshold or level at which compression begins. Its exact value is the point when the resulting output differs by 2 dB from the level which would have occurred should linear amplification have continued.

MAP	A set of measurements that define the parameters for electrical stimulation of the implant. It determines which electrodes are to be stimulated, and controls the amount of electrical current transferred to the cochlea via these electrodes. These measurements ensure that stimulation occurs between the set T- and C- levels in order for it to be comfortable and audible to the listener. The MAP is programmed into the speech processor, and can be changed or modified as often as required.
Mapping	The clinical procedure undertaken to determine a patient's MAP.
Maxima	The filterbank outputs with the greatest amplitudes at any moment in time. In some of the current Nucleus speech-processing strategies, the outputs of the filterbanks are scanned with only the amplitudes of a limited number of these maxima being converted to appropriate electrical stimulation levels.
Microphone Sensitivity	This control sets the threshold for automatic gain compression, thereby controlling the amount of gain applied to the input signal. The setting determines the input intensity level that will result in the maximal output signal used to stimulate at 'C' levels along with the minium signal level for 'T' level stimulation. Increasing sensitivity will make soft sounds more audible, but similarly makes background noise louder.
Monopolar Stimulation	The reference or ground electrode is situated remote to the active electrode. Results in a larger current spread than bipolar stimulation. In the Nucleus CI24 implants, two ground electrodes are available. One consists of a ball electrode placed under the temporalis muscle with the second being a plate electrode situated on the implanted receiver-stimulator package. Current can flow between either one or both of these electrodes and the active intracochlear electrode to provide electrical stimulation. This stimulation mode is not available in the CI22M where no extracochlear electrode is available.

Multi-Channel Cochlear Implant — Cochlear implants involving the use of multiple active electrodes where different processed information is delivered to each stimulating electrode. The use of multiple electrodes allows different neural populations to be stimulated relatively independently to exploit the tonotopicity of the cochlea.

Multi-Channel Hearing Aid — A hearing aid with two or more electrical circuit paths. The incoming acoustic signal is divided up into a number of frequency bands; modifications can then be individually made for each of these frequency bands before the signals are re-combined and delivered to the wearer.

***n-of-m* strategy** — The ‘n’ largest envelope signals from ‘m’ bandpass channels for each cycle of stimulation are presented to the selected electrodes.

Octave — Doubling of frequency (for example, 880 Hz is one-octave higher than 440 Hz). In a musical context, an octave consists of 12 consecutive steps (or semitones) using the ‘equal temperament’ or chromatic scale commonly used in western music.

Omni-directional Microphone — A microphone with a single port delivering sound to the front of the microphone diaphragm. Sounds are picked up equally from all directions in a free field situation. When worn at the ear-level, some of this omni-directionality is lost at mid to high frequencies, with greatest sensitivity at approximately 60 to 100 degrees azimuth. Also called pressure microphones.

Peak Clipping — When the input sound rises above a predetermined level, the resulting output waveform will be clipped. This can result in the distortion of sounds above the predetermined maximum level.

Peak Rounding — Non-linear amplification. A negative feedback loop is used to gradually diminish the output level as the input level increases. That is, amplification is at a ratio of less than 1:1. The amount of negative feedback can be adjusted to reduce the gain and output power. Like hard peak clipping, distortion is also created, however by commencing at a lower level, its gradual onset results in less severe distortion.

Percutaneous Link	A direct connection of the electrode wires through skin, via a plug and socket. Was only commercially available in the Symbion cochlear implant, although it has been used in some research-orientated implants.
Periodicity Pitch	One of the two main theories of pitch perception (in addition to place pitch). This theory is based on the timing patterns of neural impulses which are phase locked to the stimulus frequency; the detection of the arrival times of the resulting action potentials underlies this theory of pitch perception. However this theory is only for lower frequencies as a saturation limit is reached at the higher frequencies. Phase locking is strongest at the low to mid frequencies, and does not occur above 5 kHz. That is, synchrony of nerve impulses to the phase of the stimulating sound disappears around 4 kHz to 5 kHz. The degree of phase locking significantly decreases from the mid to high frequency range, and it is generally thought that temporal (or rate) coding is dominant for frequencies up to approximately 2500 Hz. Above this, pitch perception is predominantly by mechanisms relevant to the place coding theory.
Place Pitch	One of the two main theories of pitch perception (in addition to periodicity pitch). It is based on the finding that different frequencies excite different places on the basilar membrane relating to the tonotopic organisation of the cochlea. It can particularly help with the recognition of vowels by providing formant frequency information, and was critical in the development of multiple channel cochlear implants.
Pulsatile Stimulation	The use of constant current pulses to stimulate selected electrodes, as opposed to delivering a continuously varying analogue waveform (analogue stimulation). Most commonly, biphasic charge-balanced rectangular pulses are used in a sequential manner, however the delivery of non-rectangular waveforms is possible by varying the specified current for each electrode over a certain time-interval.

Sampling Rate The number of times per second the input signal is sampled. In order for analogue-to-digital conversions to be made, the input signal needs to be sampled (where the signal's size value is measured) at specified moments in time. Research has shown that this rate needs to be greater than twice the highest frequency component of the signal in order not to lose information. Most cochlear implant processors use rates of above 10 kHz; the highest frequencies important for speech perception are in the vicinity of 4 kHz. Although higher sampling rates would provide more information, this would require greater power consumption.

Semitone In western music using the common 'equal temperament' or chromatic scale, a semitone divides up the octave up into twelve equal steps and is representative of adjacent notes on a keyboard. It corresponds to approximately a 6% change in frequency, or a ratio of 1:1.059.

Sensorineural Hearing Loss Hearing loss arising from an impairment at or beyond the level of the inner ear. 'Sensory' refers to damage to the sensory organ or cochlear hair cells, with 'neural' referring to damage to the auditory nerve or higher structures. Unlike a conductive hearing loss, this type of hearing loss is rarely reversible and can also lead to distortion in the sound quality.

Sequential Stimulation The presentation of pulsatile stimuli to one electrode at a time, in succession, resulting in only one electrode being stimulated at a time. This provides the advantage that a current pulse from one electrode does not interfere with a pulse from another. Most commonly, biphasic rectangular charge-balanced pulses are used. Also called interleaved or non-simultaneous stimulation.

Simultaneous Stimulation Two or more electrodes are stimulated simultaneously. However, such concurrent stimulation potentially gives rise to unpredictable side-effects arising from current summation in the cochlea which leads to channel interaction. This has been reported to result in unpredictable loudness variations, and the reduced salience of channel-based cues for some patients.

Single-Channel Cochlear Implant Cochlear implants where the same information is delivered to each stimulating electrode(s). Generally only one active electrode is used.

Single-Channel Hearing Aid	A hearing aid with a single electrical circuit path. Any sound-processing parameter is applied to the entire frequency range.
Stimulation Cycle	The time frame in which a group of pulses are sent to the associated channels for sequential stimulation in cochlear implants.
Stimulation Rate	The rate at which the electrodes are stimulated. Used in association with pulsatile stimulation.
Straight Array	An intracochlear electrode array used with the Nucleus CI22 and CI24K implants. Variants of this array are available with all currently manufactured implants. The straight array is designed to be inserted into the scala tympani with differing numbers of electrodes to stimulate residual nerve fibres. The intracochlear component of the Nucleus CI24K version consists of 22 full-banded platinum electrodes designed to be inserted approximately 25 mm into the cochlea.
Telecoil	A coil of wire wound around a core of highly permeable metal. This enables the use of magnetic induction as an alternative method of providing acoustic input. Magnetic fields from telephones or induction loop systems instigate a voltage in the telecoil which is then amplified.
T-Level	The level where the patient first identifies the presence of sound sensations resulting from electrical stimulation. It is the lowest level where the sound is detected each time it is presented.
Tonotopic Organisation	The travelling wave excites the neural populations at the basal end of the basilar membrane near the oval window for high frequency stimuli, with lower frequency sounds exciting more-apical neurons.
Transcutaneous Link	Information from the sound processor along with power to run internal components is sent from an external transmitting coil to the internal receiver-stimulator package through intact skin. This is essentially achieved with electromagnetic induction via radio frequency carrier waves. All currently implanted devices utilise this type of transmission link.

***Whisper* Setting** A fast-acting compression circuit available in the Nucleus Esprit 3G speech processor. It is located prior to the automatic gain control system and aims to improve the listener's access to soft or distant sounds. It uses a kneepoint of 52 dBSPL, compression ratio of 2:1, attack time of 5 ms and release time of 100 ms. The Instantaneous Input Dynamic Range (i.e., the short term intensity range of the input signal coded by the speech processor) is increased from 30 dB for the usual microphone settings to around 40 dB with the Whisper setting with the extra 10 dB given to low level signals.

REFERENCES

Advanced Bionics. (2004). *Implant safety and reliability report*. California: Advanced Bionics Corporation.

Arehart, K. H. (1994). Effects of harmonic content on complex-tone fundamental-frequency discrimination in hearing-impaired listeners. *Journal of the Acoustical Society of America, 95*(6), 3574-3585.

Battmer, R. D., Gnadeberg, D., von Wallenberg, E., Lehnhardt, E., & Allum, D. J. (1993). A study of monopolar and bipolar stimulation modes with a modified Nucleus Mini-22 cochlear implant. *Advances in Oto Rhino Laryngology, 48*, 9-16.

Beal, A. L. (1985). The skill of recognizing musical structures. *Memory & Cognition, 13*(5), 405-412.

Bess, F. H., & Humes, L. E. (1995). *Audiology: The Fundamentals* (2nd ed.). Baltimore: Williams & Wilkins.

Blamey, P. J., Dowell, R. C., Brown, A. M., Clark, G. M., & Seligman, P. M. (1987b). Vowel and consonant recognition of cochlear implant patients using formant-estimating speech processors. *Journal of the Acoustical Society of America, 82*(1), 48-57.

Blamey, P. J., Dowell, R. C., Clark, G. M., & Seligman, P. M. (1987a). Acoustic parameters measured by a formant-estimating speech processor for a multiple-channel cochlear implant. *Journal of the Acoustical Society of America, 82*(1), 38-47.

Blamey, P. J., Dowell, R. C., Tong, Y. C., Brown, A. M., Luscombe, S. M., & Clark, G. M. (1984). Speech processing studies using an acoustic model of a multiple-channel cochlear implant. *Journal of the Acoustical Society of America, 76*(1), 104-110.

Brockmeier, S. J., Grasmeder, M. L., Vischer, M., Mawman, D., Baumgartner, W. D., Stark, T., et al. (2004). Comparison of musical activities by cochlear implant users. Paper presented at *The VIII International Cochlear Implant Conference*, Indianapolis, Indiana USA, May 10-13.

Bruhn, H. (2002). Musical development of elderly people. *Pyschomusicology, 18,* 59-75.

Buchman, C. A., Fucci, M. J., & Luxford, W. M. (1999). Cochlear implants in the geriatric population: benefits outweigh risks. *Ear Nose & Throat Journal, 78*(7), 489-494.

Burian, K., Hochmair-Desoyer, I. J., & Eisenwort, B. (1986). The Vienna cochlear implant program. *Otolaryngologic Clinics of North America*, *19*(2), 313-328.

Burns, E. M., & Viemeister, N. F. (1976). Nonspectral pitch. *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. *Journal of the Acoustical Society of America, 70*(6), 1655-1660.

Busby, P. A., Tong, Y. C., & Clark, G. M. (1993). The perception of temporal modulations by cochlear implant patients. *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. *Journal of the Acoustical Society of America, 95*(5 Pt 1), 2658-2669.

Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M., & Wouters, J. (2002). Temporal pitch mechanisms in acoustic and electric hearing. *Journal of the Acoustical Society of America*, *112*(2), 621-633.

Chasin, M. (2003). Music and hearing aids. *The Hearing Journal, 56*(7), 36-41.

Chasin, M., & Russo, F. A. (2004). Hearing aids and music. *Trends in Amplification, 8*(2), 35-47.

Chouard, C. H., Fugain, C., Meyer, B., & Chabolle, F. (1985). The Chorimac-12. A multichannel cochlear implant for total deafness. Description and clinical results. *Acta Oto-Rhino-Laryngologica Belgica*, *39*(4), 735-748.

Chouard, C. H., Fugain, C., Meyer, B., & Chabolle, F. (1986). The Chorimac 12. A multichannel intracochlear implant for total deafness. *Otolaryngologic Clinics of North America, 19*(2), 355-370.

Clark, G. M. (1995). Cochlear Implants: Historical Perspectives. In G. Plant & K. E. Spens (Eds.), *Profound Deafness and Speech Communication* (pp. 165-218). London: Whurr.

Clark, G. M. (1997). Historical Perspectives. In G. M. Clark, R. S. C. Cowan & R. C. Dowell (Eds.), *Cochlear Implantation for Infants and Children: Advances* (pp. 9-24). San Diego: Singular Publishing.

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

Cochlear Ltd. (2000). *Which Cochlear Implant System? Making the Right Choice. (Promotional Material).* Lane Cove: Cochlear Ltd.

Cohen, A. (1961). Further investigation of the effects of intensity upon the pitch of pure tones. *Journal of the Acoustical Society of America, 33*(10), 1363-1376.

Cohen, L. T., Busby, P. A., Whitford, L. A., & Clark, G. M. (1996). Cochlear implant place psychophysics 1. Pitch estimation with deeply inserted electrodes. *Audiology & Neuro-Otology, 1*(5), 265-277.

Cohen, L. T., Saunders, E., & Clark, G. M. (2001). Psychophysics of a prototype peri-modiolar cochlear implant electrode array. *Hearing Research*, *155*, 63-81.

Collins, L. M., Zwolan, T. A., & Wakefield, G. H. (1997). Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects. *Journal of the Acoustical Society of America, 101*(1), 440-455.

Co-operative Research Centre for Cochlear Implant and Hearing Aid Innovation. (2003a). *Cochlear Implant Training Workshop Notes*. Melbourne: Co-operative Research Centre for Cochlear Implant and Hearing Aid Innovation.

Co-operative Research Centre for Cochlear Implant and Hearing Aid Innovation. (2003b). *Advanced Problem Solving Cochlear Implant Workshop Notes.* Melbourne: Co-operative Research Centre for Cochlear Implant and Hearing Aid Innovation.

Crowder, R. G. (1989). Imagery for Musical Timbre. *Journal of Experimental Psychology - Human Perception and Performance*, *15*(3), 472-478.

Darrow, A. (1979). The beat reproduction response of subjects with normal and impaired hearing: An empirical comparison. *Journal of Music Therapy, 16*(2), 91-98.

Darrow, A. (1984). A comparison of rhythmic responsiveness in normal and hearing impaired children and an investigation of the relationship of rhythmic responsiveness to the suprasegmental aspects of speech perception. *Journal of Music Therapy, 21*(2), 48-66.

Darrow, A. (1987). An investigative study: The effect of hearing impairment on musical aptitude. *Journal of Music Therapy, 24*(2), 88-96.

Dillon, H. (2001). *Hearing Aids.* Sydney: Boomerang Press.

Djourno, A., & Eyries, C. (1957). Auditory prosthesis by means of a distant electrical stimulation of the sensory nerve with the use of an indwelling coil. *La Presse Medicale*, *65*(63), 1417.

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

Dorman, M. F., Basham, K., McCandless, G., & Dove, H. (1991). Speech understanding and music appreciation with the Ineraid cochlear implant. *The Hearing Journal, 44*(6), 34-37.

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. *Journal of the Acoustical Society of America*, *102*, 2403-2411.

Dorman, M. F., Loizou, P. C., Spahr, A. J., & Maloff, E. (2002). Factors that allow a high level of speech understanding by patients fit with cochlear implants. *American Journal of Audiology, 11*(2), 119-123.

Dowell, R. C., Hollow, R. D., & Winton, L. (2004). Changing the selection criteria for cochlear implants - The Melbourne experience. In *Cochlear Limited - White Paper*. Melbourne: Cochlear Limited.

Dreschler, W. A., Verschuure, H., Ludvigsen, C., & Westermann, S. (2001). ICRA noises: artificial noise signals with speech-like spectral and temporal properties for hearing instrument assessment. International Collegium for Rehabilitative Audiology. *Audiology, 40*(3), 148-157.

Durrant, J. D., & Lovrinic, J. H. (1995). *Bases of Hearing Science* (3rd ed.). Baltimore: Williams & Wilkins.

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., & Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *Annals of Otology, Rhinology and Laryngology, 87*(6, part 2), 1-39.

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

Eisen, M. D. (2003). Djourno, Eyries, and the first implanted electrical neural stimulator to restore hearing. *Otology & Neurotology, 24*(3), 500-506.

Erickson, M. L. (2003). Dissimilarity and the classification of female singing voices: A preliminary study. *Journal of Voice, 17*(2), 195-206.

Fearn, R., & Wolfe, J. (2000). Relative importance of rate and place: experiments using pitch scaling techniques with cochlear implants recipients. *Annals of Otology, Rhinology & Laryngology Supplement, 185*, 51-53.

Fletcher, H. (1940). Auditory Patterns. *Reviews of Modern Physics, 12*, 47-65.

Florentine, M., & Buus, S. (1988). Temporal and spectral processing in cochlear and partially simulated impairments. In H. Duifhuis, J. W. Horst & H. P. Witt (Eds.), *Basic Issues in Hearing* (pp. 412-419). London: Academic.

Flynn, M. C., Dowell, R. C., & Clark, G. M. (1998). Aided speech recognition abilities of adults with a severe or severe-to-profound hearing loss. *Journal of Speech, Language & Hearing Research, 41*(2), 285-299.

Fravel, R. P. (1986). Cochlear implant electronics made simple. *Otolaryngologic Clinics of North America, 19*(2), xi-xxii.

Frijns, J. H., Klop, W. M., Bonnet, R. M., & Briaire, J. J. (2003). Optimizing the number of electrodes with high-rate stimulation of the clarion CII cochlear implant. *Acta Otolaryngologica, 123*(2), 138-142.

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

Gantz, B. J. (1987). Cochlear implants: An overview. *Advances in Otolaryngology - Head and Neck Surgery, 1*, 171-200.

Gantz, B. J., & Turner, C. (2003). Combining acoustic and electrical hearing. *Laryngoscope*, *113*(10), 1726-1730.

Gantz, B. J., & Turner, C. (2004). Combining acoustic and electrical speech processing: Iowa/Nucleus hybrid implant. *Acta Otolaryngologica, 124*, 344-347.

Gantz, B. J., Turner, C., Gfeller, K. E., & Lowder, M. W. (2005). Preservation of hearing in cochlear implant surgery: Advantages of combined electrical and acoustical speech processing. *Laryngoscope, 115*(5), 796-802.

Gelfand, S. A. (1998). *Hearing: An Introduction to Psychological and Physiological Acoustics* (3rd ed.). New York: Marcel Dekker.

Geurts, L., & Wouters, J. (2001). Coding of the fundamental frequency in Continuous Interleaved Sampling processors for cochlear implants. *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. *Journal of the Acoustical Society of America, 115*(2), 844-852.

Gfeller, K. (2000). Accommodating children who use cochlear implants in music therapy or educational settings. *Music Therapy Perspectives, 18*, 122-130.

Gfeller, K. (2001). Aural rehabilitation of music listening for adult cochlear implant recipients: Addressing learner characteristics. *Music Therapy Perspectives, 19*, 88-95.

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

Gfeller, K., Christ, A., Knutson, J. F., Witt, S., Murray, K. T., & Tyler, R. S. (2000b). Musical backgrounds, listening habits, and aesthetic enjoyment of adult cochlear implant recipients. *Journal of the American Academy of Audiology, 11*(7), 390-406.

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. (1992). Musical perception of cochlear implant users as measured by the Primary Measures of Music Audiation: An item analysis. *Journal of Music Therapy, 29*(1), 18-39.

Gfeller, K., & Lansing, C. R. (1991). Melodic, rhythmic, and timbral perception of adult cochlear implant users. *Journal of Speech & Hearing Research, 34*(4), 916-920.

Gfeller, K., Olszewski, C., Rychener, M., Sena, K., Knutson, J. F., Witt, S., et al. (2005). Recognition of "real-world" musical excerpts by cochlear implant recipients and normal-hearing adults. *Ear & Hearing, 26*(3), 237-250.

Gfeller, K., Turner, C., Knutson, J., Zhang, X., Froman, R., Olszewski, C., et al. (2004). Relationships among music, cognitive, and speech in noise measures. Paper presented at *The 8th Cochlear Implant Conference*, Indianapolis, Indiana, USA, May 13, 2004.

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

Gfeller, K., Witt, S., Adamek, M., Mehr, M., Rogers, J., Stordahl, J., et al. (2002b). Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients. *Journal of the American Academy of Audiology, 13*(3), 132-145.

Gfeller, K., Witt, S., Stordahl, J., Mehr, M., & Woodworth, G. (2000a). The effects of training on melody recognition and appraisal by adult cochlear implant recipients. *Journal of the Academy of Rehabilitative Audiology, 33*, 115-138.

Gfeller, K., Witt, S., Woodworth, G., Mehr, M. A., & Knutson, J. (2002c). Effects of frequency, instrumental family, and cochlear implant type on timbre recognition and appraisal. *Annals of Otology, Rhinology & Laryngology, 111*(4), 349-356.

Gfeller, K., Woodworth, G., Robin, D. A., Witt, S., & Knutson, J. F. (1997). Perception of rhythmic and sequential pitch patterns by normally hearing adults and adult cochlear implant users. *Ear & Hearing, 18*(3), 252-260.

Glasberg, B. R., & Moore, B. C. J. (1990). Derivation of auditory filter shapes from notched-noise data. *Hearing Research, 47*, 103-138.

Glasberg, B. R., Moore, B. C. J., & Bacon, S. P. (1987). Gap detection and masking in hearing-impaired and normal-hearing subjects. *Journal of the Acoustical Society of America, 54*, 1546-1556.

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

Gordon, E. E. (1979). *Primary Measures of Music Audiation*. Chicago: GIA Publications.

Gordon, E. E. (1986). *Manual for the Primary Measures of Music Audiation and the Intermediate Measures of Music Audiation*. Chicago: GIA Publications.

Grant, K. W., Summers, V., & Leek, M. R. (1998). Modulation rate detection and discrimination by normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 104*(2), 1051-1060.

Grey, J. M. (1977). Multidimensional scaling of musical timbres. *Journal of the Acoustical Society of America, 61*, 1270-1277.

Hamzavi, J., Franz, P., Baumgartner, W. D., & Gstoettner, W. (2001). Hearing performance in noise of cochlear implant patients versus severely-profoundly hearing-impaired patients with hearing aids. *Audiology, 40*(1), 26-31.

Handel, S. (1989). *Listening: An Introduction to the Perception of Auditory Events.* Cambridge: MIT Press.

Handel, S., & Erickson, M. L. (2004). Sound source identification: The possible role of timbre transformations. *Music Perception, 21*(4), 587-610.

Hansen, M. (2002). Effects of multi-channel compression time constants on subjectively perceived sound quality and speech intelligibility. *Ear & Hearing, 23*(4), 369-380.

Harris, G. G. (1963). Periodicity Perception by Using Gated Noise. *Journal of the Acoustical Society of America, 35*(8), 1229-1233.

Hartmann, W. M. (1996). Pitch, periodicity, and auditory organization. *Journal of the Acoustical Society of America, 100*(6), 3491-3502.

Helmholtz, H. (1863). *On the Sensations of Tone as a Physiological Basis of the Theory of Music.* (A. Ellis, Trans.). New York: Dover Publications (1954).

Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *Journal of the Acoustical Society of America*, *113*(5), 2861-2873.

Henry, B. A., Turner, C. W., & Behrens, A. (2005). Spectral peak resolution and speech recognition in quiet: Normal hearing, hearing impaired, and cochlear implant listeners. *Journal of the Acoustical Society of America*, *118*(2), 1111-1121.

Holden, L. K., Skinner, M. W., Holden, T. A., & Demorest, M. E. (2002). Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear & Hearing, 23*(5), 463-476.

Hollow, R. D., Dowell, R. C., Cowan, R. S. C., Skok, M. C., Pyman, B. C., & Clark, G. M. (1995). Continuing improvements in speech processing for adult cochlear implant patients. Paper presented at *The International Cochlear Implant, Speech and Hearing Symposium*, Melbourne, 1994.

House, W. F., & Berliner, K. I. (1986). Safety and efficacy of the House/3M cochlear implant in profoundly deaf adults. *Otolaryngologic Clinics of North America, 19(*2), 275-286.

Houtsma, A. J. M., & Goldstein, J. L. (1972). The Central Origin of the Pitch of Complex Tones: Evidence from Musical Interval Recognition. *Journal of the Acoustical Society of America, 51*(2B), 520-529.

Hughes, M. L., & Abbas, P. J. (2003). The relationship between electrophysiologic and psychophysical measures of spatial spread in cochlear implants. Paper presented at *The 26th ARO (Association for Research in Otolaryngology) Midwinter Research Meeting*, Daytona Beach, Florida., Feb 22-27, 2003.

ICRA. (2005). *Specification of ICRA-Noise.* Retrieved 6 April 2005, from http://www.icra.nu/ITEMS/ICRAnoise.htm

James, C. J., Blamey, P. J., Martin, L., Swanson, B., Just, Y., & Macfarlane, D. (2002). Adaptive dynamic range optimization for cochlear implants: a preliminary study. *Ear & Hearing, 23*(1 Suppl), 49S-58S.

Jerger, J., Jerger, S., Oliver, T., & Pirozzolo, F. (1989). Speech understanding in the elderly. *Ear & Hearing, 10*(2), 79-89.

Kaernbach, C., & Bering, C. (2001). Exploring the temporal mechanism involved in the pitch of unresolved harmonics. *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. *Journal of the Acoustical Society of America, 104*(4), 2298-2306.

Kasturi, K., & Loizou, P. C. (2005a). Effect of filter spacing and correct tonotopic representation on melody recognition: Implications for cochlear implants. Paper presented at *The Conference for the Association for Research In Otolaryngology*, New Orleans, USA.

Kasturi, K., & Loizou, P. C. (2005b). Optimizing virtual channel selection. Paper presented at *The 2005 Conference on Implantable Auditory Prostheses*, Asilomar, Monterey, California, USA, July 30 - Aug 4.

Kessler, D. K. (1999). The CLARION multi-strategy cochlear implant. *Annals of Otology, Rhinology & Laryngology, 108*, 8-16.

Kiefer, J., Pok, M., Adunka, O., Sturzebecher, E., Baumgartner, W., Schmidt, M., et al. (2005). Combined electric and acoustic stimulation of the auditory system: results of a clinical study. *Audiology & Neuro-Otology, 10*(3), 134-144.

Kiefer, J., von Ilberg, C., Reimer, B., Knecht, R., Gall, V., Diller, G., et al. (1998). Results of cochlear implantation in patients with severe to profound hearing loss--implications for patient selection. *Audiology, 37*(6), 382-395.

Klajman, S., Koldej, E., & Kowalska, A. (1982). Investigation of musical abilities in hearing-impaired and normal hearing children. *Folia Phoniatrica, 34*, 229-233.

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 & Neurotology, 9*(4), 214.

Kohlrausch, A., & Houtsma, A. J. M. (1989). Description of complex sounds. In S. Nielzen & O. Olsson (Eds.), *Structure and Perception of Electroacoustic Sound and Music* (pp. 141-159). Amsterdam: Excerpta Medica.

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

Kong, Y. Y., Stickney, G. S., & Zeng, F. G. (2005). Speech and melody recognition in binaurally combined acoustic and electric hearing. *Journal of the Acoustical Society of America, 117*(3 Pt 1), 1351-1361.

Korduba, O. M. (1975). Duplicated rhythm patterns between deaf and normal hearing children. *Journal of Music Therapy, 12*, 136-146.

Krumhansl, C. L., & Iverson, P. (1992). Perceptual interactions between musical pitch and timbre. *Journal of Experimental Psychology, 18*(3), 739-751.

Laback, B., Deutsch, W. A., & Baumgartner, W. D. (2004). Coding of vowellike signals in cochlear implant listeners. *Journal of the Acoustical Society of America, 116*(2), 1208-1223.

Laneau, J., Wouters, J., & Moonen, M. (2006). Improved music perception with explicit pitch coding in cochlear implants. *Audiology & Neuro-Otology, 11*(1), 38-52.

Leal, M. C., Shin, Y. J., Laborde, M. L., Calmels, M. N., Verges, S., Lugardon, S., et al. (2003). Music perception in adult cochlear implant recipients. *Acta Otolaryngologica, 123*, 826-835.

Lentz, J. J., & Leek, M. R. (2003). Spectral shape discrimination by hearing-impaired and normal-hearing listeners. *Journal of the Acoustical Society of America, 113*(3), 1604-1616.

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

Lieberman, P., & Blumstein, S. E. (1988). *Speech Physiology, Speech Perception, and Acoustic Phonetics.* Cambridge: Cambridge University Press.

Limb, C. J. (2000). Music and cochlear implantation. In J. K. Niparko, K. I. Kirk, N. K. Mellon, A. M. Robbins, D. L. Tucci & B. S. Wilson (Eds.), *Cochlear Implants: Principles and Practices* (pp. 265-267). Philadelphia: Lippincott Williams & Wilkins.

Lobo, A. P., Toledo, F., & Loizou, P. C. (2002). Effect of envelope low pass filtering on melody recognition. Paper presented at *The 33rd Neural Prosthesis Workshop*, Bethesda, MD, October, 2002.

Loizou, P. C. (1998). Mimicking the human ear. *IEEE Signal Processing Magazine, 15,* 101-130.

Loizou, P. C., Dorman, M., & Tu, Z. (1999). On the number of channels needed to understand speech. *Journal of the Acoustical Society of America, 106*(4 Pt 1), 2097-2103.

Loizou, P. C., Poroy, O., & Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *Journal of the Acoustical Society of America, 108*(2), 790-802.

Loizou, P. C., Stickney, G., Mishra, L., & Assmann, P. (2003). Comparison of speech processing strategies used in the Clarion implant processor. *Ear & Hearing, 24*(1), 12-19.

Looi, V., & McDermott, H. J. (2002). Music Perception of Cochlear Implant Users. Paper presented at *The Garnett Passe and Rodney Williams Memorial Foundation Conference, Frontiers in Otorhinolaryngology*, Noosa Heads, July 31 - Aug 2.

Looi, V., McDermott, H. J., McKay, C. M., & Hickson, L. (2004). Pitch discrimination and melody recognition by cochlear implant users. Paper presented at *The VIII International Cochlear Implant Conference*, Indianapolis, Indiana USA, May 10-13.

Looi, V., Sucher, C., & McDermott, H. J. (2003). Melodies familiar to the Australian population across a range of hearing abilities. *The Australian and New Zealand Journal of Audiology, 25*(2), 75-83.

Luxford, W., & Brackmann, D. (1985). The history of cochlear implants. In R. Gray (Ed.), *Cochlear Implants* (pp. 1-26). San Diego: College Hill Press.

McDermott, H. J. (2004). Music perception with cochlear implants: A review. *Trends in Amplification, 8*(2), 49-82.

McDermott, H. J., Henshall, K. R., & McKay, C. M. (2002). Benefits of syllabic input compression for users of cochlear implants. *Journal of the American Academy of Audiology, 13*, 14-24.

McDermott, H. J., & Looi, V. (2004). Perception of complex signals, including musical sounds, with cochlear implants. Paper presented at *The VIII International Cochlear Implant Conference,* Indianapolis, Indiana USA., May 10-13.

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

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

McDermott, H. J., McKay, C. M., & Vandali, A. E. (1992). A new portable sound processor for the University of Melbourne/Nucleus Limited multielectrode cochlear implant. *Journal of the Acoustical Society of America, 91*(6), 3367-3371.

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. (2005). Spectral processing in cochlear implants. In M. S. Malmierca & D. R. F. Irvine (Eds.), *Auditory Spectral Processing* (Vol. 70, pp. 474-509). London: Academic Press.

McKay, C. M., & McDermott, H. J. (1993). Perceptual performance of subjects with cochlear implants using the Spectral Maxima Sound Processor (SMSP) and the Mini Speech Processor (MSP). *Ear & Hearing, 14*(5), 350-367.

McKay, C. M., & McDermott, H. J. (1996). The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. *Journal of the Acoustical Society of America, 100*(2 Pt 1), 1081-1092.

McKay, C. M., & McDermott, H. J. (1998). Loudness perception with pulsatile electrical stimulation: the effect of interpulse intervals. *Journal of the Acoustical Society of America, 104*(2 Pt 1), 1061-1074.

McKay, C. M., McDermott, H. J., & Carlyon, R. P. (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. J., & Clark, G. M. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *Journal of the Acoustical Society of America, 96*(5 Pt 1), 2664-2673.

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. *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. *Journal of the Acoustical Society of America, 99*(2), 1079-1090.

McKay, C. M., McDermott, H. J., Vandali, A. E., & Clark, G. M. (1992). A comparison of speech perception of cochlear implantees using the Spectral Maxima Sound Processor (SMSP) and the MSP (MULTIPEAK) processor. *Acta Oto-Laryngologica, 112*(5), 752-761.

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

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

Med-El. (2005). *Med-El technology update*. Innsbruck: Med-El.

Med-El. (2006). *High Definition Signal Processing*. Retrieved 21st January 2006, from http://www.medel.com/ENG/INT/20_Products/20_Speech_Processor/999_tempo_5.asp

Med-El. (n.d.). Combi 40+ Cochlear Implant System (Promotional Brochure). Innsbruck: Med-El.

Miller, G. A., & Taylor, W. G. (1948). The Perception of Repeated Bursts of Noise. *Journal of the Acoustical Society of America, 20*(2), 171-182.

Mirza, S., Douglas, S. A., Lindsey, P., Hildreth, T., & Hawthorne, M. (2003). Appreciation of music in adult patients with cochlear implants: A patient questionnaire. *Cochlear Implants International, 4*(2), 85-95.

Moore, B. C. J. (1995). *Perceptual Consequences of Cochlear Damage*. Oxford: Oxford University Press.

Moore, B. C. J. (1996). Perceptual consequences of cochlear hearing loss and their implications for the design of hearing aids. *Ear & Hearing, 17*(2), 133-161.

Moore, B. C. J. (2003a). Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otology & Neurotology, 24*(2), 243-254.

Moore, B. C. J. (2003b). *An Introduction to the Psychology of Hearing* (5th ed.). San Diego: 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. J. Plack, A. J. Oxenham, R. R. Fay & A. N. Popper (Eds.), *Pitch: Neural Coding and Perception* (Vol. 24, pp. 234-277). New York: Springer-Verlag.

Moore, B. C. J., & Glasberg, B. R. (1988a). Effects of the relative phase of the components on the pitch discrimination of complex tones by subjects with unilateral cochlear impairments. Paper presented at *Basic Issues in Hearing Proceedings of the 8th International Symposium on Hearing*, Paterswolde, Netherlands, April 5-9, 1988.

Moore, B. C. J., & Glasberg, B. R. (1988b). Gap detection with sinusoids and noise in normal, impaired, and electrically stimulated ears. *Journal of the Acoustical Society of America, 83*(3), 1093-1101.

Moore, B. C. J., & Glasberg, B. R. (2001). Temporal modulation transfer functions obtained using sinusoidal carriers with normally hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America, 110*(2), 1067-1073.

Moore, B. C. J., & Moore, G. A. (2003). Discrimination of the fundamental frequency of complex tones with fixed and shifting spectral envelopes by normally hearing and hearing-impaired subjects. *Hearing Research*, *182*(1-2), 153-163.

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

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

Moore, B. C. J., & Skrodzka, E. (2002). Detection of frequency modulation by hearing-impaired listeners: Effects of carrier frequency, modulation rate, and added amplitude modulation. *Journal of the Acoustical Society of America*, *111*(1, Part 1), 327-335.

Morgan, C. T., Garner, W. R., & Galambos, R. (1951). Pitch and intensity. *Journal of the Acoustical Society of America, 23*(6), 658-663.

Nadol, J. B., Jr., Young, Y. S., & Glynn, R. J. (1989). Survival of spiral ganglion cells in profound sensorineural hearing loss: implications for cochlear implantation. *Annals of Otology, Rhinology & Laryngology, 98*(6), 411-416.

Nakata, T., Trehub, S. E., Mitani, C., Kanda, Y., Shibasaki, A., & Schellenberg, E. G. (2005). Music recognition by Japanese children with cochlear implants. *Journal of Physiological Anthropology and Applied Human Science, 24*(1), 29-32.

National Institutes of Health Consensus Development Conference. (1995). Cochlear implants in adults and children. *Journal of the American Medical Association, 274*(24), 1955-1961.

Nelson, D. A., & Freyman, R. L. (1987). Temporal resolution in sensorineural hearing-impaired listeners. *Journal of the Acoustical Society of America, 81*, 709-720.

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.

Otte, J., Schunknecht, H. F., & Kerr, A. G. (1978). Ganglion cell populations in normal and pathological human cochleae. Implications for cochlear implantation. *Laryngoscope, 88*(8 Pt 1), 1231-1246.

Oxenham, A. J., Bernstein, J. G., & 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.

Pasanisi, E., Bacciu, A., Vincenti, V., Guida, M., Barbot, A., Berghenti, M. T., et al. (2003). Speech recognition in elderly cochlear implant recipients. *Clinical Otolaryngology & Allied Sciences, 28*(2), 154-157.

Peeters, S., Marquet, J., Offeciers, F. E., Bosiers, W., Kinsbergen, J., & Van Durme, M. (1989). Cochlear implants: The Laura prosthesis. *Journal of Medical Engineering & Technology, 13*(1-2), 76-80.

Peeters, S., Offeciers, E., Kinsbergen, J., van Durme, M., van Enis, P., Dijkmans, P., et. al. (1993). A digital speech processor and various speech encoding strategies for cochlear implants. *Progress in Brain Research, 97*, 283-289.

Pfingst, B. E., Miller, A. L., Morris, D. J., Zwolan, T. A., Spelman, F. A., & Clapton, B. M. (1995b). Effects of electrical current configuration on stimulus detection. *Annals of Otology, Rhinology & Laryngology, 166*(Supplement), 127-131.

Pfingst, B. E., Morris, D. J., & Miller, A. L. (1995a). Effects of electrode configuration on threshold functions for electrical stimulation of the cochlea. *Hearing Research, 85*, 76-84.

Pijl, S. (1995). Musical pitch perception with pulsatile stimulation of single electrodes in patients implanted with the Nucleus cochlear implant. *Annals of Otology, Rhinology & Laryngology Supplement, 166*, 224-227.

Pijl, S. (1997a). Labeling of musical interval size by cochlear implant patients and normally hearing subjects. *Ear & Hearing, 18*(5), 364-372.

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

Pijl, S., & Schwarz, D. W. (1995). Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes. *Journal of the Acoustical Society of America, 98*(2 Pt 1), 886-895.

Pitt, M. A. (1994). Perception of pitch and timbre by musically trained and untrained listeners. *Journal of Experimental Psychology-Human Perception and Performance, 20*(5), 976-986.

Pitt, M. A., & Crowder, R. G. (1992). The role of spectral and dynamic cues in imagery for musical timbre. *Journal of Experimental Psychology - Human Perception and Performance, 18*(3), 728-738.

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

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

Pollack, I. (1969). Periodicity Pitch for Interrupted White Noise - Fact or Artifact? *Journal of the Acoustical Society of America, 45*(1), 237-238.

Poroy, O., & Loizou, P. C. (2001). Pitch perception of virtual channels. Paper presented at *The 2001 Conference on Implantable Auditory Prostheses,* Asilomar, Monterey, California, USA.

Rasch, R. A., & Plomp, R. (1982). The perception of musical tones. In D. Deutsch (Ed.), *The Psychology of Music* (pp. 1-22). New York: Academic Press.

Rebscher, S. J., Snyder, R. L., & Leake, P. A. (2001). The effect of electrode configuration and duration of deafness on threshold and selectivity of responses to intracochlear electrical stimulation. *Journal of the Acoustical Society of America, 109*(5), 2035-2048.

Reed, C. M., & Delhorne, L. A. (2005). Reception of environmental sounds through cochlear implants. *Ear & Hearing, 26*(1), 48-61.

Rileigh, K. K., & Odom, P. B. (1972). Perception of rhythms by subjects with normal and deficient hearing. *Developmental Psychology, 7,* 54-61.

Rimsky-Korsakov, N. (1891). *Principles of Orchestration* (M. Steinberg, Ed.; E. Agate, Trans.). New York: Edwin F Kalmus (1912).

Ritsma, R. J. (1967). Frequencies Dominant in the Perception of the Pitch of Complex Sounds. *Journal of the Acoustical Society of America, 42*(1), 191-198.

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos Trans R Soc Lond B Biol Sci, 336*(1278), 367-373.

Rubinstein, J. T., & Hong, R. (2003). Signal coding in cochlear implants: Exploiting stochastic effects of electrical stimulation. *Annals of Otology, Rhinology & Laryngology Supplement, 191,* 14-19.

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.

Sanderson, G. (2004). An overview of adult cochlear implantation. In *Cochlear Limited - Global White Paper* (pp. 1-7). Sydney: Cochlear Ltd.

Schatzer, R., Wilson, B. S., Wolford, R. D., & Lawson, D. T. (2003). *Speech processors for auditory prostheses, NIH project NO1-DC-2-1002, 6th Quarterly Progress Report*: Centre for Auditory Prosthesis Research, Research Triangle Institute, Research Triangle Park, NC.

Schindler, R. A., Kessler, D. K., Rebscher, S. J., & Yanda, J. (1986). The University of California, San Francisco/Storz cochlear implant program. *Otolaryngologic Clinics of North America, 19*(2), 287-305.

Schulz, E., & Kerber, M. (1994). Music perception with the MED-EL implants. In I. J. Hochmair-Desoyer & E. S. Hochmair (Eds.), *Advances in Cochlear Implants* (pp. 326-332). Vienna: Datenkonvertierung, Reproduktion und Druck.

Sekuler, R., & Blake, R. (1994). *Perception* (3rd ed.). New York: McGraw-Hill.

Seligman, P. (2003a). Behind-The-Ear Speech Processors. Paper presented at *Cochlear Implant Training Workshop*, Bionic Ear Institute, Melbourne, March, 2003.

Seligman, P. (2003b). Cochlear Implant Engineering. Paper presented at *Cochlear Implant Training Workshop*, Bionic Ear Institute, Melbourne, March 2003.

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

Shannon, R. V. (1989). Detection of gaps in sinusoids and pulse trains by patients with cochlear implants. *Journal of the Acoustical Society of America, 85*(6), 2587-2592.

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

Shannon, R. V., Fu, Q. J., Galvin, J., & Friesen, L. (2004). Speech perception with cochlear implants. In F. G. Zeng, A. N. Popper & R. R. Fay (Eds.), *Cochlear Implants: Auditory Prostheses and Electric Hearing* (Vol. 20, pp. 334-376). New York: Springer.

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., & Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science, 270*, 303-304.

Shepherd, R. K., Hatsushika, S., & Clark, G. M. (1993). Electrical stimulation of the auditory nerve: The effect of electrode position on neural excitation. *Hearing Research, 66*(1), 99-107.

Skarzynski, H., Lorens, A., & Piotrowska, A. (2003). A new method of partial deafness treatment. *Medical Science Monitor, 9*(4), CS20-24.

Skinner, M. W., Clark, G. M., Whitford, L. A., Seligman, P., Staller, S. J., Shipp, D. B., et al. (1994). Evaluation of a new spectral peak coding strategy (Speak) for the Nucleus 22 channel cochlear implant system. *American Journal of Otology, 15*(Suppl 2), 15-27.

Skinner, M. W., Holden, L. K., Whitford, L. A., Plant, K. L., Psarros, C., & Holden, T. A. (2002). Speech recognition with the Nucleus 24 SPEAK, ACE, and CIS speech coding strategies in newly implanted adults. *Ear & Hearing, 23*(3), 207-223.

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

Spahr, A. J., Dorman, M., & Kruger, T. (2005). Performance with spectral resolution in CII and HiRes 90K implant users. Paper presented at *The Asia Pacific Symposium on Cochlear Implants and Related Sciences*, Hong Kong, Nov 26 - 28, 2005.

Staab, W. J. (2002). Characteristics and use of hearing aids. In J. Katz (Ed.), *Handbook of Clinical Audiology* (5th ed., pp. 631-686). Philadelphia: Lippincott Williams and Wilkins.

Stainsby, T. H. (2001). The Perception of Musical Sounds with Cochlear Implants (unpublished thesis). Department of Otolaryngology, The University of Melbourne, Melbourne.

Stainsby, T. H., McDermott, H. J., McKay, C. M., & Clark, G. M. (1997). Preliminary results on spectral shape perception and discrimination of musical sounds by normal hearing subjects and cochlear implantees. *Proceedings of the ICMC 1997*, 11-15.

Stordahl, J. (2002). Song recognition and appraisal: A comparison of children who use cochlear implants and normally hearing children. *Journal of Music Therapy, 39*, 2-19.

Summers, V., & Leek, M. R. (1994). The internal representation of spectral contrast in hearing-impaired listeners. *Journal of the Acoustical Society of America*, *95*(6), 3518-3525.

Swanson, B. (2003). Audio signal processing in the Nucleus cochlear implant system. Paper presented at *Cochlear Implant Training Workshop*, Bionic Ear Institute, Melbourne, March 2003.

Teig, E., Linderman, H. H., Tvete, O., & Rasmussen, K. (1997). Routine use of monopolar stimulation mode in totally deaf children and adults. *Advances in Otorhinolaryngology, 52*, 181-186.

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

Thurlow, W. R. (1963). Perception of low auditory pitch: A multicue, mediation theory. *Psychological Review, 70*(5), 461-470.

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. *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. *Journal of the Acoustical Society of America, 77*(5), 1881-1888.

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

Tyler, R. S., Gfeller, K., & Mehr, M. (2000). A preliminary investigation comparing one and eight channels at fast and slow rates of music appraisal in adults with cochlear implants. *Cochlear Implants International, 1*, 82-87.

Tyler, R. S., Parkinson, A. J., Wilson, B. S., Witt, S., Preece, J. P., & Noble, W. (2002). Patients utilizing a hearing aid and a cochlear implant: Speech perception and localization. *Ear & Hearing, 23*(2), 98-105.

van Buren, R. A., Festen, J. M., & Houtgast, T. (1999). Compression and expansion of the temporal envelope: Evaluation of speech intelligibility and sound quality. *Journal of the Acoustical Society of America, 105*(5), 2903-2913.

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

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 & Hearing, 21*(6), 608-624.

von Wallenberg, E. L., Battmer, R. D., Doden, I., Gnadeberg, D., Hautle, K., & Lenarz, T. (1995). Place-pitch and speech perception measures with bipolar and monopolar electrical stimulation of the cochlea. *Annals of Otology, Rhinology & Laryngology Supplement, 166*, 372-375.

von Wallenberg, E. L., Laszig, R., Gnadeberg, D., Battmer, R. D., Desloovere, C., Kiefer, J., et al. (1993). Initial findings with a modified Nucleus implant comprised of 20 active intracochlear and 2 extracochlear reference electrodes. In I. J. Hochmair-Desoyer & E. S. Hochmair (Eds.), *Advances in Cochlear Implants*. Wien: Manz.

Vongpaisal, T., Trehub, S. E., Schellenberg, E. G., & Papsin, B. (2004). Music recognition by children with cochlear implants. Paper presented at *The VIII International Cochlear Implant Conference*, Indianapolis, Indiana USA, May 10-13.

Warrier, C. M., & Zatorre, R. J. (2002). Influence of tonal context and timbral variation on perception of pitch. *Perception & Psychophysics, 64*(2), 198-207.

Wilson, B. S. (2000). Strategies for representing speech information with cochlear implants. In J. K. Niparko, K. I. Kirk, N. K. Mellon, A. M. Robbins, D. L. Tucci & B. S. Wilson (Eds.), *Cochlear Implants: Principles and Practices* (pp. 129-170). Philadelphia: Lippincott Williams & Wilkins.

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

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., & Zerbi, M. (1993). Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. *Journal of Rehabilitation Research & Development*, *30*(1), 110-116.

Wilson, B. S., Finley, C. C., Lawson, D. T., & Zerbi, M. (1997). Temporal representations with cochlear implants. *American Journal of Otology, 8*(6 (Suppl)), s30-34.

Wilson, B. S., Lawson, D. T., Muller, J. M., Tyler, R. S., & Kiefer, J. (2003a). Cochlear implants: Some likely next steps. *Annual Review of Biomedical Engineering, 5*, 207-249.

Wilson, B. S., Lawson, D. T., & Zerbi, M. (1995). Advances in coding strategies for cochlear implants. *Advances in Otolaryngology - Head and Neck Surgery, 9*, 105-129.

Wilson, B. S., Sun, X., Schatzer, R., & Wolford, R. D. (2004). Representation of fine structure or fine frequency information with cochlear implants. Paper presented at *The VIII International Cochlear Implant Conference*, Indianapolis, Indiana USA, May 10-13.

Wilson, B. S., Wolford, R. D., Schatzer, R., Sun, X., & Lawson, D. T. (2003b). *Speech processors for auditory prostheses, NIH project NO1-DC-2-1002, 7th Quarterly Progress Report*: Centre for Auditory Prosthesis Research, Research Triangle Institute, Research Triangle Park, NC.

Wolfe, J. (2002). Speech and music, acoustics and coding, and what music might be 'for'. Paper presented at *The 7th International Conference on Music Perception and Cognition*, Sydney, Australia.

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.

Zwicker, E. (1961). Subdivision of the audible range into critical bands (Frequenzgruppen). *Journal of the Acoustical Society of America, 33*, 248.

Zwolan, T. A., Kileny, P. R., Ashbaugh, C., & Telian, S. A. (1996). Patient performance with the Cochlear Corporation "20+2" implant: Bipolar versus monopolar activation. *American Journal of Otology, 17*(6), 717-723.

APPENDIX 1: HISTORICAL CI MANUFACTURERS

NOW-OBSOLETE MAJOR COCHLEAR IMPLANT COMPANIES

SINGLE-CHANNEL COCHLEAR IMPLANTS

COMPANY	DETAILS
HOUSE-3M	**Implant** • Single ball electrode inserted 6 mm into the scala tympani • Titanium package • Monopolar stimulation • Transcutaneous link • External transmitter and internal receiver held in place with magnets • Signals conveyed by magnetic induction • External components: microphone, external transmitter, signal processor
	Speech Processing • Analog transformation scheme • Incoming signal is picked up by the microphone, amplified, and sent to a 340 Hz – 2700 Hz bandpass filter • Incoming signal's amplitude used to determine the amplitude of a 16 kHz carrier • No automatic gain control – high and low clipping used • The modulated signal is transmitted directly to the implanted electrode (i.e. the signal is not demodulated before being transmitted to the electrode) • Gross temporal cues of the speech signal are contained in the envelope of the modulated signal, however little fine temporal information is preserved
	Other • Provided increased awareness of environmental sounds, and enhanced lipreading. • Few patients achieved open-set speech recognition • Originally designed by William House and Jack Urban • FDA approval in 1984 for use in adults • Implanted in approximately 3000 patients in the USA and Europe • Absorbed by Cochlear Ltd.
	For more information (Gantz, 1987; House & Berliner, 1986; Loizou, 1998)

VIENNA-3M	**Implant** • 2 electrode arrays were available: i) Monopolar: Active electrode placed adjacent to round window (extracochlear); ground placed medial to temporalis muscle ii) Bipolar: 4 bipolar channels placed in area between 20 mm – 22 mm within the scala tympani. Each channel is independent, although typically only one channel is stimulated at a time • Ceramic packaging • Transcutaneous link • Signals conveyed by magnetic induction
	Speech Processing • Same speech processor drives both arrays • Analog transformation scheme • 16 MHz or 31 MHz carrier frequency used • Incoming acoustic signal (100 Hz - 5000 Hz) picked up by microphone, amplified, then compressed to suit electrically-stimulated hearing via automatic gain compression. Extreme frequencies out of this frequency range are attenuated • Compression ratio of 6:1 used (the approximate difference between the dynamic range for acoustic and electric hearing). Signal not clipped • Signal is demodulated by the receiver-stimulator package before being sent to the active electrode • Fine temporal information of the analog waveform was preserved
	Other • Aided lipreading, and increased awareness of environmental sounds • Some reports of open-set speech recognition in a few high-performing recipients • Developed by E. S. Hochmair & I. J. Hochmair-Desoyer in the early 1980's. • Precursor to Med-El devices
	For more information (Burian et al., 1986; Gantz, 1987; Loizou, 1998)

MULTI-CHANNEL COCHLEAR IMPLANTS

COMPANY	DETAILS
LAURA	**Implant** • 16 platinum-iridium electrodes • Electrodes configured as bipolar pairs • Transcutaneous link • Allows monopolar, bipolar, multipolar, and simultaneous stimulation
	Speech Processing • Uses a bank of bandpass filters to divide the signal into 8 bipolar channels, or 15 monopolar channels • Uses automatic gain compression at the output of each filter • External components: pre-processing hearing aid & speech processor • The hearing aid acts an input amplifier – controls the amount of gain, frequency response, and compression parameters. Also contains a voice-voiceless detector and pitch extractor • Speech processor receives the output of the hearing aid which is then scanned by a microprocessor • For voiced sounds, the 4 highest spectral peaks decoded, and only the channels corresponding to these maxima are stimulated. Amplitude of current depends on the amplitude of the maxima, rate of stimulation depends on the voicing frequency • For voiceless sounds, all 8 channels are stimulated • Charge-balanced biphasic pulses used for stimulation
	Other • Recommended for patients with bilateral total deafness • From Belgium, but sold to Phillips, and then to Cochlear Ltd.
	For more information (Peeters et. al., 1989, 1993)
Chorimac 12	**Implant** • 12 platinum electrodes • Transcutaneous link • Bipolar stimulation
	Speech Processing • Uses a bank of 12 bandpass filters to divide the signal into 12 channels • Only sounds in range 300 Hz to 3000 Hz presented. • Bandwidth each filter: 1/3 octave • Only the signal components between 40 dB – 100dB are presented; compressed into a 6 dB range • Input sound energy represented by varying pulse duration • Rate of stimulation depends on F0; for noise, stimulation is at 300 Hz • Charge-balanced biphasic pulses used for stimulation
	Other • Predominantly developed by C. Chouard • For total bilateral hearing loss. • Pre-cursor to the current-day MXM Digisonic implant
	For more information (Chouard et al., 1985, 1986)

UCSF – Storz	**Implant** • 16 electrodes arranged as 8 bipolar pairs • Transcutaneous link • Monopolar stimulation **Speech Processing** • Uses a bank of bandpass filters to divide the signal into 4 independent channels • Allowed control over individual channels to set different compression and processing parameters for each • The apical channel is used to represent F1 (200-800 Hz), the basal channel for the higher formants above F2 (2.5-6 kHz), and the middle two channels for F2 (800-2500 Hz) • Compressed Analog strategy • Simultaneous stimulation **Other** • First implanted 1981, originating from San Francisco • Reported to assist speech reading, improved vowel and consonant perception results and enabled some open-set word recognition • Preceded the current Advanced Bionic Corporation's Clarion device **For more information** (Gantz, 1987; Loizou, 1998; Schindler et al., 1986)
INERAID - Symbion	**Implant** • 6 platinum electrodes, although only 4 of these are used for any individual • Monopolar stimulation, with an extracochlear reference electrode • Percutaneous link with hardwired transmission of signals from external to internal components **Speech Processing** • Uses a bank of 4 bandpass filters to divide the signal into 4 channels • The 4 channels are used to represent the 1st, 2nd, 3rd, and higher formant regions • Compressed Analog strategy • Simultaneous stimulation **Other** • From Utah; led by D. K. Eddington, and D. E. Brackmann. • Reported to provide better speech perception in noise results than previous implants from other manufacturers • Device still has research applications, due to its percutaneous plug • Absorbed by Cochlear Ltd. **For more information** (Eddington et al., 1978; Gantz, 1987)

APPENDIX 2: MTEQ

MUSIC TRAINING AND EXPERIENCE QUESTIONNAIRE

The following questions relate to your previous music training and experiences. Please complete the questionnaire as honestly, and in as much detail as possible. The information you provide will allow us to better evaluate the results of the music tests, and will be kept confidential. Please feel free to add in additional comments, statements or anything else which you think may be relevant or helpful. If insufficient space is provided, please use the back of each page. If you are unsure or unclear about any of the questions, please ask for further clarification.

Thank you for your time.

FOR ALL SUBJECTS:

Name: ____________________________________ Date: _____________

Date of Birth: ____________________ Age: _____________

FOR SUBJECTS USING A COCHLEAR IMPLANT (CI):

Type of Implant: __________________ Ear Implanted: Left Right

Strategy used (if known – eg. ACE, Speak etc): ______________________________

Do you use a different program or setting for listening to music: Yes No
If yes, please specify (if known): ____

Date of Implant: __________________ Length of time with CI: _______

Do you use your CI everyday: Yes No
If no, when, or how often, do you use your CI? _______________________________

Duration of bilateral severe to profound hearing loss before implant operation (years): ____

Do you wear a hearing aid in the other ear? Yes No
If yes, type of aid: ______________________________

FOR SUBJECTS USING HEARING AID(S) ONLY (ie. No Cochlear Implant):

Do you wear a hearing aid in your *Right* Ear: Yes No Type: _____________

Do you wear a hearing aid in your *Left* Ear: Yes No Type: _____________

How long have you worn a hearing aid in your *Right* Ear (years): ________________

How long have you worn a hearing aid in your L*eft* Ear (years): _________________

Do you use your Hearing Aid(s) everyday: Yes No
If no, when, or how often, do you use your aid(s)? ____________________________

Duration of bilateral severe to profound hearing loss (years): ___________________

Do you use a different program or setting for listening to music: Yes No
If yes, please detail (if known): ____

Have you ever been assessed for a cochlear implant: Yes No
Result/Decision: ____

The Following Questions Refer From The Time Prior To Your Hearing Loss Through To The Present Day

1) a: Have you ever had instrumental (or practical) music lessons (ie. specifically for a music instrument or voice/singing)?

_____Yes _____No *If yes, please detail:*

Instrument	*Number of years of lessons*	*Age received lessons*
____________________	____________________	__________________
________________________	____________________	__________________
________________________	____________________	__________________

b: Did you complete formal music exams in the above instrument(s) or voice?

_____Yes _____No *If yes, please detail:*

Instrument	*Grade level achieved*
________________________	____________________
________________________	____________________
________________________	____________________

2) Did you ever do music, as a subject, at school, university, TAFE, adult colleges or any other post-school learning institution(s)?

_____Yes _____No *If yes, please detail:*

	Place	*Number of Years*	*Age involved in class(es)*
☐	Primary School	__________	___________________
☐	High School	__________	___________________
☐	University	__________	___________________
☐	TAFE	__________	___________________
☐	Adult College	__________	___________________
☐	Other (*specify*)	__________	___________________

3) Have you ever been involved in a music group or ensemble (eg. band, choir, orchestra etc.)?

_____Yes _____No *If yes, please detail:*

Group	*Number of years*	*Age at which involved*
______________________________	___________	___________
______________________________	___________	___________
______________________________	___________	___________

4) Have you ever participated in music appreciation, music theory or music history classes (eg. learning about composers, styles, harmony, composition, keys etc.)?

_____Yes _____No *If yes, please detail:*

Type of class	*Number of years*	*Age at which involved*
______________________________	___________	___________
______________________________	___________	___________
______________________________	___________	___________

5) Have you ever been involved in any other formal music classes, experiences, activities etc., not covered above?

_____Yes _____No *If yes, please detail:*

Type	*Number of years*	*Age at which involved*
______________________________	___________	___________
______________________________	___________	___________
______________________________	___________	___________

6) Please detail any informal music classes, activities, experiences etc. that you have been involved in (eg. "self-taught" musician, learning an instrument "by ear" or with friends, own "music training program", personal research for self interest and information etc).

Please include detail regarding number of years and age at which the activity(s) was undertaken.

____

____

____

____

7) On a scale of 1-5, please rate the following:

(1=None or Not Able; 2=Limited; 3=Average; 4=Above Average; 5=Extensive or Very Able).

a) Knowledge of music history:	1	2	3	4	5
b) Knowledge of music theory:	1	2	3	4	5
c) Ability to read music:	1	2	3	4	5
d) Ability to play an instrument or sing:	1	2	3	4	5
e) Overall music ability:	1	2	3	4	5

Comments:

____

____

____

The Following Question Refers To The Time Period Since You Received Your Cochlear Implant, or (for Hearing Aid-only users), since you were fitted with your current Hearing Aid(s).

8) Since you received your implant or hearing aid(s), have you:

a) Ever had formal instrumental (or vocal) music lessons: ___ Yes ___ No

If yes, please detail: ____

b) Ever attended music appreciation, music history or music theory lessons:

___ Yes ___ No

If yes, please detail: ____

c) Ever participated in a music group or ensemble (eg. choir, band, orchestra etc.):

___ Yes ___ No

If yes, please detail: ____

d) Ever taught yourself a music instrument, singing or music theory? ___ Yes ___ No

If yes, please detail: ____

e) Ever tried to improve your music perception ability? ___ Yes ___ No

If yes, please detail: ____

9) Any other information or comments?

THANK YOU FOR YOUR TIME

APPENDIX 3: MLEQ

MUSIC LISTENING AND ENJOYMENT QUESTIONNAIRE

The following questions relate to your musical preferences, and music listening experience. Please complete the questionnaire as honestly, and in as much detail as possible. The information you provide will allow us to design practical and useful strategies or programs to potentially assist with music perception. Your details and information will be kept confidential. Feel free to add in additional comments, information or anything else which you think may be relevant or helpful. If insufficient space is provided, please use the back of each page. If you are unsure or unclear about any of the questions or terms, please ask for further clarification.

Where you are asked to make comparisons between 'prior to hearing loss' and 'present time', please make comparisons from what you remember before you lost your hearing (to the best of your ability), to the present day while using your current listening device(s). Thank you for your time.

FOR ALL SUBJECTS:

Name: ____ Date: ______________

Date of Birth: ____________________ Age: ______________

Do you listen to music through a Direct Audio Input (ie., a direct connection or plug into your hearing aid or implant, from the sound source).

Yes No Detail: ____

FOR SUBJECTS USING A COCHLEAR IMPLANT (CI):

Type of Implant: __________________ Ear Implanted: Left Right

Strategy used (if known – eg. ACE, Speak etc): ________________________________

Do you wear a hearing aid in the other ear? Yes No
If yes, type of aid:_______

FOR SUBJECTS USING HEARING AID(S) ONLY (ie. No Cochlear Implant) (HA):

Do you wear a hearing aid in your *Right* Ear: Yes No Type: __________

Do you wear a hearing aid in your *Left* Ear: Yes No Type: __________

CI=Cochlear Implant; HA=Hearing Aid(s); HL=Hearing Loss

1) a: Prior to your hearing loss (HL), how often did you choose to listen to music (eg. radio, tape, CD, concerts etc.)?

___ Very Often ___ Often ___ Sometimes ___ Occasionally ___ Never

Approximately ____ hours per week

b: Since you received your CI or HA, how often do you choose to listen to music?

___ Very Often ___ Often ___ Sometimes ___ Occasionally ___ Never

Approximately ____ hours per week

2) The following is a list of common situations where you may encounter music.

Please rate, on a scale of 1 – 5, how much you enjoy listening to music in these situations; both prior to your HL as well as currently, when using your CI and/or HA.

(1=Do not enjoy at all; 5=Very much enjoy)
(NA=Have not had sufficient experience or exposure with this situation).

SITUATION	*Prior to hearing loss*	*Present Time*
Radio – *Car*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *At home or work*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Via Direct Audio Input (if applicable)*	1 2 3 4 5 NA	1 2 3 4 5 NA
Preferred radio stations:		
1) ____________________		
2) ____________________		
3) ____________________		
Tape or Record or CD – *Car*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *At home or work*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Via Direct Audio Input (if applicable)*	1 2 3 4 5 NA	1 2 3 4 5 NA
'Favourite' CD, Tape or Record:		
1) ____________________		
2) ____________________		
3) ____________________		
Place of Worship – *Choir*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Organ or other Instruments*	1 2 3 4 5 NA	1 2 3 4 5 NA

SITUATION (Cont.)	*Prior to hearing loss*	*Present Time*
Live Concert – *Classical*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Jazz*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Popular or Rock*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Other (specify)* ________________	1 2 3 4 5 NA	1 2 3 4 5 NA
Live Music (eg. playing an instrument, listening to another family member playing an instrument etc.)		
- *Instrumental*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Singing*	1 2 3 4 5 NA	1 2 3 4 5 NA
Background Music – *TV*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Movie*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Social Event*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *PA system (eg. restaurant, shops etc.)*	1 2 3 4 5 NA	1 2 3 4 5 NA

3) Please indicate which statement below best describes how your enjoyment of music has changed from prior to your HL to the present day (with your CI and/or HA).

___ I never really listened to music before my hearing loss, and I do not listen to it now.

___ Music is not as pleasant as I recall before my hearing loss, and I do not enjoy it anymore.

___ Music is not as pleasant as I recall before my hearing loss, but it is better than nothing.

___ Music is not as pleasant as I recall before my hearing loss, but I still enjoy it now.

___ Music sounds different to what I recall, but is no less enjoyable.

___ Music does not sound any different to what I recall it to be, before my hearing loss.

___ Music is more pleasant sounding than I recall before my hearing loss.

4) Please indicate which statement below best describes how your music listening habits have changed from pre-HL to the present day (with your CI and/or HA).

___ No change – I did not listen to music before my hearing loss, and do not do so now.

___ No change – I listened to music occasionally before my hearing loss, and listen to it occasionally now.

___ No change – I listened to music frequently before my hearing loss, and listen to it frequently now.

___ I listened to music more before my hearing loss, than now.

___ I listen to music more now, than before my hearing loss.

5) a: At present, with your CI and/or HA - By listening only (ie. with no visual cues or prior knowledge), do you think you would be able to distinguish between:

i)	Male and female speaker:	___ Yes	___ No
ii)	Male and female singer:	___ Yes	___ No
iii)	Different music instruments:	___ Yes	___ No
iv)	Different music ensembles (eg. band from orchestra):	___ Yes	___ No
v)	Different speakers of the same gender:	___ Yes	___ No
vi)	A singer from a speaker (of either gender):	___ Yes	___ No
vii)	Band/ensemble *with* a singer from a band/ensemble *without* a singer:	___ Yes	___ No

Comments:

____

____

b: Would you prefer to listen to a male or female singer?

___ Male ___ Female ___ No preference

6) The following are a list of common musical instruments which you may be familiar with. Currently, would you remember how this instrument sounded before you lost your hearing?

(Y=Yes, I remember the sound of this instrument
½=I think I remember it
N=No, I don't remember the sound of this instrument).
(NA=Do not know this instrument, or have not had sufficient exposure to this instrument).

INSTRUMENT	*Do you remember the sound of this instrument?*
Bass Drum or Timpani	Y ½ N *NA*
Cello	Y ½ N *NA*
Clarinet	Y ½ N *NA*
Drum Kit	Y ½ N *NA*
Female Singer	Y ½ N *NA*
Flute	Y ½ N *NA*
Guitar	Y ½ N *NA*
Male Singer	Y ½ N *NA*
Organ	Y ½ N *NA*
Piano	Y ½ N *NA*
Tambourine	Y ½ N *NA*

INSTRUMENT Cont.	*Do you remember the sound of this inst?*			
Trombone	Y	½	N	*NA*
Trumpet	Y	½	N	*NA*
Violin	Y	½	N	*NA*
Xylophone	Y	½	N	*NA*

7) The following is the same list of common musical instruments.

Please rate, on a scale of 1-5, how easy it would be to recognize these instruments by sound only (ie. no visual cues); both prior to your HL as well as now with your CI and/or HA.

(1=Impossible to recognize by sound only – I would never recognize this instrument;
5=Very easy to recognize by sound only – I would always recognize this instrument).
(NA=Do not know this instrument, or have not had sufficient exposure to this instrument)

INSTRUMENT	*Prior to Hearing Loss*	*Present Time*
Bass Drum or Timpani	1 2 3 4 5 NA	1 2 3 4 5 NA
Cello	1 2 3 4 5 NA	1 2 3 4 5 NA
Clarinet	1 2 3 4 5 NA	1 2 3 4 5 NA
Drum Kit	1 2 3 4 5 NA	1 2 3 4 5 NA
Female Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Flute	1 2 3 4 5 NA	1 2 3 4 5 NA
Guitar	1 2 3 4 5 NA	1 2 3 4 5 NA
Male Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Organ	1 2 3 4 5 NA	1 2 3 4 5 NA
Piano	1 2 3 4 5 NA	1 2 3 4 5 NA
Tambourine	1 2 3 4 5 NA	1 2 3 4 5 NA
Trombone	1 2 3 4 5 NA	1 2 3 4 5 NA
Trumpet	1 2 3 4 5 NA	1 2 3 4 5 NA
Violin	1 2 3 4 5 NA	1 2 3 4 5 NA
Xylophone	1 2 3 4 5 NA	1 2 3 4 5 NA

8) The following is the same list of musical instruments.

Please rate, on a scale of 1 – 5, how much you enjoy(ed) listening to these musical instruments, both prior to your HL as well as now with your CI and/or HA.

(1=Do not enjoy at all; 5=Very much enjoy)
(NA=Do not know this instrument, or have not had sufficient exposure to this instrument)

INSTRUMENT	*Prior to Hearing Loss*	*Present Time*
Bass Drum or Timpani	1 2 3 4 5 NA	1 2 3 4 5 NA
Cello	1 2 3 4 5 NA	1 2 3 4 5 NA
Clarinet	1 2 3 4 5 NA	1 2 3 4 5 NA
Drum Kit	1 2 3 4 5 NA	1 2 3 4 5 NA
Female Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Flute	1 2 3 4 5 NA	1 2 3 4 5 NA
Guitar	1 2 3 4 5 NA	1 2 3 4 5 NA
Male Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Organ	1 2 3 4 5 NA	1 2 3 4 5 NA
Piano	1 2 3 4 5 NA	1 2 3 4 5 NA
Tambourine	1 2 3 4 5 NA	1 2 3 4 5 NA
Trombone	1 2 3 4 5 NA	1 2 3 4 5 NA
Trumpet	1 2 3 4 5 NA	1 2 3 4 5 NA
Violin	1 2 3 4 5 NA	1 2 3 4 5 NA
Xylophone	1 2 3 4 5 NA	1 2 3 4 5 NA

9) The following are a list of common music styles which you may be familiar with. Please rate, on a scale of 1-5, how easy it would be to recognize these styles by sound only (ie. no visual cues); both prior to your HL as well as now with your CI and/or HA.

(1=Impossible to recognize by sound only – I would not recognize this music style;
5=Very easy to recognize by sound only – I would always recognize this music style).
(NA=Do not know this style, or have not had sufficient exposure to this style of music).

STYLE	*Prior to Hearing Loss*	*Present Time*
Recent Pop or Rock (eg. “top 40”, 80’s, 90’s).	1 2 3 4 5 NA	1 2 3 4 5 NA
Hard Rock (eg. “heavy metal”).	1 2 3 4 5 NA	1 2 3 4 5 NA
Rap	1 2 3 4 5 NA	1 2 3 4 5 NA
60’s – 70’s music	1 2 3 4 5 NA	1 2 3 4 5 NA
‘Old time’ songs (eg. 20’s-50’s, war songs etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA

STYLE Cont.	*Prior to Hearing Loss*	*Present Time*
Blues	1 2 3 4 5 NA	1 2 3 4 5 NA
Jazz (eg. swing band, ragtime etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
Country and Western	1 2 3 4 5 NA	1 2 3 4 5 NA
Musicals or Movie music	1 2 3 4 5 NA	1 2 3 4 5 NA
Classical	1 2 3 4 5 NA	1 2 3 4 5 NA
Easy Listening (eg. "relaxation", instrumental)	1 2 3 4 5 NA	1 2 3 4 5 NA
Religious (eg. hymns, gospel music etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
World Music (specify) _________________	1 2 3 4 5 NA	1 2 3 4 5 NA
Other (specify) ________________________	1 2 3 4 5 NA	1 2 3 4 5 NA

10) The following is the same list of music styles.

Please rate, on a scale of 1 – 5, how much you enjoy(ed) listening to these styles of music; both prior to your HL as well as now with your CI and/or HA.

(1=Do not enjoy at all;
5=Very much enjoy).
(NA=Have not had sufficient experience or exposure with this situation)

STYLE	*Prior to Hearing Loss*	*Present Time*
Recent Pop or Rock (eg. "top 40", 80's, 90's)	1 2 3 4 5 NA	1 2 3 4 5 NA
Hard Rock (eg. "heavy metal")	1 2 3 4 5 NA	1 2 3 4 5 NA
Rap	1 2 3 4 5 NA	1 2 3 4 5 NA
60's – 70's music	1 2 3 4 5 NA	1 2 3 4 5 NA
'Old time' songs (eg. 20's-50's, war songs etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
Blues	1 2 3 4 5 NA	1 2 3 4 5 NA
Jazz (eg. swing band, ragtime etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
Country and Western	1 2 3 4 5 NA	1 2 3 4 5 NA
Musicals or Movie music	1 2 3 4 5 NA	1 2 3 4 5 NA
Classical	1 2 3 4 5 NA	1 2 3 4 5 NA
Easy Listening (eg. "relaxation", instrumental)	1 2 3 4 5 NA	1 2 3 4 5 NA
Religious (eg. hymns, gospel music etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
World Music (specify) _________________	1 2 3 4 5 NA	1 2 3 4 5 NA
Other (specify) ________________________	1 2 3 4 5 NA	1 2 3 4 5 NA

11) Overall, please rate in order of frequency from 1-9, which of the following methods for listening to music you most utilize at present, with your CI and/or HA.

(1=most frequently listened to; 9=least frequently listened to).
NA= Do not listen to music using this method at all.
*If you use Direct Audio Input for any of the following, please place a * next to the item.*

___ Radio

___ TV (including videos & DVDs)

___ Compact Disk (CD)

___ Tape

___ Record/LP

___ Computer

___ Live performance (such as concerts, church, theatres etc.)

___ Playing an instrument

___ Other (*specify*) ____________________________________

12) The following is a list of items that may impact upon your music listening experience. For each, please indicate (by circling "+", "O", "-") how this item currently affects your music listening, when utilizing your CI and/or HA.

\+ *Generally makes listening to music more enjoyable or easier*
0 *Generally makes no difference in listening to music*
- *Generally makes listening to music less enjoyable or more difficult*

	Less enjoyable	No difference	More enjoyable
Familiar with song or piece (ie. knew the music before hearing loss)	**-**	**0**	+
Have played the piece on an instrument, or sung it	**-**	**0**	+
Know the title of the song or music	**-**	**0**	+
Know the style or genre of the music	**-**	**0**	+
Know the performer(s)	**-**	**0**	+
Familiar with the context or storyline (eg. storyline of a musical or movie, or a social situation)	**-**	**0**	+
Ability to follow along with the words or music score	**-**	**0**	+
Being able to watch the performer(s) (ie. watching the singer's face, or the instrument(s))	**-**	**0**	+
Optimal seating in a live performance	**-**	**0**	+
Direct Audio Input	**-**	**0**	+
Quiet Listening Environment	**-**	**0**	+

(Cont).	Less enjoyable	No difference	More enjoyable
"Echoey" room (ie. room with a lot of reverberation)	**-**	**0**	+
Noisy room	**-**	**0**	+
Good quality recordings	**-**	**0**	+
Music with a simple melody	**-**	**0**	+
Music with a simple harmony	**-**	**0**	+
Music with a simple rhythm	**-**	**0**	+
Music with a distinctive or clear rhythm	**-**	**0**	+
Music with words	**-**	**0**	+
Music without words	**-**	**0**	+
Separate coding strategy or listening program that has been programmed into your CI or HA for listening to music (if applicable)	**-**	**0** **N/A**	+
Solo instrument or performer	**-**	**0**	+
Small group of performers (eg. duet/trio, or singer with a few accompanying instruments)	**-**	**0**	+
Large group or ensemble (eg. band, orchestra, choir)	**-**	**0**	+
Experience with the CI or HA If so, how much experience? ______________________	**-**	**0**	+
Music listening practice If so, how much & what type? _____________________	**-**	**0**	+

13) Do you utilize any specific tactics to help you when listening to music?

If yes, please detail: ____

____

____

14) Please specify if there are any types of music, or particular music styles, that you would like to hear better.

____

____

____

15) Please specify if there are any particular situations where you would like to hear music better, along with the difficulties you currently experience in this situation(s).

____

____

____

16) Any other comments or information?

____

____

____

____

____

____

THANK YOU FOR YOUR TIME

APPENDIX 4: MLEQ (WL POST)

MUSIC LISTENING AND ENJOYMENT QUESTIONNAIRE (WL Post)

The following questions are a follow up to the questionnaire you completed before you received your implant. This questionnaire asks very similar questions, but aims to compare your opinions of hearing aids to cochlear implants for listening to music. Therefore, we would like you to make the comparisons between pre-implant music listening (with hearing aids) to now, with the cochlear implant. Thank you again for your time.

Name: ____ Date: ____________

Date of Birth: ____________________ Age: _____________

Date of Implant: __________________ Type of Implant: ____________________

Date of Switch on: ________________ Ear Implanted: Left Right

Strategy used (if known – eg. ACE, Speak etc): ____________________________________

Listening Program used for everyday listening: (eg P1, P2 etc): __________________

Do you still wear a hearing aid in the other ear? Yes No If yes, type of aid:_______

Have you used any of the following features with your CI:

Direct Audio Input: Yes No Telecoil: Yes No

In your opinion, how much difference has the implant made for your speech perception:

1	2	3	4	5
Worse	No change	A little better	Somewhat better	Much better

How satisfied are you with the cochlear implant so far:

1	2	3	4	5
Unsatisfied	Indifferent	A little satisfied	Somewhat satisfied	Very satisfied

Overall, has the cochlear implant met your expectations?

No Yes Unsure

Please comment: ____

CI=Cochlear Implant; HA=Hearing Aid(s); HL=Hearing Loss

1) a: Prior to your CI, how often did you choose to listen to music (eg. radio, tape, CD, concerts etc.)?

___ Very Often ___ Often ___ Sometimes ___ Occasionally ___ Never

Approximately ____ hours per week

b: Now with your CI, how often do you choose to listen to music?

___ Very Often ___ Often ___ Sometimes ___ Occasionally ___ Never

Approximately ____ hours per week

2) The following is a list of common situations where you may encounter music.

Please rate, on a scale of 1 – 5, how much you enjoy listening to music in these situations – prior to the implant, with Has, and now with the CI

(1=Do not enjoy at all; 5=Very much enjoy)
(NA=Have not had sufficient experience or exposure with this situation).

SITUATION	*Previously With HA(s)*	*Present Time with CI*
Radio – *Car*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *At home or work*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Via Direct Audio Input (if applicable)*	1 2 3 4 5 NA	1 2 3 4 5 NA
Preferred radio stations:		
1) ________________		
2) ________________		
3) ________________		
Tape or Record or CD – *Car*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *At home or work*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Via Direct Audio Input (if applicable)*	1 2 3 4 5 NA	1 2 3 4 5 NA
‘Favourite’ CD, Tape or Record:		
1) ________________		
2) ________________		
3) ________________		
Place of Worship – *Choir*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Organ or other Instruments*	1 2 3 4 5 NA	1 2 3 4 5 NA

SITUATION (Cont.)	*Previously With HA(s)*	*Present Time with CI*
Live Concert – *Classical*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Jazz*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Popular or Rock*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Other (specify)* ________________	1 2 3 4 5 NA	1 2 3 4 5 NA
Live Music (eg. playing an instrument, listening to another family member playing an instrument etc.)		
- *Instrumental*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Singing*	1 2 3 4 5 NA	1 2 3 4 5 NA
Background Music – *TV*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Movie*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *Social Event*	1 2 3 4 5 NA	1 2 3 4 5 NA
- *PA system (eg. restaurant, shops etc.)*	1 2 3 4 5 NA	1 2 3 4 5 NA

3) Please indicate which statement below best describes how you have found music to sound using your CI, when compared to using HAs.

___ Music sounds better with the CI than it did with HAs, although I don't like the sound with either device.

___ Music sounds better with the CI than it did with HAs, it now sounds more pleasant.

___ No difference – I didn't like the sound of music through the HAs, and still don't like the sound of music now with the CI.

___ No difference – I liked the sound of music through the HAs, and like the sound of music now with the CI.

___ Music sounds worse with the CI than it did with HAs, although I don't like the sound of with either device.

___ Music sounds worse with the CI than it did with HAs, I didn't mind the sound of music through HAs, but don't like the sound through the CI.

4) a: At present, with your CI - By listening only (ie. with no visual cues or prior knowledge), do you think you would be able to distinguish between:

i)	Male and female speaker:	___ Yes	___ No
ii)	Male and female singer:	___ Yes	___ No
iii)	Different music instruments:	___ Yes	___ No
iv)	Different music ensembles (eg. band from orchestra):	___ Yes	___ No
v)	Different speakers of the same gender:	___ Yes	___ No

vi) A singer from a speaker (of either gender): ___ Yes ___ No

vii) Band/ensemble *with* a singer from a band/ensemble *without* a singer: ___ Yes ___ No

Comments: ____

____

b: Would you prefer to listen to a male or female singer?

___ Male ___ Female ___ No preference

5) The following is the same list of common musical instruments.

Please rate, on a scale of 1-5, how easy it would be to recognize these instruments by sound only (ie. no visual cues) prior to your implant (with HAs) compared to now, with your CI.

(1=Impossible to recognize by sound only – I would never recognize this instrument;
5=Very easy to recognize by sound only – I would always recognize this instrument).
(NA=Do not know this instrument, or have not had sufficient exposure to this instrument)

INSTRUMENT	*Previously With HA(s)*	*Present Time with CI*
Bass Drum or Timpani	1 2 3 4 5 NA	1 2 3 4 5 NA
Cello	1 2 3 4 5 NA	1 2 3 4 5 NA
Clarinet	1 2 3 4 5 NA	1 2 3 4 5 NA
Drum Kit	1 2 3 4 5 NA	1 2 3 4 5 NA
Female Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Flute	1 2 3 4 5 NA	1 2 3 4 5 NA
Guitar	1 2 3 4 5 NA	1 2 3 4 5 NA
Male Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Organ	1 2 3 4 5 NA	1 2 3 4 5 NA
Piano	1 2 3 4 5 NA	1 2 3 4 5 NA
Tambourine	1 2 3 4 5 NA	1 2 3 4 5 NA
Trombone	1 2 3 4 5 NA	1 2 3 4 5 NA
Trumpet	1 2 3 4 5 NA	1 2 3 4 5 NA
Violin	1 2 3 4 5 NA	1 2 3 4 5 NA
Xylophone	1 2 3 4 5 NA	1 2 3 4 5 NA

6) The following is the same list of musical instruments.

Please rate, on a scale of 1 – 5, how *pleasant* these musical instruments sound prior to your CI (with HAs) as compared to now with your CI.

(1=Do not enjoy at all; 5=Very much enjoy)
(NA=Do not know this instrument, or have not had sufficient exposure to this instrument)

INSTRUMENT	*Previously With HA(s)*	*Present Time with CI*
Bass Drum or Timpani	1 2 3 4 5 NA	1 2 3 4 5 NA
Cello	1 2 3 4 5 NA	1 2 3 4 5 NA
Clarinet	1 2 3 4 5 NA	1 2 3 4 5 NA
Drum Kit	1 2 3 4 5 NA	1 2 3 4 5 NA
Female Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Flute	1 2 3 4 5 NA	1 2 3 4 5 NA
Guitar	1 2 3 4 5 NA	1 2 3 4 5 NA
Male Singer	1 2 3 4 5 NA	1 2 3 4 5 NA
Organ	1 2 3 4 5 NA	1 2 3 4 5 NA
Piano	1 2 3 4 5 NA	1 2 3 4 5 NA
Tambourine	1 2 3 4 5 NA	1 2 3 4 5 NA
Trombone	1 2 3 4 5 NA	1 2 3 4 5 NA
Trumpet	1 2 3 4 5 NA	1 2 3 4 5 NA
Violin	1 2 3 4 5 NA	1 2 3 4 5 NA
Xylophone	1 2 3 4 5 NA	1 2 3 4 5 NA

7) The following are a list of common music styles which you may be familiar with. Please rate, on a scale of 1-5, how easy it would be to *recognize* these styles by sound only (ie. no visual cues); prior to your CI (when using HA(s)), as well as now with your CI.

(1=Impossible to recognize by sound only – I would not recognize this music style;
5=Very easy to recognize by sound only – I would always recognize this music style).
(NA=Do not know this style, or have not had sufficient exposure to this style of music).

STYLE	*Previously With HA(s)*	*Present Time with CI*
Recent Pop or Rock (eg. "top 40", 80's, 90's).	1 2 3 4 5 NA	1 2 3 4 5 NA
Hard Rock (eg. "heavy metal").	1 2 3 4 5 NA	1 2 3 4 5 NA
Rap	1 2 3 4 5 NA	1 2 3 4 5 NA
60's – 70's music	1 2 3 4 5 NA	1 2 3 4 5 NA
'Old time' songs (eg. 20's-50's, war songs etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA

STYLE Cont.	*Previously With HA(s)*	*Present Time with CI*
Blues	1 2 3 4 5 NA	1 2 3 4 5 NA
Jazz (eg. swing band, ragtime etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
Country and Western	1 2 3 4 5 NA	1 2 3 4 5 NA
Musicals or Movie music	1 2 3 4 5 NA	1 2 3 4 5 NA
Classical	1 2 3 4 5 NA	1 2 3 4 5 NA
Easy Listening (eg. "relaxation", instrumental)	1 2 3 4 5 NA	1 2 3 4 5 NA
Religious (eg. hymns, gospel music etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
World Music (specify) __________________	1 2 3 4 5 NA	1 2 3 4 5 NA
Other (specify) _________________________	1 2 3 4 5 NA	1 2 3 4 5 NA

8) The following is the same list of music styles.

Please rate, on a scale of 1 – 5, how *pleasant* these styles sound - prior to your CI (using HA(s)), as well as now with your CI.

(1=Do not enjoy at all;
5=Very much enjoy).
(NA=Have not had sufficient experience or exposure with this situation)

STYLE	*Previously With HA(s)*	*Present Time with CI*
Recent Pop or Rock (eg. "top 40", 80's, 90's)	1 2 3 4 5 NA	1 2 3 4 5 NA
Hard Rock (eg. "heavy metal")	1 2 3 4 5 NA	1 2 3 4 5 NA
Rap	1 2 3 4 5 NA	1 2 3 4 5 NA
60's – 70's music	1 2 3 4 5 NA	1 2 3 4 5 NA
'Old time' songs (eg. 20's-50's, war songs etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
Blues	1 2 3 4 5 NA	1 2 3 4 5 NA
Jazz (eg. swing band, ragtime etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
Country and Western	1 2 3 4 5 NA	1 2 3 4 5 NA
Musicals or Movie music	1 2 3 4 5 NA	1 2 3 4 5 NA
Classical	1 2 3 4 5 NA	1 2 3 4 5 NA
Easy Listening (eg. "relaxation", instrumental)	1 2 3 4 5 NA	1 2 3 4 5 NA
Religious (eg. hymns, gospel music etc.)	1 2 3 4 5 NA	1 2 3 4 5 NA
World Music (specify) __________________	1 2 3 4 5 NA	1 2 3 4 5 NA
Other (specify) _________________________	1 2 3 4 5 NA	1 2 3 4 5 NA

9) Overall, which device do you think is better for listening to music?

___ Hearing Aids ___ Cochlear Implant ___ No difference

10) Could you describe/make some comments comparing the sound of music with HAs as opposed to with the CI?

____

____

____

____

____

____

____

____

____

____

THANK YOU FOR YOUR TIME