

Faculty of Medicine and Health Sciences
Department of Oto-rhino-laryngology and logopaedic-audiologic sciences

Pitch Perception and Cochlear Implants

Richard Penninger

Thesis submitted to fulfill the requirements of
Doctor of Biomedical Sciences

Ghent, 2013



Promoter

Supervisor: Prof. Dr. Ingeborg Dhooge, Ghent University, Ghent Belgium

Co-Supervisor: Prof. Dr. Marc Leman, Ghent University, Ghent Belgium

Examination Board

Prof. Dr. Johan Vande Walle, Ghent University, Ghent Belgium (chairman)

Prof. Dr. Els De Leenheer, Ghent University, Ghent Belgium

Prof. Dr. Bert Vanheel, Ghent University, Ghent Belgium

Prof. Dr. Paul Corthals, Ghent University, Ghent Belgium

Prof. Dr. Andy Beynon, Radboud University, Nijmegen Medical Centre,
Nijmegen Netherlands

Prof. Dr. Andreas Büchner, Medical University Hannover (MHH), Hannover
Germany

Dr. Filiep Van Poucke, Cochlear Technology Center (CTC), Mechelen Belgium

Supervisory Committee

Prof. Dr. Charles Limb, Johns Hopkins University, Baltimore MD USA

Prof. Dr. Charles Della Santina, Johns Hopkins University, Baltimore MD USA

Dr. Katrien Vermeire, Thomas More, Antwerp Belgium

The research leading to these results has received funding from the European Community's Seventh Framework Program under the EBRAMUS project - grant agreement number 238157

No part of this work may be reproduced in any form or by any means, electronically, mechanically, by print or otherwise without prior written permission of the author.

Richard Penninger

Email: Richard.penninger@ugent.be

Contents

List of publications.....	8
Summary	10
Samenvatting.....	12
Abbreviations	15
1 Introduction.....	17
1.1 Anatomy of the ear	17
1.1.1 Organ of Corti (OC).....	18
1.1.2 Mechanotransduction of hair bundles.....	20
1.2 Pitch.....	22
1.2.1 What is pitch and what is it used for?.....	22
1.2.2 Link between pitch perception and speech perception.....	23
1.3 Pitch Processing for Normal Hearing subjects	27
1.3.1 Place pitch	27
1.3.2 Rate pitch.....	28
1.3.3 Place and Rate Pitch work together	28
1.3.4 Polyphonic Pitch	29
1.4 Loudness.....	29
1.4.1 Effect of loudness on pitch.....	30
1.4.2 Effect of changing pitch on loudness	30
1.5 What are Cochlear Implants.....	31
1.5.1 Signal processing basics of a Cochlear Implant.....	32
1.6 Pitch for Cochlear Implants	35
1.6.1 Place pitch	36
1.6.2 Rate pitch.....	39
1.6.3 Polyphonic Pitch	41

1.7	Loudness.....	43
2	Aim.....	44
2.1	Chapter 1.....	45
2.1.1	Research question.....	45
2.2	Chapter 2.....	45
2.2.1	Research questions.....	45
2.3	Chapter 3.....	45
2.3.1	Research questions.....	46
2.4	Chapter 4.....	46
2.4.1	Research questions.....	46
2.5	Chapter 5.....	46
2.5.1	Research questions.....	46
3	Publications.....	47
3.1	Publication 1.....	48
3.1.1	Abstract.....	49
3.1.2	Introduction.....	49
3.1.3	Methods.....	51
3.1.4	Results.....	56
3.1.5	Discussion.....	60
3.2	Publication 2.....	63
3.2.1	Abstract.....	64
3.2.2	Introduction.....	64
3.2.3	Methods.....	66
3.2.4	Results.....	70
3.2.5	Discussion.....	74
3.2.6	Conclusion.....	80

3.3	Publication 3.....	81
3.3.1	Abstract	82
3.3.2	Introduction.....	82
3.3.3	Methods	84
3.3.4	Results	88
3.3.5	Discussion	93
3.3.6	Conclusion	96
3.4	Publication 4.....	97
3.4.1	Abstract	98
3.4.2	Introduction.....	98
3.4.3	Materials and Methods	100
3.4.4	Results	103
3.4.5	Discussion	106
3.4.6	Conclusion	108
3.5	Publication 5.....	109
3.5.1	Abstract	110
3.5.2	Introduction.....	111
3.5.3	Materials and Methods	113
3.5.4	Results	116
3.5.5	Discussion	119
3.5.6	Conclusion	123
4	Discussion	124
4.1	General discussion.....	124
4.1.1	Pitch Processing in Cochlear Implants related to signal processing.....	127
4.2	Future Perspectives	133
4.2.1	Current focusing	133

4.2.2	Use of atypical biphasic pulses	134
4.2.3	Laser stimulation	135
4.2.4	Drug eluting Cochlear Implants	135
4.2.5	Pitch processing problems arising from the limitation of the auditory nervous system	136
4.2.6	Plasticity and training	137
4.3	Strength and limitations	138
4.3.1	Power Analysis.....	139
4.4	Final conclusions and clinical relevance	140
5	References.....	142
	Acknowledgement.....	176

List of publications

This doctoral thesis is based on the following articles published in or submitted to international peer reviewed journals:

Cone-Beam Volumetric Tomography for Applications in the Temporal bone.

PENNINGER RT, TAVASSOLIE TS, CAREY JP

Otology&Neurotology, Volume: 32, Issue: 3, Pages: 453-460, Published: 2011

DOI:10.1097/MAO.0b013e31820d962c

Perception of Pure Tones and Iterated Rippled Noise for Normal Hearing and Cochlear Implant Users. PENNINGER RT, CHIEN WW, JIRADEJVONG P, BOEKE E, CARVER CL, LIMB CJ

Trends in Amplification, Volume: 17, Issue: 1, Pages: 45-53, Published: 2013

DOI: 10.1177/1084713813482759

Effect of location and dynamic range on temporal pitch perception of cochlear implant users.

PENNINGER RT, MORRIS D, LIMB CJ, VERMEIRE K, LEMAN M, DHOOGHE I

Trends in Amplification (under review)

Experimental Assessment of Polyphonic Tones with Cochlear Implants.

PENNINGER RT, LIMB JC, VERMEIRE K, LEMAN M, DHOOGHE I

Otology&Neurotology, Volume: 34, Issue: 7, Pages: 1267-71, Published: 2013

DOI: 10.1097/MAO.0b013e3182923f04.

Experimental Assessment of complex Polyphonic Tones with Cochlear Implants.

PENNINGER RT, KLUDT E, DHOOGHE I, LEMAN M, BUECHNER A

Otology&Neurotology (accepted for publication)

Summary

A cochlear implant (CI) was developed for patients with bilateral severe to profound sensorineural hearing loss. It is the first commercially available device that replaces a human sense. CIs try to recreate the pattern of the auditory nerve firing by directly stimulating the endings of the auditory nerve. CIs perform very well with speech perception in quiet. However, CI users have difficulties with speech perception in noise and music perception. These difficulties in language and music perception are largely attributable to poor perception of pitch, the psychoacoustic correlate of stimulus frequency. In the human auditory system, pitch can be processed in two ways. In the cochlea, the basilar membrane acts as a frequency analyzer and activates the hair cells and auditory nerve fibers that are specifically tuned to the frequency of the incoming pitch and located spatially along the tonotopic gradient of the cochlea. This type of processing is referred to as “place pitch”, and is presumably critical for processing of a pure tone (PT). It has also been shown that the firing of auditory nerve fibers can “phase lock” to the frequency of the incoming pitch signal up to around 5000 Hz, and pitch information can be encoded by the rate of auditory nerve firing. This is referred to as “rate pitch”. It is controversial whether pitch is processed primarily using place or rate pitch, because the place and temporal codes usually co-vary with stimulus frequency in acoustic hearing.

Publication 1 describes imaging of the temporal bone. The ability of two recent CT scanners was quantified. High resolution CT images of the cochlea can help to estimate the predicted insertion depth of a CI pre-operatively and to check the insertion during the operation. It further helps to verify if the electrode array is positioned correctly post-operatively. Cone beam computer tomography provides optimal resolution for imaging of the temporal bone. Full insertion of a CI is a requirement for optimal speech and music perception.

Publication 2 describes a study conducted in the sound field where pitch ranking and melody recognition were performed. It was tried to measure the performance of CI subjects based on the spectral components of the sounds.

It was found that CI subjects are severely impaired compared to normal hearing (NH) subjects in pitch related tasks.

Publication 3 describes a study conducted with direct electrical stimulation. The CI speech processor was replaced with a research unit. Pitch ranking was tested on several electrodes and the goal was to find out on which electrode performance was best. Across subjects, performance was best on middle and wide dynamic range (WDR) electrodes.

CI subjects are severely impaired in the perception of polyphonic tones in sound field studies. They cannot discriminate between a single or two and three simultaneous tones. Instead, CI subjects show perceptual fusion meaning that they perceive two and three simultaneous pitches as a single pitch.

Publication 4 is a study about perception of one single tone and two simultaneous tones. It was performed with direct electrical stimulation. It was found that CI subjects are able to perceive polyphony when the two tones are applied on two different electrodes.

Publication 5 extends the findings of Publication 4. Perception of one, two and three superimposed tones was analyzed with direct electrical stimulation. Subjects were able to identify the number of pitches for one single or two and three simultaneous tones

Samenvatting

Een cochleair implantaat (CI) is ontwikkeld voor patiënten met een bilateraal ernstig tot zeer ernstig sensorieel gehoorverlies. Het is het eerste commercieel beschikbare apparaat dat een menselijk zintuig vervangt. Cochleaire implantaten herstellen partieel de gehoorfunctie door een rechtstreeks stimulatie van de zenuwuiteinden. De meeste geïmplanteerde patiënten vertonen een goed spraakverstaan in stilte. CI-gebruikers hebben echter problemen met spraakverstaan in lawaai en muziekbeleving. Deze beperkingen zijn grotendeels toe te schrijven aan een slechte perceptie van toonhoogte, het psycho-akoestisch correlaat van stimulusfrequentie. In het menselijk oor kan toonhoogte verwerkt worden op twee manieren. In het slakkenhuis werkt het basilaire membraan als een frequentie-analyser en activeert de (inwendige) haarcellen die op hun beurt de auditieve zenuwvezels prikkelen die specifiek zijn afgestemd op de frequentie van de binnenkomende toon. Zo ontstaat er een plaats-frequentie analyse, doorgaans omschreven als de tonotopie van het binnenoor. Dit type verwerking wordt aangeduid als "place pitch" en is waarschijnlijk essentieel voor het verwerken van een zuivere toon (PT). Ook is gebleken dat het afvuren van auditieve zenuwvezels een zgn. "phase locking" ('faze-vergrendeling') vertoont met de frequentie van de inkomende signaaltoonhoogte, dit tot ongeveer 5000 Hz. Zodoende kan toonhoogte-informatie worden gecodeerd door de snelheid van de neurale ontladingen. Dit wordt aangeduid als "rate pitch". Of de neurale encoding van toonhoogte-percepten hoofdzakelijk berust op plaats-mechanismen dan wel temporele mechanismen blijft een controversieel item omdat plaats-en temporele codes meestal covariëren met de stimulusfrequentie in een normale akoestische setting.

Publicatie 1 beschrijft beeldvorming van de temporale bot. Het vermogen van twee recente CT-scanners werd gekwantificeerd. Hoge resolutie CT-beelden van het slakkenhuis kunnen helpen om pre-operatief een goede inschatting te maken van de afmetingen van het slakkenhuis. Dit kan belangrijk zijn in de keuze van de te implanteren elektrode. Het helpt verder om post-operatief te controleren of de elektrode-drager correct is

gepositioneerd. Cone beam computertomografie biedt een optimale resolutie voor beeldvorming van het temporale bot.

Publicatie 2 beschrijft een studie uitgevoerd in het akoestisch veld waar toonhoogte ranking en melodie herkenning werden bevraagd bij CI patiënten door het aanbieden van geluidstimuli. Er werd geprobeerd om de prestaties van de CI proefpersonen op basis van de spectrale componenten van het geluid te meten. Het bleek dat de toonhoogte perceptie bij geïmplanteerde patiënten in belangrijke mate slechter was in vergelijking met normaal horenden.

Publicatie 3 beschrijft een studie waarbij gebruik werd gemaakt van directe elektrische stimulatie van het implantaat. De CI spraakprocessor werd vervangen door een research processor. Pitch ranking werd getest op verschillende elektroden en het doel was om uit te maken op welke elektrode de prestaties het beste waren. Globaal gezien waren de prestaties hetbest op de mediale elektroden en die met een groot dynamisch bereik (WDR).

Geïmplanteerde patiënten vertonen een slechte perceptie van polyfone tonen in vrij veldstudies. Ze kunnen geen onderscheid maken tussen een, twee of drie simultaan aangeboden tonen. CI proefpersonen vertonen perceptuele fusie wat betekent dat zij twee en drie simultaan aangeboden tonen als een enkele toonhoogte waarnemen.

Publicatie 4 is een studie over de perceptie van een simultaan aangeboden enkele toon met twee gesuperponeerde tonen. Ze werd uitgevoerd met directe elektrische stimulatie. Het bleek dat CI proefpersonen in staat zijn polyfonie waar te nemen wanneer de twee tonen op twee verschillende elektroden worden aangeboden.

Publicatie 5 breidt de bevindingen van Publicatie 4 uit. Perceptie van een, twee en drie gesuperponeerde tonen aangeboden via directe elektrische stimulatie werd geanalyseerd. Proefpersonen waren in staat om het aantal toonhoogten te identificeren voor een of twee en drie gelijktijdige tonen.

Abbreviations

ACE	Advanced Combination Encoder
AC	Alternating Current
AGC	Automatic gain control
ASR	Accumulated Semitone Range
CBCT	Cone beam computer tomography
CBVT	Cone-beam volumetric tomography
CI	Cochlear Implant
CIS	Continuous interleaved sampling
CT	Computer Tomography
CSF	Cerebrospinal fluid
DR	Dynamic range
F0	Fundamental frequency
FSP	Fine Structure Processing
HL	Hearing Loss
IHC	Inner Hair Cell
IRN	Iterated Rippled Noise
LF	Lower frequency
MDT	Modulation Detection Threshold
MF	Modulation Frequency
NDR	Narrow dynamic range
NH	Normal Hearing
OC	Organ of Corti
OHC	Outer Hair Cell
PET	Positron Emission Tomography
PT	Pure Tone
SAM	Sinusoidal Amplitude Modulation
SCDS	Superior Canal Dehiscence Syndrome
SP	Speech Processor
SPL	Sound Pressure Level
TFS	Temporal Fine Structure
UF	Upper Frequency
WDR	Wide Dynamic Range

1 Introduction

1.1 Anatomy of the ear

Sound is a pressure wave that is transmitted to the ear in the form of vibrations of the surrounding air. The ear constantly picks up sound waves and changes them into information that the brain can interpret. The sound reaches the outer ear, travels along the ear canal and induces constant vibrations on the tympanic membrane (eardrum). The eardrum is the barrier between the outer and the middle ear. Attached to the inner side of the eardrum is the malleus which transports the vibrations further to the incus and the stapes. They constitute the three middle ear ossicles. The stapes footplate is attached to the oval window, which is the barrier between the middle and the inner ear. Since the process of the malleus is longer than the incus, a movement of the eardrum produces a shorter but more powerful displacement of the stapes (1). The piston like movement of the stapes, which is attached with its footplate to the oval window, initiates a pressure wave in the fluids of the inner ear. The converted sound pressure wave travels along the cochlea from the oval window to the apex and down to the round window. This phenomenon has been described in detail for several types of mammalian cochleas (2) including squirrel monkeys (3), guinea pigs (4), chinchillas (4) and cats (5). The cochlea is a spiral-shaped structure that has about 2.5 turns around its axis. It contains three compartments, the scala vestibuli, scala media and scala tympani. Between scala media and tympani is the basilar membrane. It acts as a frequency analyzer and the converted sound pressure wave has a peak of oscillation at the characteristic frequency of the input sound. The auditory receptor cells, which are called hair cells, due to their hair-like stereocilia, which protrude from their apical surfaces, are arranged along the organ of Corti (OC) in the scala media (1;6). First the outer hair cells (OHCs) modify the signal. OHCs are essential for cochlear amplification by augmenting basilar membrane motion (2;7). Then the signal is transduced through endolymph movement to the inner hair cells (IHCs), which act as transmitters. The pulse train caused by the IHCs is transmitted along the auditory nerve to the brainstem. High frequency sounds generate action potentials at the basal region of the cochlea whereas low frequency

1 Introduction

sounds activate apical regions. This place-frequency transformation is called tonotopy of the cochlea (8;9). The tonotopic organization is preserved throughout the auditory system (10). This is one of the mechanisms encoding pitch. The loudness of an input is related to the energy of the sound. Early studies have hypothesized that it is coded in the normal hearing (NH) ear by the number of spikes per second (11). More recent studies show that the perceived loudness of a pure tone (PT) appears to be linked to both the number of spikes fired by a single neuron and to the recruitment of more afferent nerve fibers (12). Figure 1 shows an overview of the different components of the ear.

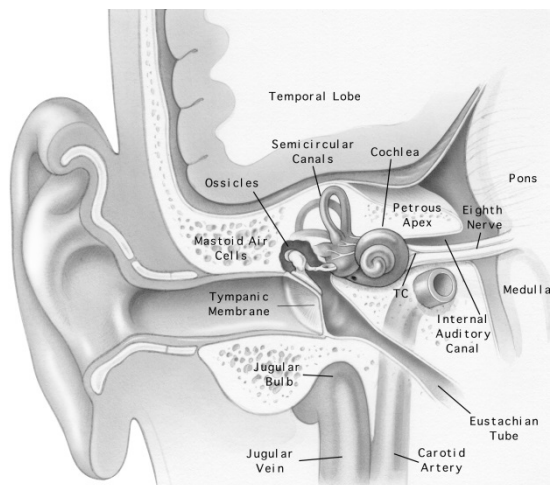


Figure 1 shows the anatomy of the human ear (copyright RK Jackler).

1.1.1 Organ of Corti (OC)

The inner ear is a coiled spiral that contains three fluid filled compartments: scala vestibuli, scala media and scala tympani (1). The scala vestibuli begins at the oval window and extends to the most apical part of the cochlea. At the level of the helicotrema, the most apical part of the cochlea, the scala vestibuli and tympani are joined together. The scala tympani extends from the helicotrema to the basal part of the cochlea where the round window is located. Scala tympani and vestibuli contain the same fluid (perilymph), which is an extracellular ionic fluid. The scala media is located in the middle of the cochlear duct and it is separated from the scala vestibuli through

Reissner's membrane. The basilar membrane separates it from scala tympani (Figure 2). The scala media is filled with endolymph. It has a high potassium ion concentration. Inside of the scala media on the basilar membrane is the OC, the actual hearing organ. The length of the OC averaged for nine Cochlear specimens is $33.13 \text{ mm} \pm 2.11 \text{ mm}$ (13).

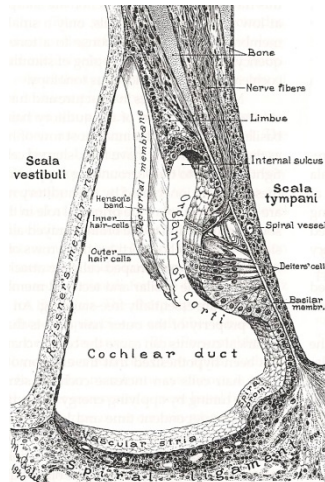


Figure 2 shows a cross-section of the organ of Corti (courtesy of Max Brödel Archives, The Johns Hopkins School of Medicine, Baltimore, MD (1)).

The OC, which is located on the basilar membrane, contains hair cells. They are organized into three rows of OHCs and a single row of IHCs. In total humans have about 15 000-20 000 hair cells. The IHCs which are less numerous than the OHCs are the primary sensory cells of the human auditory system. They stimulate 90-95% of the afferents in the cochlear nerve and are therefore the main source of auditory input (6). The IHCs send the signal to the dendrite that goes to the cell body of the auditory nerve and then to the axons leading to the brainstem (14). The IHCs have a characteristic frequency to which they are most responsive and it is defined by the mechanical properties of the basilar membrane and the mechanical and electrical properties of the hair cells themselves (6). The shape of the hair bundles is different between hair cells. IHCs have flat or U shaped bundles whereas OHCs are connected with V or W shaped bundles (6)

1 Introduction

(Figure 3). The OHCs are equipped with efferent synapses. They play a crucial role in signal amplification (2;6).

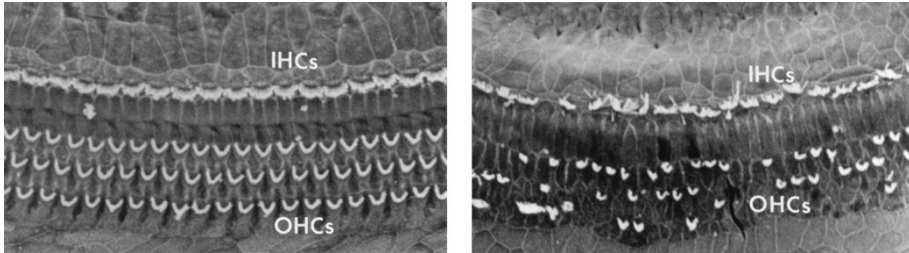


Figure 3 shows OHCs and IHCs in a NH (left) and in a damaged OC (right). Exact source of damage unknown. Damage could have occurred mechanically, chemically or from a genetic defect. (adapted from Ryan (15))

1.1.2 Mechanotransduction of hair bundles

The basilar membrane vibration stimulates the hair cells and initializes mechanotransduction. The hair bundle is directly involved in this process and it usually consists of a large number of stereocilia (16). The hair cells detect mechanical stimuli by submicron displacements of their stereocilia bundles (17). They are arranged in multiple rows with increasing heights and tip links extend from the tip of the shorter hair cell bundle to its larger neighbor (6;18). Deflections of the stereocilia result in sliding of neighboring stereocilia and the tip links are strained due to these movements. This movement has an effect on mechanical transduction channels. Beurg et al. used fast confocal imaging of fluorescence changes reflecting calcium entry during bundle stimulation to localize the transducer channels. It has been found recently that the channels are located at the tips of the smaller stereocilia and that the tallest stereocilia have no transducer channels (Figure 4), (17;19). The deflection of the bundle triggers the opening of ion channels (20). Only the movement of the hair cell bundles towards the tallest stereocilia results in opening of ion channels (21). Most of the current is based on potassium ions (20;22). As the hair cells depolarize, voltage dependent calcium channels open and this stimulates neurotransmitter release at synapses. This results in signal transduction towards the afferent neurons (23). The auditory nerve connects the cochlea to the brainstem and

transmits the mechano-electrically transformed stimulus to the central auditory system.

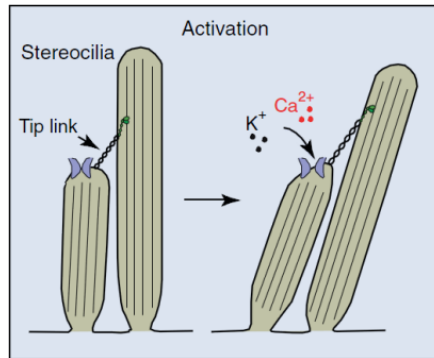


Figure 4 shows the activation of hair cells. The transduction channels are located close to the lower insertion site of the tip links. The tip links are, however, not directly attached to them. The channels are opened by the deflection of the hair cell bundle in the direction of the longest stereocilia (adapted from Kazmierczak et al. (19)).

1.1.2.1 Phase locking on afferent auditory neurons

The source of phase locking of the auditory neurons to the input stimulus is the cyclic increase and decrease of glutamate release from the IHC. This is caused by the alternating current (AC) receptor potential on the IHC membrane. Changes in this endocochlear potential prepare the hearing organ for optimal sound reception by fine-tuning its geometry and its mechanical properties (24). This AC receptor potential, which has the same frequency as the stimulus, is caused by cyclic opening of the transducer channels, whenever the bundle is deflected towards the tallest stereocilia. When the bundle is deflected in opposite direction, it depolarizes the membrane and closes the channels which hyperpolarizes the membrane. Hyperpolarization and closing is possible because at rest (zero deflection), the open probability of each transducer channel is about 0.15. Stated in another way, at every moment in the absence of a stimulus, 15% of all transducer channels present in a given IHC are open (25). This is due to the basic tension in the tip links, caused by the operation of the adaptation motor proteins such as some myosin isoforms in the upper or lower tip link density. Deflection in negative direction closes the channels, producing the hyperpolarization phase of the AC receptor potential, decreasing glutamate

1 Introduction

release and decreasing the spike rate below the spontaneous spike rate on the afferent neuron. The AC receptor potential decreases at higher frequencies and vanishes at about 5 KHz, due to the properties of the IHC membrane. Rose et al. measured phase locking to low-frequency tones in single auditory nerve fibers of the squirrel monkey. They found that the discharges of the auditory nerve fibers are usually locked to the phase of a sinusoidal stimulus for lower frequency tones (26). The recorded potentials were positive spikes whose amplitudes varied between several hundred microvolts to about 10 mv. The responses of primary neurons to the vowel in the syllable /da/ are shown in Figure 5.

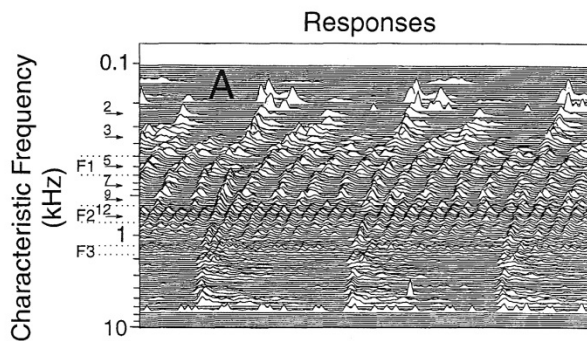


Figure 5 shows the responses of a large group of primary afferent neurons which were evoked by a segment of the syllable /da/. The responses are ordered vertically according to their respective characteristic frequency (27).

1.2 Pitch

1.2.1 What is pitch and what is it used for?

Pitch sensation can be described as the attribute of auditory sensation by which sounds may be ordered on a scale from low to high (28). Another definition of pitch is that it is the perceptual attribute of sound that can be used to produce melodies (29).

Pitch describes the absolute frequency of a musical note framed within the context of a musical scale (30). The range of frequencies that can be perceived by NH listeners ranges from 20 Hz to 20 kHz.

For harmonic, repetitive sounds, e.g. pure tone sinusoids, there is a one to one relationship between the perceived pitch and the frequency of a tone. For such complex harmonic sounds (i.e. not noise) the most important determinant of pitch in the spectral composition is the density of harmonics. This can be demonstrated by the psychoacoustic experiment entitled “the case of the missing fundamental”. This demonstrates that, apart from and better than the fundamental itself, the separation of the harmonics on the frequency axis, which is equal to the fundamental itself, dictates the sensation of pitch. If the fundamental frequency is not present, pitch is perceived either as the missing fundamental frequency or as the spectral pitch (31). For instance, a tone complex of 1800 Hz, 2000 Hz and 2200 Hz elicits a 200 Hz pitch sensation. It can be explained by a pattern recognition mechanism in the central auditory pathway that interpolates the series of harmonics until the lowest (and common) frequency component is found. Functional imaging studies suggest the existence of a “pitch processing center” anterolateral to the primary auditory cortex for all kinds of pitch processing including pitch direction analysis (32).

Noisy signals e.g. obstruent speech sounds or traffic noise, cannot convey the same sensation of pitch as pure tones or harmonic sounds. Nevertheless they can be ordered on an auditory scale from low to high based on the energy distribution over the spectrum. This “center of gravity” in the spectrum is a cue for consonant identification.

1.2.2 Link between pitch perception and speech perception

The fundamental frequency of voiced pitch can be a useful cue to segregate competing speech sounds (33). When two voices are presented at the same time it is easier for NH users to segregate them if the competing voice has a different pitch or pitch range (34). Therefore the two research fields (pitch perception and speech perception in noise) may be related with each other. NH users are able to use pitch cues to segregate competing speech sounds whereas CI users show no benefit when two competing sentences are spoken by talkers with different pitch (e.g. different genders) (35).

1 Introduction

1.2.2.1 Vowels

A vowel is a sound in a spoken language, which is pronounced with an open vocal tract, so that there is no built-up of air pressure at any time above the vocal chords (1). Examples are the English [a:] or [o:]. Vowel recognition relies in perceiving their formant structure. (36). Formants are the high-energy zones in the vowel spectrum. Formants in a vowel each have a frequency value (or rather a center frequency within a spectral zone), that depends on the contours of the vocal tract facilitating particular resonance effects (Figure 6). However, formants do not elicit (a) sensation(s) of pitch. The sensation elicited by the formant structure is not pitch, but rather timbre and, based on that, vowel identity. Pitch in vowels is determined by the fundamental frequency produced by the vocal folds. The same vowel (e.g. the same formant pattern) can be produced with a different pitch by simply varying vocal fold tension while at the same time keeping vocal tract contours stable. Formant patterns are frequency-based cues but not voice-pitch related cues. Consonants are speech sounds that are articulated with complete or partial closure of the vocal tract. Formant transitions are well known cues for consonant identification and they form another aspect of spectral energy distribution carrying information.

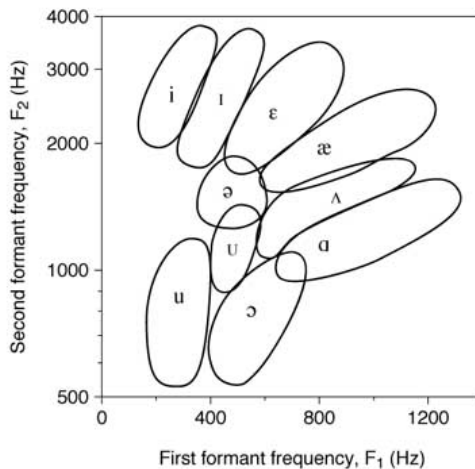


Figure 6 shows the frequency range of the first and second formants in the English language. (Copyright www.ncvs.org).

1.2.2.2 Prosody

Human speech does not only convey syntactic and semantic content. Emotional, non-verbal cues are also important information carriers (37). Prosody is described as the non-propositional cue of a language including intonations, stresses and accents (38). Accurate representation of fundamental frequencies (f_0) is important for understanding emotions in a sentence. The most important cue for emotional speech is the f_0 , followed by duration and intensity (39). Therefore pitch is also needed to perceive the intonation of a sentence. The change of the f_0 determines if a sentence is a question or a statement. It also expresses affection (e.g. in parentese (40)) and if a speaker is male or female (e.g. female speakers use more pitch variation). Furthermore it helps listeners (including children in the stage of language development) to parse incoming flow of speech into syntactic constituents. For phoneme recognition in the English language, the most important modulation frequencies are around 16-20 Hz (41-44).

1.2.2.3 Tonal languages

In tonal languages like Mandarin Chinese, the change of the frequency within a word determines its meaning. Mandarin is the most common spoken language worldwide (45). The meanings of the Mandarin words “ma” and “wang” are based on their pitch contours (Table 1). Therefore Mandarin word recognition strongly relates to the perception of pitch (46). Current Cochlear Implants (CIs) cannot deliver satisfactory performance in Mandarin tone recognition (47) and this may result in poor tone production of pre-lingually deafened children with CIs (48). The most important cue for Mandarin tone recognition is a change in voice pitch and this is manifested acoustically by changes in the f_0 and its associated harmonics (49;50).

妈, 'mā' mother	马, 'mǎ' horse	骂, 'mà' curse
王, 'wáng' king	网, 'wǎng' network	汪, 'wāng' pool

Table 1 shows some of the meanings of the words ‘ma’ and ‘wang’ which are defined by their individual pitch contour.

1 Introduction

1.2.2.4 Melody Recognition

Pitch perception is also important to perceive and enjoy melodies. Pitch is related to the repetition rate of a sound and is therefore related to the frequency of a PT sinusoid or to the f_0 of a more complex tone. In Western Music, each note corresponds to exactly one frequency. Multiplying the frequency with a factor of two represents the raise of an octave. Each octave consists of twelve semitones. An increase of one semitone is equivalent to the multiplication of the base tone with $2^{\frac{1}{12}}$. This is equivalent to an increase of 5.95%. Therefore an increase of “n” semitones is equivalent to multiplying the base tone with $2^{\frac{n}{12}}$.

Note	Frequency (Hz)	Semitone distance to C4	Multiplication factor related to C4
C4	262	0	$2^{\frac{0}{12}}$
D4	294	2	$2^{\frac{2}{12}}$
E4	330	4	$2^{\frac{4}{12}}$
F4	350	5	$2^{\frac{5}{12}}$
G4	392	7	$2^{\frac{7}{12}}$
A4	440	9	$2^{\frac{9}{12}}$
B4	495	11	$2^{\frac{11}{12}}$
C5	524	12	$2^{\frac{12}{12}}$

Table 2 shows frequencies of notes within an octave.

1.2.2.5 *Speaker segregation*

The heavy vocal folds of a man move at relatively low fundamental frequencies which produces a low voice pitch. Lighter and shorter vocal folds of women and children produce higher pitched sounds (1). Accurate representation of f_0 s, the depth of voice pitch modulations and the number of modulations per unit of time are important for separation of auditory streams from different sources and gender identification of speaker. This capacity is most important in order to track a certain voice which is overlapped by background noise. Whenever several persons are speaking at the same time, pitch information is needed to be able to extract each of the individual talkers. These “cocktail party effects” are very challenging for CI users mainly due to a lack of accurate pitch perception (51). The f_0 of the human voice ranges from as low as 75 Hz in a male to a height of 350 Hz in an adult female voice. The f_0 of young children may go up to 600 Hz (52).

1.3 Pitch Processing for Normal Hearing subjects

Pitch is coded by the NH ear with two fundamental mechanisms. One is called place pitch, which is related to the mechanics of the basilar membrane and the tonotopy of the cochlea. The second mechanism is called rate pitch and it resembles the firing rate of the auditory nerve. The auditory neurons phase lock to the oscillations of the auditory system (53). In a NH ear, the pitch ranges from about 20 Hz to about 20 kHz according to the Greenwood function (8). It determines the characteristic frequency at a given point along the scala tympani according to Equation 1.

$$f = A \cdot (10^{ax} - k) \text{ [Hz]} \text{ (Equation 1)}$$

The factor A equals 165.4 and the exponent a equals 0.06. The summand k equals 0.88 and x is the distance in mm from the apical end of the cochlea.

1.3.1 Place pitch

When a PT is applied to a NH ear it causes a travelling wave along the cochlea. It creates a maximum point of oscillation on a specific location on the cochlea based on the specific properties of the basilar membrane. It is narrow and stiff at the base, then widens and increases in compliance towards the apex (6). Its stiffness gradually decreases and the width

1 Introduction

gradually increases from the base (0.04 mm) to the apex (0.5 mm). The change in stiffness along the basilar membrane defines its motion. The region of maximum oscillation depends on the frequency of the input signal. The basilar membrane acts as a passive filter which splits up the input sounds according to their frequency components (3;54-57). Higher frequencies have peaks closer to the base of the cochlea whereas lower frequencies have them closer to the apical end of the cochlea. This is referred to as place pitch (58-60) and it is one mechanism that the cochlea uses to process pitch (1).

1.3.2 Rate pitch

An additional phenomenon for pitch perception is the rate pitch theory. The spikes of the auditory neurons elicited by a PT correspond to the periodic peaks in the amplitude of the input signal. This is referred to as rate pitch (61;62). The pitch of a PT is estimated by the phase locking of the auditory nerve (63;64). However, a single neuron does not fire on every peak of the signal due to refractory effects. In the refractory period the individual neuron cannot fire because it is still recuperating from the previous firing. The brain sums up the responses of all neurons and receives a pattern of firing that resembles the characteristic frequency of the input signal. This is called the volley theory (65). Rate pitch works up to around 4-5 kHz for NH (26). This is because above these frequencies the AC receptor potential of the IHCs is filtered out and the DC component is being left over. For pitches beyond 5 kHz, a place pitch cue has to be used. Several studies have shown that the ability to recognize melodies, music intervals and tones is drastically degraded for frequencies higher than 4-5 kHz (66).

1.3.3 Place and Rate Pitch work together

Complex tones consist of an f_0 and harmonics which are integer multiples of f_0 . The response of the auditory nerve becomes more complex but the principle of processing is the same as for a PT sinusoid. Both the place and the rate pitch cue are important for the NH ear to extract the f_0 pitch of a tone. Below 4-5 kHz rate pitch is most important and above place pitch becomes the dominant cue. These low frequencies where rate pitch is involved are most important for music and speech. Above 5 kHz tones lose their musical quality, and therefore it is not possible to transmit music

without a rate pitch cue. The just noticeable difference (JND) is the smallest difference in pitch that a person can identify. JNDs for pitches at lower frequencies are around 0.2% whereas they rise up to 1% at higher frequencies. Tuning curves constructed from single units originating at the base of the cochlea are less frequency specific (broader) than tuning curves from units at the apex of the cochlea. This means that the human ear is worse at detecting differences in high frequencies because there is no rate pitch cue anymore.

There is still scientific debate about place and rate pitch. (67). It is however commonly accepted that the neural coding of pitch arises from the precise timing of spikes within the auditory nerve up to 4-5 kHz (68;69). Higher frequencies seem to be based on a place pitch mechanism (70-73).

1.3.4 Polyphonic Pitch

Most Western music is polyphonic. Musical elements such as harmony, consonance, dissonance, and tonality are polyphonic (74). They enrich the sound just like polyphonic ringtones did when they replaced monophonic ringtones around 2000. A study by Donnelly et al. tested the perception of polyphonic pitches for NH subjects. Subjects listened to acoustically presented stimuli consisting of one, two and three simultaneous tones with different f_0 including PT and piano tones. NH subjects were generally able to correctly identify the number of simultaneous pitches (75).

1.4 Loudness

Loudness is primarily a psychoacoustical strength which is related to the amplitude of a sound. It is the subjective perception of sound pressure level (SPL). It is formally defined as “that attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud” (76). For tones that are above 40 dB SPL a loudness ratio of 2:1 is produced by a pair of stimuli that differ in 10 dB SPL. This relation seems to hold over the entire range of audible intensities (77). This phenomenon was confirmed in more recent studies (78;79). Coding of loudness is believed to be related to two different phenomena. Firstly the discharge rate of auditory nerve fibers (11;80-83). What contributes besides the discharge rate on one afferent nerve fiber is the additional recruitment of high threshold fibers

1 Introduction

contacting to the IHC (84). Liberman measured single unit recordings in cats and found that the afferent fibers innervating the IHCs have gradual thresholds. These range from low to medium until high number of spikes per second. The majority (61% in the cat) are fibers with a high spontaneous rate (e.g. high rate fibers), (85). Low spontaneous rate fibers have a higher threshold and these are additionally recruited. However, the increasing spike rate and the recruitment of more afferents cannot fully account for the 120 dB dynamic range (12;86); possibly because also more centrally located neuronal processing contributes. For instance, in data obtained from guinea pig experiments, the central projections of low rate and high rate primary afferent fibers are different and they target to different sub regions in the cochlear nucleus (87).

1.4.1 Effect of loudness on pitch

To study if the pitch of a tone is changed by varying its intensity, Stevens et al. asked NH subjects to adjust the intensity of a tone until it had the same pitch of a standard tone which had a different frequency. The intensity of a tone was increased and its perceived pitch was measured. For frequencies above 3000 Hz its pitch increased and for frequencies below 1000 Hz the pitch was lowered. Between 1000 and 3000 Hz there was little or no pitch change (77). A more recent study by Cohen found no significant differences in pitch as intensity was increased for low frequencies. For high frequencies, pitch changes were only 2% (88). It may be assumed that the relationship between pitch and intensity cannot be generally described and that changes in pitch that occur caused by loudness changes are rather small (89).

1.4.2 Effect of changing pitch on loudness

In order to give a listener a percept of equal loudness as the pitch of a tone changes, its sound pressure level has to be adjusted. The reason for this phenomenon is that the human ear is not equally sensitive to sounds which are presented in the same sound pressure level but with different frequencies. Equal loudness contours are used to give the listener a constant loudness when presented with a pure, steady tone.

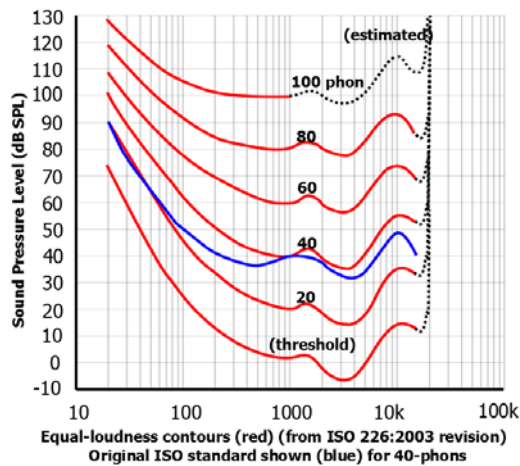


Figure 7 shows the equal loudness contours of several loudness levels.
(adapted from www.wikipedia.org)

1.5 What are Cochlear Implants

CIs were developed for patients with bilateral severe to profound sensorineural hearing loss. They are the first commercially available devices that replace a human sense. Performance for speech perception in quiet is generally good, but performance drops significantly in background noise or with competing speakers (90). Music appreciation is generally poor (91). CIs try to recreate the pattern of the auditory nerve firing by directly stimulating the endings of the auditory nerve. The neural elements, which are likely to be stimulated by the CI, are the spiral ganglion cells (92). The number of these cells is similar between deaf and NH subjects. CIs consist of an external and an internal component. The external component (speech processor) is usually worn behind the ear. It picks up the sound via a microphone and it contains a power source. It creates a radiofrequency signal that is sent via the coil through the skin to the internal component in the form of a pattern of electrical pulses, which stimulate different electrodes. The internal component consists of a receiver-stimulator and an electrode array. The receiver-stimulator is surgically placed under the skin. The electrode array is placed in the scala tympani of the cochlea. Signals are sent from the speech processor via the coil to the implant. With a CI, the non-functioning part of

1 Introduction

the ear (e.g. outer ear, middle ear and parts of inner ear structures) is bypassed and the sound signal is directly passed on to the functioning nerve. In theory, the brain is not able to know whether the nerve firing comes from basilar membrane motion or from an artificial electrical pulse created by an electrode. Whenever a certain electrode is stimulated the signal is interpreted by the brain as sound. Its location is linked to a specific frequency according to the tonotopy of the cochlea. CIs have been very successful in replacing the human sense of hearing throughout the world (93).

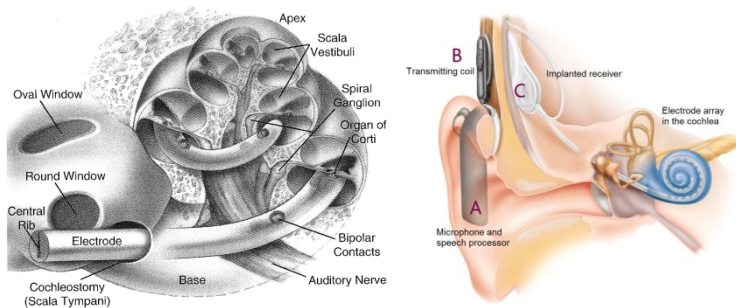


Figure 8A (left) shows a magnification of the inserted CI with cochleostomy approach (adapted from Zeng et al. (94))

Figure 8B (right) shows the main components of a CI (adapted from www.cochlear.com)

1.5.1 Signal processing basics of a Cochlear Implant

The basic concept of signal processing of a CI looks as follows: The input sound gets picked up by the microphone which is located on the speech processor. The input is pre-processed in the automatic gain control (AGC) unit. The AGC effectively reduces the volume if the input sound is strong and it augments the sound if the present input sound level is weaker than the previous sound. The average output signal level is fed back to adjust the gain of the next input level. The output of the AGC is processed by a pre-emphasis filter, which attenuates frequencies below 1200 Hz at 6dB/octave in Nucleus CIs. All other CI systems have similar AGCs. This filter increases the contribution of strong low frequency components that may contain important information for speech (e.g. weak consonants). The output of the

filter bank is forwarded to a set of bandpass filters which have filter frequencies that try to mimic the basilar membrane. The number of filters is equal to the number of channels (Table 3).

Electrode 22		Electrode 21		Electrode 20		Electrode 19		Electrode 18	
LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)
188	313	313	438	438	563	563	813	813	938

Table 3 shows the assigned frequency range of the bandpass filters of the five most apical electrodes in a Nucleus CI. Lower and upper frequencies (LF and UF) are shown in line three.

Then the envelope of each filter output is low-pass filtered and half wave rectified (95). The extracted envelope is then compressed on a logarithmic scale which maps the wide input dynamic range (DR) on the narrow electrical DR. The electrical DR is between 6-20 dB which is a lot lower compared to the DR of speech which is 50 dB (96;97). Finally, the amplitude envelopes are used to modulate charge balanced pulses at a rate of around 900 pps or higher (98-101). Sequential stimulation is mostly used by current CIs due to the uncontrollable loudness artifacts which are caused by parallel stimulation due to overlaps of currents of neighboring electrodes. Furthermore it prevents damaging electrochemical reactions (102). Short biphasic pulses are consecutively applied to each electrode. The reason for using biphasic pulses is to have a charge balanced signal (Figure 9). The biphasic pulse has a specific characteristic. Normally the phase width and the amplitude of the positive and the negative phase are similar. Recent experiments have been conducted with asymmetric waveforms which consisted of a short, high-amplitude phase followed by a longer low amplitude phase with opposite polarity. Using such pulses did improve pitch perception on apical electrodes in some testing conditions (103). These pulses were not used in the present experiments due to several reasons. First, the research interface from Cochlear Ltd. (Sydney Australia) does not permit the creation of these pulses. Only the research interface from Advanced Bionics Ltd. (Valencia, USA) allows the use of such pulses. Secondly, at University Ghent, the majority of subjects were implanted with

1 Introduction

CIs from Cochlear Ltd. Thirdly, the usage of a different research platform would have required a whole redesign of all the source codes and a resubmission of the test protocol to the Ethics committee. This would have delayed the publications significantly. Lastly, this paper came out in the end of 2011 and at that time several studies were already in progress.

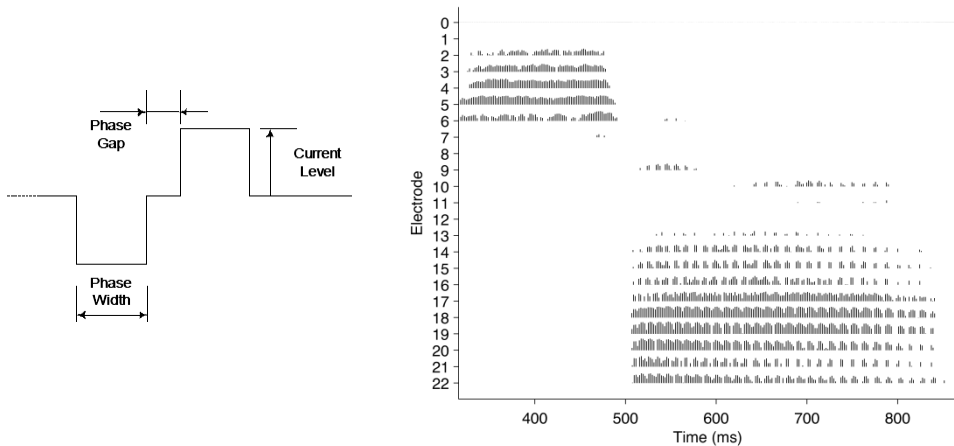


Figure 9A (left) shows a magnification of one of the pulses from the right side. The summed integral over each individual pulse has to be zero (charge balanced).

Figure 9B (right) shows the stimulation pattern of a CI. From 300-500 ms it processes the sound “s” and from 500ms until 900 ms it processes the sound “a”. For “s”, only electrodes that are in the basal region of the cochlea are activated. For “a” lower frequencies are activated.

1.5.1.1 Continuous interleaved sampling (CIS) and Advanced Combination Encoder (ACE)

The input signal which is picked up by the microphone is decomposed into a small number of bands by the speech processor (usually 16-22). This process is performed with fast Fourier transformation or a bank of bandpass filters. Then the envelopes are extracted in each band and they are used to modulate biphasic pulses that are sent to the electrodes for stimulation. The number of envelopes, which is equivalent to the number of stimulation sites, that are selected in each stimulation cycle, differs between continuous interleaved sampling (CIS) stimulation and advanced combination encoder

(ACE). CIS was first introduced by Wilson et al. (104) and it was originally designed to address the problem of channel interactions through the use of interleaved non simultaneous stimuli. The ACE strategy does not stimulate all electrodes in each stimulation cycle. Instead, it selects the electrodes which have the highest amplitudes within a stimulation cycle. This process is called “maxima selection” (105). The ACE strategy is based on the CIS strategy and it is used frequently in Nucleus CIs. In the CIS strategy, a fixed number of envelopes (8–10) is computed, and only the corresponding electrode sites (8–10) are used for stimulation (106).

1.5.1.2 Fine structure processing (FSP)

A speech signal can be decomposed into a slowly varying envelope and a high frequency carrier. Fine structure processing (FSP), which was developed by MED-EL (Innsbruck, Austria), was designed to overcome the limitations of envelope-based coding strategies. Channel specific sampling frequencies are used in low-frequency (apical) channels (107) which has proven to give better speech perception in noise compared to CIS strategies in at least some subject populations (108). FS4-p provides increased temporal resolution which could possibly improve performance for bilaterally implanted subjects.

1.5.1.3 High Resolution (HiRes120)

One of the challenges in today’s CIs is to increase the number of frequency information with a limited number of fixed electrodes. By stimulating two neighboring electrodes at the same time an intermediate pitch percept can be created. This process is called “current steering”. Advanced Bionic’s (Valencia, USA) High Resolution 120 (HiRes 120) strategy is based on the implementation of active current steering in a CI system. Using current steering did improve appreciation of music subjectively in some subject populations (109).

1.6 Pitch for Cochlear Implants

As for NH subjects, pitch enables CI subjects to classify if a signal is higher or lower. However, the creation of the sensation of pitch differs fundamentally from that of NH subjects due to the fact that the outer, the middle and parts of the inner ear are no longer involved in the transmission of sound.

1 Introduction

1.6.1 Place pitch

Just like for NH subjects, stimulating basal locations of the cochlea induces a perception which is high-pitched. The more apical the electrodes are located, the more the perception gradually becomes lower-pitched. This pitch percept is entitled “place pitch” (58;60). Place pitch enables CI users to hear a pitch difference when different electrodes are stimulated. The number of discriminable pitches is limited by several factors: the length of the electrode array, the insertion depth, the number of electrodes, which goes along with the spread of excitation, and the condition of the spiral ganglion cells. Due to spread of excitation, stimuli need to be separated by more than one electrode to be reliably discriminated from each other (110-112). The place pitch perception of CI users affects speech perception. There is a clear link between the electrode discrimination capability and speech perception (113). In the following chapters the main limitations of place pitch will be described.

1.6.1.1 Electrode length and position

Anatomic variations of the cochlea influence the position of the CI (114-116). Greater clinical benefit might result from optimization of the insertion and location of scala tympani electrodes (117). The size of the human cochlea does not change significantly with age but it varies highly between subjects. The difference in size ranges up to several millimeters and this has an effect on the insertion depth of the implant array (117). Implants that are longer can cover a wider frequency range in large cochleas but they may be too long for short cochleas. For optimal hearing restoration it would be necessary to measure pre operatively how long the predicted insertion depth will be and then use longer or shorter implants according to the cochlear duct length of each patient. Stimulating a greater region of the cochlea can improve speech understanding and sound quality (118;119). Accurate Computer Tomography (CT) imaging of the temporal bone might help predict the insertion depth of a CI before the actual surgery (120). The small size of the cochlea and the density of the bone surrounding it make it impossible to exactly predict the insertion depth of a CI. Cone beam CTs (CBCT)s generally perform better than multi-slice CTs (121;122). CBCTs provide sufficient resolution to check the electrode insertion during and after surgery and to give a reasonable

prediction of the insertion depth pre operatively (123). The insertion depth does not only depend on the electrode array length but also on the type of the electrode array (124). The array consists of straight arrays which change their shape as they are inserted and are usually positioned close to the lateral cochlear wall. Perimodiolar arrays are pre-curved and they are usually positioned close to the medial wall of the cochlea. Therefore perimodiolar electrode arrays may lead to deeper insertion angle even though they are generally shorter in length due to the fact that the path that they take is shorter along the medial wall compared to the lateral wall.

1.6.1.2 Selectivity of the place of stimulation due to spread of excitation

Ideally it would be best to stimulate single neurons or small groups of neurons. But the scala tympani, in which the CI gets inserted, is filled up with perilymph. The perilymphatic fluid consists of ions and it is therefore highly conductive. The current used to stimulate the auditory nerve spreads out to different neighboring nerve fibers which make the stimulation broader and less focused. The results of spread of excitation studies indicate that only about seven or eight channels can therefore be stimulated per stimulation cycle (125-132). Spread of excitation is also thought to be a major contributor to the huge difference in pitch perception observed for CI subjects which ranges in JNDs from 1 to 24 semitones (133). As the number of electrodes increases beyond e.g. 22 contacts/array, the amount of current which spreads out to neighboring electrode contacts increases. This leads to situations where the subjects are unable to hear a place pitch difference between the stimulation of two neighboring electrodes. Therefore improving the number of place pitches by increasing the number of electrodes is not an option.

1.6.1.3 Frequency mismatch

Each bandpass filter defines the frequency range for each electrode. A huge portion of the acoustic frequency range (5 Hz to 5 kHz) has to be mapped to only e.g. 22 electrodes. Table 4 shows the characteristic frequencies of five electrodes on a standard electrode array from Cochlear Ltd.

1 Introduction

Electrode 22		Electrode 17		Electrode 12		Electrode 6		Electrode 1	
LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)	LF (Hz)	UF (Hz)
188	313	813	938	1563	1813	3563	4063	6935	7938

Table 4 shows the assigned frequency range of the bandpass filters of some electrodes in a Nucleus CI. Lower and upper frequencies (LF and UF) are shown in line three.

To analyze the place pitch perception studies were conducted with unilateral CI subjects who had a NH contralateral ear. These CI subjects were asked to pitch match their most apical electrode to a sine wave played on the contralateral ear (134-138). It would be intuitive to think that the most apical electrodes create a pitch percept which is generally at around 1-1.5 kHz based on the tonotopy of the cochlea according to Greenwood (8). The perceived place pitch (through a CI) can differ by as much as from one to two octaves from acoustically stimulated pitch on the contralateral ear (134-138). CI subjects show a huge frequency mismatch between the actual input sound and their perceived tone. This confirms anecdotal observations from CI subjects who describe that voices through the CI initially sound high pitched. However, electric pitch perception often shifts in frequency, sometimes by as much as two octaves, during the first few years of implant use (139).

1.6.1.4 Non stochasticity of stimulation

The CI stimulates groups of neurons in a deterministic manner. Ideally the stimulation would be stochastic and individual on each neuron. This is not possible due to spread of excitation. Recent studies have tried to create artificial stochasticity and tested its effect on pitch perception. Due to spread of excitation no real stochasticity could be reproduced. Instead it was tried to overlap the actual signal with a random conditioner pulse. No significant advantage has been shown over conventional stimulation by Carlyon et al. (140). Rubinstein et al. suggested that pseudospontaneous activity of auditory nerve fibers can enhance neural representation of temporal detail (141).

1.6.2 Rate pitch

As mentioned previously in chapter 1.3.2, the responses of auditory nerve fibers to PTs are synchronized or phase locked to the frequency of the input signal. The temporal pattern in the stimulation train of an electrode results in very accurate phase locking of the auditory nerve to the frequency of the input (142-145). The brain interprets the frequency of the phase locking also as a pitch percept (26). The perceived pitch is similar to the NH pitch percepts of amplitude modulated noise (61). When CI subjects are asked to rank rate pitches they can do this generally well for frequencies below 300 Hz (140;146-151). Beyond this limit there is no more pitch percept available. There are a few subjects, however, who can perceive rate changes which go up to around 1 kHz (152-154). There are several factors which may help to explain the lack of temporal coding in CI subjects for higher frequencies. Although phase locking to high frequencies has been observed in recently deafened animals (155), creating it in deaf humans may be more challenging. Firstly because the appearance of phase locking is negatively correlated with duration of deafness (143). Increased duration of deafness together with the higher level of neural degradation may impair the triggering of a phase locking effect in humans. The duration of deafness of animals used in studies is generally shorter than for usual CI candidates. This results in little neural degradation of the tested animals. CI candidates, however, usually have a relatively long period of deafness. Therefore it may be challenging to restore phase locking in CI subjects. Furthermore, pitch perception is also affected by variations in levels (147;150;156-158). Therefore variations in current levels are likely to result in uncontrolled variations in pitch which furthermore impede performance.

Secondly, in NH subjects, increasing the intensity of a stimulus leads to a shift in place of excitation. In electrical stimulation the place remains constant and only the current level is increased.

Lastly, the incapability of CIs to encode phase differences among different neurons due to low carrier rates (and not to provide a global temporal cue) could be another explanation for the poor rate pitch performance (159;160).

1 Introduction

1.6.2.1 Ways of creating a sensation of rate pitch

While using their speech processor, the stimulation rate per electrode is usually constant at e.g. 900 Hz. By replacing the speech processor with a research unit, the rate per channel can be adjusted. The rate at which biphasic pulses are applied to a specific location on the implant array is proportional to the frequency perceived (at least for low frequencies up to around 300 Hz). Changing the rate of pulses is called “pulse frequency modulation”. Explicit coding of pitch can be accomplished by varying the carrier rate on one or more electrodes (62;146;161).

A rate pitch sensation can also be created by modulating the amplitude of a high frequency pulse train. The rate of pulses of the pulse train is called “carrier rate” (61;162;163) (Figure 10). This type of pitch encoding is called “Sinusoidal Amplitude modulation”.

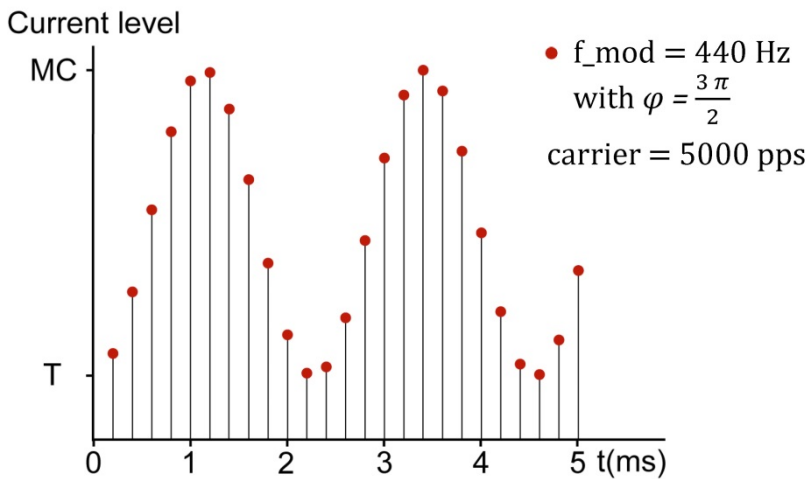


Figure 10 shows the principle of sinusoidal amplitude modulation. The current amplitude modulates between the threshold (T) level, the lowest detectable amplitude, and the maximum comfortable (MC) level, the loudest tolerable current level.

A basic concept of sampling theory is the Nyquist theorem (164). It describes that the carrier rate has to be at least two times as high as the maximum modulation frequency. It is therefore required that the carrier rate designed

to transmits pulses up to 50 Hz has to be at least 100 Hz. For creating a rate pitch sensation in the inner ear another factor has to be considered. A low carrier rate of e.g. 100 Hz used to transmit a 50 Hz amplitude modulated sine may give rise to two pitch percepts at the same time: the 100 Hz pitch from the carrier and the 50 Hz pitch from the modulation frequency. To avoid abnormalities in the relationship between the pitch of the carrier and the modulation pitch, the carrier rate must be at least four times higher than the modulation frequency (165). Amplitude modulated rate pitch sensations can also be created by the speech processor as it extracts the envelope of a signal and maps it on a corresponding electrode. With a carrier rate of e.g. 900 pps per channel, amplitude modulations until around 225 Hz can be sent to the implant. Landsberger compared pitch discrimination with sine, sawtooth, modified sawtooth and square modulations on pitch perception. It was concluded that the sine and the two sawtooth waveforms provided the same JNDs. Frequency discrimination performance with the square modulation was worse compared to the other waveforms (166). Since it is unsure which exact phenomenon causes these effects the author suggests that when designing a speech processing strategy in which modulations are used to convey F0, sine and sawtooth waveforms are interchangeable. The choice in waveform may be driven by concerns other than frequency discrimination (e.g. interaction on multiple electrodes within a speech processing strategy or power consumption).

1.6.3 Polyphonic Pitch

Most studies on pitch perception have focused on pitch discrimination where subjects are required to detect whether two sounds differ in pitch. Other pitch tests are related to pitch ranking where subjects are asked to listen to two sounds in sequence and judge which one has the higher pitch. However, various elements of music often occur simultaneously (e.g. melody, harmony, rhythm and timbre) (75). The study by Donnelly et al. found that CI users are severely impaired in the perception of two and three simultaneous pitches. CI subjects demonstrated perceptual fusion meaning that they frequently perceived two and three simultaneous pitches as a single pitch (75). A polyphonic pitch sensation can be created for CI subjects based on polyphonic place pitch and polyphonic rate pitch.

1 Introduction

1.6.3.1 Polyphonic Place Pitch

By stimulating two different electrodes at the same time with the same sinusoidally amplitude modulated frequency, a polyphonic place pitch sensation is created. The polyphonic place pitch is made stronger by increasing the distance between the two stimulated electrodes.

1.6.3.2 Polyphonic Rate Pitch

A Polyphonic rate pitch sensation is created when two different sinusoidally amplitude modulated frequencies are applied to two different electrodes. Increasing the difference between the two frequencies increases the strength of the polyphonic rate pitch. A sensation of polyphonic rate pitch only can be created by stimulating one electrode with two amplitude modulated stimuli at the same time. To do this, the carrier rate has to be increased to e.g. 10 000 pps and the pulses for each e.g. 5 000 pps carrier are then presented alternatively.

1.7 Loudness

The loudness in CIs is coded by the amount of charge applied to a location in the cochlea. Increasing the amount of current can be done either by increasing the amplitude of a biphasic pulse or by increasing the phase width (see Figure 9). Both of these two manipulations lead to increased spike rate and therefore increased loudness. Ideally, the input DR of the speech processor would be 120 dB which is comparable to NH users' DR. CI users however have a DR which covers only about 10-20 dB of their electrical DR (167;168). Therefore, the acoustic DR has to be compressed to fit to the greatly reduced DR of CI users. It has long been assumed that speech has a DR around 30 dB (169-171). However, more recent studies show that the DR of speech ranges around 40-50 dB (172) and goes up to 70 dB for German sentences (173). Perceptual studies also confirm that the DR of speech is higher than 30 dB. Word and sentence recognition increases when the speech presentation level is raised from 64 to 99 dB SPL (174). Current speech processors provide between 20 and 80 dB SPL and the most frequently used speech processors provide only 30 dB (97).

Several methods have been evaluated to increase the DR for CI users. These include high rate conditioning pulses that can increase the spontaneous activity of the auditory nerve. By adding a 5000 pps pulse with spontaneous activity Hong and Rubinstein have showed that the DR increased significantly (175). Similar results can be obtained by adding of background noise (176). Other studies have shown that an increase in stimulation rate decreases threshold and comfortable levels which would promote current focusing (99;177-183). Other studies have suggested to manipulate the pulse shape to increase DR (184;184-187) which helped to increase DR in some cases. There is still a lot of debate about whether or not high stimulation rate or use of conditioning pulses have a practical benefit. They may enhance neural representation and promote speech understanding (99;188-198). Other studies suggest that they may lead to increased temporal interaction and therefore reduce speech understanding or have no effect at all (101;140;181;199-204).

2 Aim

The aim of the presented thesis is to analyze pitch perception for CI users from a multidisciplinary perspective. In Publication 1 it was tried to find out most appropriate ways for imaging of the temporal bone. In Publication 2 it was tried to tease apart envelope and temporal fine structure (TFS) cues in a psychophysical experiment. Pitch ranking and melody recognition were measured in the sound field with pure tones (PT) and iterated rippled noise (IRN) stimuli. Based on Publication 2 it was concluded that it would be best to replace the speech processor with a research unit to have more exact control of the stimulation pattern on the implant array. In Publication 3 it was found that most subjects performed well in a pitch ranking task based on rate pitch with at least one of the tested electrodes. These results served as a basis for the following polyphonic pitch perception studies. One single or two simultaneous tones were presented in Publication 4 and the complexity of the experiment was increased in Publication 5 by adding another three pitch condition.

2.1 Chapter 1

Publication 1 describes imaging of the temporal bone. The ability of two recent CT scanners was quantified. High resolution CT images of the cochlea can help to

- 1) Estimate the predicted insertion depth of a CI pre-operatively
- 2) Check the insertion during the operation and
- 3) Verify if the electrode array is positioned correctly post-operatively

2.1.1 Research question

Does a CBCT scanner have better spatial resolution compared to a multi slice CT scanner for temporal bone imaging?

2.2 Chapter 2

Publication 2 describes a study conducted in the sound field where pitch ranking and melody recognition were performed with pure tones and iterated rippled noise stimuli. It was found that CI subjects are severely impaired compared to NH subjects in these tasks. Especially iterated rippled noise stimuli are very hard to perceive with current CI processing strategies.

2.2.1 Research questions

- Do CI subjects demonstrate deterioration in performance for pitch ranking and melody recognition with iterated rippled noise compared to pure tone stimuli?
- Do CI subjects perform worse than NH subjects in both tasks and with iterated rippled noise and/or pure tone stimuli?
- Do NH subjects show fewer differences in performance between iterated rippled noise and pure tones in pitch ranking and melody recognition compared to NH subjects?

2.3 Chapter 3

Publication 3 describes a study conducted with direct electrical stimulation. The CI speech processor was replaced with a research unit. Pitch ranking was tested on several electrodes and the goal was to find out on which electrode performance was best.

2 Aim

2.3.1 Research questions

- Do electrodes with a WDR perform better on a pitch ranking task than electrodes with a narrow dynamic range (NDR)
- Does the performance gradually improve as the site of stimulation is moved from the base to the apex?

2.4 Chapter 4

CI subjects are severely impaired in the perception of polyphonic tones in a sound field study. Publication 4 is a study about perception of one single tone and two superimposed tones.

2.4.1 Research questions

- Does identification of two tones improve if the distance on the CI array between the electrode pair is increased?
- Does identification of one tone improve as the stimulation site on the electrode array is moved from base to the apex due to a greater match between the modulation frequency of the stimulus and the characteristic frequency of the neurons?
- Is a sole polyphonic rate pitch cue sufficient to perceive a polyphonic pitch cue?

2.5 Chapter 5

Publication 5 extends the findings of Publication 4. It was however performed in a different country with a different study population. Perception of one, two and three superimposed tones was analyzed.

2.5.1 Research questions

- Does the location on the implant array have an impact on performance?
- Does the difference in frequency between the tones have an impact on performance?
- Is a sole polyphonic place pitch cue sufficient to perceive a polyphonic pitch cue?

3 Publications

3 Publications

3.1 Publication 1

Title

Cone-Beam Volumetric Tomography for Applications in the Temporal Bone

Authors

Richard T Penninger

Tanya S Tavassolie

John P Carey

Journal

Otology&Neurotology 32:453-460 (2011)

3.1.1 Abstract

Hypothesis: Cone Beam Volumetric Tomography (CBVT) has better spatial resolution compared to Multi Slice Computed Tomography (MSCT) in temporal bone imaging for superior canal dehiscence (SCD).

Background: Imaging of SCD has traditionally utilized MSCT, but the ability to resolve thin bone next to low-radiodensity brain and inner ear fluids at the interface of the superior canal (SC) with the middle cranial fossa can be adversely affected by partial volume averaging, errors in registration of successive slices, and other factors. CBVT may offer advantages in these regards and may have better spatial resolution for this application.

Methods: Five cadaveric temporal bones were scanned using both CBVT and MSCT. The information content at the interface of the SC and the middle cranial fossa was measured for each method using spatial differential transformations. The ability of each method to resolve progressively smaller interfaces between bone and fluid was measured by creating a spatial grating model from a human temporal bone.

Results: The information content and spatial resolution were superior for CBVT compared to MSCT.

Conclusion: The gold standard for diagnosis of SCD has been MSCT, but CBVT offers improvements in information content and spatial resolution at the interface of the SC and the middle cranial fossa.

3.1.2 Introduction

Modern diagnosis of diseases of the human temporal bone depends heavily on high-resolution computed tomography (CT). At most major centers, high-resolution CT is performed with scanners that have linearly-arrayed emitters and detectors that travel around a circular gantry. The patient is moved along the z-axis either stepwise between image acquisition periods (step-scan mode) or continually (spiral mode). Image processing “stacks” the planar datasets to create volumetric datasets that can be rotated and projected into any informative plane. For example, projections in the plane of the superior semicircular canal (SC) and orthogonal to it have become essential for the diagnosis of superior canal dehiscence (SCD) (205).

Collimation of effective x-ray beam widths to 0.5 mm have allowed axial CT imaging to generate nearly isotropic voxels in volumetric datasets. Isotropy

3 Publications

of voxels is essential to generating smooth images in reconstructions of planes outside of the direct imaging plane (206). This has largely eliminated the “step-offs” seen in reconstructions of images in vertical planes after axial acquisition. Nevertheless, stacking of axial images still creates irregular transitions and potential loss of spatial resolution in vertical planes. In the case of SCD, this loss of resolution occurs exactly at the area of interest – the top of the SC – and can negatively impact diagnostic reliability. In the extreme case, MSCT might make a dehiscence appear where there is none, which would constitute a diagnostic error. As noted by Hendee et al. (207), “Every imaging examination exposes patients to some element of risk. That risk comes from unwarranted exposure to radiation, as well as from false-positive or false-negative examination results.”

A relatively new CT technology is that of cone-beam volumetric tomography (CBVT). In CBVT, the emitter and detector are not linear arrays. Rather, an emitter casts a cone of photons while a 2-dimensional plate detects the transmitted ones. The arrangement is similar to that used in fluoroscopy, except that in CBVT, the emitter and detector are rotated around the subject as in MSCT. A 3-D volume of data is collected directly, not created by stacking up planar datasets. In theory, this should improve resolution and prevent the degradation that occurs due to imperfect registration of planes in MSCT.

CBVT units that deliver small radiation doses (e.g., 5mA and 120kV) compared to conventional MSCT (250mA and 135kV) have been gaining popularity because such CBVT units can be deployed directly to otolaryngology and dental offices. A disadvantage of these CBVT units is that the use of low energy photons results in less image contrast. However, there is potentially better spatial resolution compared to MSCT when inherent tissue contrast is high. This should be the case with SCD, where thin bone must be detected between low radiodensity brain and perilymph, as well as with delineation of the stapes and cochlear implant arrays.

There has been one other study comparing the resolution of CBVT vs. MSCT scanners for evaluation of temporal bone structures (208). The investigators in that study scanned cadaveric temporal bones using both technologies and

asked three experts to perform qualitative assessment of the anatomic structures in the images on a three-point scale. They found that the flat panel prototype provided “better definition of fine osseous structures of temporal bone than that of currently available MSCT scanners.”

In the present study, image data from cadaveric temporal bones scanned with both CBVT and MSCT are compared quantitatively with measures of information content and spatial resolution.

3.1.3 Methods

3.1.3.1 CT Hardware

Comparisons were made between two commercially-available CT units. The Toshiba Aquilion 64–slice MSCT scanner (Toshiba America Medical systems, Inc., Tustin, CA) was used in step-scan mode at 120 kV, 300 mA and 0.5 mm collimation with a field of view of 18 cm. For CBVT, the i-CAT 17/19 office scanner (Imaging Sciences, Hatfield, PA) was used at 120 kV, 5 mA with a field of view of 16.5 cm width x 13.5 cm height.

3.1.3.2 Computer Tomography Principle

To construct a 3D volumetric dataset in CBVT, various 2-dimensional planar scans have to be created as demonstrated in Figure 1. The x-ray source and the detector rotate in small steps 360 degrees around the patient. The acquired data are transformed with inverse radon transformation to obtain the radiodensity values of the intervening tissues.

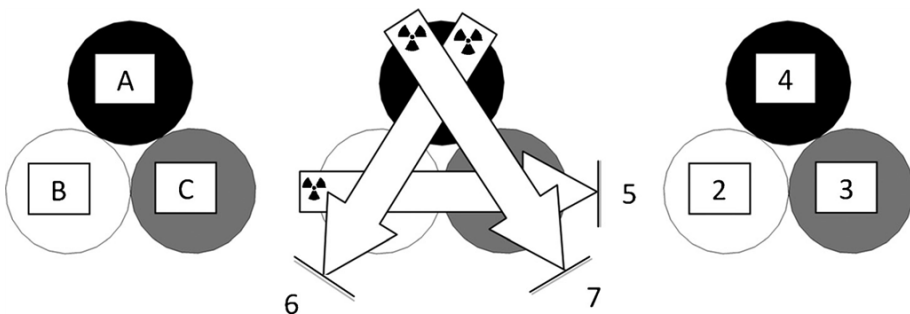


Figure 1: This shows the steps for processing the images.

3 Publications

To construct the full 3D dataset, the steps in Figure 1 have to be repeated over many slices with very small slice thickness. The result is a volumetric dataset that consists of 3D voxels. That is necessary for 3D reconstructions and therefore essential for rotating images into any plane (Figure 2).

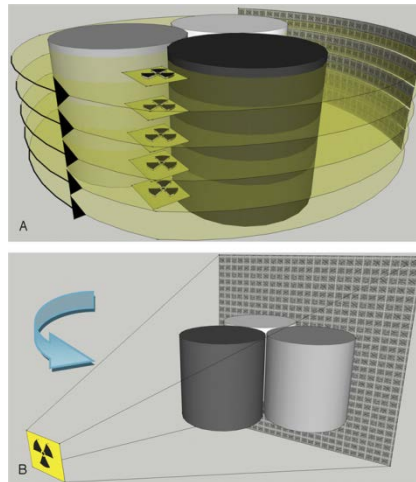


Figure 2A: The principle of the MSCT. Multiple detectors are used, and each is much narrower. In one rotation, a limited area is scanned, and slices must be stacked up to acquire a whole volume dataset.

B: The principle of CBVT. The x-ray receptor is much larger compared with the MSCT, and the image acquisition is completed in one rotation.

3.1.3.3 Acquisition and processing of the images

Each of 5 cadaveric temporal bones and a bony “phantom” were scanned with MSCT and CBVT, and after extracting the image information in DICOM format, they were post processed in Vitrea (Vital Images, Minnetonka, MN) to adjust windowing and rotation angle for reconstructions in the plane of the SC. A proprietary edge detection algorithm was used in post-processing by each unit. DICOM data were finally extracted to MATLAB 7.9.0.529 (R2009b) (The MathWorks Inc., Natick, MA) for analysis (Figure 3).

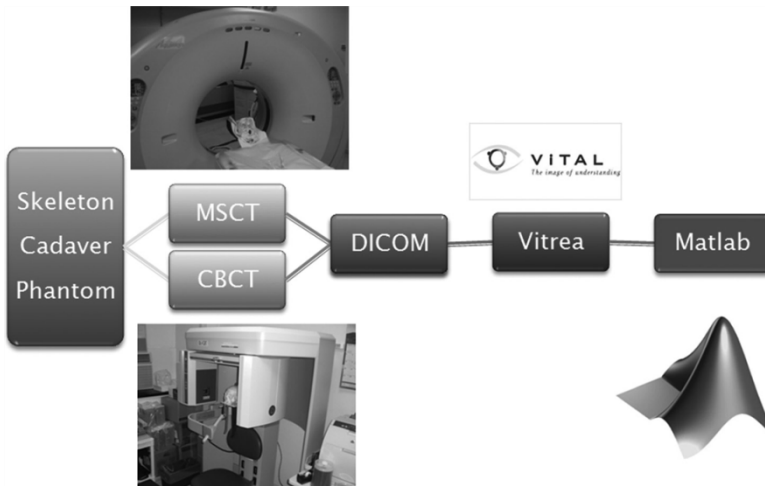


Figure 3 shows the steps for processing the images.

For each scanning method, the image that best captured the entire SC in its plane was selected in the reconstructions using Vitrea. In order to make images of the SC from the two methods comparable, the data were then transformed in MATLAB so that the canals would appear to have the same size and would be centered at the same location in each image (Figure 4). The transformations were as follows: Grayscale images (Figure 4a) were converted into black and white images (Figure 4b) by defining a threshold between black and white based on the distribution of grayscale values of the image histogram. Then the boundaries between black and white were calculated and the areas that were completely white were colored similarly (Figure 4c). MATLAB's *roipoly* function was then used to detect the center of the SC, which was the roundest object in the image data (Figure 4d, yellow area). Using the coordinates of this center that were just determined, the original grayscale image (Figure 4a) was then aligned on this center, and the image was cropped with a square box measuring 7 times the radius of the canal. This created a normalized image of the SC that could be quantitatively compared across the different scanning methods (Figure 5). For images that were to be compared for the smoothness of transitions between bone and

3 Publications

surrounding materials, the images were interpolated in order to equalize the image matrices across techniques.

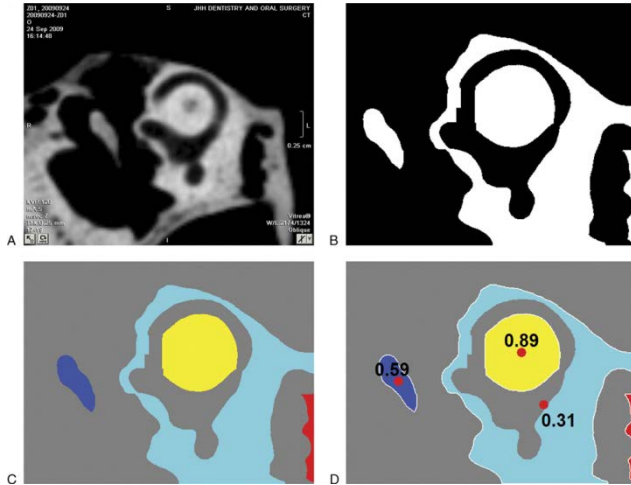


Figure 4A: Original image produced by Vitrea rotated to demonstrate the plane of the SC with a minimum region of interest. **B:** Black and white image after applying `imopen` and `imclose` with MATLAB. **C:** Contiguous bony areas are segmented and similarly colored. (Note: This function requires a closed geometric region, but in this example, a region of the SC lumen was not captured in the image (A) and still appears to have bone, probably because this canal is not entirely planar. The bone here was replaced with black pixels (in B) to complete a round SC only for the purpose of calculating the center of the canal in the next step. These altered pixels were not considered in the quantitative comparisons of resolution.) **D:** The dots are the area centers of the yellow, blue, and turquoise areas. If the areas were perfectly round, the associated index would be 1.

3.1.3.4 Comparing images with spatial differentials

To magnify differences in spatial resolution, we calculated the spatial derivative of each pair of pixels in the horizontal axis of each image. This means that a matrix of grayscale differences was created: $[X(2)-X(1), X(3)-X(2), \dots, X(n)-X(n-1)]$, where $X(i)$ is the grayscale value of the i^{th} pixel moving across a line in the image from left to right. We multiplied each of these small grayscale difference values by 2^8 in order to make them visible, and we then created a new image of the spatial derivatives (Figure 6). Sharp transitions from light to dark in the original images – i.e., differences in

adjacent pixels - manifest themselves as bright zones in these derivative images.

3.1.3.5 Temporal bone phantom

In order to show the limits of the spatial resolution for the tissue characteristics of the temporal bone, a “phantom” was created from a dry temporal bone by making a spatial grating with microscopic holes drilled into the squamous portion ranging from 0.8 mm diameter in 0.05 mm steps to 0.3 mm. The centers were 1.5 mm apart and there were 5 or 6 holes in each vertical set. The phantom was put into water in order to mimic the radiodensity of neighboring cerebrospinal fluid. The phantom was scanned with both MSCT (Figure 9b) and CBVT (Figure 9c).

3.1.3.6 Measuring the smoothness of radiodensity transitions in the CT images

The grayscale values encountered while moving along a vertical bar through each image from top to bottom was analyzed (Figure 7). The bar was chosen to pass through the SC where the bone overlying it was the thinnest. Grayscale values along this vertical line moving from top to bottom were plotted as y-values above the x-axis moving left to right. High values in the plot show white pixels and low values in the plot show black pixels. The higher the graphed value, the whiter is the original image at that specific location.

3.1.4 Results

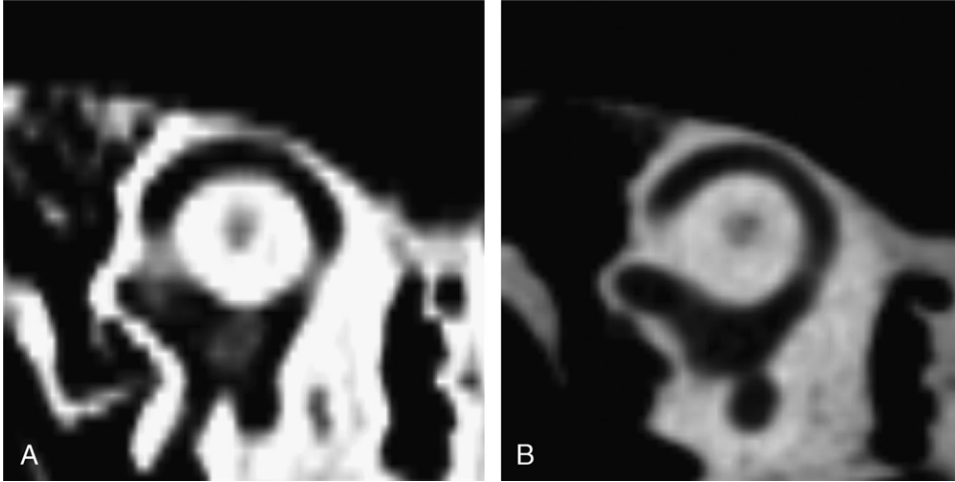


Figure 5 shows the normalized images from MSCT (A) and CBVT (B).

Figure 5 demonstrates that when images from MSCT (a) and CBVT (b) of the same temporal bone are scaled and centered, it appears that there are smoother transitions in the CBVT image (b) between bone and surrounding soft tissue or air than in the MSCT image (a). This allows the thin bone overlying the SC to be more readily seen in the CBVT image. The data presented here are for one temporal bone in which there was thin bone overlying the SC; however, similar findings were obtained from the other four specimens. The main reason for the difference in transition smoothness derives from the fact that the flat-panel detector used in CBVT has inherently greater resolution than the linear detectors used in MSCT. For example, the image in Figure 5a has an underlying image matrix measuring 322 X 322 (0.10 megapixels), whereas the image in Figure 5b has an image matrix of 437 X 437 (0.19 megapixels).

The transitions between bone and surrounding soft tissue or air are seen more clearly after application of the differential transformation (Figure 6). In the CBVT image (b) the transitions are more finely graded than in the MSCT image (a). The subtlety of these transitions is portrayed graphically in Figure 7. The transitions going from top to bottom along a line through the SC are

displayed in blue for MSCT and in red for CBVT. The graph of the MSCT has constant derivatives in certain intervals whereas the CBVT has a smoother inclination of its graph. Grayscale transitions are smoother for CBVT (red) than they are for MSCT (blue), and in some regions MSCT derivatives reach a constant maximum value over successive pixels whereas CBVT derivatives continue to show pixel-to-pixel changes. Thus, more information about radiodensity is conveyed by the CBVT transitions than by the MSCT transitions.

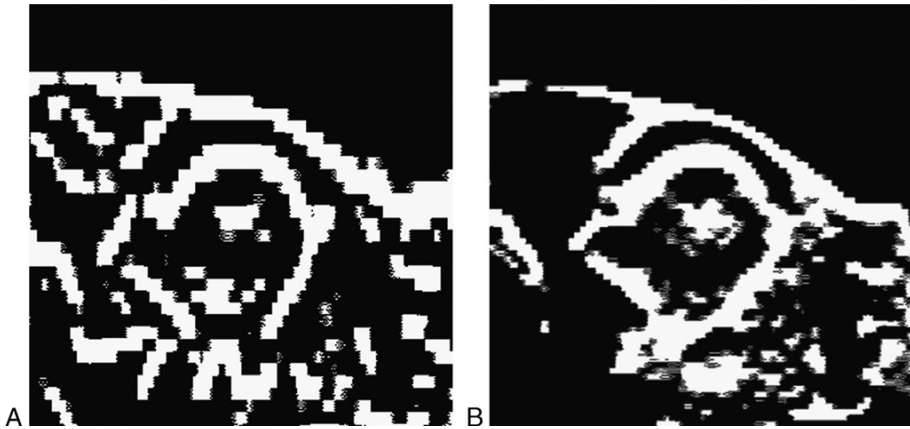


Figure 6 shows The derivative of the MSCT image (A) and the CBVT image (B).

3 Publications

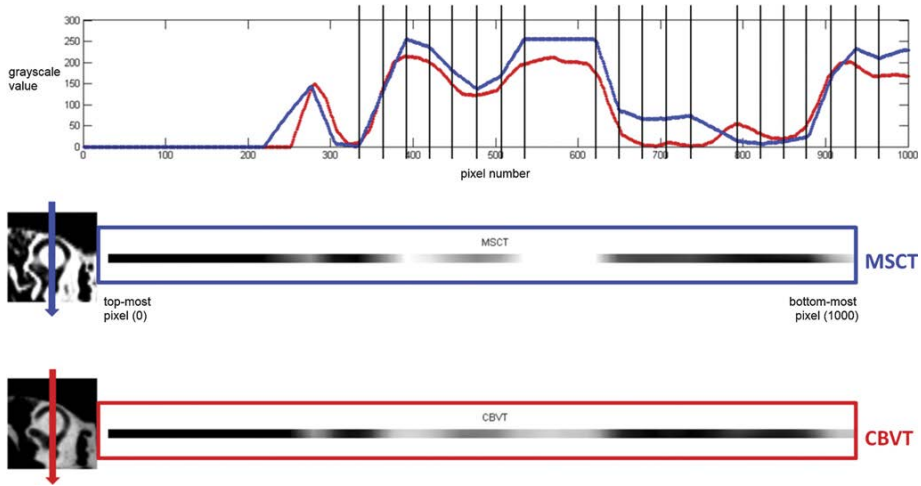


Figure 7 shows grayscale transitions along a vertical line through the SC for MSCT (blue) and CBVT (red).

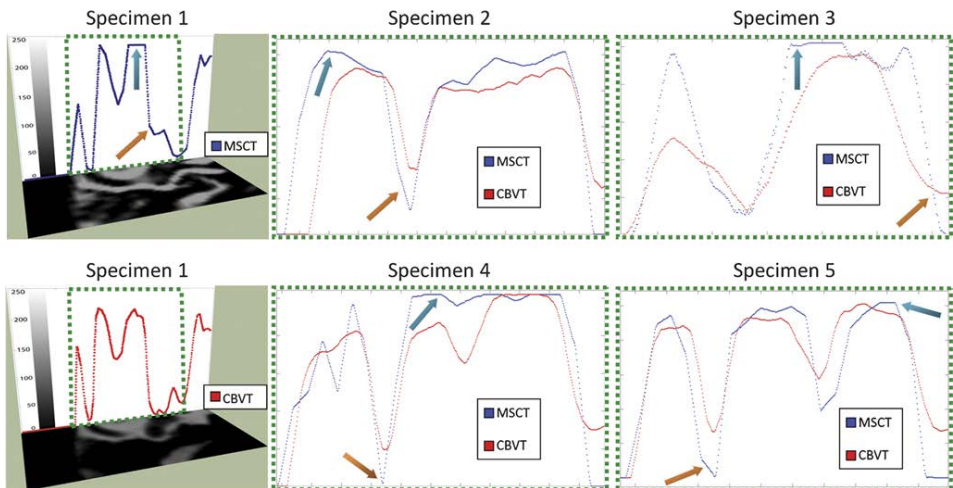


Figure 8: Left panels: transitions in grayscale intensity encountered on a linear path through the image of the SC in the plane of the canal for images reconstructed from MSCT (blue) and CBVT (red). The first (leftmost) transition is where the opening of the SCD normally occurs. Right panels: graphical representations of the grayscale transitions from 4 other cadaveric specimens. Despite some variability between specimens, a consistent finding is that transitions in MSCT (blue traces) are more abrupt (orange arrows) and have more regions of saturation (blue arrows) than the transitions in CBVT. This is due to the fact that the spatial resolution in the CBVT is almost twice as high as in the MSCT.

Spatial resolution in CT is typically determined empirically in line pairs per cm using industry-standard “phantoms,” objects of defined radiodensity marked with gratings of different radiodensity. However, the actual resolution of a CT scanner in a given diagnostic application depends on the radiodensities of adjacent materials. Thus, the ideal method for determining resolution of bone overlying the SC would be to use materials with radiodensities like those of bone, brain, dura, and cerebrospinal fluid. In order to model this, we used the squamous portion of a cadaveric temporal bone as a template, micro drilled holes in it, and filled them with water to mimic the density of the soft tissue (Figure 8a).

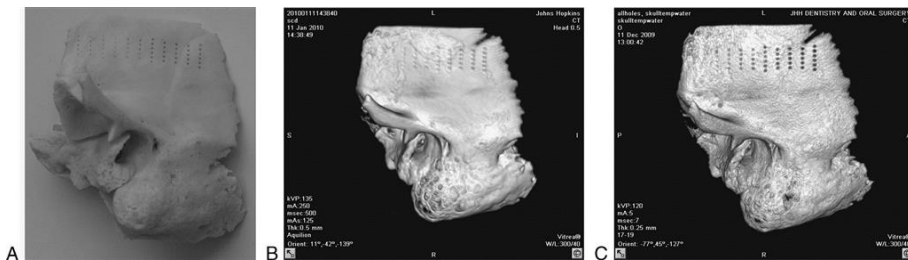


Figure 9:

The temporal bone phantom (a) and corresponding reconstructions from MSCT (b) and the CBVT (c). In creating the phantom, holes were drilled of identical diameter in each column.

The right-most column began with six holes measuring 0.8 mm in diameter. The next column to the left has five holes measuring 0.75 mm in diameter. Subsequent columns have holes decreasing in diameter by 0.05 mm every column. The number of holes follows the pattern: 6-5-5-5-6.

Figure 9b and 9c show that MSCT has less spatial resolution compared to CBVT for detecting these small apertures containing fluid in the temporal bone. The reconstructions from the CBVT scan demonstrate that the micro holes can be counted accurately down to the 7th column of holes (0.45 mm in diameter). In contrast, the reconstruction from MSCT only allows accurate detection of the holes down to the 5th column (0.60 mm in diameter). Furthermore, it can be seen from the largest holes in the CBVT

3 Publications

reconstruction that these holes actually penetrate fully through the squamous temporal bone, but this is not apparent from any of the holes depicted by the MSCT reconstruction.

3.1.5 Discussion

The results of this study suggest that CBVT may have advantages over MSCT for the diagnosis of SCD because CBVT demonstrates greater information content in the transitions from soft tissue to thin bone as well as greater absolute resolution of small apertures in bone that contain low-radiodensity material like water. There are some caveats to note before extrapolating the present findings to the intact and living human temporal bone. First, the present results were obtained from cadaveric temporal bones with intact surface soft tissues but without the full skull or adjacent brain. Thus, the absorption of photons by these other structures is not accounted for in our model, and the reader should be cautious not to conclude that the exact parameters used in this study of isolated temporal bones could be immediately applied for clinical scanning of the whole head. Second, the detection challenge in SCD is to detect thin bone bounded by fluid (perilymph) on one side and dura, cerebrospinal fluid, and brain on the other. Our cadaveric temporal bones may have had air inside the canals at the time of scanning, and they had only dura, not brain or CSF, overlying the canals. Finally, the phantom is a test of the detection of small holes filled with fluid surrounded by dense bone, which is opposite the task of detecting thin bone next to soft tissue as in SCD. But these limitations of the methods apply to the tests of both CBVT and MSCT, yet notable differences between the methods were nevertheless found. It may be the case that the cadaveric temporal bone model would overestimate the differences that would be obtained between CBVT and MSCT in the intact and living human temporal bone. Still, it is worthwhile to begin with this in vitro model to know whether or not it is worthwhile to pursue these studies in patients, especially considering that it may not be possible to do a comparative study of two different CT scanning modalities in the same individuals because of the additional radiation exposure that it would require.

Our first finding was that CBVT has greater information content than MSCT in the grayscale values representing the transitions between bone and

surrounding soft tissue (or perhaps air) in our temporal bone model. The visible manifestation of this is that CBVT reconstructions have less “step” artifacts at the transitions between bone and adjacent tissue than do the MSCT reconstructions. The first explanation for these step artifacts in the MSCT reconstruction is the inadequate size of its image matrix. In Figures 5 and 6, for example, the image matrix for the CBVT reconstruction was 90% larger than the image matrix for the MSCT reconstruction. The effective image matrix can be improved for MSCT by decreasing the field of view, but the field of view used for this study (120 mm) was that typically used in our high resolution temporal bone MSCTs. Another contributing factor for the step artifacts in MSCT could be errors in the registration of planar datasets acquired each time the scanner stepped along the z-axis. In contrast, data are acquired along the entire length of the z-axis in one revolution of the cone beam emitter and flat-panel detector. There is no need to register planar data sets acquired at separate times; thus, there is far less probability for step and motion artifacts.

Our second finding was that CBVT demonstrates better spatial resolution compared to MSCT for small apertures filled with water in the temporal bone. Although standard industry phantoms are typically used to define resolution for CT scanner technology, the actual resolution for a particular application depends upon the relative radiodensities of the adjacent materials. By creating our own phantom with a spatial grating drilled into an actual temporal bone, and by filling the small apertures with water to mimic the radiodensity of adjacent soft tissue, we attempted to reproduce as closely as possible the challenge for resolution of bone overlying the SC. CBVT was able to resolve apertures down to 0.5 mm in diameter, whereas MSCT could only resolve apertures down to 0.6 mm in diameter. This improvement in resolution could reduce the probability of error in detecting thin bone overlying the SC, an error which could potentially lead to a recommendation for surgery for SCD.

CBVT has other potential benefits in imaging the temporal bone. First, CBVT has a lower total radiation dose than MSCT (121). Second, the time for image acquisition is faster due to the fact that the volumetric data are acquired in only one rotation. Third, due to the lower power requirements

3 Publications

and smaller size of CBVT machines, they can be located in offices for potential “turnkey” utilization. It should be borne in mind that there are inherent disadvantages of CBVT. The most important of these is CBVT has a lower dynamic range of the photons, which results in lower contrast resolution. Furthermore, the slower scintillation material (CsI) that is used gives a poorer signal-to-noise ratio in very thin slices (208).

As in all imaging applications, the ideal technology must be matched to the question to be answered. For the detection of bone interfacing with soft tissue or air, CBVT has a number of advantages because of the inherently high contrast in the radiodensities of the tissues involved. Our results suggest that the use of CBVT should be further explored for imaging of the temporal bone, and that it may have particular promise for the problem of detecting bone overlying the superior semicircular canal.

3.2 Publication 2

Title

Perception of Pure Tones and Iterated Rippled Noise for Normal Hearing and Cochlear Implant Users

Authors

Richard T Penninger

Wade W Chien

Patpong Jiradejvong

Emily Boeke

Courtney L Carver

Charles J Limb

Journal

Trends in Amplification 17 (1) 45-53 (2013)

3 Publications

3.2.1 Abstract

Cochlear Implant (CI) users typically perform poorly on musical tasks, especially those based on pitch ranking and melody recognition. It was hypothesized that CI users would demonstrate deterioration in performance for a pitch ranking and a melody recognition task presented with iterated rippled noise (IRN) in comparison to pure tones (PT). Additionally it was hypothesized that normal hearing (NH) listeners would show fewer differences in performance between IRN and PT for these two tasks.

In this study, the ability of CI users and NH subjects to rank pitches and to identify melodies created with IRN and PT was assessed in free field in a sound isolated room. CI subjects scored significantly above chance level with PT stimuli in both tasks. With IRN stimuli their performance was around chance level. NH subjects scored significantly above chance level in both tasks and with all stimuli. NH subjects performed significantly better than CI subjects in both tasks. These results illustrate the difficulties of CI subjects to rank pitches and to identify melodies.

3.2.2 Introduction

Cochlear implant (CI) users often have difficulties with music perception (209), despite relative overall success with speech perception. These difficulties in music perception are largely attributable to poor perception of pitch, the psychoacoustic correlate of stimulus frequency. In the human auditory system, pitch can be processed in two ways. In the cochlea, the basilar membrane acts as a frequency analyzer and activates the hair cells and auditory nerve fibers that are specifically tuned to the frequency of the incoming pitch and located spatially along the tonotopic gradient of the cochlea. This type of processing is referred to as “place pitch” (58-60), and is presumably critical for processing of a pure tone (PT). It has also been shown that the firing of auditory nerve fibers can “phase lock” to the frequency of the incoming pitch signal up to around 5000 Hz (26), and pitch information can be encoded by the rate of auditory nerve firing. This is referred to as “rate pitch” (61;62). It is controversial whether pitch is processed primarily using place or rate pitch because the place and temporal codes usually co-vary with stimulus frequency in acoustic hearing (210-212).

Iterated rippled noise (IRN) is created by a cascade of add and delay cycles (213). The pitch of the IRN is shown by performing an autocorrelation with the signal (Figure 1A). The delay of the first peak of the autocorrelation which is not at delay zero (highlighted) is the reciprocal of the pitch frequency. The strength of the pitch is determined by the relative height of this highlighted peak (214) and it increases with stimulus duration (215).

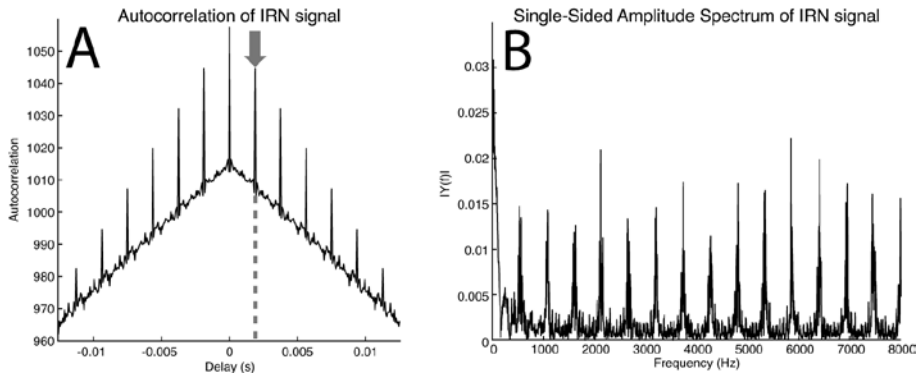


Figure 1. Figure 1A shows the autocorrelation and the spectrum of a 523.25 Hz iterated rippled noise (IRN) stimulus. The pitch of the IRN is indicated by the location of the first peak of the autocorrelation next to lag zero. It occurs at lag 1.91 ms. The reciprocal of the lag is the pitch frequency. Figure 1B shows the spectrum of the same stimulus. The IRN has linear spacing of the spectral peaks in the frequency domain.

The spectrum of IRN has equal amplitude peaks at integer multiples of the fundamental frequency of the pitch (Figure 1B). The spectrum of PT sinusoids has peaks at the fundamental frequency of the pitch only. NH listeners can discriminate between two IRN stimuli up to around 5 kHz (216). Most current CI speech processing strategies only process the envelope of the input signal, whereas temporal fine structure is the basis for IRN pitch processing (216). Although there has been considerable data published on spectral ripple discrimination with CI users, there has been no publication using IRN. This differs from spectral ripple discrimination in that IRN has a distinctive temporal property which is absent in the spectral ripple stimuli used in previous studies (217-219).

3 Publications

In this study, IRN and PT were used to investigate the differences in place and rate pitch perception for CI and NH users in a pitch ranking and a melody recognition task. It was hypothesized that CI subjects would demonstrate deterioration in performance for pitch ranking and melody recognition with IRN compared to PT. It was further hypothesized that normal hearing (NH) users would perform better than CI users and would show fewer differences in performance between IRN and PT in the two tasks.

3.2.3 Methods

Ten NH subjects and ten CI subjects participated in the study. The age range was 36-75 years (mean=53, SD=11). The biographical information of all CI subjects in this study is shown in Table 1. All persons enrolled were native English speakers and the CI subjects had unilateral CI implantation. Although the CI subjects had a huge range of hearing loss and CI exposure, they were all performing well at speech perception in quiet. This research was approved by the Institutional Review Board of the Johns Hopkins University School of Medicine. Written consent was obtained from each participant. Each participant underwent pitch ranking and melody recognition tasks as described below.

Table 1 CI Subject demographics.

ID	Age	Sex	Cause of hearing loss (HL)	Duration hearing loss (years)	CI exposure (months)	Device
1	60	M	Unknown	25	36	Nucleus Freedom
2	48	F	Unknown	25	96	Clarion
3	36	F	Unknown	3	19	Nucleus Freedom
4	57	F	Otosclerosis	39	48	Nucleus Freedom
5	75	F	Autoimmune	28	140	Nucleus 22
6	54	M	Sudden HL	3	24	Sonata
7	39	M	Unknown	2	12	Hi-Res 90K
8	52	F	Unknown	6	72	Hi-Res 90K
9	50	F	Unknown	2	15	Hi-Res 90K
10	59	M	Meniere	24	24	Hi-Res 90K

3 Publications

Table 2 NH Subject demographics.

ID	Age	Sex
1	49	F
2	30	F
3	26	F
4	36	F
5	30	F
6	33	M
7	31	F
8	27	M
9	24	F
10	46	F

3.2.3.1 Stimulus Generation

IRN was generated by delaying and adding white noise to itself. The IRN tones were created with 8 iterations and with gain of one (see Figure 2). The output waveform of one delay and add stage served as the input to the next stage (“add original” configuration) and for noise-delays (d) between 2 and 30 ms, IRN stimuli have a pitch corresponding to $1/d$ kHz (213).

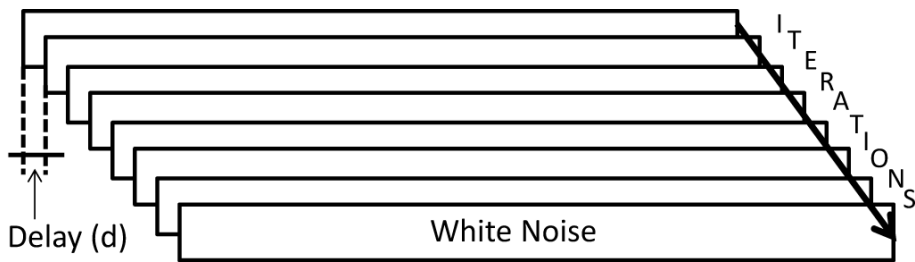


Figure 2: The process of making iterated rippled noise (IRN). IRN was generated by delaying white noise by a certain amount of ms and adding it to itself. The perceived pitch was $1/\text{delay}$. This process was repeated eight times (eight iterations).

All pitches (IRN and PT) were generated using Audacity 1.2.5 (Dominic Mazzoni, open source) at a sampling rate of 44.1 kHz. The IRNs were then filtered using 4th order Butterworth filters between 150 Hz and 4.5 kHz to minimize any spectral cues. Stimuli were randomly presented in a soundproof booth through a single calibrated loudspeaker (Sony SS-MB150H) at a presentation level of 75 dB sound pressure level through an OB822 clinical audiometer (Madsen Electronics). The speaker was positioned directly in front of the listener. For CI subjects, the contralateral ear (which was profoundly impaired in all individuals) was occluded with an earplug to diminish the effects of any minimal residual hearing, and no hearing aids were used. Each pitch was constructed such that an eighth note was exactly 250 ms in duration. They were presented at a tempo of 120 beats per minute. Each note was given linear rise/decay ramps of 50 ms to reduce onset clicks and to minimize transients in the filter bank outputs.

3.2.3.2 Pitch ranking

The pitch ranking task was implemented using a two interval, two alternative, forced choice (2I2AFC) test. On each presentation, two pitches were played sequentially. The listener was asked to identify which of the two pitches was higher in frequency. The minimum tested interval was one semitone, and the maximum was twelve semitones. The pitch pairs used consisted of semitone steps within an octave ranging from 261.63Hz to 523.25 Hz. Each interval was tested six times per subject using either PT or

3 Publications

IRN in a randomized fashion. PT and IRN stimuli were run in intermixed blocks with randomized intervals.

3.2.3.3 Melody recognition

In this test, the listeners were asked to identify the recordings of twelve common melodies from a closed set. Individual pitches were combined to create isochronous, eighth notes melodies in order to reduce potential rhythm cues that might be used for melody identification. The accumulated semitone range (ASR) of all notes of each melody was calculated. ASR ranged from 26 to 73 semitones. The following melodies were selected for their general familiarity. In parenthesis is the corresponding ASR. “Auld Lang Syne” (47), “Deck The Halls” (53), “Frère Jaques” (58), “Frosty The Snow Man” (54), “London Bridges” (32), “Mary Had A Little Lamb” (26), “Ode Of Joy” (26), “Somewhere Over The Rainbow” (73), the opening theme of “The Sound Of Music” (29), “Swing Low Sweet Chariot” (51), “Twinkle, Twinkle Little Star” (30) and “Yankee Doodle” (46). All melodies were presented for twelve seconds to prevent the use of melody length as a cue. Prior to testing, all listeners were given a list of the twelve melodies and were asked to indicate their familiarity with each melody. Unfamiliar melodies were included in the test but were removed from the final analysis. Each melody was presented three times using PTs and IRNs in a randomized fashion. PT and IRN stimuli were run in intermixed blocks with randomized intervals.

3.2.4 Results

Results from both experiments were found to be not normally distributed. The Kolmogorov-Smirnov test was used to check the data distribution. Both of the experiments in this study used forced choice procedures, therefore the results can be analyzed with the binomial probability distribution. A binomial experiment consists of repeated trials where the outcome of each trial is labeled either success or failure. The probability of success remains constant from trial to trial. In evaluating the result of a forced-choice experiment, the first question is whether the subjects were merely guessing. For the pitch ranking procedure it was considered that the null hypothesis of the probability of success on each trial was 50%. For the melody recognition task there were twelve melodies to select meaning that the probability of success in each trial was 8.33% (1/12). If the resulting probability (p) is less

than the criterion value $\alpha = 0.05$ that is generally accepted for statistical significance. It was then concluded that it was unlikely that the null hypothesis is true, i.e. the subjects were most likely using some cue in the stimuli to obtain a good score. The second question is whether the subjects performed better with IRN or PT stimuli. Simon (1997) advocated the unorthodox approach of simulating the experiment on a computer, sometimes known as the Monte Carlo method (220). It is based on pseudo-random numbers with a binomial distribution. In the present study 100 000 runs were stimulated in the Monte Carlo simulation to check if the subjects performed better in one of the two conditions (IRN or PT).

3.2.4.1 Pitch Ranking

Averaged across all pitches, NH users scored $89.58\% \pm 6.28\%$ correct (mean \pm SD) for the IRN stimuli. For the PT stimuli they scored $93.19\% \pm 5.55\%$ correct. Performance with both listening conditions was significantly above chance level (both $p = 0.00$ ($p_{PT} = 6.16e-141$ and $p_{IRN} = 3.03e-114$)). Figure 3A shows details about the performance of NH subjects.

CI subjects scored on average $57.08\% \pm 7.76\%$ for the IRN stimuli and $80.00\% \pm 11.90\%$ for the PT stimuli. Performance was not significantly above chance level for IRN stimuli ($p = 0.31$) but significantly above chance level for PT stimuli ($p < 0.01$). Figure 3B shows details about the performance of the CI subjects.

3 Publications

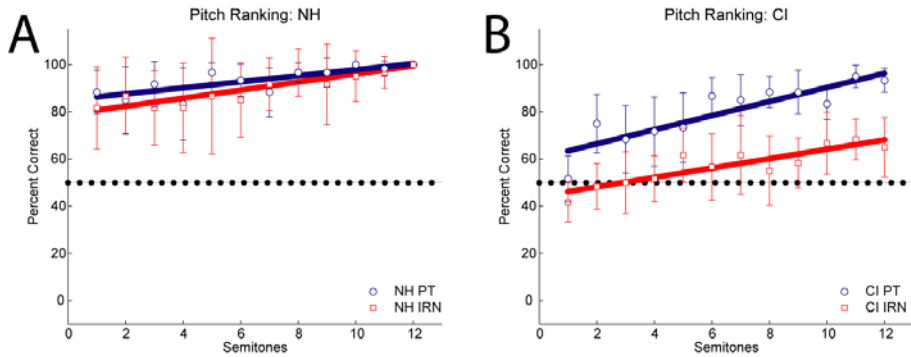


Figure 3: Performance of both listeners in the pitch ranking task. The error bars represent the 95% confidence interval of the mean. The dotted lines which are at 50% in both graphs show chance performance level. Linear regression was used on both graphs to fit a line to show the increase in performance with increasing semitone distance.

For NH subjects no significant difference in performance was observed between both listening conditions ($p = 0.31$). CI subjects performed significantly better with PTs compared to IRN stimuli ($p < 0.01$). The performance of the NH subjects was significantly better than the performance of the CI subjects in both listening conditions ($p < 0.01$).

Linear regression was used to fit a line to the results using the equation $y(t) = k \times t + d$. The factor 'k' is the inclination and the summand 'd' is the offset of the fitted line. The inclination 'k' helps to determine if there is a performance increase with increasing semitone distance. Inclination and offset lie with a probability of 95% within their corresponding interval range. For NH subjects inclination and offset (each \pm intervals) are [$k_{NH\ PT} = 1.3\% \pm 0.6\%$ and $d_{NH\ PT} = 84.8\% \pm 4.9\%$] for the PT stimuli and [$k_{NH\ IRN} = 1.7\% \pm 0.5\%$ and $d_{NH\ IRN} = 78.9\% \pm 3.9\%$] for the IRN stimuli. Both inclinations are positive, meaning that the performance increased on average 1.3% per semitone for the PT and with 1.7% per semitone for the IRN stimuli (Figure 3A). For CI subjects, performance of pitch ranking also increased as semitone distance increased. Inclination and offset were [$k_{CI\ IRN} = 2.0\% \pm 0.7\%$ and $d_{CI\ IRN} = 44.0\% \pm 5.3$] for IRN stimuli and [$k_{CI\ PT} = 3.0\% \pm 1.1\%$ and $d_{CI\ PT} = 60.2\% \pm 8.4\%$] for the PT stimuli. Both inclinations are positive, meaning that performance

increased on average with 2.0% per semitone for the IRN and with 3.0% per semitone for the PT stimuli (Figure 3B).

NH listeners perform equally well regardless of listening condition. This is due to a ceiling effect in the results. CI users perform significantly better with PT stimuli.

3.2.4.2 Melody Recognition

CI subjects scored on average $15.56\% \pm 4.78\%$ for the IRN stimuli and $19.17\% \pm 7.47\%$ for the PT stimuli. Performance was not significantly above chance level for the IRN ($p = 0.09$) but only just significantly above chance level for the PT stimuli ($p = 0.03$).

No significant difference in performance between both listening conditions was observed for NH subjects ($p = 0.26$). The performance of the NH subjects was significantly better than the performance of the CI subjects in both listening conditions ($p < 0.01$). Figure 4 shows the performance in the melody recognition task for both groups of listeners. Linear regression was used again to test if increasing accumulated semitone range (ASR) improves performance. For NH subjects, inclinations of the performance graph for IRN and PT were $k_{\text{NH IRN}} = 0.19\% \pm 0.36\%$ and $k_{\text{NH PT}} = 0.08\% \pm 0.30\%$. The corresponding offsets were $d_{\text{NH IRN}} = 80.3\% \pm 15.3\%$ and $d_{\text{NH PT}} = 89.4\% \pm 12.21\%$. For CI subjects, IRN stimuli inclination and offset were $k_{\text{CI IRN}} = 0\% \pm 0.16\%$ and $d_{\text{CI IRN}} = 16.6\% \pm 7.1\%$. For the PT stimuli, inclination was $k_{\text{CI PT}} = 0\% \pm 0.26\%$ and offset was $d_{\text{CI PT}} = 18.9\% \pm 11.2\%$. For NH subjects, performance increased with 0.19% per ASRs for the IRN and with 0.08% per ASRs for the PT. For CI subjects inclinations averaged around 0% per ASR for both conditions.

3 Publications

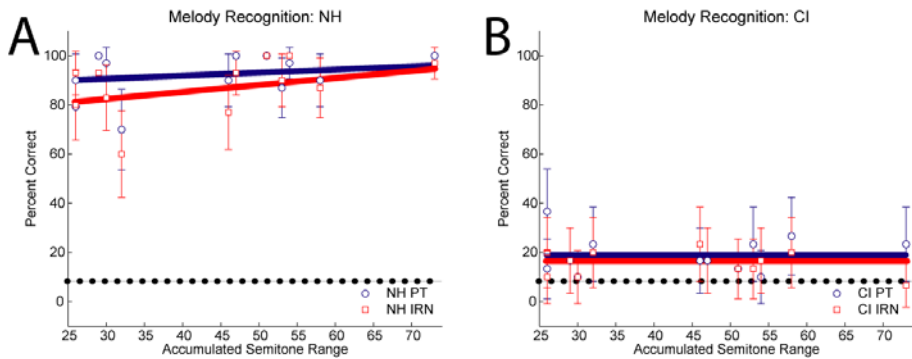


Figure 4: Performance of both listeners in the melody recognition task. The error bars represent the 95% confidence intervals. The dotted lines which are at 8.3% in both graphs show chance performance level.

There was no significant difference observed between performance of Subject S6 whose CI system is optimized for fine structure processing and all the other subjects in the pitch ranking and the melody recognition task. This is most likely due to the fact that CI subjects require multiple semitones until they are able to detect a pitch change (see Figure 3).

3.2.5 Discussion

NH subjects showed no difference in performance between IRN and PT stimuli in the pitch ranking and in the melody recognition task. The reason could be a ceiling effect in both tasks which might have washed out the differences in performance. It could however be that pitch ranking and melody discrimination performance for IRN and PT is the same for NH subjects. Further studies with NH subjects which investigate just noticeable differences in pitch discrimination with IRN and PT stimuli would be needed to test this assumption.

CI subjects performed significantly above chance level with the PT stimuli in the pitch ranking and the melody recognition task. With IRN stimuli performance was around chance level in both tasks. For both listening conditions and tasks, the NH subjects performed significantly better than the CI users. Pitch perception for PT and IRN differs from one another in several ways that are particularly relevant for CI mediated listening.

3.2.5.1 Temporal fine structure and envelope

Speech is a temporally complex signal, containing both slow amplitude modulations of the temporal envelope and fast frequency oscillations of the temporal fine structure (TFS) within each frequency band (221). The envelope information which is transmitted primarily by most current CI processing strategies is sufficient for understanding speech in quiet conditions (98;104). TFS also plays an important role in pitch and speech perception and it enhances pitch and sound quality. Behavioral studies in humans show that sensorineural hearing loss (SNHL) decreases sensitivity to TFS of sound (222-224). Henry and Heinz found that SNHL reduces the strength of temporal coding in noise at the most peripheral level of auditory processing (225). TFS is not transmitted well by any CI system. However, attempts to deliver TFS information have been made in some current sound processing strategies (107;163).

3.2.5.2 Amplitude modulation detection problem due to signal processing

In the process of vocoding, the stimuli is filtered and TFS is effectively tossed out and replaced with a constant rate pulse train in each channel. TFS can be expressed by frequency modulation and as the frequency modulations move in and out of the CI filters the process created amplitude modulations. TFS cues may also be perceived as a within-channel pitch cue if a broadband, flat envelope stimulus is supplied as input for the speech processor (226;227). Modulations can also occur as a response to a complex tone: If more than one harmonic falls within the bandwidth of a filter, the envelope modulation frequency will be the fundamental frequency of the input sound (228). No matter how the TFS cues are created, the CI subjects seem to have a temporal pitch limit around 300 Hz (210). The maximum rate per channel is e.g. 900 Hz for implants from Cochlear Ltd. and it could serve as a carrier of the amplitude modulation-frequency. There is a factor of around 3-4 between the highest modulation frequency and the carrier rate (61). Due to this low carrier rate the maximum modulation frequency is around 300 Hz which is lower than most pitches that were used in the present study.

3.2.5.3 Amplitude modulation detection problem due to background noise

The periodic peaks in the time domain of an IRN signal are accompanied by the presence of a high background noise. This background noise overlaps with the signal and decreases the modulation depth of the periodic peaks. It severely impedes CI users because they generally have a small dynamic range (DR). Normal acoustic hearing can process sounds over a range of 120 dB, and instantaneous amplitudes in normal speech cover a 30 to 60 dB range (229). Implant listeners typically have DRs of only 6 to 15 dB in electric current, requiring the larger acoustic range to be compressed into the smaller electrical range (230). This DR compression might be another explanation for the poor performance of CI subjects in ranking IRN stimuli. While certain stimulus manipulations such as increasing the duration of the stimulus may strengthen the pitch percept induced by IRN (215), the present study found that NH controls could rank IRN pitches with only 250 ms duration well. CI users demonstrated great difficulty in these tasks. DR, along with other factors was also found to significantly affect spectral-ripple discrimination for CI users (218). Reducing DR also lowers phoneme recognition significantly, particularly in noise and for vowels (168)

3.2.5.4 Comparison spectral ripples and iterated rippled noise (IRN)

Won et al performed a study with spectral ripple discrimination for CI users (231). Their spectral ripples were logarithmically spaced in the frequency domain with an amplitude envelope determined by a sinusoid in a decibel scale. They found that spectral-ripple resolution correlates with speech reception in noise for CI users and could serve as a tool to evaluate CI performance with different speech processing strategies. In another study Won et al found that that temporal modulation detection measured with the sound processor can serve as a useful measure of the ability of clinical sound processing strategies to deliver clinically pertinent temporal information (232). Without TFS present, IRN and spectral ripples might look much more alike in their spectral properties. CI subjects have great difficulty with TFS perception, therefore results should be similar for both stimulations. As noted in the introduction (Figure 1), IRN have linear spacing of the peaks in the frequency spectrum whereas summed sinusoid spectral ripples are

usually done with logarithmical spacing of the peaks in the frequency spectrum(231). There is no clear benefit whether IRN or spectral ripple stimuli should be used to test CI subjects. Both could be an effective test to quantify CI performance.

3.2.5.5 Temporal fine structure (TFS) discrimination of the hearing impaired

Several studies have been conducted to analyze TFS discrimination for hearing impaired subject groups. Drennan et al investigated the effect of randomized TFS presented with vocoded speech on NH subjects. They found that improved delivery of TFS improves speech understanding in noise for implant recipients and that bilateral implant recipients might benefit from temporal envelope interaural time differences (233). Henry et al found a relationship between the spectral-ripple threshold and vowel and consonant recognition in quiet in NH, hearing impaired and CI subjects (217). More recently, Imennov et al investigated the perception of acoustic TFS with single channel and multiple channel strategies. Although both strategies were capable of delivering acoustic TFS cues, a single channel analog signal performed better under challenging discrimination condition (226). Without TFS cues the spectral properties of IRN and spectral ripples look very similar. Improving transmission of IRN stimuli could as well be beneficial for speech understanding in noise because (1) it has been shown that improving TFS improved speech perception in noise (233) and (2) because Won et al. showed that spectral-ripple resolution correlates with speech perception in noise (IRN and spectral ripples have a very similar frequency spectrum) (231).

3.2.5.6 Effect of age

The age of the CI group was 53 ± 11 years and the NH group was 33 ± 8 years old (mean \pm standard deviation). The NH group was not matched to control age effects. A study investigating the effect of age on pitch perception found that some older adults with normal audiograms perform at levels that are typical of younger adults (234). But it was also found that age-related declines in temporal processing contribute to deficits in melodic pitch perception. The complexity of the stimuli in that previous study was a lot higher than in the presented study. To extract age effects, the range of

3 Publications

itches was within less than six semitones (compared to up to 12 semitones in the present study).

Therefore we assume that the pure lack of accurate sound coding for pitch as shown in Figure 5 is the reason for the bad performance of the CI subjects and not their age. Furthermore, no significant effects of duration of deafness or speech coding strategies were found in the present study.

3.2.5.7 Place Pitch

An illustration of the difference in processing between IRN and PT-pitch is provided by plotting the output current on each electrode of a CI speech processor over time. The tones in this example were processed by the ACE strategy and at a pulse rate of 500 pps per channel, implemented in Nucleus MATLAB Toolbox (NMT) from Cochlear Ltd. Both electrograms in Figure 5 show the output current of two proceeding tones with 250 ms in duration (separated by a 250 ms pause). For PT frequency 261.63 Hz, the majority of the current is on electrode 21 and 22, the most apical electrodes. A little bit of current is also applied to electrode 20. For PT frequency 523.25 Hz the peak of the stimulation current is shifted towards electrodes 19-21 with a bit of current on electrodes 18 and 22 (Figure 5A). The higher the pitch, the more the stimulation current shifts towards the middle of the cochlea. The lower the fundamental frequency, the more apical are the groups of electrodes that get stimulated. In this example the semitone distance between the two tones is 12 (one octave). These two tones had the maximum semitone distance played in the pitch ranking experiment. It best illustrates the difference in processing in the electrograms for two tones. A certain distance in fundamental frequencies is required to activate different channels. Therefore greater semitone distance leads to better performance in CI users. The good pitch ranking and melody recognition results with PT seem to be based mainly on a place pitch cue.

IRN stimuli lead to stimulation of all active electrodes of the CI. Just by looking at the electrograms in Figure 5B it is hard to tell which of the two tones is higher in pitch. Any place pitch cue is eliminated completely. The sound files which served as input for these two stimuli were directly forwarded to NMT with 100% input-output dynamic range. This eliminates

any potential background noise which could appear in the free field sound-isolated room and gives the best possible output of the sound processor.

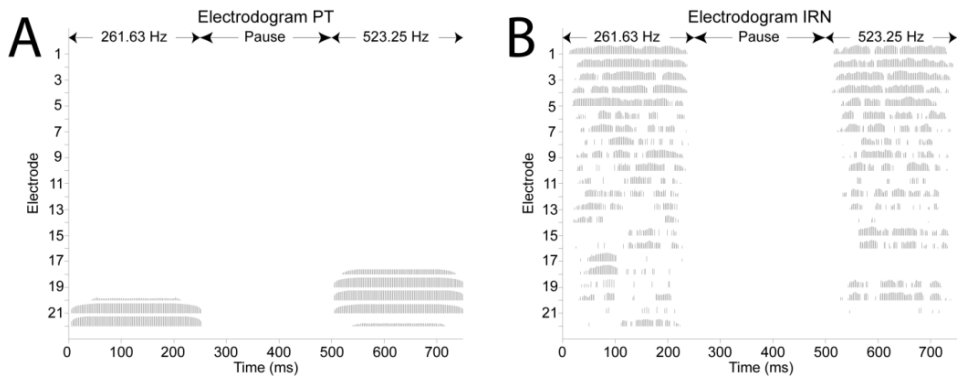


Figure 5 shows the change of current level over time on each electrode processed with Advanced Combination Encoders, or ACE, in a speech processor by Cochlear Ltd. (Cochlear Corp., Sydney, Australia). The output of two sequential tones with 261.63 and 523.25 Hz. PT (left), IRN (right) for the ACE processing scheme in a Nucleus implant is plotted. Each tone in the pair has a length of 250ms with linear rise/decay ramps of 50ms and they are interrupted by a silence pause with 250ms length. This processing scheme is used very frequently in the subjects tested.

These findings suggest that several factors account for the bad performance of CI subjects with IRN stimuli. 1) The lack of accurate TFS 2) the background noise which obscures the dips in the amplitude modulations, and 3) the lack of a place pitch cue in in the IRN stimuli. The lack of place pitch is probably the most important difference between the processing of IRN and PT stimuli for CI subjects. Although pitch ranking performance increases with increasing semitone distance, there was no effect of increasing ASR on performance for CI subjects. CI subjects are able to rank IRN pitch significantly above chance level but melody recognition makes the task too complex for them. For NH users, there is still a small increase in performance with increasing ASR for the PT and almost no effect for the IRN stimuli. Improving the processing for IRN stimuli could not only help to improve music perception, it could also help to improve speech perception in noise because it would be a confirmation for more accurate TFS transmission in CI subjects.

3.2.5.8 Effect of training

Due to the fact that the coding of sounds at the most distant pitches (e.g. 261.63 and 523.25 Hz) in Figure 5 already looks very similar we are unsure to which extent training might help to improve pitch perception. In order to trigger a training effect it would be required to play two stimuli that differ to a certain extent in the way they get processed by the CI speech processor. In other words we believe that for the pitch differences that were tested it may be likely that there will be no training effect observed because CI subjects simply hear two times the same tone based on the way it gets processed in the speech processor. Studies which send signals to the CI through direct electrical stimulation (by replacing the speech processor) offer more optimal conditions for training. Ray Goldsworthy, a CI user who is a researcher studying the effect of training on temporal pitch perception observed that training pitch perception would require huge dedication and practice. Even after working with subjects for 20 hours or more, he is uncertain if they have been sufficiently trained to improve in pitch perception tasks. He himself had listened to different pitches for 100 hours himself and still wonders how much he can improve upon the task (235). These rates were, however presented optimally through direct electrical stimulation. The speech processor cannot even represent the stimuli that optimally and this decreases the potential effect of training even further.

3.2.6 Conclusion

CI subjects are able to rank pitch and to identify melodies only with PT. IRN pitches and IRN melodies are impossible to be ranked or identified by CI subjects mainly due to the lack of a place pitch cue. Furthermore, the input is smeared with a high background noise. The limited DR and the lack of accurate TFS seems to impede CI users further in filtering out the high background noise of the IRN stimuli. CI users are severely impaired compared to NH subjects in perceiving IRN pitches and melodies. Improving the processing of IRN stimuli could not only help to improve music perception - which was the primary goal of the present study - it could also improve language perception in noisy environments.

3.3 Publication 3

Title

Effect of dynamic range and location on temporal pitch perception of cochlear implant users.

Authors

Richard T Penninger

David Morris

Charles Limb

Katrien Vermeire

Marc Leman

Ingeborg Dhooge

Journal

Trends in Amplification: submitted

3 Publications

3.3.1 Abstract

The coding of pitch in modern cochlear implant devices is challenging for recipients as it is mainly based on place pitch, which is subject to the location of a finite number of electrodes on an array. The coding of pitch via temporal cues in the stimulation may help to improve pitch perception since most cochlear implant subjects are able to perceive rate changes on a single electrode up to around 300 Hz. Some optimally performing subjects are able to perceive temporal pitch up to 1000 Hz. However, performance varies highly between subjects and depends on the selected electrodes.

The first objective of this study was to quantify the effect of electrical dynamic range on temporal pitch perception. Therefore pitch ranking was measured on electrodes with the widest and the narrowest dynamic ranges. The second objective was to investigate the effect of the location of the electrodes on pitch ranking and therefore a basal, a middle and an apical electrode were additionally tested for each subject.

The results show a high degree of intersubject variability. However, the electrical dynamic range was found to correlate with the pitch ranking score. No consistent effect was observed based on the location on the implant array.

Improving CI performance based on individual subject and electrode performance might be the key to a better understanding of the mechanisms behind temporal pitch perception and it could lead to improved music and pitch perception in general.

3.3.2 Introduction

For cochlear implant (CI) users music sounds unpleasant mainly due to a lack of accurate pitch perception (74). Pitch is an important attribute of many types of music and CI listeners generally perform poorly at pitch discrimination tasks. Just noticeable differences (JND) in pitch for CI subjects are a lot worse compared to normal hearing (NH) control groups (133;236). In addition to their poor pitch discrimination capabilities, CI subjects frequently confuse the direction of pitch change and this phenomenon is known as pitch reversal (237). These pitch perception problems arise from the way how the CI encodes pitch information. The placement of the electrodes makes use of the tonotopic arrangement of the cochlea (place

pitch) which varies from the basal end which responds optimally to high frequencies, to the apical end which responds optimally to low frequencies (8). The endolymphatic fluid in the cochlea is highly conductive and therefore the pitch resolution of the CI is limited by current spread. The pitch range in the low frequency region (apical end) is cut off by the insertion depth of the electrode (238).

Changing the time structure of the electrical stimulation can also lead to changes in pitch perception and this is referred to as temporal pitch. It is similar to NH-pitch perception observed for amplitude modulated noise (61). Temporal pitch can be created by sinusoidal amplitude modulation (SAM) of a fixed carrier rate on a single electrode (61;162;163). There is an upper rate pitch limit of around 300 Hz on most (146;147) but not all subjects (153).

The electrical dynamic range (DR) of an electrode refers to the current level between that which is required to elicit a just noticeable sensation and that which is perceived as being loud but tolerable. In the fitting of a CI to a recipient the DR can be individually adjusted for each electrode. The size of the DR determines the depth of modulation for temporal pitch based on SAM. The perception of SAM pitch becomes exponentially weaker as the modulation depth decreases until it is perceptually similar in pitch to that of an unmodulated pulse train (239). Narrow dynamic range (NDR) electrodes have been found to be linked to poor electrode discrimination, poor place-pitch perception and poor speech recognition (110;240;241).

Stimulation of apical electrodes with SAM below 1 kHz could result in better pitch discrimination compared to basal electrodes due to a better match between temporal and place pitch. Kong et al observed that subjects' ability to discriminate rate differences varied significantly depending on the electrode site stimulated (153). The fact that rate discrimination performance depends on electrode location has also been shown by Zeng and by Baumann and Nobbe (146;151). Performance seems to depend heavily on the selection of the appropriate electrode for each subject.

3 Publications

The first goal of the present study was to evaluate the effect of DR on temporal pitch perception. We hypothesized that electrodes with a wide dynamic range (WDR) may perform better than NDR electrodes. Our second goal was to evaluate the effect of location on pitch ranking performance. Our second hypothesis was that performance might gradually improve as the site of stimulation on the implant array is moved towards the apex due to a better match between SAM and the place of stimulation. Knowledge of the characteristics and locations of electrodes where subjects perform well on temporal pitch tasks has the potential to improve pitch processing strategies, which in turn may benefit the perception of music and speech intonation.

3.3.3 Methods

Ten CI users participated in the experiment. Their age range was 30 to 76 years (mean = 60, SD=16). Subjects used Cochlear devices (Cochlear Ltd., Sydney, Australia) and their clinically-assigned speech processing strategies were either ACE or MP3000. Relevant subject details are shown in Table 1. All subjects had more than six months of experience with their implant system. The experiments were performed at the Department of Otolaryngology, Head and Neck Surgery of the University Hospital Ghent, Belgium and at the Medical University Hannover, Germany under a research protocol approved by their Ethical Committees. Written consent was obtained from all participants.

Table 1 shows details about subject demographics

Subject	Age (yrs)	Processor Type	CI use (month)	Implant type	DR	
					LDR (μ A)	HDR (μ A)
S1	71	CP 810 SP	17	CI512	38	57
S2	35	CP 810 SP	25	CI512	31	45
S3	73	Freedom SP	39	CI24R(CA)	27	47
S4	30	CP 810 SP	8	CI512	42	61
S5	69	Freedom SP	62	CI24R(CA)	38	62
S6	76	CP 810 SP	7	CI512	24	47
S7	73	Freedom SP	52	CI24R(CA)	37	54
S8	57	CP 810 SP	18	CI512	29	59
S9	55	CP 810 SP	87	CI512	46	48
S10	63	CP 810 SP	16	CI24R(CA)	52	105

All stimuli were delivered via the L34 research sound processor using Nucleus Implant Communicator software (Cochlear Ltd.). The software for the experiment was written locally using MATLAB R2009b (The MathWorks Inc., Natick, MA, USA) and run on a personal computer. All stimuli were biphasic pulse trains of 300 ms duration and were based on a carrier rate (=stimulation rate) of 5000 pps. The high carrier rate was chosen because it has been found that CI listeners can rank SAM stimuli if the carrier rate and the highest frequency of the SAM differ by a factor of approximately four (61). To convey pitch information SAM was applied to a carrier pulse train

3 Publications

using the equation $SAM(t) = f(t) + d \times \sin(2\pi f_m \times t + \frac{3\pi}{2})$ $SAM(t) =$

$$f(t) + d \times \sin\left(2\pi f_m \times t + \frac{3\pi}{2}\right)$$

where $f(t)$ $F(t)$ was the unmodulated pulse train at 5000 pps presented at threshold level and d was the depth of the modulation. The factor f_m F_m was the modulation frequency and it had a starting phase of $\frac{3\pi}{2}$. The maxima and minima of the SAM corresponded to the subjects' maximum comfortable (MC) level and the threshold (T) level of an unmodulated 5000 pps pulse train.

To quantify the effect of DR on pitch ranking performance, the DR of each electrode was estimated during a standard clinical follow-up session. Then a WDR and a NDR electrode were individually selected based on the widest and the narrowest DR. These two electrodes were on a different location on the implant array for each subject. To evaluate if different locations on the implant array had an influence on performance, three further electrodes were selected. These were on the same location on the implant array for each subject. A basal electrode (electrode 4), a middle electrode (electrode 11) and an apical electrode (electrode 18) were tested.

Monopolar stimulation was used in all cases. This involved current flow between an electrode on the implant array and two extracochlear returns in parallel: one on the case of the receiver stimulator and one on a ball electrode placed under the temporalis muscle. The experiment started with loudness balancing using the method described by Landsberger and McKay (242). Initially new T and MC levels had to be defined for an equal amplitude pulse train at 5000 pps for each participating electrode. To equalize loudness levels across experimental conditions, presentation levels of all SAM stimuli were loudness balanced to a reference level of 261.63 Hz on each electrode. Prior to this the amplitude of 261.63 Hz was verified to be comfortable for each subject. Each subject reported that they would be able to listen to this loudness level for at least an hour without discomfort. First the baselines

were balanced with each other, and then the signals were loudness balanced with the corresponding baselines (see Table 2).

Modulation Rates						
Baseline	131 Hz	185 Hz	262 Hz	370 Hz	523 Hz	693 Hz
Signal (26% higher)	165 Hz	233 Hz	330 Hz	466 Hz	659 Hz	880 Hz

Table 2 shows all modulation rates presented in the pitch discrimination task.

The electrical signals investigated in this study were SAM electrical pulse trains delivered to a single electrode position. The modulation frequencies of the six baseline stimuli were between 130.81 Hz (musical note C3) and 693.46 Hz (musical note F5). Each baseline stimulus had a corresponding signal stimulus with a modulation frequency which was four semitones (26%) higher.

Pitch ranking was performed using a two-interval-two-alternative-forced-choice test (2I2AFC) “mixed-block” procedure described by Kong et al (153). All five electrodes were tested in the same block. For most of the subjects, 1500 responses were collected (50 per frequency per electrode). Subject (S) 3 did 900 trials (30 per frequency and electrode) and S9 and S10 only did 600 trials (20 per frequency and electrode). In each trial, the order of the signal and the baseline stimuli was randomized and the subjects were asked to choose the stimulus with the higher pitch. No feedback was given as to whether the responses were correct. Drawings of animals making low pitched sounds (e.g. a bear) and high pitched sounds (e.g. a mouse) were introduced to illustrate what low pitches and high pitches are. Furthermore, musical notations of high and low notes were shown to the subjects. The buttons on the input device were labeled with these illustrations. Subjects were permitted a short training block where feedback was provided after each response to confirm that they had understood the task. The output of all presented stimuli was verified prior to the experiment with the CIC4 Decoder Implant Emulator (Cochlear Ltd.) to ensure that the experimental software operated correctly.

3 Publications

3.3.4 Results

Our experiment used forced choice procedures, therefore the results can be analyzed with the binomial probability distribution. A binomial experiment consists of repeated trials where the outcome of each trial is labeled either success or failure. The probability of success remains constant from trial to trial. In evaluating the result of a forced-choice experiment, we were initially interested in ascertaining whether the subjects had been merely guessing. For our 2I2AFC procedure, we consider the null hypothesis that the probability of success on each trial was 50%. If the resulting probability (p) is less than the criterion value $\alpha = 0.05$ that is generally accepted for statistical significance, we can conclude that it is unlikely that the null hypothesis is true. That indicates that the subjects were likely to have used some cue in the stimuli as a basis for their responses. Scores significantly above chance level mean that the subjects ranked the pitch correctly. Most subjects (S1, S2, S4 and S7-10) scored significantly above chance level on at least one of the tested electrodes. Performance significantly below chance level indicated that the subjects' responses were pitch inverted and this was found in four of the ten subjects (S3, S5, S6 and S8).

The individual results from the subjects tested in this study shows disparity in their pitch ranking performance. Averaged across all modulation frequencies the following results were obtained for each subject: S1 and S9 scored significantly above chance level on all tested electrodes while S5, S6 and S8 performed significantly below chance level on all electrodes. S4 performed significantly above chance level on the basal, the middle and the apical electrode while S2 performed significantly above chance level only on the WDR electrode. S3 performed significantly below chance level on the basal, the middle and the NDR electrode and S10 scored significantly above chance level on the middle and on the apical electrode. S7 performed significantly above chance on the WDR electrode and significantly below chance on the basal, the apical and the NDR electrode.

The Kolmogorov-Smirnov test was used to check the data distribution. Only the DR was normally distributed ($p > 0.05$). Non-parametric analysis with Spearman's rho showed the following: DR was significantly positively

correlated with score ($p < 0.05$) and location ($p < 0.01$). Details about the performance for each subject in part one of the experiment are shown in Figure 1.

3 Publications

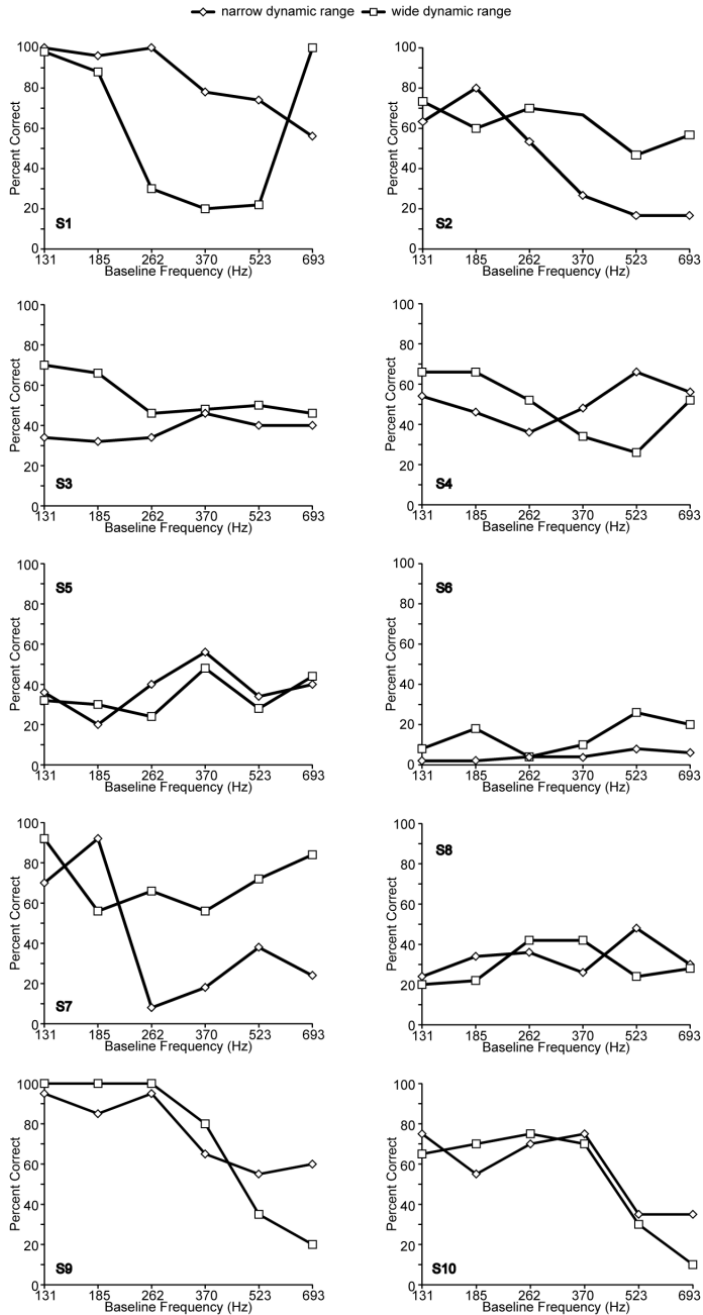


Figure 1. Performance based on dynamic range. Dynamic range was positively correlated with the percent correct score.

For the second part of the experiment the following results were acquired: Electrode location and score were not significantly correlated with each other ($p = 0.58$). The baseline frequency and the percent correct score showed a tendency towards correlation but this was not found to be significant ($p = 0.09$). Details about the performance for each subject in part two of the experiment are shown in Figure 2.

3 Publications

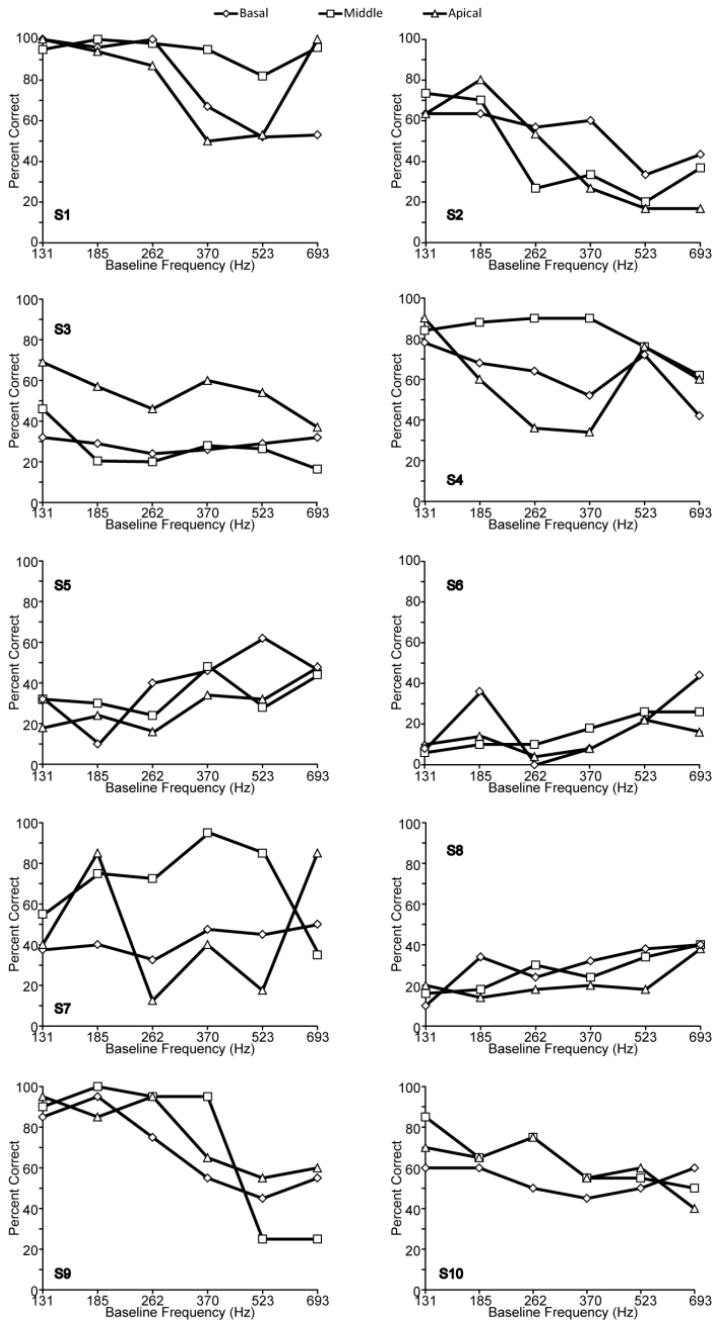


Figure 2: Performance based on location on the implant array. Location and percent correct score were not correlated with each other.

To assess the dependent variables of T and MC on the SAM score we performed a linear regression analysis which showed that the MC was positively related to the SAM score whereas T was negatively related to the score.

3.3.5 Discussion

3.3.5.1 *Effect of DR on pitch ranking performance*

The electrical DRs observed in CI subjects are usually small and they can vary across different electrodes. Our first hypothesis was that SAMs with wider modulation depths are easier to perceive than SAMs with narrower modulation depths. This hypothesis was confirmed by our study where subjects showed a positive correlation between DR and performance on pitch ranking. A previous study by McKay et al. investigated a similar effect. They decreased the DR of SAM stimuli and matched the decreased SAM to the pitch of an equal amplitude pulse train. They started with zero modulation depth which matched the pitch of the carrier rate. As they increased the modulation depth the matched rate came closer to the value of the modulation frequency. It was tested for frequency regions that partly overlap with the baseline frequencies tested in the current experiment (80-300 Hz) (239). Green et al found that WDR was not associated with better temporal pitch because of larger modulation depth. They suggested that the potential for developing strategies delivering enhanced pitch perception is limited (243). In a study by Pfingst et al electrode place discrimination was evaluated. First, regions on the electrode array were determined where electrode-place discrimination was the best or the poorest. Then electrode place discrimination was tested in these regions at different loudness levels in a 2I2AFC test. Electrode pairs with poor discrimination typically had narrower DRs than those with good discrimination (241). In a study by Galvin and Fu modulation detection thresholds (MDTs) for a 20 Hz SAM pulse train were measured as functions of stimulation rate, mode, and level. It was found that MDTs were sensitive to stimulation rate and not sensitive to stimulation mode. Modulation sensitivity seems to be related to intensity resolution which is related to DR. (244). The negative effect of NDR electrodes on other aspects of CI listening has also been reported. NDR electrodes are linked with poor electrode discrimination, poor place-pitch

3 Publications

perception and poor speech recognition (241;245;246). It could be suggested that the same mechanism is active behind poor pitch and speech perception of NDR electrodes. An alternative interpretation could be that the NDR deficit is causing bad speech perception by interfering with amplitude or loudness cues. It could even be a combination of both causal mechanisms: NDR is electing poor pitch perception as well as poor amplitude cue perception and both deficits in turn interact to cause bad speech perception outcomes. Pfungst et al. suggests that a large number of variables influences speech perception. Therefore the effect of any specific variable on speech recognition (e.g. DR) may be masked by the intersubject differences caused by other variables (241).

3.3.5.2 Effect of location on the implant array on pitch ranking performance

The SAM frequencies that were tested in this study were below 1000 Hz. Such low frequencies are tonotopically located in the apical part of the cochlea. If place and rate pitch perception are linked the expectation would be that apical electrodes should perform best, followed by middle and basal electrodes. However in the present study no significant correlation was observed between location of stimulation on the implant array and pitch ranking score. Several other studies have investigated the effect of location on the implant array on performance. In a study by Kong et al four subjects were tested with pulse rate modulation to evaluate the limits of temporal pitch at different cochlear locations. The baseline rates ranged from 100 to 500 Hz. There was no consistent pattern of how the performance changed as the electrode position moved from the basal to the apical location (153). The lack of a consistent electrode effect on rate discrimination performance is similar to the findings of Baumann and Nobbe and Zeng. For pulse rate modulation discrimination there was no significant difference between basal and apical electrodes. Amplitude modulation was only tested on an apical electrode with baseline rates ranging from 200 to 800 Hz with a 5081 pps carrier (146;151). A study of Middlebrooks and Snyder showed that intra cochlear neurons with a low characteristic frequency had a higher “limiting” rate than intracochlear neurons with a high characteristic frequency. This study suggests the existence of a high-temporal acuity brainstem pathway

starting in the cochlear apex which is characterized by low characteristic frequencies, short latencies and high-fidelity transmission of periodic stimulation (145). Macherey et al. showed that stimulation at an apical site of the cochlea yields better rate discrimination at high rates when asymmetric waveforms are used (103).

3.3.5.3 Performance based on modulation frequency

The pitch ranking ability of most subjects in this study was found to decrease as the baseline frequency increased. This tendency was confirmed with statistical analysis but was not found to be significant ($p = 0.09$).

Performance was observed to be significantly better on the two lowest SAM pairs. It has been previously described in the literature that CI subjects seem to have an upper limit of temporal pitch which is around 300 Hz (146;147). It remains at chance levels for frequencies greater than 300 Hz and therefore pitch perception follows a low pass characteristic. Some subjects who participated in this experiment were able to rank pitch up to around 700 Hz (e.g. S1, S4 and S7). Such high performers have also been observed in previous studies. Two subjects were observed in the study by Kong et al, to be able to follow rate changes up to around 900 pps and changes in pulse rate over the range of 500–840 pps were perceived along a perceptual dimension that was orthogonal to the place of excitation (152). Most of our subjects seem to be able to perceive a difference between the two stimuli that is only based on a change in the temporal pattern on at least one electrode in the tested frequency range.

3.3.5.4 Applications

The introduction of SAM with a high frequency pulse train may be beneficial for the perception of complex tones (247). Another approach that has been taken in implementing temporal pitch information into processing strategies is to apply the cue within each channel separately as opposed to providing a global cue across all stimulation channels. One such strategy which provides a channel specific temporal cue is MED-EL's Fine Structure Processing (FSP). They select the two most apical electrodes to provide a channel specific sampling sequence (107;248). The results of the present study show that performance varies considerably according to the site of the stimulation on

3 Publications

the electrode array. Identifying sites where the psychoacoustic performance is best and increasing their contribution via the speech processor map could be beneficial to pitch perception (249-252). However routine performance of such procedures is labor and cost intensive. Detailed studies of cochlear morphology have the potential to assist in electrode insertion to optimize neural and hair cell preservation and ultimately guide electrode location (117). However current multi slice computed tomography scanners may have insufficient spatial resolution to provide enough anatomic details about the location of the in-situ electrode array (253). The data from the present study suggests that the pitch ranking ability of CI listeners is better on NDR compared to WDR electrodes although there was considerable variation between subjects. This may be useful in the selection of electrodes during the clinical programming of implants and it may improve the listening experience of CI users.

3.3.6 Conclusion

This study investigated the pitch ranking ability of CI listeners in response to SAM stimulation on different electrodes. We found that DR and score were significantly correlated with each other. We also analyzed the effect of three positions along the electrode array but we did not find a significant effect for electrodes from the locations that we tested. It is difficult to assess the impact of these results on speech processing of CIs, since the electrical stimulations of the CI speech processor are continuously changing over time and place of stimulation. The development of processing strategies directed towards a more accurate consideration of inter-subject and inter-electrode variability for pitch discrimination may improve the ability of CI users to perceive pitch. Improving CI performance based on individual subject and electrode performance might be the key to better understanding the mechanisms behind temporal pitch perception.

3.4 Publication 4

Title

Experimental Assessment of Polyphonic Tones With Cochlear Implants

Authors

Richard T Penninger

Charles J Limb

Katrien Vermeire

Marc Leman

Ingeborg Dhooge

Journal

Otology&Neurotology 34(7): 1267-71 (2013)

3 Publications

3.4.1 Abstract

Hypothesis: We hypothesized that cochlear implant (CI) users are able to discriminate tones consisting of 1 and 2 modulation frequencies.

Background: Music perception is a very challenging task for CI users. In music, multiple tones often occur simultaneously, an essential feature of harmony. Proper encoding of simultaneous tones is crucial to musical perception and appreciation. With current implant processing strategies, CI users are severely impaired in the perception of pitch and polyphony.

Methods: The ability of CI users to identify the number of simultaneous tones was assessed. Stimuli were applied with direct electrical stimulation. Stimuli with 1 modulation frequency were applied on a basal, a middle, and an apical electrode to determine if there was an effect of cochlear region. Stimuli with 2 modulation frequencies were applied on combinations of an apical electrode together with a basal or a middle electrode. Additionally, 2 modulations frequencies were presented at the same time on an apical electrode only.

Results: Results demonstrate that subjects were generally able to identify the number of modulation frequencies in the presented stimuli. Performance for 1 modulation frequency stimuli was significantly above chance level on all 3 electrodes tested. Performance was best on the apical and the middle electrode, followed by the basal electrode. Subjects were also able to identify 2 modulation frequencies significantly above chance level on all 3 combinations tested. Performance was best on combination apical-basal followed by apical-middle. Performance was worst when 2 modulation frequencies were applied on an apical electrode only, but it was still significantly above chance level.

Conclusion: If polyphonic sound coding would use concurrent modulation frequencies on multiple or single electrodes, then polyphonic tones would be better perceived by CI users yielding better music perception.

3.4.2 Introduction

Accurate music perception remains one of the difficult auditory tasks for cochlear implant (CI) users. Music represents the next frontier for CI-mediated listening beyond speech, and given the enormous acoustic complexity of musical stimuli, possibly the most challenging sounds that any

listener will ever encounter. The primary impairment for accurate music perception is pitch perception, which remains very poor for most CI users (74). Previous research has shown that CI users rarely perceive pitch differences of less than several semitones, a level of performance that pales in comparison to normal hearing (NH) subjects (133). Perception of pitch, which is the psychophysical correlate of a sound's fundamental frequency, is severely altered in CI users due to the nature of electrical hearing, in which a periodic pattern of electrical pulses is used to stimulate spiral ganglion neurons in a manner that differs significantly from NH. A partial explanation is related to the fact that the number of discriminable pitches, which can be heard as current is delivered to distinct locations along the cochlea, is limited by the number of available electrodes (254). Another factor results from the patients' specific cochlear anatomy that impedes device manufacturers in designing electrodes that minimize cochlear damage and optimize the insertion depth to reach the low frequency regions in the apical end of the cochlea (117). Beyond delivery of individual pitches to the inner ear, the presentation of multiple simultaneous pitches-referred to as polyphony-represents a major obstacle for implant-mediated transmission of musical information.

Very few studies have examined how CI users perceive polyphony, an acoustic property that is essential to almost all forms of music. More specifically, the relationship between acoustic polyphony (in a complex auditory waveform) and the electrical representation of that polyphony (in the pattern of electrode stimulation) remains unclear. In a study of acoustic polyphony, Donnelly et al. presented free-field polyphonic auditory stimuli consisting of one, two or three simultaneous tones to CI and NH users (75). Listeners were asked to identify how many separate pitches they heard. The authors found that CI users demonstrated perceptual fusion for polyphonic pitches, meaning that they frequently perceived two and three separate acoustically presented pitches as a single pitch. NH users performed significantly better than CI users in this task. The explanation for this perceptual fusion in CI subjects remains incompletely understood, particularly with regards to how polyphonic acoustic information is transmitted electrically by CI devices.

3 Publications

Changing the periodicity of stimulation can lead to changes in the perceived pitch in the low frequency range for CI users. The perceived pitch resembles the perception of NH individuals for amplitude modulated noise (61). In the present study, the amplitude of the current on each electrode was modulated with a sine wave at a frequency referred to as the modulation frequency. The modulations were applied on one single electrode or simultaneously on two different electrodes. McKay and McDermott investigated how the perception of pitch changes when two different modulation frequencies are presented on two electrodes: As the distance in the array between these two electrodes was smaller than 1.5 mm, some CI subjects perceive the aggregated temporal pattern (255). This would mean that if the two electrodes are more than 1.5 mm apart, most CI users will perceive a polyphonic stimulus.

The goal of the present study was to characterize polyphonic pitch identification ability of CI users using direct electrical stimulation. The goal of this study was to evaluate the ability of post-lingually deafened adult CI subjects to discriminate between one and two modulation frequencies presented on one or two electrodes. We hypothesized that the subjects would perform better in the two modulation frequency condition as the distance on the CI array between the electrode pair was increased, leading to minimal overlap between neural populations stimulated by each electrode. We further hypothesized that the subjects' performance for the one modulation frequency condition would improve as the stimulation site on the electrode array moved from the base towards the apex due to a greater match between the modulation frequency of the stimulus and the characteristic frequency of the neurons. Our goal was to gain a better understanding of how limitations in the ability to perceive simultaneous pitch may ultimately affect music perception.

3.4.3 Materials and Methods

Seven monaurally implanted CI users aged 30 to 76 years (mean = 64 years, standard deviation (SD) = 16 years) participated in the study (Table 1). The experiments were performed at the Department of Otolaryngology, Head and Neck Surgery of the University Hospital Ghent, Belgium under a research

protocol approved by the Ethical Committee (protocol number 2011/324). Informed consent was obtained from all participants.

Table 1: Cochlear Implant subject demographics.

Subject	Age	Speech Processor	Cl experience	Implant type	Coding
	(years)		(months)		Strategy
S1	71	CP 810	17	CI512	ACE
S2	73	Freedom	39	CI24R(CA)	ACE
S3	30	CP 810	8	CI512	ACE
S4	69	Freedom	62	CI24R(CA)	MP3000
S5	76	CP 810	6	CI512	ACE
S6	73	Freedom	52	CI24R(CA)	ACE
S7	57	CP 810	18	CI512	ACE

All stimuli were delivered via the L34 research processor using the Nucleus Implant Communicator software (Cochlear Ltd.). The software for the experiment was written locally and run on a personal computer. All stimuli were biphasic pulse trains with 2.5 s in duration. The pulse trains were based on a carrier rate of 5000 pps. To convey pitch information sinusoidal amplitude modulation (SAM) was applied to a carrier pulse train using the equation $SAM(t) = f(t) + d \times \sin(2\pi f_m \times t + \frac{3\pi}{2})$. The summand $f(t)$ was the unmodulated pulse train at 5000 pps presented at threshold level and d was the depth of the modulation. The multiplier f_m was the modulation frequency and it had a starting phase of $\frac{3\pi}{2}$. The maxima and minima of the SAM corresponded to the subjects' maximal comfortable level and their threshold as measured by an unmodulated 5000 pps pulse train. All stimuli consisted of modulation frequencies within an octave ranging from 261.63 Hz (C4) to 523.25 Hz (C5) (Table 2).

3 Publications

Table 2: Modulation frequencies of condition one and condition two.

One Modulation Frequency		Two Modulation Frequencies		
Modulation Frequency number	Frequency (Hz)	Semitone Spacing (Pair Number)	Tone 1: Frequency (Hz)	Tone 2: Frequency (Hz)
1	261.63	1	369.99	391.99
2	277.18	2	349.23	391.99
3	293.66	3	349.23	415.30
4	311.13	4	329.63	415.30
5	329.63	5	329.63	440.00
6	349.23	6	311.13	440.00
7	369.99	7	311.13	466.16
8	391.99	8	293.66	466.16
9	415.30	9	293.66	493.88
10	440.00	10	277.18	493.88
11	466.16	11	277.18	523.25
12	493.88	12	261.63	523.25
13	523.25			

The stimuli in condition one (one modulation frequency) consisted of 13 unique pitches applied on a basal (electrode 4), a middle (electrode 11) and an apical electrode (electrode 18). The stimuli of condition two (two modulation frequencies) consisted of two SAM stimuli on three different electrode combinations. Combination one consisted of a basal electrode along with an apical electrode. Combination two consisted of a middle electrode together with an apical electrode. In combination one and two each of the two electrodes was stimulated with a different modulation

frequency. It is not possible to stimulate two electrodes at the same time with the present hardware. Therefore quasi-parallel stimulation was used meaning that the two modulation frequencies alternated at an overall rate of 10 kHz between the electrode pair (each of the two individual modulation frequencies had a carrier rate of 5000 pps). Combination three involved the application of two modulation rates to a single apical electrode. The two modulation frequencies alternated again with 10 kHz.

All stimuli were loudness balanced and randomly presented in a two alternative forced-choice (2AFC) procedure in which the subjects were instructed to choose whether the given stimulus consisted of one single or two simultaneous tones. In total, 720 stimuli were presented to each subject for identification. Subjects were familiarized with the stimuli prior to formal testing. No feedback was given regarding the correctness of the responses.

3.4.3.1 Statistical analysis

In the analysis of the data the first question was whether the subjects were merely guessing. This was analyzed with the binomial probability distribution. The second question was whether the subjects performed in a different way on the corresponding electrodes or electrode pairs within each condition and it was analyzed with the Monte Carlo Method (220). If the resulting probability (p) was less than 0.05 that was accepted for statistical significance in both analyses.

3.4.4 Results

The overall mean score was $82.34\% \pm 37.82\%$ for condition one and $76.31\% \pm 39.20\%$ for condition two (mean \pm SD). The mean scores for both conditions are shown in Figure 1. Confusion matrices are presented in Table 3.

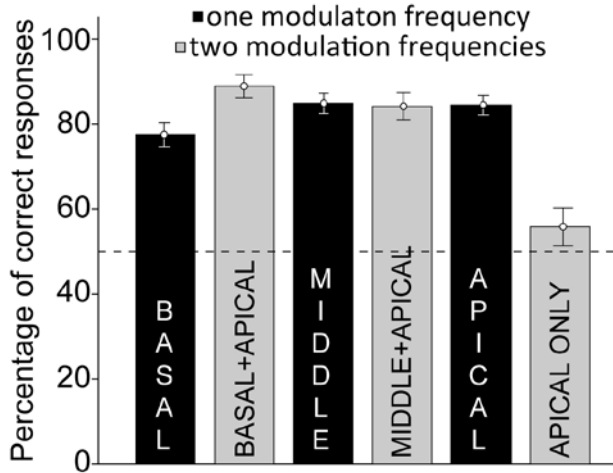


Figure 1: Mean performance accuracy across subjects and combinations for condition one (one modulation frequency) and condition two (two modulation frequencies). The error bars represent the 95% confidence intervals of the mean and the dashed line shows chance performance level (50% correct).

Table 3: Confusion matrix.

		Identified	
		One modulation frequency	Two modulation frequencies
Presented	One modulation frequency	2075 82.34%	445 17.66%
	Two modulation frequencies	597 23.69%	1923 76.31%

For condition one the following results were obtained: Averaged across all subjects and modulation frequencies performance was significantly above chance level on basal, middle and apical electrodes ($p < 0.01$). For each

individual modulation frequency the subjects also performed significantly above chance level on all tested electrodes (all $p < 0.01$). Performance on the middle and the apical electrodes was significantly better than on the basal electrode (both $p < 0.01$). Performance on the middle and the apical electrodes was not significantly different from each other ($p = 0.46$). Performance accuracy for condition one for all tested electrodes across modulation frequencies is shown in Figure 2A.

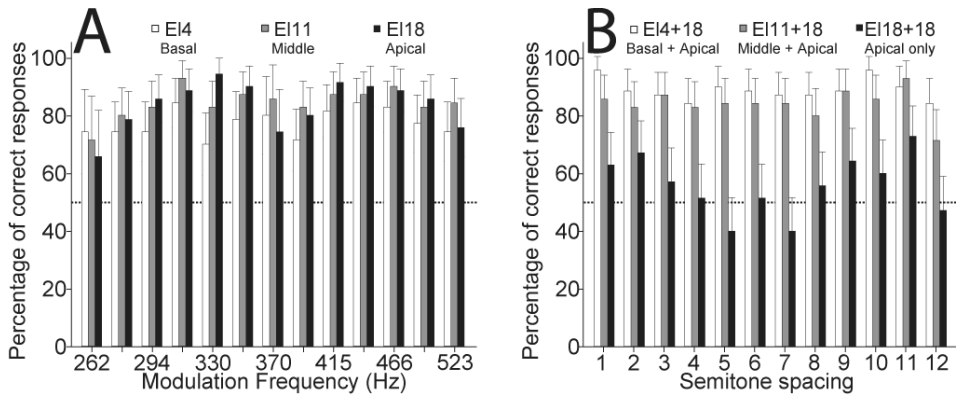


Figure 2A: Percentage of correct responses (number of correct responses/number of stimuli presented) across different combinations as a function of modulation frequency.

Figure 2B: Percentage of correct responses across semitone spacing in condition two.

The error bars represent the 95% confidence intervals of the mean and the dashed line shows chance performance level (50% correct) in Figure 2A and 2B.

Averaged across all subjects and modulation frequency pairs the subjects scored significantly above chance level on combination one, two and three ($p < 0.01$). Analysis of individual modulation frequency pairs revealed the following: Subjects performed significantly above chance level on combination one and two on all pairs (all $p < 0.01$). On combination three, the subjects scored significantly above chance level in only four of the twelve pairs tested. And these were pair one, two, nine and eleven. The respective p values were $p = [0.02, 0.00, 0.01, 0.00]$. Performance on combination one and two was significantly better than on combination three (both $p < 0.01$).

3 Publications

Performance on combination one was significantly better than on combination two ($p < 0.01$). The performance accuracy for combination two for all three combinations as a function of modulation frequency difference (semitone spacing) is shown in Figure 2B. Increased semitone spacing did not consistently lead to better performance for identification of two modulation frequencies, but greater physical distance between the electrodes on the implant array did.

3.4.5 Discussion

In condition one CI subjects were able to correctly identify one modulation frequency on all electrodes tested. Performance on the middle and the apical electrode was observed to be significantly better than on the basal electrode. Performance on the apical electrode was the same as on the middle electrode. We originally hypothesized that subjects might gradually perform better in condition one as the stimulation site on the electrode array is moved towards the apex. This assumption was made since the mechanical properties of the basilar membrane vary progressively along its length from the basal (high frequency end) to the apical (low frequency end) (8). There is a closer match between rate and place of stimulation on the apical electrodes. One of the reasons why the performance on the middle and the apical electrodes is the same in our experiment could be due to a ceiling effect in the results. Macherey et al. showed that stimulating at an apical site of the cochlea yielded better rate discrimination at high rates (154). No consistent electrode effect on the rate discrimination performance was found by Baumann and Nobbe (151).

In condition two, CI subjects performed significantly above chance level on all combinations. It has been observed in clinical practice that subjects are reasonably accurate at indicating whether a stimulus consists of one or two active electrodes. In combination one and two, the subjects are able to use a place pitch cue to identify the two modulation frequencies. In combination three where two modulation frequencies are applied on the same electrode the effect of place pitch is completely removed. The subjects were forced to base their decision solely on the difference in the modulation frequencies to identify them as two frequencies. Performance dropped significantly compared to combination one and two but it was still significantly better

than chance level. Analysis of individual frequency pairs for combination three shows that subjects perform only significantly above chance level on four out of the twelve modulation frequency pairs, namely pair one, two, nine and eleven (Table 2). Experiments with acoustic input signals may help to explain this effect. Donnelly et al. tested identification of one, two and three simultaneous pitches via acoustic input for CI subjects. The stimuli consisted of pure tones and piano tones which were all processed through the speech processor of the subjects' CI. Subjects were most accurate in identifying two-pitch stimuli for both piano and pure tones when the two pitches were one semitone apart (75). This could explain why our subjects performed significantly above chance level on pair one. Oxenham (2008) found that when the frequency separation between the two tones is small, listeners typically hear a single stream of alternating tones. Performance increases with increasing semitone spacing between the pitches (256). This would explain the good performance observed on pair nine and eleven. Performance was found to drop again on pair twelve. In this pair the higher modulation frequency is exactly two times as high as the lower modulation frequency. It leads to overlaps in every second periodic peak of the modulation frequencies which could be the reason why they are more likely to be identified as one modulation frequency. Two modulation frequencies can be best identified when applied on two different electrodes but for some pairs of frequencies they can also be identified on only one electrode.

3.4.5.1 Implications and applications

Our results show that CI subjects are able to identify two pitches which are presented electrically at the same time and to discriminate them from one pitch stimuli. This finding is intriguing in light of previous findings that CI users display perceptual fusion for acoustic polyphonic stimuli (75). Therefore, using stimulations presented in this study could possibly improve discrimination for acoustic polyphonic stimuli. Anecdotal observations during the present experiment showed that the subjects generally felt that the polyphonic tones sounded pleasant and reminded them of how they heard music before their deafness. These observations emphasize the fact that perception of polyphony is critically important for a satisfactory musical experience. Improvement in musical polyphony would also likely improve

3 Publications

auditory streaming for two independent signals, which would have broadly favorable effects on implant user experiences for a wide range of complex listening situations. Further studies would be required to investigate how implementing such stimuli into current speech processing strategies would impact speech understanding. Implementation of such stimuli may increase current spread which could result in concomitant diminution of speech perception ability but this has to be quantified in further studies. Lastly, it is worth mentioning that pre-operative prediction of electrode insertion depth with high resolution computer tomography together with patient specific electrode design may help improve the range of pitches for CI users (253), leading to a substantial improvement in sound quality.

3.4.6 Conclusion

Subjects can correctly identify one modulation frequency on all electrodes tested. Subjects can also identify two modulation frequencies on one or two electrodes. Performance is best when the two modulation frequencies are applied to a pair consisting of a basal and an apical electrode or a middle and an apical electrode. Due to the fact that CI subjects are unable to perceive any form of polyphony with free field stimuli, the development of processing strategies directed towards polyphonic pitch perception could improve the ability of CI users to perceive complex musical stimuli.

3.5 Publication 5

Title

Experimental Assessment of Complex Polyphonic Tones with Cochlear Implants

Authors

Richard T Penninger

Eugen Kludt

Marc Leman

Ingeborg Dhooge

Andreas Büchner

Journal

Accepted for publication in *Otology&Neurotology*

3 Publications

3.5.1 Abstract

Hypothesis: We hypothesized that cochlear implant (CI) subjects would be able to correctly identify one, two and three simultaneous. We further hypothesized that the location on the implant array and the fundamental frequency of the pitches would have an impact on the performance.

Background: “They gave me back speech but not music.” is a sentence commonly heard by CI subjects. One of the reasons is that in music, multiple streams are frequently played at the same time which is an essential feature of harmony. Current CI speech processors do not allow CI users to perceive such complex polyphonic sounds.

Methods: In the present study the authors assessed the ability of CI subjects to perceive simultaneous modulation frequencies based on direct electrical stimulation. Ten CI subjects were asked to identify one, two and three simultaneous pitches applied on different electrodes using sinusoidal amplitude modulation. All stimuli were loudness balanced before the actual identification task.

Results: Subjects were able to identify one, two and three simultaneous pitches. The further the distance between the two electrodes, the better was the performance in the two modulation frequency condition. The distance between the modulation frequencies had a significant effect on the performance in the two and three pitch condition.

Conclusion: Subjects are able to identify complex polyphonic stimuli based on the number of active electrodes. The additional polyphonic rate pitch cue improves performance in some conditions.

3.5.2 Introduction

Music perception is one of the greatest challenges for cochlear implant (CI) subjects, mainly due to a lack of accurate pitch perception. Pitch is the psychoacoustic correlate of a sound's fundamental frequency, a percept which can also be influenced by encoding of harmonic information for complex tones. Although purely monophonic music exists, the overwhelming majority of most Western music is polyphonic in nature, with multiple melody lines, instruments and voices superimposed over supporting harmonic and bass information simultaneously (257). Perception of musical harmony, which is the relationship between musical notes presented concurrently, is absolutely dependent upon the ability to perceive polyphony. This ability is also directly responsible for the ability to perceive differences between musical intervals (two notes) or chords (at least three notes) that are consonant vs. dissonant in nature, ultimately affecting a listener's aesthetic evaluation of the music being heard. In light of the critical importance of polyphonic pitch information during music perception, it is important to establish the ability of CI users to perceive differences between stimuli that are monophonic, intervallic or chordal in nature. In this study, we therefore sought to examine the ability of CI users to perceive differences between electrically presented stimuli consisting of one, two or three simultaneously presented pitches.

The ability of the cochlea to discriminate pitch depends on two mechanisms, place and rate pitch. Place pitch is based on the mechanical properties of the basilar membrane. It acts as a frequency analyzer and activates the hair cells and auditory nerve fibers which are located spatially along the tonotopic gradient of the cochlea (8). Rate pitch is based on the firing rate of the auditory nerve fibers. They phase lock with the variation of sound pressure over time (258). One way of quantifying pitch perception is a pitch ranking task in which subjects are presented with pairs of short acoustic tones. The subjects have to decide which of two tones is higher in pitch. Several studies have shown that CI subjects are severely impaired in pitch ranking compared to normal hearing (NH) control groups (133;236). In a CI, a polyphonic pitch perception can be created with place pitch, rate pitch or based on a combination of both. Polyphonic place pitch results from simultaneous

3 Publications

stimulation of distinct neuron populations along the basilar membrane. Polyphonic rate pitch is based on the firing rate of auditory neurons.

Surprisingly few studies have investigated the perception of polyphonic pitch by CI users. Donnelly et al. investigated polyphonic pitch perception for NH and CI users using acoustically presented free-field stimuli. Whereas NH subjects were able to differentiate between one, two and three simultaneous pitches, CI users demonstrated severe impairment on this task, with perceptual fusion of multiple tones into a single tone (75).

By using direct electrical stimulation the speech processor is replaced with a research unit. Pitch is encoded in the form of current amplitude modulation. Using direct electrical stimulation allows to select the electrodes which are stimulated and to adjust their amplitude modulation rate. That way it is possible to tease apart rate and place pitch perception. In that regard, direct electrical stimulation differs significantly from normal speech processing strategies.

In the one-pitch condition of the present experiment, stimulating electrodes which are located closer to the apical end of the cochlea may improve performance due to a smaller difference between the rate pitch of the frequencies tested and place pitch compared to basal and middle electrodes. Furthermore it has been shown that apical neurons perceive pitch better than basal neurons in the applicable low frequency range (145). In the two-pitch condition the distance between two electrodes influences the overlap in the current fields of each electrode. Therefore performance might improve as the distance between the electrode pair is increased due to a reduction of the field overlaps.

In a previous study, Penninger et al. investigated the ability of CI users to differentiate between one and two musical pitches only. This study showed that subjects were generally able to discriminate between one- and two-pitch stimuli. Performance was best when the two-pitch stimuli were based on a polyphonic place and rate pitch cue (e.g. the two pitches were presented on two different electrodes with different modulation frequencies) (259). In this study, we presented polyphonic sound stimuli

consisting of one, two or three pitches to the implant via direct electrical stimulation and examined whether or not polyphonic place pitch cues are sufficient to perceive polyphony. The purpose of this study was to gain insight into the ability of CI users to perceive musical chords, rather than just isolated notes or pairs of notes (intervals). It was hypothesized that the location of the electrodes in the one- and two-pitch condition would affect performance on a polyphonic pitch identification task. We further hypothesized that the difference in frequency between simultaneous tones would impact the ability of a CI user to detect polyphony.

3.5.3 Materials and Methods

Ten CI users aged 38 to 76 years (mean = 59 years, standard deviation (SD) = 12 years) participated in the study (Table 1). The experiments were performed at the Department of Otolaryngology, Head and Neck Surgery of the Medical University Hannover, Germany under a research protocol approved by the Ethical Committee. Written consent was obtained from all participants.

3 Publications

Table 1: Cochlear Implant (CI) subject demographics.

Subject	Age (years)	CI experience (months)	Implant type (Nucleus)	Coding Strategy
S1	76	14	CI512	ACE
S2	71	76	RE-CI24R CA	ACE
S3	63	51	RE-CI24R CA	ACE
S4	64	68	RE-CI24R CA	ACE
S5	52	47	RE-CI24R CA	ACE
S6	38	85	RE-CI24R CA	MP3000
S7	64	47	RE-CI24R CA	MP3000
S8	47	15	CI512	ACE
S9	50	50	RE-CI24R CA	ACE
S10	61	96	RE-CI24R CA	ACE

All stimuli were delivered via the L34 research sound processor using Nucleus Implant Communicator software (Cochlear Ltd., Sydney, Australia). The software for the experiment was written locally using MATLAB R2010b (The Math Works Inc., Natick, MA, USA) and run on a personal computer. All stimuli were biphasic pulse trains with exactly 2.5 s in duration. All stimuli were based on a carrier rate of 5000 pps. To convey pitch information sinusoidal amplitude modulation (SAM) was applied using the following equation $SAM(t) = f(t) + d \times \sin(2\pi f_m \times t + \frac{3\pi}{2})$. $f(t)$ was the unmodulated pulse train at 5000 pps presented at threshold level and d was the depth of the modulation. F_m was the modulation frequency of the SAM and it had a

starting phase of $\frac{3\pi}{2}$. All stimuli consisted of modulation frequencies within an octave ranging in fundamental frequencies from 261.63 Hz (C4) to 523.25 Hz (C5) (Table 2). In total, 585 stimuli were presented to each subject.

Table 2: Modulation frequencies of condition one and condition two.

One Modulation Frequency		Two Modulation Frequencies			Three Modulation Frequencies			
SAM Number	Frequency (Hz)	Semitone Spacing	Tone 1 (Hz)	Tone 2 (Hz)	Semitone Spacing	Tone 1 (Hz)	Tone 2 (Hz)	Tone 3 (Hz)
1	262	0	370	370	0	370	370	370
2	277	1	370	392	1	466	494	523
3	294	2	349	392	1	277	294	311
4	311	3	349	415	2	277	311	349
5	330	4	330	415	2	392	440	494
6	349	5	330	440	3	311	370	440
7	370	6	311	440	3	311	370	440
8	392	7	311	466	4	262	330	415
9	415	8	294	466	4	330	415	523
10	440	9	294	494	5	277	391	523
11	466	10	277	494	5	262	349	466
12	494	11	277	523	6	262	370	523
13	523	12	262	523	6	262	370	523

3 Publications

One pitch stimuli consisted of 13 unique pitches on a basal (electrode 4), a middle (electrode 11) and an apical electrode (electrode 18). Two pitch stimuli consisted of 13 unique polyphonic pitches on three different electrode combinations. Combination one consisted of a basal together with an apical electrode. Combination two consisted of a middle together with an apical electrode and combination three consisted of a basal electrode together with a middle electrode. In all two pitch combinations the two SAM pulses alternated at an overall rate of 10 kHz between each corresponding electrode pair. Three pitch stimuli were presented on a basal, together with a middle and an apical electrode. The pulses alternated at an overall rate of 15 kHz between the three electrodes.

All stimuli were loudness balanced and randomly presented in a one interval, three alternative forced-choice procedure in which the subjects were instructed to choose whether the given stimulus consisted of one, two or three pitches. Subjects were familiarized with the stimuli and procedure prior to formal testing. No feedback was given regarding the correctness of responses.

3.5.4 Results

The first question was whether the subjects were merely guessing. This was analyzed with the binomial probability distribution. The second question was whether the subjects performed in a different way on the tested electrode configuration and this was analyzed with the Monte Carlo Method (8). If the resulting probability (p) was less than 0.05 that was accepted for statistical significance in both analyses.

For all conditions the subjects were able to identify one, two and three pitch stimuli significantly above chance level. The mean scores are shown in Figure 1.

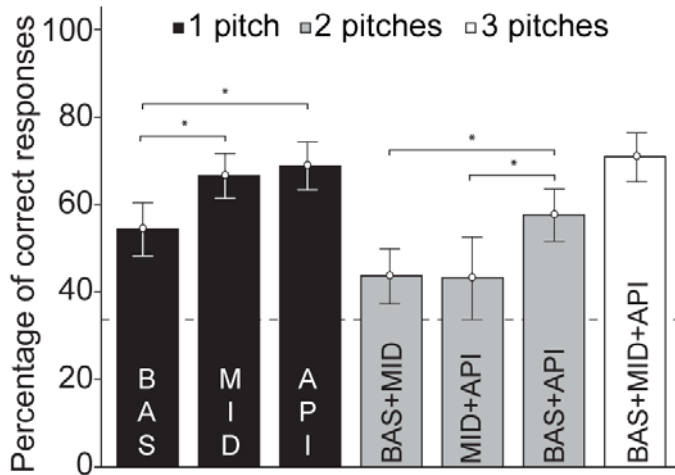


Figure 1: Mean performance accuracy across subjects and combinations for the one-, two-, and three pitch stimuli. The error bars represent standard deviation of the mean and the dashed line shows chance performance level (33% correct).

In the one pitch condition the subjects performed significantly above chance level on all electrodes (all $p < 0.001$). The mean performance over all frequencies was 54.6% on the basal electrode, 66.8% on the middle electrode and 69.1% on the apical electrode. Performance on the middle and the apical electrode was significantly better than performance on the basal electrode (both $p < 0.001$). Performance on the apical electrode was not significantly better than on the middle electrode ($p = 0.71$). Performance on each individual frequency was significantly above chance level on the middle and apical electrode. On the basal electrode performance was significantly above chance level on all except on frequency one which was 261.63 Hz. (Figure 2A)

3 Publications

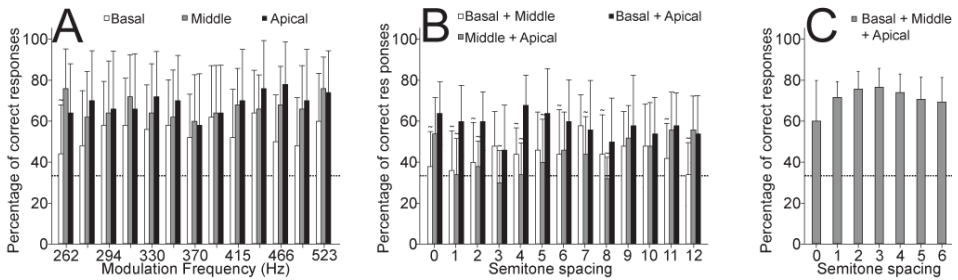


Figure 2A: Percentage of correct responses (number of correct responses/number of stimuli presented) across different combinations as a function of modulation frequency. The “~” signs indicate frequencies where the performance was not significantly above chance level.
Figure 2B: Percentage of correct responses across semitone spacing in two-pitch stimuli. The dashed line shows chance performance (33% correct) in Figure 2A and 2B. The “~” signs indicate frequency pairs where the performance was not significantly above chance level.
Figure 2C: Percentage of correct responses (number of correct responses/number of stimuli presented) as a function of semitone spacing for the three pitch stimuli.

In the two pitch condition the subjects performed significantly above chance level on all electrode pairs (all $p < 0.001$). The mean performance was 43.8% on combination three (basal+middle), 43.4% on combination two (middle+apical) and 57.8% on combination one (basal+apical). Performance on combination one was significantly better than on combination two and three (both $p < 0.001$). Performance on combination three was not significantly different from performance on combination two ($p = 0.45$). Analysis of individual frequency pairs revealed the following: On combination three performance was significantly above chance level on frequency pair 3, 5, 7, 9 and 10. On combination two performance was significantly above chance level on frequency pair 0, 6 and 9-12. On combination one, performance was significantly above chance level on all pairs. (Figure 2B).

In the three pitch condition the subjects were able to identify the stimuli with performance significantly above chance level ($p < 0.001$) with mean performance of 71.1%. Performance was significantly above chance level on all individual frequency pairs. Performance was significantly worse on semitone spacing zero compared to all the other tested semitones (all $p < 0.01$). (Figure 2c). Confusion matrixes are presented in Table 3.

Table 3: Confusion matrix.

		Identified		
		One pitch	Two pitches	Three pitches
Presented	One pitch	1238	568	144
		63.49%	29.13%	7.38%
	Two pitches	372	943	635
		19.08%	48.36%	32.56
	Three pitches	63	483	1404
		3.23%	24.77%	72.00%

3.5.5 Discussion

Despite the difficulty that CI subjects demonstrate in identifying acoustic polyphonic tone complexes (75), in this study we showed that CI subjects were able to identify one, two and three polyphonic pitches when presented electrically.

Of particular relevance is the finding that CI subjects made their judgments predominantly based on a polyphonic place pitch cue. In the one-pitch condition performance gradually increased as the stimulation site was moved towards the apex. Furthermore it was observed that performance was best on middle and apical electrodes. In condition two the distance increase between the electrode pair did improve performance significantly due to reduced current overlap. This confirmed our original hypothesis that the location on the implant array would have an effect on performance. In a previous study it was also found that when two different amplitude modulated pulse trains are presented to two electrodes separated by less than 1.5 mm, some CI subjects perceive an aggregated (one-pitch) cue (255). Place pitch seems to be the most important cue for perceiving two and three

3 Publications

simultaneous pitches and the additional rate cue (e.g. the difference in stimulation frequency), which was hypothesized to effect performance as well, only led to significant performance increase in the three-pitch condition.

The main goal of the presented experiment was to test the extent to which polyphonic pitch can be perceived when presented electrically. As mentioned earlier in the introduction, a polyphonic place and/or rate pitch cue can be used to create the sensation of polyphony. To maximize the performance of the subjects the majority of the stimuli used a combination of rate and place pitch cues. Subject's ability to identify polyphonic tones consisting only of a rate pitch cue was tested by Penninger et al. (259) and it showed a significant performance decrease with scores just above chance level. In the present study it was evaluated if a place pitch cue alone would be sufficient to perceive polyphonic tones (e.g. in the zero semitone spacing option in condition two and three). Ideally, a specific rate pitch should be presented on the corresponding place pitch location along the cochlea to optimize pitch coding. However, the fundamental frequencies that were used in the present experiment were all below 600 Hz. Locations where such low frequencies would be perceived based on place pitch are beyond the maximum insertion depth of the implant arrays of the subjects tested. One of the reasons for using short electrode arrays is that the high inter-subject variability in cochlear anatomies makes it hard to design implants that cover the entire cochlear duct length (117). Therefore, based on the anatomy of a NH cochlea, the rate and the place pitch did not match in any case of this experiment. In the cochlea of a CI subject, the frequency perception reorganizes over time. The most apical electrodes are perceived lower than they would be based on the tonotopy of NH subjects. Therefore a closer match between the stimulated rate pitch along with the reorganized perceived place pitch leads to good one-pitch performance on the apical electrodes. Similar results have been obtained in a comparable study where the perception of one and two pitches was investigated (259). On the basal electrode, which stimulated locations that are tuned to high frequencies, performance was worst. On the extreme example, the lowest of all the tested frequencies (261.63 Hz) on the most basal electrode, performance

dropped down to chance level (Figure 3A). It was unsure if the good one-pitch performance on middle and apical electrodes would additionally improve the perception of two pitches which were applied on these two electrodes. No such effect was observed. Performance was significantly above chance level for all two-pitch combinations on semitone spacing 9 and 10. This is consistent with previous work with NH subjects from Oxenham et al. who showed that at larger frequency separations, a tone sequence splits into two auditory streams, each comprising one of the two tone frequencies. (256).

To maximize the chance of good performance the polyphonic pitches in condition three were mainly based on a polyphonic place and rate pitch cue. CI subjects already show great impairment for perception of two pitches based on a polyphonic rate pitch cue only with performance just above chance level (259). It is therefore unlikely that any polyphony would be perceived with three rate pitches on one electrode. It was however unsure if reducing the polyphonic rate pitch cue would decrease performance. Performance was significantly above chance level on all semitone spacings. In the zero semitone spacing option, the rate pitch cue was completely removed. We originally assumed that performance might drop down to chance level at these points. However, subjects were still able to perceive polyphony based on the number of activated electrodes. Performance dropped significantly compared to all the other frequencies when the three pitches were the same. In this condition the benefit of the additional rate pitch cue becomes most apparent. It increases until semitone spacing three and from then on it gradually decreases again (Figure 3C). In the two-pitch condition no such effect was observed.

There are several major engineering challenges in current speech processors which cause the presented problems. First, the speech processor which is oriented towards language processing has a pitch resolution which is a lot worse compared to NH subjects. It uses mainly envelope cues to transmit the signals. Temporal fine structure is not transmitted accurately which has a huge negative effect on pitch and speech perception in general (260-262) and particularly in the low frequency range (51;263). Secondly, the ability to encode input frequencies is severely impaired based on current spread.

3 Publications

Reducing current spread and increasing the number of electrodes at the same time may help to design implants which have a greater and more focused frequency resolution. Other promising approaches are to use drug eluting implants which may help to make neurons grow closer to the implant which could lead to lower currents required to induce stimulation (and therefore decrease current spread).

Subjects performed better in the three-pitch condition compared to combination three in the two-pitch condition (Figure 2). The confusion matrix in Table 3 shows that three-pitch stimuli were perceived mostly as three-pitched (with 72% correctness) but 24.77% of them were perceived as two-pitched. Almost none of them (3.23%) were perceived as one-pitched. Donnelly et al. performed a very similar polyphonic pitch experiment in the sound field with sound stimuli. Their confusion matrix for NH subjects shows that three-pitched stimuli are perceived with 49.6% accuracy as three-pitched and with 42.4% accuracy as two-pitched (75). 49.6% accuracy is close to the accuracy of the zero semitone spacing in Figure 3C. The additional rate and place pitch cue on three electrodes in the present experiment seems to improve performance in the three-pitch compared to combination one in the two-pitch condition. Furthermore the two-pitch stimuli are based on one interval between the two tones whereas three-pitch stimuli are based on three intervals. These consist of the first tone together with the second or third tone and the second tone together with the third tone. This could explain why mean performance was better in the three-pitch condition than in the two-pitch condition using combination one.

3.5.5.1 Implications and Applications

The stimulation patterns used in direct electrical stimulation are more selective compared to CI speech processors. Therefore, bypassing the speech processor and mapping the frequencies on selected individual electrodes enables CI subjects to correctly identify one, two and three pitches. Subjects reported during the present study that the stimuli sounded very pleasant and reminded them of how they used to hear music before their deafness. With present CI processing strategies such sensation could be reproduced by adopting the way the speech processor encodes sound. Implementing stimuli presented here into current CI processing strategies with focus on music

perception may greatly improve the music listening experience for CI users. Besides music, simultaneous speech signals often occur in everyday life. CI subjects have great difficulty understanding multiple speakers (264). Improving polyphonic pitch perception may also be helpful in this aspect.

This study has investigated CI subjects' ability to distinguish multiple concurrent temporal pitches. Another compelling question would be how many concurrent temporal pitches can be perceived by CI users.

3.5.6 Conclusion

Subjects are able to identify complex polyphonic stimuli based on the number of active electrodes. The additional temporal pitch cue helps to improve their performance in some conditions. The smaller difference between the place pitch at the stimulation site and the SAM frequency improves performance in the one-pitch condition. The distance between the active electrodes has a significant effect on performance for the perception of two pitches. In the three pitch condition increasing distance between the modulation frequencies improves performance.

If a sound processing strategy was to use polyphonic stimulation on multiple or single electrodes, then possibly polyphonic tones will be better perceived by implant users yielding improved music perception.

4 Discussion

The major aim of this doctoral thesis was to analyze pitch perception for CI subjects from a multidisciplinary perspective. Publication one showed that the CBCT scanners provide better resolution than MSCT. Pitch ranking was measured in study two in the sound field. Two sound inputs with different frequency spectrums were used to tease apart the two fundamental components of speech, the envelope and the TFS. It was shown that CI subjects are severely impaired in the perception of IRN stimuli which are only perceivable with a TFS cue (Publication 2). All of the following studies (Publication 3 - Publication 5) were conducted with direct electrical stimulation. The speech processor was replaced with the L34 research sound processor from Cochlear Ltd. This allows precise control of the stimulation patterns of the CI. All subjects who were used as participants in Publications 2-5 had fully inserted CIs. A necessary requirement for full insertion is to have accurate CT imaging modalities that can be used to plan and confirm the successful implantation of a CI (Publication 1).

4.1 General discussion

The major aim, pitch perception, was subdivided into four studies: Firstly the perception of pitch was measured in the sound field (Publication 2). Secondly the perception of pitch with direct electrical stimulation was measured (Publication 3). Perception of pitch is a requirement for perception of simultaneous (polyphonic) pitches. The last two studies analyzed the perception of simultaneous pitches (Publication 4 and Publication 5).

Perception of pitch has always been a challenge for CI subjects. Trying to improve pitch perception has been enthusiastically embraced by researchers all over the world as it would significantly improve CI users' perception of sounds. Pitch perception is an essential requirement for perception of language and music. In tonal languages, the pitch contour of a word determines its meaning and in non-tonal languages pitch information is essential for e.g. perception of vowels, prosody and speaker segregation (see Introduction). Pitch is one of the main aspects of music as it is the basis for perception of melodies. Both speech and music are semi-periodic signals consisting of complex sound waves. These have specific frequency,

amplitude and timbral components that are presented in an organized way. Both language and music have spectral and temporal envelopes that vary in time (265;266). However, speech and music also differ in several ways. The range of F0 and the loudness levels for music are significantly greater than for speech. In music perception, accurate perception of F0s is essential for the recognition of a melody. F0s are however not imperative for understanding basic context of non-tonal languages. (266).

Many studies have previously assessed pitch perception in the sound field (see Publication 2) and concluded that pitch perception for CI users is very bad compared to NH subjects (for a review see (267)). By using the speech processor in pitch perception studies, it is often not entirely clear which exact conclusion can be drawn from the results. The speech processor is frequently used as a “black box” without knowing what exactly happens during the processing. Therefore, many studies which are conducted in free field do not give exact explanations where the lack of accurate pitch perception arises from and how it could be improved in further speech processing strategies.

In Publication 2 it was tried to give explanations where the poor pitch ranking and melody recognition of CI subjects arises from. First it confirms that pitch perception for CI subjects is bad compared to NH subjects. It was found that CI subjects were only able to perform well with PT stimuli. Using IRN tones, which are mainly based on a TFS cue, did not provide a sufficient pitch cue to the CI subjects. Testing purely rate pitch cues through the speech processor is very challenging because of the nature of the sampling of the input signal. For example, the rate per channel for implants from e.g. Cochlear Ltd. is not high enough to provide a rate pitch cue above e.g. 300 Hz. Previous studies have investigated pitch perception with direct electrical stimulation and huge inter subject differences were observed. In order to understand better where these differences arise from, direct electrical stimulation has to be used. By using direct electrical stimulation the speech processor is replaced by a research unit and the sound input is created in a laptop and directly forwarded into the implant.

4 Discussion

In Publication 3 it was shown that most subjects were able to perform pitch ranking tasks purely based on rate pitch cues up to around 600 Hz. Only one electrode is required to perform a rate pitch task. Analyzing pitch ranking on all electrodes would require excessive testing time. Therefore the selection of electrodes had to be narrowed down to two components: the location on the implant array and the DR. These two factors showed generally acceptable reliability to improve pitch ranking but some caution must be taken into account for a few subjects which showed results that were inconsistent with the rest of the tested group (e.g. pitch reversals). It was shown that the pitch ranking capability across subjects improves if electrodes with a WDR are chosen. It is important to notice that by using direct electrical stimulation the pitch ranking capabilities of CI subjects can be improved compared to stimulation in the sound field. But performance was still a lot worse in comparison to NH subjects.

In Publication 4 polyphonic pitch perception was measured. Polyphony is a crucial aspect of basic musical elements such as harmony, consonance and dissonance. The ability to perceive e.g. harmony depends on identification of the relative relationships between multiple pitches (268). Most Western music that CI users encounter in everyday life is polyphonic. Relatively few studies have investigated the perception of musical polyphony. By designing a polyphonic experiment it had to be decided in which way the polyphony should be created. With place pitch or with rate pitch or with both pitch percepts. The rate pitch sensation was created with amplitude modulation of a high rate carrier. By using two different electrodes a different polyphonic place pitch perception can be created for each of the tones. A polyphonic rate pitch sensation can also be created by stimulating with two different modulation frequencies. By increasing the spatial distance between two electrodes the polyphonic place pitch sensation is increased. Decreasing the special distance between the electrodes decreases the place pitch percept. The ultimate decrease of polyphonic place pitch was the test combination where the two amplitude modulations were applied on one electrode only (Publication 4). This condition only leaves a polyphonic rate pitch sensation to identify that the tone consists of two pitches. The one-pitch tones were applied to a single electrode. Three electrodes were tested in total.

Performance on the basal electrode was worse compared to the middle and the apical electrodes. Neuron activation in the basal end of the Cochlea is perceived as high pitched and the tones that we played were low pitched. It was therefore concluded that the huge distance between rate and place pitch on the basal electrode creates a pitch percept which is less likely to be identified as one pitch.

The findings of Publication 4 revealed that CI subjects were generally able to identify simultaneous pitches. Performance was significantly above chance level for all combinations of electrodes. The distance between the electrodes had a significant effect on performance. It was therefore concluded that by adding a place pitch cue to the perception performance significantly increases. However, a rate pitch cue alone is sufficient to perceive two pitches. Based on the results of the one pitch condition it was concluded that by decreasing the link between rate and place pitch in an extreme way (by stimulating low rate pitches on basal electrodes) performance decreases significantly.

The results of Publication 4 were implemented in the study design of Publication 5. In this study it was first tested if subjects are able to identify the simultaneous stimulation of three electrodes as three pitches. Furthermore it was tested how the subjects would perceive the same modulation frequencies on two or three different electrodes. They perceived them as two or three pitched tones. Therefore it can be concluded from Publication 5 that if there is a polyphonic place pitch cue the subjects identify the stimuli as polyphonic. The additional rate pitch cue improves performance but not significantly. The polyphonic place pitch cue is stronger than the polyphonic rate pitch cue. If the rate pitch cue is non-polyphonic and the place pitch cue is polyphonic then the subjects perceive the tones as polyphonic.

4.1.1 Pitch Processing in Cochlear Implants related to signal processing

In the following chapters the basic pitch processing constraints of CIs are discussed. It should help to understand basic mechanisms of processing and how they relate to the presented studies.

4.1.1.1 Front end

There are several technological constraints that reduce pitch processing capabilities of a CI. Some of them occur already at the level of the CI speech processor. Modern multi-channel CIs bandpass filter the input signals. Each channel covers a specific frequency range. The acoustical input is split up by the speech processor into several frequency components which are sent to the electrodes according to the tonotopy of the cochlea. In the NH ear there are about 3500 inner hair cells that detect the sound and process it. CIs, however only have (at most) 22 electrodes to replace the function of thousands of hair cells. Therefore it is impossible to transmit such finely graded frequency information to the auditory nerve.

CI devices were primarily designed to convey speech cues. Therefore the bandpass filtering of the input signal occurs in a specific way to emphasize speech cues. Only frequency components that range from 188 Hz – 7938 Hz are processed by the CI. These are device specific values from Cochlear Ltd., but they are in a range that is typical for all CI manufacturers. The pitch information is not only degraded due to the splitting up in relatively few channels, it is also degraded based on the fact that high and low frequency components are removed. This has significant practical limitations for the perception of pitch. Roy et al. presented musical pieces to CI subjects which had different sound quality. The low frequency components were systematically cut off in the input signal. CI and NH subjects were asked to rank the quality of the musical sound. The low frequency components were cut off in 200 Hz steps from zero (unaltered reference) to 1 kHz (most degraded signal). NH subjects showed a linear decay in sound quality ranking ranging from about 100 % in the unaltered reference to about 20 % in the most degraded signal. CI subjects also ranked the signal quality more or less according to the signal's degradation. However, they did not perceive a significant difference in sound quality until up to 400 Hz were cut off in the range of the signal. For NH subjects, playing the signals with 200 Hz cutoff frequency already leads to significant sound quality decay (269;270). Pitch perception disruptions already happen at the front end of the CI. Most severe limitations result from the auditory nerve interface. The results of Publication 2 are impaired by these limitations. Publication 3 - Publication 5

are not influenced by the limitations based on the processing of the speech processor. There are however several other effects that occur on the level of the implant and they are discussed in the following chapters.

4.1.1.2 Auditory nerve interface

The frequency components of the input signal are mapped to the electrodes of the CI. Any current results in imprecise stimulation of huge populations of nerve fibers. This imprecise auditory nerve stimulation represents another limiting factor for pitch perception (254;271).

4.1.1.3 Surgical and anatomical factors

Optimal electrode placement is a prerequisite for maximizing CI success and full atraumatic scala tympani insertion is the goal. The insertion of the CI can be performed with two different techniques: Round window approach or cochleostomy approach. Several other variables affect the insertion including anatomic abnormalities, characteristics of the electrode array and the experience of the surgeon. Especially in case of cochlear malformations surgical experience can affect optimal electrode insertion and avoid complications (272). Accurate CT imaging of the temporal bone plays a crucial role in detecting abnormalities and helps to optimize the placement of the CI. Most accurate imaging techniques for temporal bone imaging are shown in Publication 1. Less ideal electrode placement may lead to increased current levels required for stimulation (e.g. due to the greater distance to the nerve endings) and this in turn increases current spread which decreases frequency selectivity (268).

4.1.1.4 Electrode design

Most implanted electrodes do not reach the apical turns of the cochlea where the low frequencies are tonotopically represented. This limits the range of neurons that can be activated and it prevents the utilization of the normal cochlear response. The length of the cochlea averages around 33 mm (273). No implant which is currently on the market can be inserted beyond 30 mm measured from the round window. One of the reasons is that the diameter of the scala tympani decreases consistently towards the apex and therefore the electrode may be too thick to fit in the end. Too far insertion increases the risk of insertion trauma particularly for short cochleas. Since

4 Discussion

the sizes of the cochleas are highly variable it may be useful to use longer electrodes for subjects with bigger cochleas (117). The average CI insertion depths currently averages around 20 mm which covers only about 60% of the range of NH subjects (274;275). Due to the short insertion, the frequencies of the most apical electrode have a value around 2 kHz according to the prediction of Greenwood (8;274). Apical regions of the cochlea are therefore not stimulated although they provide important low frequency information. Recent studies have shown that patients who have deeper insertion perform better in speech recognition (119;276;277). There is however a downside of such deep insertions. It can also result in damage of the cochlea from the electrode array (278;279). Overall, deep insertion is of benefit for CI subjects however their specific cochlear duct length might make full insertion of a longer electrode impossible. Publication 1 presents CT imaging modalities which help to determine the optimal insertion depth pre operatively. Measuring the length of the cochlear duct or at least the length and the width of the cochlea may give important cues for selecting the size of the electrode array.

4.1.1.5 Frequency mismatch

In CIs there are frequency mismatches between the electrode place maps and the original frequency that would have been perceived at a specific point based on NH. This mismatch between the characteristic frequency of the auditory nerve fibers and the spectral information delivered by the electrodes causes a large decrease in performance in the short term (280;281). In a study by Di Nardo et al. it was investigated whether the correction of the mismatch through individual fitting of the processor's frequency map can improve music and speech understanding. No significant effects on speech understanding were observed. Most subjects said however that the new map sounded better for music listening by making the sound more pleasant, natural and easier to follow. None of the subjects noted a discomfort with the never map, only two out of the ten tested subjects preferred to keep their old map instead (282). Several other studies indicate that speech recognition is highly correlated with frequency to electrode mapping (281;283-285) Reducing the frequency mismatch may help several subjects. However such modification takes up a huge amount of time and

cannot be performed in a clinical setting due to time constraints. Another practical limitation is that frequency matching is only possible for subjects which have substantial residual acoustic hearing on the contralateral side. However, due to the plasticity of the brain, the CI subjects are able to compensate the frequency mismatch on the long term (286-289;289;290). Electric pitch perception often shifts in frequency by up to two octaves in the first years of implant use (139). In addition to the frequency to electrode mismatch, CI subjects also have a compressed frequency range which is upshifted for high frequencies and downshifted for low frequencies (291;292).

4.1.1.6 Virtual channels

Virtual channels are created by stimulating two neighboring electrodes simultaneously or interleaved which creates a pitch percept in between the two electrodes. Advanced Bionics Corp. uses such concepts in their sound coding strategies. However, previous studies have shown that the interaction between channels may be even more significant for pitch perception than the number of actual or virtual channels present (293;294). These channel interactions arise from various sources. These include the spread of excitation, the neural and the perceptual levels (268;295). Furthermore, previous CI studies have shown that speech recognition in quiet requires only four spectral channels (296) whereas speech recognition in noise requires only about eight 8 channels (127;297;298). CI users perform as if they have only 8-10 effective channels which is a lot fewer than the actual number of electrodes used by all manufacturers (294). Commercial CI processing strategies provide up to 22 physical channels (Cochlear Ltd.) or 120 virtual channels (Advanced Bionics Corp.). Many more channels are required to segregate competing talkers or to perceive music (299). Smith et al estimated that at least 64 functional channels would be required for a satisfying music listening experience (298). More than 16 channels would be required for speech understanding in noisy conditions (127;132;296-298;300). Crew et al. investigated the effect of channel interaction on melodic contour identification for NH subjects who listened to stimuli created by a vocoder which simulated channel interactions in CI signal processing. They suggest that although a greater number of spectral

4 Discussion

channels is sufficient, it is also important to improve the independence among spectral channels (301). In other words, using virtual electrodes without increasing their independence does not help for pitch perception for CI users. Therefore it is essential to focus on reduction of channel interaction to be able to utilize the increased number of channels that can be created by newer devices.

4.1.1.7 Limitation in rate pitch for CI subjects

In NH subjects the auditory nerve synchronizes to the amplitude peaks of the input signal. This phenomenon is possible up to around 5 kHz (302) and it is described in details in the chapters entitled “Rate pitch” (1.3.1 and 1.6.2). Rate pitch can be encoded for CI subjects by varying the rate of stimulation or by amplitude modulation of a high carrier rate. Both mechanisms are very limited in electrical hearing and they show a saturation effect of about 300 Hz (146). CI devices typically process only the temporal envelope of the input signal and the TFS information is lost entirely. Therefore, signals that depend heavily on TFS cues are poorly perceived by CI subjects. One of such signals are IRN stimuli. As described in Publication 2, CI subjects are severely impaired in their perception of TFS and this explains their poor performance with IRN compared to PT stimuli. Electrical pulses are generally delivered to the CI at a fixed rate. The pulse rates of older devices are usually too slow to carry amplitude modulated pitch information (303). As rate pitch information can be delivered with amplitude modulation of high rate carriers, it may be beneficial to improve stimulation rates on the electrodes. However, in addition to the higher power consumption which would be required by using higher carrier rates they might also cause greater channel interaction which would lead to diminished spectral sensitivity (295). A rate pitch cue based on amplitude modulation can be created with newer CI devices. However, the way how the rate pitch percept is created still differs fundamentally from the way how it is created based on NH. For NH subjects the frequency of the oscillation of the basilar membrane creates the sensation of rate pitch. In CI subjects the cue is based on gradual variations in current level.

As confirmed in Publication 2, musical pitch perception depends heavily on the perception of TFS (298;304). One of such processing strategies that tries to integrate TFS cues is Med-El’s Fine Structure Processing (FSP) strategy.

Two to three most apical electrodes are used to send TFS information. The instantaneous stimulation rate matches the instantaneous fine structure frequency of the signal within the frequency range of the channel (107). Another coding strategy by Laneau et al. also extracts the F0 of the input sound. All electrodes are amplitude modulated with the F0 frequency (163). Subjects with residual hearing in the low frequency range are able to use their residual low frequency hearing. Shorter implants only stimulate the higher frequency component of the cochlea. By using such electro-acoustic stimulation the low frequency acoustic input provides essential TFS information that has a huge benefit on speech and music perception (305;306). Leaving functioning low frequency hearing cues intact by using shorter implants substantially improves pitch perception for CI subjects but it can of course only be used for subjects with residual low frequency hearing.

4.2 Future Perspectives

The coding of pitch has been described as “the most important unresolved problem in CIs” by Bob Shannon who has been conducting research in this field for more than 25 years. This sentence was made exactly 10 years ago. Since then no huge improvements were achieved for pitch perception although many research groups have been investigating this issue. It is likely that the pitch perception limitations are a result from the nature of the signal that is used to directly stimulate the neurons. It might be that spread of excitation (see chapter 1.6.1.2 entitled “Selectivity of the place of stimulation due to spread of excitation”) will always impede focused stimulation of small groups of neurons. The following chapters briefly explain concepts that appear to be most relevant in the future.

4.2.1 Current focusing

The current applied to an electrode in the CI spreads out to neighboring populations of neurons. This limits the access to all the spectral information of the CI (307). Attempts of current focusing can be applied by modifying the shape of the stimulating electrodes or by reducing the distance between the stimulating electrodes and the afferent fibers of the auditory nerve (308-316). Current focusing can be applied to different stimulation modes. In

4 Discussion

monopolar stimulation current is delivered to an active electrode and the same amount of current in opposite phase is simultaneously delivered to an extracochlear electrode. This reduces the current spread but it needs higher current amplitudes and phase durations to achieve adequate loudness (294). Litvak et al. for instance found that the required pulse phase durations needed to be around 150 μ s to achieve adequate loudness (317). Tripolar stimulation has been proposed to provide current focusing with greater loudness. Several research groups have tried to implement tripolar, quadrupolar or even more complex current focusing (318-320) in order to receive more focused place pitch percepts (245;321;322). The main challenge is again that the higher the focus of the current the harder it is to receive adequate loudness. Loudness and current spread are somehow directly proportional to each other. Keeping the loudness constant and reducing only the spread of excitation would be ideal but it has not been achieved to our knowledge.

4.2.2 Use of atypical biphasic pulses

It has been shown that monophasic stimuli result in lower loudness thresholds than biphasic current pulses (323). Neural membranes integrate the externally applied current and therefore the effective strength of the first phase can be reduced by the second phase of a biphasic pulse. However, using pulses that are not charge balanced poses a health risk for subjects due to resulting electrochemical reactions. A solution to this is the use of pseudomonophasic biphasic pulses where one of the two phases has a different duration and amplitude than the other one. Therefore pseudomonophasic stimuli can be as efficient as monophasic stimuli while being charge balanced (184). Macherey et al. used asymmetric waveforms which had a short, high amplitude phase followed by a longer, low amplitude phase. When such pulses were presented in a narrow bipolar mode with the first phase on the most apical electrode, the perceived place pitch was lower than with standard symmetric biphasic pulses. This suggests that such pulses can extend the pitch range of CI listeners by focusing the spread of excitation in a more apical region. Indeed it was shown that the upper limit of temporal pitch was significantly higher than that for biphasic pulses averaging around 713pps (154).

4.2.3 Laser stimulation

Optical stimulation provides more precise neural stimulation compared to electrical stimulation particularly with regard to spatial precision (324). Pulsed near-infrared lasers have been suggested as a means of increasing the number of independent channels used in cochlear stimulation (325). Such stimulation could improve the DR and the frequency resolution of the implant (326). However, the physiological mechanisms of optical stimulation are different from that of electrical stimulation and are to date not exactly understood. Two different excitation mechanisms were suggested recently. One of them is linked to the modulation of the membrane capacity by heat due to the absorption of the laser light in water (327). Other explanations are related to hair cell stimulation via optoacoustic effects (328-330). Patents have already been signed about e.g. enhancing music perception with laser stimulation (331). CIs based on laser stimulation are still in development. It should take several years until the first clinical trial can be started.

4.2.4 Drug eluting Cochlear Implants

In some subjects the performance of a CI is suboptimal due to apoptosis or necrosis of the nervous tissue surrounding the electrode. Such 'dead regions' have been described previously (332) and they are negatively correlated with speech understanding in CI users (320;333). The encapsulation of the electrode array caused by a fibrous membrane also increases the impedance of the electrode. Both of these phenomena require greater current levels to reach satisfactory loudness and this has a negative effect on frequency sensitivity due to higher current spread. Therefore there is a need for CIs that can elute substances following implantation that are pharmacologically active. It is believed that this could improve the signal quality of a CI. Such active substances could be for instance antibiotics, anti-inflammatory drugs or neurotropic factors. They could prevent the formation of fibrous tissue around the electrode and infection or necrosis of the nervous tissue. Another form of drug delivery could be the coating of the electrodes which increases their lubricity to decrease the insertion forces during the implantation. Neurotrophic factors may also stimulate the healing of the auditory nerve endings or to decrease the distance between the nerves and the implant (334). A recent study by Farhadi et al. assessed dexamethasone delivery via a

drug eluting CI in guinea pigs. They found significant inflammation reduction compared to a control group (335).

4.2.5 Pitch processing problems arising from the limitation of the auditory nervous system

No matter how a new processing strategy will work in the future, there will always be limitations for its success based on the biological properties of each individual subject. Such limitations were not covered in the presented thesis. These will be shortly described in the next chapters.

4.2.5.1 Peripheral deficits

Hardy et al. confirmed that that a profound bilateral or unilateral sensorineural hearing loss in cats results in extensive degeneration of the auditory nerve (336). The amount of degradation depends on whether the hearing loss is unilateral or bilateral (337). The loss of the electrical activity in the auditory nerve results in a degenerated cellular response in the auditory brainstem (338). A profound sensorineural hearing loss induces significant pathological and atrophic changes within the cochlea and the central auditory pathway (339). CIs have no impact on the signal transmission from the auditory nerve endings to the brain. The use of drugs that can prevent nerve cell degeneration and promote their regeneration may improve clinical outcomes for CIs (340). Other studies show that nerve survival is not the limiting factor of auditory performance. In a study by Fayad et al. no significant correlations were found between speech recognition and surviving cell populations (341). These studies suggest that the redundancy within the auditory nerve is great enough that spiral ganglion cell survival is not the bottleneck in performance with CIs. It is explained by the relatively crude spacing of the implant electrodes (342).

4.2.5.2 Cortical deficits

In order to measure cortical processes involved in sound processing it is necessary to observe the cortical changes in brain activity when the sound stimulus is present. CIs possess an internal magnet and therefore the most appropriate functional imaging technology that can be used on implantees is positron emission tomography (PET). PET has been used to assess the cortical activity of CI subjects and to compare it with NH subjects. It has been

shown that non-implanted deaf individuals display decreased levels of metabolic activity in auditory cortices in comparison to NH individuals (343-345). CI users show greater brain activation than NH subjects and the recruitment of brain areas that are not traditionally utilized for auditory processing (346-348). This process is called cross modal plasticity. CI users who have unilateral implants seem to recruit additional regions to process speech to compensate for the degraded signal from the implant. For bilateral CI subjects no exact conclusions can be drawn (349). In general, CI users may need greater intensity of activation than NH subjects to achieve similar language perception results (268). This greater activation results from the difficulty to interpret the degraded signal. Cortical measures help to show differences in processing between CI and NH subjects but do not necessary give any practical cues on how to improve pitch perception for CI users. They can however be a limiting factor in the performance of some subjects.

4.2.6 Plasticity and training

Duration of deafness in post lingually deafened adults is negatively correlated with auditory performance (246). In prelingually deafened children there is a critical time window until the age of 3.5-4 years before the cross modal plasticity hampers the successful use of the auditory input. Therefore it is important to implant prelingually deafened children as early as possible. Children implanted at younger ages perform better on all clinical tests than children implanted later in life but even adults with long-term prelingual deafness benefit from implantation (268;350;351). There is a huge difference in music appreciation in post lingually deafened adults (who generally do not like music because they remember how it sounded before their deafness) versus prelingually deafened children who in most cases do enjoy music. Music rehabilitation seems to have a substantial benefit for CI subjects. It can help their perception of melodic contour and timbre (352-354). However, musical training takes up a huge amount of time and the engagement of the participant has a significant influence on the results. Music can however be interpreted as the highest form of hearing that humans possess. It provides therefore an effective tool to assess the limitations in CI mediated listening (268).

4.3 Strength and limitations

This doctoral thesis is based on five individual manuscripts. All of them have their specific strengths and limitations.

Publication 1 was important to understand how CT imaging works which is an integral tool to define the quality of the CI implantation. It showed a clear advantage of CBCT compared to MSCT for imaging the temporal bone. However, all scans were performed with cadaveric temporal bones. None of the scans were performed in vivo. Nevertheless, earlier studies have confirmed that CBCT provides sufficient resolution to confirm electrode insertion or to predict insertion depth (121).

Publication 2 tried to split up envelope and TSF cues in free field. The study shows that CI subjects perform relatively well with PT compared to IRN stimuli. For IRN stimuli there is no place pitch cue due to the linear spacing of the peaks in the amplitude spectrum over all frequencies. This should in theory reduce any place pitch cue for CI subjects. The reduction of place pitch can be confirmed by looking at the electrodiagram of the CI. The rate pitch percept cannot be transferred that easily into the CI. It would require higher stimulation rates than used in present CI speech processors.

Publication 3 performed pitch ranking with direct electrical stimulation. The results show that there is a significant correlation between dynamic range and performance. One of the main weaknesses in this study is that the frequency differences were kept constant. The frequent occurrence of pitch reversals is surprising. Furthermore the good performance in the high frequency range is unexpected. Increasing the number of subjects might provide clearer results. A pitch ranking experiment which ranks tones on multiple dimensions could help to confirm that the subjects performed their task only based on a pitch cue only.

Publication 4 shows that subjects are able to differentiate one- from two-pitch stimuli. Two pitch stimuli can also be perceived when applied to only one electrode. Publication 5 shows that subjects can also perceive one- from two- and three-pitch stimuli. The main weaknesses of Publication 4 and Publication 5 are to decide if the subjects based their decision on a rate

pitch, on a place pitch or on some other cue. Other cues could be e.g. sharpness or brightness of the sound (228). Some earlier studies show that CI subjects perceive rate pitch only up to around 300 Hz, therefore almost all the pitches used in the two experiments are too high to be perceived as pitch. However more recent studies show that some CI subjects are able to perceive rate pitch changes beyond 300 Hz (152). Publication 3 also confirms that for many subjects the upper pitch limit is higher than 300 Hz. And all subjects who participated in Publication 4 also participated in Publication 3.

4.3.1 Power Analysis

Statistical power is defined as the ability to find a difference when a real difference exists and it depends in general on a number of factors (355;356).

- The statistical significance criterion
- The magnitude of the effect of interest in the population
- The sample size used to detect the effect

The number of CI recipients who are willing and available to participate in mostly unpaid experiments is limited. Therefore we should not waste their time and instead design tests which have the least number of trials needed to demonstrate a significant effect. We should not do more trials just for the convenience of using normal approximation which is the requirement for many statistical tests (e.g. t-tests and ANOVAS). Instead we should apply statistical methods that cope with relatively small numbers of trials. If the magnitude of the effect of interest is so low that 32 trials or more are needed to show a statistical significance, then this effect is of little practical importance. Most of the experiments that were used in this thesis were forced choice experiments. Therefore we applied the binomial distribution directly instead of the normal approximation. The binomial distribution tells us the occurrence of each outcome. If the probability of success on a single trial is “p” then the probability of obtaining “x” successes out of n trials is given by equation two:

$$b(x, n, p) = C(n, x) \cdot p^x (1 - p)^{n-x} \text{ (Equation two)}$$

4 Discussion

Where $C(n,x)$ is the number of combinations of “n” objects taken “x” at a time. It is referred to as the binomial probability density function and its variance is:

$$\sigma^2 = np(1 - p) \text{ (Equation three)}$$

In Publication 5, for instance, five stimulus representations were applied on each frequency on e.g. the apical electrode in the one pitch condition. If the subject was guessing then the probability of each of the 5 possible outcomes with 30% success is given by Table 6.

Number of correct outcomes	0	1	2	3	4	5
Probability of outcome	0.17	0.36	0.31	0.13	0.02	0.00

Table 5 shows all possible outcomes for a binomial experiment with 5 trials with probability of 30%.

Let’s assume the subject gets a score of 4 or more correct out of 5 trials. The probability of getting this score by guessing only is 0.03. Since this probability is less than the statistical significance criterion value of $\alpha = 0.05$ that is generally accepted for statistical significance.

4.4 Final conclusions and clinical relevance

Publication 1 has most applicable clinical relevance. The other described papers are basis science reports. Plain radiographic images in Stenver’s projection are sufficient to analyze the postoperative radiographic appearance of a multichannel CI. However, in about 3.7% of the cases described by Shpizner et al. (357), additional CT imaging was necessary to confirm correct electrode placement. MSCT images allow assessment of the precise intracochlear position of the electrode and the visualization of individual electrode contacts. Such images are especially useful for the fitting of the speech processor in difficult cases (358). It helps to understand the wide variability in fitting parameters (e.g. T levels). However, MSCT imaging requires on average about 53 times more effective radiation dose than CBCT imaging (121). Furthermore, CBCT provides high-resolution and almost artifact free multiplanar reconstruction images which allow the assessment of the precise intracochlear position of the electrode and its individual

contacts. Furthermore, it is important to know about cochlear malformations before the surgery to know if the chosen CI can be fully inserted into the cochlea. If electrode contacts are left outside of the cochlea they cannot contribute to sound perception of the subject. Publication 1 quantifies the increase in spatial resolution of CBCT compared to MSCT. Using such scanners pre-operatively helps to estimate the insertion depth of the CI. Using them post operatively helps to confirm the correct insertion of the CI.

Publication 2 showed the huge impairment of CI subjects for pitch perception. It was found that CI pitch perception is limited due to lack of TFS cue. In Publication 3 it was shown that pitch ranking performance was best on WDR electrodes. In Publication 4 and Publication 5 it was shown that CI subjects are able to perceive polyphony either based on rate or place pitch only when the stimulation patterns are adjusted accordingly.

5 References

1. May, BJ and Niparko, J. K. Auditory Physiology and Perception. Cochlear Implants Principles and Practices. Lippincott Williams & Wilkins, 2009.
2. Dallos P. The active cochlea. *J Neurosci* 1992;12:4575-4585.
3. Rhode WS. Observations of the vibration of the basilar membrane in squirrel monkeys using the Mossbauer technique. *J Acoust Soc Am* 1971;49:Suppl.
4. Robles L, Ruggero MA, Rich NC. Basilar membrane mechanics at the base of the chinchilla cochlea. I. Input-output functions, tuning curves, and response phases. *J Acoust Soc Am* 1986;80:1364-1374.
5. Hudspeth AJ. How the ear's works work: mechano-electrical transduction and amplification by hair cells. *C R Biol* 2005;328:155-162.
6. LeMasurier M, Gillespie PG. Hair-cell mechanotransduction and cochlear amplification. *Neuron* 2005;48:403-415.
7. Brownell WE, Bader CR, Bertrand D et al. Evoked mechanical responses of isolated cochlear outer hair cells. *Science* 1985;227:194-196.
8. Greenwood DD. A cochlear frequency-position function for several species--29 years later. *J Acoust Soc Am* 1990;87:2592-2605.
9. Simon E, Perrot X, Mertens P. [Functional anatomy of the cochlear nerve and the central auditory system]. *Neurochirurgie* 2009;55:120-126.
10. Moore BC. Temporal analysis in normal and impaired hearing. *Ann N Y Acad Sci* 1993;682:119-136.
11. Fletcher H, Munson W. Loudness, its definition, measurement and calculation. *J Acoust Soc Am* 1933;82-108.
12. Relkin EM, Doucet JR. Is loudness simply proportional to the auditory nerve spike count? *J Acoust Soc Am* 1997;101:2735-2740.

13. Stakhovskaya O, Sridhar D, Bonham BH et al. Frequency map for the human cochlear spiral ganglion: implications for cochlear implants. *J Assoc Res Otolaryngol* 2007;8:220-233.
14. Chan DK, Hudspeth AJ. Ca²⁺ current-driven nonlinear amplification by the mammalian cochlea in vitro. *Nat Neurosci* 2005;8:149-155.
15. Ryan AF. Protection of auditory receptors and neurons: evidence for interactive damage. *Proc Natl Acad Sci U S A* 2000;97:6939-6940.
16. Hudspeth AJ, Jacobs R. Stereocilia mediate transduction in vertebrate hair cells (auditory system/cilium/vestibular system). *Proc Natl Acad Sci U S A* 1979;76:1506-1509.
17. Beurg M, Fettiplace R, Nam JH et al. Localization of inner hair cell mechanotransducer channels using high-speed calcium imaging. *Nat Neurosci* 2009;12:553-558.
18. Goodyear RJ, Marcotti W, Kros CJ et al. Development and properties of stereociliary link types in hair cells of the mouse cochlea. *J Comp Neurol* 2005;485:75-85.
19. Kazmierczak P, Muller U. Sensing sound: molecules that orchestrate mechanotransduction by hair cells. *Trends Neurosci* 2012;35:220-229.
20. Corey DP, Hudspeth AJ. Ionic basis of the receptor potential in a vertebrate hair cell. *Nature* 1979;281:675-677.
21. Shotwell SL, Jacobs R, Hudspeth AJ. Directional sensitivity of individual vertebrate hair cells to controlled deflection of their hair bundles. *Ann N Y Acad Sci* 1981;374:1-10.
22. Ohmori H. Mechano-electrical transduction currents in isolated vestibular hair cells of the chick. *J Physiol* 1985;359:189-217.
23. Takumi Y, Bergersen L, Landsend AS et al. Synaptic arrangement of glutamate receptors. *Prog Brain Res* 1998;116:105-121.
24. Jacob S, Pienkowski M, Fridberger A. The endocochlear potential alters cochlear micromechanics. *Biophys J* 2011;100:2586-2594.

5 References

25. Patterson RD, Allerhand MH, Giguere C. Time-domain modeling of peripheral auditory processing: a modular architecture and a software platform. *J Acoust Soc Am* 1995;98:1890-1894.
26. Rose JE, Brugge JF, Anderson DJ et al. Phase-locked response to low-frequency tones in single auditory nerve fibers of the squirrel monkey. *J Neurophysiol* 1967;30:769-793.
27. Geisler C. *From Sound to Synapse*. New York: Oxford University Press, 1998.
28. American National Standards Institute. American National Standard Acoustical Terminology for physiological and psychological acoustics. ANSI 1994;S1.1-1994.
29. Plack, C. and Oxenham, A. J. *Pitch - Neural coding and perception*. Springer-Verlag, 2005.
30. Donnelly, P. J. and Limb, C. J. Music Perception In Cochlear Implant Users. In: Niparko JK, ed. *Cochlear Implants: Principles & Practices*. Lippincott Williams & Wilkins, 2009.
31. Schneider P, Sluming V, Roberts N et al. Structural and functional asymmetry of lateral Heschl's gyrus reflects pitch perception preference. *Nat Neurosci* 2005;8:1241-1247.
32. Johnsrude IS, Penhune VB, Zatorre RJ. Functional specificity in the right human auditory cortex for perceiving pitch direction. *Brain* 2000;123 (Pt 1):155-163.
33. Bregman A. *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge: MIT Press, 1990.
34. Darwin CJ, Hukin RW. Effectiveness of spatial cues, prosody, and talker characteristics in selective attention. *J Acoust Soc Am* 2000;107:970-977.
35. Stickney GS, Zeng FG, Litovsky R et al. Cochlear implant speech recognition with speech maskers. *J Acoust Soc Am* 2004;116:1081-1091.

36. Peterson G, Barney H. Control Methods Used in a Study of the Vowels. *J Acoust Soc Am* 1952;24.
37. Agrawal D, Timm L, Viola FC et al. ERP evidence for the recognition of emotional prosody through simulated cochlear implant strategies. *BMC Neurosci* 2012;13:113.
38. Ross ED. The aprosodias. Functional-anatomic organization of the affective components of language in the right hemisphere. *Arch Neurol* 1981;38:561-569.
39. Murray IR, Arnott JL. Toward the simulation of emotion in synthetic speech: a review of the literature on human vocal emotion. *J Acoust Soc Am* 1993;93:1097-1108.
40. Corthals, P. Speaker gender, child age, syntax, and the prosody of parentese. 2012 Annual convention of the American Speech-language Hearing Association (ASHA). 2012-11-15-2012-11-17. 2012. Ref Type: Conference Proceeding
41. Drullman R, Festen JM, Plomp R. Effect of reducing slow temporal modulations on speech reception. *J Acoust Soc Am* 1994;95:2670-2680.
42. Drullman R, Festen JM, Plomp R. Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am* 1994;95:1053-1064.
43. Fu QJ, Shannon RV. Effect of stimulation rate on phoneme recognition by nucleus-22 cochlear implant listeners. *J Acoust Soc Am* 2000;107:589-597.
44. Xu L, Thompson CS, Pfingst BE. Relative contributions of spectral and temporal cues for phoneme recognition. *J Acoust Soc Am* 2005;117:3255-3267.
45. Gordon R, Grimes B. *Ethnologue: Languages of the World*. Summer Institute of Linguistics International 2005.
46. Milczynski M, Wouters J, van WA. Improved fundamental frequency coding in cochlear implant signal processing. *J Acoust Soc Am* 2009;125:2260-2271.

5 References

47. Wei CG, Cao K, Zeng FG. Mandarin tone recognition in cochlear-implant subjects. *Hear Res* 2004;197:87-95.
48. Xu L, Li Y, Hao J et al. Tone production in Mandarin-speaking children with cochlear implants: a preliminary study. *Acta Otolaryngol* 2004;124:363-367.
49. Wei C, Cao K, Jin X et al. Psychophysical performance and Mandarin tone recognition in noise by cochlear implant users. *Ear Hear* 2007;28:62S-65S.
50. Liang Z. Auditory perceptual cues in Mandarin tones. *Acta Physiologica Sinica* 1963;26(2):85-91.
51. Buchner A, Schussler M, Battmer RD et al. Impact of low-frequency hearing. *Audiol Neurootol* 2009;14 Suppl 1:8-13.
52. Rhode WS, Greenberg S. Encoding of amplitude modulation in the cochlear nucleus of the cat. *J Neurophysiol* 1994;71:1797-1825.
53. Langner G. Periodicity coding in the auditory system. *Hear Res* 1992;60:115-142.
54. Sellick PM, Patuzzi R, Johnstone BM. Measurement of basilar membrane motion in the guinea pig using the Mossbauer technique. *J Acoust Soc Am* 1982;72:131-141.
55. Moore BC. Coding of sounds in the auditory system and its relevance to signal processing and coding in cochlear implants. *Otol Neurotol* 2003;24:243-254.
56. Oghalai JS. The cochlear amplifier: augmentation of the traveling wave within the inner ear. *Curr Opin Otolaryngol Head Neck Surg* 2004;12:431-438.
57. White RD, Grosh K. Microengineered hydromechanical cochlear model. *Proc Natl Acad Sci U S A* 2005;102:1296-1301.
58. Goldstein JL. An optimum processor theory for the central formation of the pitch of complex tones. *J Acoust Soc Am* 1973;54:1496-1516.

59. Terhardt E. Pitch, consonance, and harmony. *J Acoust Soc Am* 1974;55:1061-1069.
60. Wightman FL. The pattern-transformation model of pitch. *J Acoust Soc Am* 1973;54:407-416.
61. McKay CM, McDermott HJ, Clark GM. Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *J Acoust Soc Am* 1994;96:2664-2673.
62. Pijl S, Schwarz DW. Melody recognition and musical interval perception by deaf subjects stimulated with electrical pulse trains through single cochlear implant electrodes. *J Acoust Soc Am* 1995;98:886-895.
63. Srulovicz P, Goldstein JL. A central spectrum model: a synthesis of auditory-nerve timing and place cues in monaural communication of frequency spectrum. *J Acoust Soc Am* 1983;73:1266-1276.
64. Heinz MG, Colburn HS, Carney LH. Evaluating auditory performance limits: i. one-parameter discrimination using a computational model for the auditory nerve. *Neural Comput* 2001;13:2273-2316.
65. Joris PX, Smith PH. The volley theory and the spherical cell puzzle. *Neuroscience* 2008;154:65-76.
66. Ward WD. Subjective Musical Pitch. *J Acoust Soc Am* 1954;26:369-380.
67. Oxenham AJ, Micheyl C, Keebler MV et al. Pitch perception beyond the traditional existence region of pitch. *Proc Natl Acad Sci U S A* 2011;108:7629-7634.
68. Cariani PA, Delgutte B. Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *J Neurophysiol* 1996;76:1698-1716.
69. Meddis R, O'Mard L. A unitary model of pitch perception. *J Acoust Soc Am* 1997;102:1811-1820.

5 References

70. Moore BC, Ernst SM. Frequency difference limens at high frequencies: evidence for a transition from a temporal to a place code. *J Acoust Soc Am* 2012;132:1542-1547.
71. Attneave F, Olson RK. Pitch as a medium: a new approach to psychophysical scaling. *Am J Psychol* 1971;84:147-166.
72. Moore BC. Frequency difference limens for short-duration tones. *J Acoust Soc Am* 1973;54:610-619.
73. Semal C, Demany L. The Upper Limit of "Musical" Pitch. *Music Perception: An Interdisciplinary Journal* 1990;8:165-175.
74. Limb CJ, Rubinstein JT. Current research on music perception in cochlear implant users. *Otolaryngol Clin North Am* 2012;45:129-140.
75. Donnelly PJ, Guo BZ, Limb CJ. Perceptual fusion of polyphonic pitch in cochlear implant users. *J Acoust Soc Am* 2009;126:128-133.
76. Sonn M. American National Standard Psychoacoustical Terminology. New York: American National Standards Institute, 1973.
77. Stevens SS. The measurement of Loudness. *J Acoust Soc Am* 1955;27:815-829.
78. Hellman RP. Growth of loudness at 1000 and 3000 Hz. *J Acoust Soc Am* 1976;60:672-679.
79. Moore BC. *An Introduction to the Psychology of Hearing*. Emerald Group Publishing, 2012.
80. Howes W. Loudness function derived from data on electrical discharge rates in auditory nerve fibers. *J Acoust Soc Am* 1974;30:247-259.
81. Teich MC, Lachs G. A neural-counting model incorporating refractoriness and spread of excitation. I. Application to intensity discrimination. *J Acoust Soc Am* 1979;66:1738-1749.

82. Lachs G, Teich MC. A neural-counting model incorporating refractoriness and spread of excitation. II. Application to loudness estimation. *J Acoust Soc Am* 1981;69:774-782.
83. Lachs G, Al-Shaikh R, Bi Q et al. A neural-counting model based on physiological characteristics of the peripheral auditory system. V. Application to loudness estimation and intensity discrimination. *Systems, Man and Cybernetics, IEEE Transactions on* 1984;14:819-836.
84. Chatterjee M, Zwislocki JJ. Cochlear mechanisms of frequency and intensity coding. II. Dynamic range and the code for loudness. *Hear Res* 1998;124:170-181.
85. Liberman MC. Auditory-nerve response from cats raised in a low-noise chamber. *J Acoust Soc Am* 1978;63:442-455.
86. Heinz MG, Issa JB, Young ED. Auditory-nerve rate responses are inconsistent with common hypotheses for the neural correlates of loudness recruitment. *J Assoc Res Otolaryngol* 2005;6:91-105.
87. Tsuji J, Liberman MC. Intracellular labeling of auditory nerve fibers in guinea pig: central and peripheral projections. *J Comp Neurol* 1997;381:188-202.
88. Cohen A. Further Investigation of the Effects of Intensity upon the Pitch of Pure Tones. *J Acoust Soc Am* 1961;33:1363-1376.
89. Gelfand S. *Hearing An Introduction to the Psychological and Physiological Acoustics*. New York: Marcel Dekker Inc., 1998.
90. Spahr AJ, Dorman MF, Loisel LH. Performance of patients using different cochlear implant systems: effects of input dynamic range. *Ear Hear* 2007;28:260-275.
91. Spahr AJ, Dorman MF. Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices. *Arch Otolaryngol Head Neck Surg* 2004;130:624-628.
92. Hinojosa R, Marion M. Histopathology of profound sensorineural deafness. *Ann N Y Acad Sci* 1983;405:459-484.

5 References

93. Clark GM. Personal reflections on the multichannel cochlear implant and a view of the future. *J Rehabil Res Dev* 2008;45:651-693.
94. Zeng FG, Popper AN, Fay RR. *Auditory Prosthesis: Cochlear Implants and beyond*. New York: Springer Verlag, 2004.
95. Loizou PC. Mimicking the human ear. *IEEE Signal Processing Magazine* 1998;15:101-130.
96. Zeng FG, Shannon RV. Loudness balance between electric and acoustic stimulation. *Hear Res* 1992;60:231-235.
97. Zeng FG, Grant G, Niparko J et al. Speech dynamic range and its effect on cochlear implant performance. *J Acoust Soc Am* 2002;111:377-386.
98. Wilson BS, Lawson DT, Zerbi M et al. New processing strategies in cochlear implantation. *Am J Otol* 1995;16:669-675.
99. Holden LK, Skinner MW, Holden TA et al. Effects of stimulation rate with the Nucleus 24 ACE speech coding strategy. *Ear Hear* 2002;23:463-476.
100. Frijns JH, Klop WM, Bonnet RM et al. Optimizing the number of electrodes with high-rate stimulation of the clarion CII cochlear implant. *Acta Otolaryngol* 2003;123:138-142.
101. Bonnet RM, Boermans PP, Avenarius OF et al. Effects of pulse width, pulse rate and paired electrode stimulation on psychophysical measures of dynamic range and speech recognition in cochlear implants. *Ear Hear* 2012;33:489-496.
102. Brummer SB, Turner MJ. Electrochemical considerations for safe electrical stimulation of the nervous system with platinum electrodes. *IEEE Trans Biomed Eng* 1977;24:59-63.
103. Macherey O, Deeks JM, Carlyon RP. Extending the limits of place and temporal pitch perception in cochlear implant users. *J Assoc Res Otolaryngol* 2011;12:233-251.

104. Wilson BS, Finley CC, Lawson DT et al. Better speech recognition with cochlear implants. *Nature* 1991;352:236-238.
105. Kiefer J, Hohl S, Sturzebecher E et al. Comparison of speech recognition with different speech coding strategies (SPEAK, CIS, and ACE) and their relationship to telemetric measures of compound action potentials in the nucleus CI 24M cochlear implant system. *Audiology* 2001;40:32-42.
106. Hu Y, Loizou PC. A new sound coding strategy for suppressing noise in cochlear implants. *J Acoust Soc Am* 2008;124:498-509.
107. Hochmair I, Nopp P, Jolly C et al. MED-EL Cochlear implants: state of the art and a glimpse into the future. *Trends Amplif* 2006;10:201-219.
108. Vermeire K, Punte AK, Van de Heyning P. Better speech recognition in noise with the fine structure processing coding strategy. *ORL J Otorhinolaryngol Relat Spec* 2010;72:305-311.
109. Filipo R, Ballantyne D, Mancini P et al. Music perception in cochlear implant recipients: comparison of findings between HiRes90 and HiRes120. *Acta Otolaryngol* 2008;128:378-381.
110. Donaldson GS, Nelson DA. Place-pitch sensitivity and its relation to consonant recognition by cochlear implant listeners using the MPEAK and SPEAK speech processing strategies. *J Acoust Soc Am* 2000;107:1645-1658.
111. Kwon BJ, Perry TT, Olmstead VL. Effects of stimulation configurations on place pitch discrimination in cochlear implants. *J Acoust Soc Am* 2011;129:3818-3826.
112. Collins LM, Zwolan TA, Wakefield GH. Comparison of electrode discrimination, pitch ranking, and pitch scaling data in postlingually deafened adult cochlear implant subjects. *J Acoust Soc Am* 1997;101:440-455.
113. Henry BA, McKay CM, McDermott HJ et al. The relationship between speech perception and electrode discrimination in cochlear implantees. *J Acoust Soc Am* 2000;108:1269-1280.

5 References

114. Erixon E, Hogstorp H, Wadin K et al. Variational anatomy of the human cochlea: implications for cochlear implantation. *Otol Neurotol* 2009;30:14-22.
115. Meshik X, Holden TA, Chole RA et al. Optimal cochlear implant insertion vectors. *Otol Neurotol* 2010;31:58-63.
116. Helbig S, Settevendemie C, Mack M et al. Evaluation of an electrode prototype for atraumatic cochlear implantation in hearing preservation candidates: preliminary results from a temporal bone study. *Otol Neurotol* 2011;32:419-423.
117. Martinez-Monedero R, Niparko JK, Aygun N. Cochlear coiling pattern and orientation differences in cochlear implant candidates. *Otol Neurotol* 2011;32:1086-1093.
118. Hamzavi J, Arnoldner C. Effect of deep insertion of the cochlear implant electrode array on pitch estimation and speech perception. *Acta Otolaryngol* 2006;126:1182-1187.
119. Hochmair I, Arnold W, Nopp P et al. Deep electrode insertion in cochlear implants: apical morphology, electrodes and speech perception results. *Acta Otolaryngol* 2003;123:612-617.
120. Adunka O, Unkelbach MH, Mack MG et al. Predicting basal cochlear length for electric-acoustic stimulation. *Arch Otolaryngol Head Neck Surg* 2005;131:488-492.
121. Ruivo J, Mermuys K, Bacher K et al. Cone beam computed tomography, a low-dose imaging technique in the postoperative assessment of cochlear implantation. *Otol Neurotol* 2009;30:299-303.
122. Tavassolie TS, Penninger RT, Zuniga MG et al. Multislice computed tomography in the diagnosis of superior canal dehiscence: how much error, and how to minimize it? *Otol Neurotol* 2012;33:215-222.
123. Arweiler-Harbeck D, Monninghoff C, Greve J et al. Imaging of Electrode Position after Cochlear Implantation with Flat Panel CT. *ISRN Otolaryngol* 2012;2012:728205.

124. Tykocinski M, Cohen LT, Pyman BC et al. Comparison of electrode position in the human cochlea using various perimodiolar electrode arrays. *Am J Otol* 2000;21:205-211.
125. Dorman MF, Loizou PC. Speech intelligibility as a function of the number of channels of stimulation for normal-hearing listeners and patients with cochlear implants. *Am J Otol* 1997;18:S113-S114.
126. Fishman KE, Shannon RV, Slattery WH. Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor. *J Speech Lang Hear Res* 1997;40:1201-1215.
127. Friesen LM, Shannon RV, Baskent D et al. Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants. *J Acoust Soc Am* 2001;110:1150-1163.
128. Holmes AE, Kemker FJ, Merwin GE. The effects of varying the number of cochlear implant electrodes on speech perception. *Am J Otol* 1987;8:240-246.
129. Dorman M, Dankowski K, McCandless G et al. Consonant recognition as a function of the number of channels of stimulation by patients who use the Symbion cochlear implant. *Ear Hear* 1989;10:288-291.
130. Kileny PR, Zimmerman-Phillips S, Zwolan TA et al. Effects of channel number and place of stimulation on performance with the Cochlear Corporation multichannel implant. *Am J Otol* 1992;13:117-123.
131. Geier LL, Norton SJ. The effects of limiting the number of Nucleus 22 cochlear implant electrodes programmed on speech perception. *Ear Hear* 1992;13:340-348.
132. Dorman MF, Loizou PC, Kemp LL et al. Word recognition by children listening to speech processed into a small number of channels: data from normal-hearing children and children with cochlear implants. *Ear Hear* 2000;21:590-596.
133. Wang W, Zhou N, Xu L. Musical pitch and lexical tone perception with cochlear implants. *Int J Audiol* 2011;50:270-278.

5 References

134. McDermott H, Sucher C, Simpson A. Electro-acoustic stimulation. Acoustic and electric pitch comparisons. *Audiol Neurootol* 2009;14 Suppl 1:2-7.
135. Baumann U, Nobbe A. The cochlear implant electrode-pitch function. *Hear Res* 2006;213:34-42.
136. Blamey PJ, Dooley GJ, Parisi ES et al. Pitch comparisons of acoustically and electrically evoked auditory sensations. *Hear Res* 1996;99:139-150.
137. Boex C, Baud L, Cosendai G et al. Acoustic to electric pitch comparisons in cochlear implant subjects with residual hearing. *J Assoc Res Otolaryngol* 2006;7:110-124.
138. Dorman MF, Spahr T, Gifford R et al. An electric frequency-to-place map for a cochlear implant patient with hearing in the nonimplanted ear. *J Assoc Res Otolaryngol* 2007;8:234-240.
139. Reiss LA, Turner CW, Erenberg SR et al. Changes in pitch with a cochlear implant over time. *J Assoc Res Otolaryngol* 2007;8:241-257.
140. Carlyon RP, Deeks JM, McKay CM. The upper limit of temporal pitch for cochlear-implant listeners: stimulus duration, conditioner pulses, and the number of electrodes stimulated. *J Acoust Soc Am* 2010;127:1469-1478.
141. Rubinstein JT, Wilson BS, Finley CC et al. Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation. *Hear Res* 1999;127:108-118.
142. van den Honert C, Stypulkowski PH. Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. *Hear Res* 1987;29:207-222.
143. Shepherd RK, Javel E. Electrical stimulation of the auditory nerve. I. Correlation of physiological responses with cochlear status. *Hear Res* 1997;108:112-144.

144. Javel E, Shepherd RK. Electrical stimulation of the auditory nerve. III. Response initiation sites and temporal fine structure. *Hear Res* 2000;140:45-76.
145. Middlebrooks JC, Snyder RL. Selective electrical stimulation of the auditory nerve activates a pathway specialized for high temporal acuity. *J Neurosci* 2010;30:1937-1946.
146. Zeng FG. Temporal pitch in electric hearing. *Hear Res* 2002;174:101-106.
147. Shannon RV. Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. *Hear Res* 1983;11:157-189.
148. Shannon RV. Multichannel electrical stimulation of the auditory nerve in man. II. Channel interaction. *Hear Res* 1983;12:1-16.
149. Tong YC, Clark GM. Absolute identification of electric pulse rates and electrode positions by cochlear implant patients. *J Acoust Soc Am* 1985;77:1881-1888.
150. Townshend B, Cotter N, Van CD et al. Pitch perception by cochlear implant subjects. *J Acoust Soc Am* 1987;82:106-115.
151. Baumann U, Nobbe A. Pulse rate discrimination with deeply inserted electrode arrays. *Hear Res* 2004;196:49-57.
152. Kong YY, Carlyon RP. Temporal pitch perception at high rates in cochlear implants. *J Acoust Soc Am* 2010;127:3114-3123.
153. Kong YY, Deeks JM, Axon PR et al. Limits of temporal pitch in cochlear implants. *J Acoust Soc Am* 2009;125:1649-1657.
154. Macherey O, Deeks JM, Carlyon RP. Extending the limits of place and temporal pitch perception in cochlear implant users. *J Assoc Res Otolaryngol* 2011;12:233-251.
155. Parkins CW. Temporal response patterns of auditory nerve fibers to electrical stimulation in deafened squirrel monkeys. *Hear Res* 1989;41:137-168.

5 References

156. Pijl S. Pulse rate matching by cochlear implant patients: effects of loudness randomization and electrode position. *Ear Hear* 1997;18:316-325.
157. Arnoldner C, Kaider A, Hamzavi J. The role of intensity upon pitch perception in cochlear implant recipients. *Laryngoscope* 2006;116:1760-1765.
158. Carlyon RP, Lynch C, Deeks JM. Effect of stimulus level and place of stimulation on temporal pitch perception by cochlear implant users. *J Acoust Soc Am* 2010;127:2997-3008.
159. Loeb GE. Are cochlear implant patients suffering from perceptual dissonance? *Ear Hear* 2005;26:435-450.
160. Carlyon RP, Long CJ, Micheyl C. Across-channel timing differences as a potential code for the frequency of pure tones. *J Assoc Res Otolaryngol* 2012;13:159-171.
161. Fearn R, Wolfe J. Relative importance of rate and place: experiments using pitch scaling techniques with cochlear implants recipients. *Ann Otol Rhinol Laryngol Suppl* 2000;185:51-53.
162. Geurts L, Wouters J. Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *J Acoust Soc Am* 2001;109:713-726.
163. Laneau J, Wouters J, Moonen M. Improved music perception with explicit pitch coding in cochlear implants. *Audiol Neurootol* 2006;11:38-52.
164. Marks RJ. *Introduction to Shannon Sampling and Interpolation Theory*. New York: Springer Verlag, 1991.
165. McDermott HJ. Music perception with cochlear implants: a review. *Trends Amplif* 2004;8:49-82.
166. Landsberger DM. Effects of modulation wave shape on modulation frequency discrimination with electrical hearing. *J Acoust Soc Am* 2008;124:EL21-EL27.

167. Skinner MW, Holden LK, Holden TA et al. Speech recognition at simulated soft, conversational, and raised-to-loud vocal efforts by adults with cochlear implants. *J Acoust Soc Am* 1997;101:3766-3782.
168. Zeng FG, Galvin JJ, III. Amplitude mapping and phoneme recognition in cochlear implant listeners. *Ear Hear* 1999;20:60-74.
169. Dunn HK, White SD. Statistical measurements on conventional speech. *J Assoc Res Otolaryngol* 1940;11:278-288.
170. Beranek LL. The design of speech communication systems. *Proc Inst Radio Eng* 1947;35:880-890.
171. Fletcher H. *Speech and Hearing in Communication*. New York: van Nostrand, 1953.
172. Cox RM, Matesich JS, Moore JN. Distribution of short-term rms levels in conversational speech. *J Acoust Soc Am* 1988;84:1100-1104.
173. Stobich B, Zierhofer CM, Hochmair ES. Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. *Ear Hear* 1999;20:104-116.
174. Studebaker GA, Sherbecoe RL, McDaniel DM et al. Monosyllabic word recognition at higher-than-normal speech and noise levels. *J Acoust Soc Am* 1999;105:2431-2444.
175. Hong RS, Rubinstein JT, Wehner D et al. Dynamic range enhancement for cochlear implants. *Otol Neurotol* 2003;24:590-595.
176. Morse RP, Morse PF, Nunn TB et al. The effect of Gaussian noise on the threshold, dynamic range, and loudness of analogue cochlear implant stimuli. *J Assoc Res Otolaryngol* 2007;8:42-53.
177. Shannon RV. Threshold and loudness functions for pulsatile stimulation of cochlear implants. *Hear Res* 1985;18:135-143.
178. Shannon RV. A model of threshold for pulsatile electrical stimulation of cochlear implants. *Hear Res* 1989;40:197-204.

5 References

179. Shannon RV. Threshold functions for electrical stimulation of the human cochlear nucleus. *Hear Res* 1989;40:173-177.
180. Skinner MW, Holden LK, Holden TA et al. Effect of stimulation rate on cochlear implant recipients' thresholds and maximum acceptable loudness levels. *J Am Acad Audiol* 2000;11:203-213.
181. Vandali AE, Whitford LA, Plant KL et al. Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system. *Ear Hear* 2000;21:608-624.
182. van WA, Carlyon RP, Macherey O et al. Effects of pulse rate on thresholds and loudness of biphasic and alternating monophasic pulse trains in electrical hearing. *Hear Res* 2006;220:49-60.
183. Kreft HA, Donaldson GS, Nelson DA. Effects of pulse rate and electrode array design on intensity discrimination in cochlear implant users. *J Acoust Soc Am* 2004;116:2258-2268.
184. van WA, Carlyon RP, Laneau J et al. Effects of waveform shape on human sensitivity to electrical stimulation of the inner ear. *Hear Res* 2005;200:73-86.
185. Macherey O, van WA, Carlyon RP et al. Asymmetric pulses in cochlear implants: effects of pulse shape, polarity, and rate. *J Assoc Res Otolaryngol* 2006;7:253-266.
186. Macherey O, Carlyon RP, van WA et al. Higher sensitivity of human auditory nerve fibers to positive electrical currents. *J Assoc Res Otolaryngol* 2008;9:241-251.
187. van WA, Macherey O, Carlyon RP et al. Alternative pulse shapes in electrical hearing. *Hear Res* 2008;242:154-163.
188. Brill SM, Gstottner W, Helms J et al. Optimization of channel number and stimulation rate for the fast continuous interleaved sampling strategy in the COMBI 40+. *Am J Otol* 1997;18:S104-S106.
189. Loizou PC, Poroy O, Dorman M. The effect of parametric variations of cochlear implant processors on speech understanding. *J Acoust Soc Am* 2000;108:790-802.

190. Litvak LM, Delgutte B, Eddington DK. Improved temporal coding of sinusoids in electric stimulation of the auditory nerve using desynchronizing pulse trains. *J Acoust Soc Am* 2003;114:2079-2098.
191. Litvak LM, Smith ZM, Delgutte B et al. Desynchronization of electrically evoked auditory-nerve activity by high-frequency pulse trains of long duration. *J Acoust Soc Am* 2003;114:2066-2078.
192. Runge-Samuels CL, Abbas PJ, Rubinstein JT et al. Response of the auditory nerve to sinusoidal electrical stimulation: effects of high-rate pulse trains. *Hear Res* 2004;194:1-13.
193. Mino H. Information Rate of Neural Spike Trains in Response to Sinusoidal Electric Stimuli in the Presence of a Pseudo-spontaneous Activity. *Conf Proc IEEE Eng Med Biol Soc* 2005;1:417-420.
194. Verschuur CA. Effect of stimulation rate on speech perception in adult users of the Med-El CIS speech processing strategy. *Int J Audiol* 2005;44:58-63.
195. Hong RS, Turner CW. Pure-tone auditory stream segregation and speech perception in noise in cochlear implant recipients. *J Acoust Soc Am* 2006;120:360-374.
196. Nie K, Barco A, Zeng FG. Spectral and temporal cues in cochlear implant speech perception. *Ear Hear* 2006;27:208-217.
197. Runge-Samuels CL. Effects of high-rate pulse trains on electrode discrimination in cochlear implant users. *Trends Amplif* 2009;13:76-86.
198. Hughes ML, Castioni EE, Goehring JL et al. Temporal response properties of the auditory nerve: data from human cochlear-implant recipients. *Hear Res* 2012;285:46-57.
199. Kreft HA, Donaldson GS, Nelson DA. Effects of pulse rate on threshold and dynamic range in Clarion cochlear-implant users. *J Acoust Soc Am* 2004;115:1885-1888.

5 References

200. Middlebrooks JC. Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *J Acoust Soc Am* 2004;116:452-468.
201. Plant K, Holden L, Skinner M et al. Clinical evaluation of higher stimulation rates in the nucleus research platform 8 system. *Ear Hear* 2007;28:381-393.
202. Middlebrooks JC. Cochlear-implant high pulse rate and narrow electrode configuration impair transmission of temporal information to the auditory cortex. *J Neurophysiol* 2008;100:92-107.
203. Shannon RV, Cruz RJ, Galvin JJ. Effect of stimulation rate on cochlear implant users' phoneme, word and sentence recognition in quiet and in noise. *Audiol Neurootol* 2011;16:113-123.
204. Green T, Faulkner A, Rosen S. Variations in carrier pulse rate and the perception of amplitude modulation in cochlear implant users. *Ear Hear* 2012;33:221-230.
205. Belden CJ, Weg N, Minor LB et al. CT Evaluation of Bone Dehiscence of the Superior Semicircular Canal as a Cause of Sound- and/or Pressure-induced Vertigo. *Radiology* 2003;226:337-343.
206. Ueda T, Mori K, Minami M et al. Trends in oncological CT imaging: clinical application of multidetector-row CT and 3D-CT imaging. *Int J Clin Oncol* 2006;11:268-277.
207. Hendee WR. An opportunity for radiology. *Radiology* 2006;238:389-394.
208. Gupta R, Bartling SH, Basu SK et al. Experimental flat-panel high-spatial-resolution volume CT of the temporal bone. *AJNR Am J Neuroradiol* 2004;25:1417-1424.
209. Nimmons GL, Kang RS, Drennan WR et al. Clinical assessment of music perception in cochlear implant listeners. *Otol Neurotol* 2008;29:149-155.
210. Zeng FG. Temporal pitch in electric hearing. *Hear Res* 2002;174:101-106.

211. Moller AR. Review of the roles of temporal and place coding of frequency in speech discrimination. *Acta Otolaryngol* 1999;119:424-430.
212. Chatterjee M, Zwislocki JJ. Cochlear mechanisms of frequency and intensity coding. I. The place code for pitch. *Hear Res* 1997;111:65-75.
213. Yost WA. Pitch of iterated rippled noise. *J Acoust Soc Am* 1996;100:511-518.
214. Yost WA, Patterson R, Sheft S. A time domain description for the pitch strength of iterated rippled noise. *J Acoust Soc Am* 1996;99:1066-1078.
215. Yost WA. Iterated rippled noise discrimination at long durations. *J Acoust Soc Am* 2009;126:1336-1341.
216. Yost WA, Patterson R, Sheft S. The role of the envelope in processing iterated rippled noise. *J Acoust Soc Am* 1998;104:2349-2361.
217. Henry BA, Turner CW, Behrens A. Spectral peak resolution and speech recognition in quiet: normal hearing, hearing impaired, and cochlear implant listeners. *J Acoust Soc Am* 2005;118:1111-1121.
218. Won JH, Jones GL, Drennan WR et al. Evidence of across-channel processing for spectral-ripple discrimination in cochlear implant listeners. *J Acoust Soc Am* 2011;130:2088-2097.
219. Litvak LM, Spahr AJ, Saoji AA et al. Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners. *J Acoust Soc Am* 2007;122:982-991.
220. Simon T. Two-Dimensional Monte Carlo Simulation and Beyond: A Comparison of Several Probabilistic Risk Assessment Methods Applied to a Superfund Site. *Human and Ecological Risk Assessment: An International Journal* 1997.
221. Fogerty D, Humes LE. A correlational method to concurrently measure envelope and temporal fine structure weights: effects of

5 References

- age, cochlear pathology, and spectral shaping. *J Acoust Soc Am* 2012;132:1679-1689.
222. Lorenzi C, Debruelle L, Garnier S et al. Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal. *J Acoust Soc Am* 2009;125:27-30.
223. Lorenzi C, Gilbert G, Carn H et al. Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proc Natl Acad Sci U S A* 2006;103:18866-18869.
224. Moore BC, Glasberg BR, Hopkins K. Frequency discrimination of complex tones by hearing-impaired subjects: Evidence for loss of ability to use temporal fine structure. *Hear Res* 2006;222:16-27.
225. Henry KS, Heinz MG. Diminished temporal coding with sensorineural hearing loss emerges in background noise. *Nat Neurosci* 2012;15:1362-1364.
226. Imenov NS, Won JH, Drennan WR et al. Detection of acoustic temporal fine structure by cochlear implant listeners: Behavioral results and computational modeling. *Hear Res* 2013.
227. Gilbert G, Lorenzi C. The ability of listeners to use recovered envelope cues from speech fine structure. *J Acoust Soc Am* 2006;119:2438-2444.
228. Swanson, B. Pitch Perception with Cochlear Implants. The University of Melbourne . 2008.
- Ref Type: Thesis/Dissertation
229. Boothroyd A, Erickson FN, Medwetsky L. The hearing aid input: a phonemic approach to assessing the spectral distribution of speech. *Ear Hear* 1994;15:432-442.
230. Fu QJ, Shannon RV. Effect of acoustic dynamic range on phoneme recognition in quiet and noise by cochlear implant users. *J Acoust Soc Am* 1999;106:L65-L70.

231. Won JH, Drennan WR, Rubinstein JT. Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J Assoc Res Otolaryngol* 2007;8:384-392.
 232. Won JH, Drennan WR, Nie K et al. Acoustic temporal modulation detection and speech perception in cochlear implant listeners. *J Acoust Soc Am* 2011;130:376-388.
 233. Drennan WR, Won JH, Dasika VK et al. Effects of temporal fine structure on the lateralization of speech and on speech understanding in noise. *J Assoc Res Otolaryngol* 2007;8:373-383.
 234. Russo FA, Ives DT, Goy H et al. Age-related difference in melodic pitch perception is probably mediated by temporal processing: empirical and computational evidence. *Ear Hear* 2012;33:177-186.
 235. Goldsworthy, R. Personal Communication about Pitch Perception. 10-10-2011.
- Ref Type: Unpublished Work
236. Gfeller K, Turner C, Mehr M et al. Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. *Cochlear Implants Int* 2002;3:29-53.
 237. Sucher CM, McDermott HJ. Pitch ranking of complex tones by normally hearing subjects and cochlear implant users. *Hear Res* 2007;230:80-87.
 238. Rubinstein JT. How cochlear implants encode speech. *Curr Opin Otolaryngol Head Neck Surg* 2004;12:444-448.
 239. McKay CM, McDermott HJ, Clark GM. Pitch matching of amplitude-modulated current pulse trains by cochlear implantees: the effect of modulation depth. *J Acoust Soc Am* 1995;97:1777-1785.
 240. Blamey PJ, Pyman BC, Gordon M et al. Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients. *Ann Otol Rhinol Laryngol* 1992;101:342-348.

5 References

241. Pfingst BE, Holloway LA, Zwolan TA et al. Effects of stimulus level on electrode-place discrimination in human subjects with cochlear implants. *Hear Res* 1999;134:105-115.
242. Landsberger DM, McKay CM. Perceptual differences between low and high rates of stimulation on single electrodes for cochlear implantees. *J Acoust Soc Am* 2005;117:319-327.
243. Green T, Faulkner A, Rosen S. Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *J Acoust Soc Am* 2004;116:2298-2310.
244. Galvin JJ, Fu QJ. Effects of stimulation rate, mode and level on modulation detection by cochlear implant users. *J Assoc Res Otolaryngol* 2005;6:269-279.
245. Donaldson GS, Kreft HA, Litvak L. Place-pitch discrimination of single-versus dual-electrode stimuli by cochlear implant users (L). *J Acoust Soc Am* 2005;118:623-626.
246. Blamey P, Arndt P, Bergeron F et al. Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants. *Audiol Neurootol* 1996;1:293-306.
247. Drennan WR, Rubinstein JT. Music perception in cochlear implant users and its relationship with psychophysical capabilities. *J Rehabil Res Dev* 2008;45:779-789.
248. Arnoldner C, Riss D, Brunner M et al. Speech and music perception with the new fine structure speech coding strategy: preliminary results. *Acta Otolaryngol* 2007;127:1298-1303.
249. Zwolan TA, Collins LM, Wakefield GH. Electrode discrimination and speech recognition in postlingually deafened adult cochlear implant subjects. *J Acoust Soc Am* 1997;102:3673-3685.
250. Pfingst BE, Burkholder-Juhasz RA, Xu L et al. Across-site patterns of modulation detection in listeners with cochlear implants. *J Acoust Soc Am* 2008;123:1054-1062.

251. Pfingst BE, Burkholder-Juhasz RA, Zwolan TA et al. Psychophysical assessment of stimulation sites in auditory prosthesis electrode arrays. *Hear Res* 2008;242:172-183.
252. Bierer JA, Faulkner KF. Identifying cochlear implant channels with poor electrode-neuron interface: partial tripolar, single-channel thresholds and psychophysical tuning curves. *Ear Hear* 2010;31:247-258.
253. Penninger RT, Tavassolie TS, Carey JP. Cone-Beam Volumetric Tomography for Applications in the Temporal Bone. *Otol Neurotol* 2011.
254. Firszt JB, Koch DB, Downing M et al. Current steering creates additional pitch percepts in adult cochlear implant recipients. *Otol Neurotol* 2007;28:629-636.
255. McKay CM, McDermott HJ. The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. *J Acoust Soc Am* 1996;100:1081-1092.
256. Oxenham AJ. Pitch perception and auditory stream segregation: implications for hearing loss and cochlear implants. *Trends Amplif* 2008;12:316-331.
257. Temperley D. *The Cognition of Basic Musical Structures*. Cambridge: Massachusetts Institute of Technology, 2001.
258. Moore BC, Sek A. Detection of frequency modulation at low modulation rates: evidence for a mechanism based on phase locking. *J Acoust Soc Am* 1996;100:2320-2331.
259. Penninger RT, Limb CJ, Vermeire K et al. Experimental Assessment of Polyphonic Tones with Cochlear Implants. *Otol Neurotol* 2013.
260. Penninger RT, Chien WW, Jiradejyong P et al. Perception of pure tones and iterated rippled noise for normal hearing and cochlear implant users. *Trends Amplif* 2013;17:45-53.

5 References

261. Rubinstein JT, Hong R. Signal coding in cochlear implants: exploiting stochastic effects of electrical stimulation. *Ann Otol Rhinol Laryngol Suppl* 2003;191:14-19.
262. Kong YY, Stickney GS, Zeng FG. Speech and melody recognition in binaurally combined acoustic and electric hearing. *J Acoust Soc Am* 2005;117:1351-1361.
263. Turner CW, Gantz BJ, Vidal C et al. Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing. *J Acoust Soc Am* 2004;115:1729-1735.
264. Tobey EA, Shin S, Sundarajan M et al. Spoken word recognition in adolescent cochlear implant users during quiet and multispeaker babble conditions. *Otol Neurotol* 2011;32:413-418.
265. Krumhansl CL, Iverson P. Perceptual interactions between musical pitch and timbre. *J Exp Psychol Hum Percept Perform* 1992;18:739-751.
266. Looi V, Gfeller K, Driscoll V. MUSIC APPRECIATION AND TRAINING FOR COCHLEAR IMPLANT RECIPIENTS: A REVIEW. *Semin Hear* 2012;33:307-334.
267. Hsiao F, Gfeller K. Music Perception of Cochlear Implant Recipients with Implications for Music Instruction: A Review of Literature. *Update Univ S C Dep Music* 2012;30:5-10.
268. Limb CJ, Roy AT. Technological, biological, and acoustical constraints to music perception in cochlear implant users. *Hear Res* 2013.
269. Roy AT, Jiradejvong P, Carver C et al. Assessment of sound quality perception in cochlear implant users during music listening. *Otol Neurotol* 2012;33:319-327.
270. Roy AT, Jiradejvong P, Carver C et al. Musical sound quality impairments in cochlear implant (CI) users as a function of limited high-frequency perception. *Trends Amplif* 2012;16:191-200.

271. Snel-Bongers J, Briaire JJ, Vanpoucke FJ et al. Spread of excitation and channel interaction in single- and dual-electrode cochlear implant stimulation. *Ear Hear* 2012;33:367-376.
272. Cosetti MK, Waltzman SB. Outcomes in cochlear implantation: variables affecting performance in adults and children. *Otolaryngol Clin North Am* 2012;45:155-171.
273. Wright A, Davis A, Bredberg G et al. Hair cell distributions in the normal human cochlea. *Acta Otolaryngol Suppl* 1987;444:1-48.
274. Ketten DR, Skinner MW, Wang G et al. In vivo measures of cochlear length and insertion depth of nucleus cochlear implant electrode arrays. *Ann Otol Rhinol Laryngol Suppl* 1998;175:1-16.
275. Adunka O, Unkelbach MH, Mack M et al. Cochlear implantation via the round window membrane minimizes trauma to cochlear structures: a histologically controlled insertion study. *Acta Otolaryngol* 2004;124:807-812.
276. Yukawa K, Cohen L, Blamey P et al. Effects of insertion depth of cochlear implant electrodes upon speech perception. *Audiol Neurootol* 2004;9:163-172.
277. Landwehr M, Furstenberg D, Walger M et al. Effects of various electrode configurations on music perception, intonation and speaker gender identification. *Cochlear Implants Int* 2013.
278. Adunka O, Kiefer J. Impact of electrode insertion depth on intracochlear trauma. *Otolaryngol Head Neck Surg* 2006;135:374-382.
279. Zeng FG, Rebscher S, Harrison W et al. Cochlear implants: system design, integration, and evaluation. *IEEE Rev Biomed Eng* 2008;1:115-142.
280. Shannon RV, Zeng FG, Wygonski J. Speech recognition with altered spectral distribution of envelope cues. *J Acoust Soc Am* 1998;104:2467-2476.

5 References

281. Fu QJ, Shannon RV. Recognition of spectrally degraded and frequency-shifted vowels in acoustic and electric hearing. *J Acoust Soc Am* 1999;105:1889-1900.
282. Nardo WD, Cantore I, Marchese MR et al. Electric to acoustic pitch matching: a possible way to improve individual cochlear implant fitting. *Eur Arch Otorhinolaryngol* 2008;265:1321-1328.
283. Fu QJ, Shannon RV. Effects of electrode configuration and frequency allocation on vowel recognition with the Nucleus-22 cochlear implant. *Ear Hear* 1999;20:332-344.
284. Fu QJ, Shannon RV. Effects of electrode location and spacing on phoneme recognition with the Nucleus-22 cochlear implant. *Ear Hear* 1999;20:321-331.
285. Fu QJ, Shannon RV. Frequency mapping in cochlear implants. *Ear Hear* 2002;23:339-348.
286. Rosen S, Faulkner A, Wilkinson L. Adaptation by normal listeners to upward spectral shifts of speech: implications for cochlear implants. *J Acoust Soc Am* 1999;106:3629-3636.
287. Dorman MF, Ketten D. Adaptation by a cochlear-implant patient to upward shifts in the frequency representation of speech. *Ear Hear* 2003;24:457-460.
288. Li T, Fu QJ. Perceptual adaptation to spectrally shifted vowels: training with nonlexical labels. *J Assoc Res Otolaryngol* 2007;8:32-41.
289. Nogaki G, Fu QJ, Galvin JJ, III. Effect of training rate on recognition of spectrally shifted speech. *Ear Hear* 2007;28:132-140.
290. Faulkner A. Adaptation to distorted frequency-to-place maps: implications of simulations in normal listeners for cochlear implants and electroacoustic stimulation. *Audiol Neurootol* 2006;11 Suppl 1:21-26.
291. Baskent D, Shannon RV. Speech recognition under conditions of frequency-place compression and expansion. *J Acoust Soc Am* 2003;113:2064-2076.

292. Baskent D, Shannon RV. Frequency-place compression and expansion in cochlear implant listeners. *J Acoust Soc Am* 2004;116:3130-3140.
293. Koch DB, Downing M, Osberger MJ et al. Using current steering to increase spectral resolution in CII and HiRes 90K users. *Ear Hear* 2007;28:385-415.
294. Srinivasan AG, Shannon RV, Landsberger DM. Improving virtual channel discrimination in a multi-channel context. *Hear Res* 2012;286:19-29.
295. Tang Q, Benitez R, Zeng FG. Spatial channel interactions in cochlear implants. *J Neural Eng* 2011;8:046029.
296. Shannon RV, Zeng FG, Kamath V et al. Speech recognition with primarily temporal cues. *Science* 1995;270:303-304.
297. Fu QJ, Shannon RV, Wang X. Effects of noise and spectral resolution on vowel and consonant recognition: acoustic and electric hearing. *J Acoust Soc Am* 1998;104:3586-3596.
298. Smith ZM, Delgutte B, Oxenham AJ. Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 2002;416:87-90.
299. Shannon RV, Fu QJ, Galvin J, III. The number of spectral channels required for speech recognition depends on the difficulty of the listening situation. *Acta Otolaryngol Suppl* 2004;50-54.
300. Loizou PC, Dorman M, Tu Z. On the number of channels needed to understand speech. *J Acoust Soc Am* 1999;106:2097-2103.
301. Crew JD, Galvin JJ, III, Fu QJ. Channel interaction limits melodic pitch perception in simulated cochlear implants. *J Acoust Soc Am* 2012;132:EL429-EL435.
302. Johnson DH. The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones. *J Acoust Soc Am* 1980;68:1115-1122.

5 References

303. Hong RS, Turner CW. Sequential stream segregation using temporal periodicity cues in cochlear implant recipients. *J Acoust Soc Am* 2009;126:291-299.
304. Kong YY, Zeng FG. Temporal and spectral cues in Mandarin tone recognition. *J Acoust Soc Am* 2006;120:2830-2840.
305. Gfeller K, Turner C, Oleson J et al. Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise. *Ear Hear* 2007;28:412-423.
306. Sucher CM, McDermott HJ. Bimodal stimulation: benefits for music perception and sound quality. *Cochlear Implants Int* 2009;10 Suppl 1:96-99.
307. Srinivasan AG, Landsberger DM, Shannon RV. Current focusing sharpens local peaks of excitation in cochlear implant stimulation. *Hear Res* 2010;270:89-100.
308. Shepherd RK, Hatsushika S, Clark GM. Electrical stimulation of the auditory nerve: the effect of electrode position on neural excitation. *Hear Res* 1993;66:108-120.
309. Cohen LT, Saunders E, Clark GM. Psychophysics of a prototype perimodiolar cochlear implant electrode array. *Hear Res* 2001;155:63-81.
310. Frijns JH, Briaire JJ, de Laat JA et al. Initial evaluation of the Clarion CII cochlear implant: speech perception and neural response imaging. *Ear Hear* 2002;23:184-197.
311. Wackym PA, Firszt JB, Gaggl W et al. Electrophysiologic effects of placing cochlear implant electrodes in a perimodiolar position in young children. *Laryngoscope* 2004;114:71-76.
312. van WS, Stokroos RJ, Rikers MM et al. Effect of peri-modiolar cochlear implant positioning on auditory nerve responses: a neural response telemetry study. *Acta Otolaryngol* 2005;125:725-731.
313. Cohen LT, Saunders E, Knight MR et al. Psychophysical measures in patients fitted with Contour and straight Nucleus electrode arrays. *Hear Res* 2006;212:160-175.

314. Stickney GS, Loizou PC, Mishra LN et al. Effects of electrode design and configuration on channel interactions. *Hear Res* 2006;211:33-45.
315. Snyder RL, Middlebrooks JC, Bonham BH. Cochlear implant electrode configuration effects on activation threshold and tonotopic selectivity. *Hear Res* 2008;235:23-38.
316. Runge-Samuels C, Firszt JB, Gaggl W et al. Electrically evoked auditory brainstem responses in adults and children: effects of lateral to medial placement of the nucleus 24 contour electrode array. *Otol Neurotol* 2009;30:464-470.
317. Litvak LM, Spahr AJ, Emadi G. Loudness growth observed under partially tripolar stimulation: model and data from cochlear implant listeners. *J Acoust Soc Am* 2007;122:967-981.
318. Jolly CN, Spelman FA, Clopton BM. Quadrupolar stimulation for Cochlear prostheses: modeling and experimental data. *IEEE Trans Biomed Eng* 1996;43:857-865.
319. van den Honert C, Kelsall DC. Focused intracochlear electric stimulation with phased array channels. *J Acoust Soc Am* 2007;121:3703-3716.
320. Bierer JA. Threshold and channel interaction in cochlear implant users: evaluation of the tripolar electrode configuration. *J Acoust Soc Am* 2007;121:1642-1653.
321. Kwon BJ, van den Honert C. Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users. *J Acoust Soc Am* 2006;120:EL1-EL6.
322. Saoji AA, Litvak LM. Use of "phantom electrode" technique to extend the range of pitches available through a cochlear implant. *Ear Hear* 2010;31:693-701.
323. Miller CA, Robinson BK, Rubinstein JT et al. Auditory nerve responses to monophasic and biphasic electric stimuli. *Hear Res* 2001;151:79-94.

5 References

324. Kral A, Hartmann R, Mortazavi D et al. Spatial resolution of cochlear implants: the electrical field and excitation of auditory afferents. *Hear Res* 1998;121:11-28.
325. Izzo AD, Richter CP, Jansen ED et al. Laser stimulation of the auditory nerve. *Lasers Surg Med* 2006;38:745-753.
326. Schultz M, Baumhoff P, Maier H et al. Nanosecond laser pulse stimulation of the inner ear-a wavelength study. *Biomed Opt Express* 2012;3:3332-3345.
327. Shapiro MG, Homma K, Villarreal S et al. Infrared light excites cells by changing their electrical capacitance. *Nat Commun* 2012;3:736.
328. Teudt IU, Maier H, Richter CP et al. Acoustic events and "optophonic" cochlear responses induced by pulsed near-infrared laser. *IEEE Trans Biomed Eng* 2011;58:1648-1655.
329. Wenzel GI, Balster S, Zhang K et al. Green laser light activates the inner ear. *J Biomed Opt* 2009;14:044007.
330. Zhang KY, Wenzel GI, Balster S et al. Optoacoustic induced vibrations within the inner ear. *Opt Express* 2009;17:23037-23043.
331. Stafford, RC and Stafford, JW. Cochlear Implant and method enabling enhances music perception. Lockheed Martin Corporation. 2012. USA.

Ref Type: Patent

332. Moore BC, Glasberg B, Schlueter A. Detection of dead regions in the cochlea: relevance for combined electric and acoustic stimulation. *Adv Otorhinolaryngol* 2010;67:43-50.
333. Pfungst BE, Xu L, Thompson CS. Across-site threshold variation in cochlear implants: relation to speech recognition. *Audiol Neurootol* 2004;9:341-352.
334. Lenarz, T., Schmitz, KP., Behrend, D., Sternberg, K., Williams, S., and Martin, D. Drug eluting Cochlear Implants. 12-5-2012.

Ref Type: Patent

335. Farhadi M, Jalessi M, Salehian P et al. Dexamethasone eluting cochlear implant: Histological study in animal model. *Cochlear Implants Int* 2013;14:45-50.
336. Hardie NA, Shepherd RK. Sensorineural hearing loss during development: morphological and physiological response of the cochlea and auditory brainstem. *Hear Res* 1999;128:147-165.
337. Moore DR. Auditory brainstem of the ferret: bilateral cochlear lesions in infancy do not affect the number of neurons projecting from the cochlear nucleus to the inferior colliculus. *Brain Res Dev Brain Res* 1990;54:125-130.
338. Pasic TR, Rubel EW. Rapid changes in cochlear nucleus cell size following blockade of auditory nerve electrical activity in gerbils. *J Comp Neurol* 1989;283:474-480.
339. Shepherd RK, Hardie NA. Deafness-induced changes in the auditory pathway: implications for cochlear implants. *Audiol Neurootol* 2001;6:305-318.
340. Wise AK, Gillespie LN. Drug delivery to the inner ear. *J Neural Eng* 2012;9:065002.
341. Fayad J, Linthicum FH, Jr., Otto SR et al. Cochlear implants: histopathologic findings related to performance in 16 human temporal bones. *Ann Otol Rhinol Laryngol* 1991;100:807-811.
342. Peterson NR, Pisoni DB, Miyamoto RT. Cochlear implants and spoken language processing abilities: review and assessment of the literature. *Restor Neurol Neurosci* 2010;28:237-250.
343. Green T, Faulkner A, Rosen S et al. Enhancement of temporal periodicity cues in cochlear implants: effects on prosodic perception and vowel identification. *J Acoust Soc Am* 2005;118:375-385.
344. Ito J, Sakakibara J, Iwasaki Y et al. Positron emission tomography of auditory sensation in deaf patients and patients with cochlear implants. *Ann Otol Rhinol Laryngol* 1993;102:797-801.

5 References

345. Lee HJ, Giraud AL, Kang E et al. Cortical activity at rest predicts cochlear implantation outcome. *Cereb Cortex* 2007;17:909-917.
346. Limb CJ, Molloy AT, Jiradejvong P et al. Auditory cortical activity during cochlear implant-mediated perception of spoken language, melody, and rhythm. *J Assoc Res Otolaryngol* 2010;11:133-143.
347. Naito Y, Tateya I, Fujiki N et al. Increased cortical activation during hearing of speech in cochlear implant users. *Hear Res* 2000;143:139-146.
348. Wong D, Miyamoto RT, Pisoni DB et al. PET imaging of cochlear-implant and normal-hearing subjects listening to speech and nonspeech. *Hear Res* 1999;132:34-42.
349. Aggarwal R, Green KM. Cochlear implants and positron emission tomography. *J Laryngol Otol* 2012;126:1200-1203.
350. Teoh SW, Pisoni DB, Miyamoto RT. Cochlear implantation in adults with prelingual deafness. Part II. Underlying constraints that affect audiological outcomes. *Laryngoscope* 2004;114:1714-1719.
351. Waltzman SB, Roland JT, Jr., Cohen NL. Delayed implantation in congenitally deaf children and adults. *Otol Neurotol* 2002;23:333-340.
352. Galvin JJ, Fu QJ, Nogaki G. Melodic contour identification by cochlear implant listeners. *Ear Hear* 2007;28:302-319.
353. Galvin JJ, III, Fu QJ, Shannon RV. Melodic contour identification and music perception by cochlear implant users. *Ann N Y Acad Sci* 2009;1169:518-533.
354. Gfeller K, Witt S, Adamek M et al. Effects of training on timbre recognition and appraisal by postlingually deafened cochlear implant recipients. *J Am Acad Audiol* 2002;13:132-145.
355. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. Lawrence Erlbaum Associates, Publishers, 1988.

356. Aberson C. Applied Power Analysis for the Behavioral Science. New York: Taylor & Francis Group, 2010.
357. Shpizner BA, Holliday RA, Roland JT et al. Postoperative imaging of the multichannel cochlear implant. *AJNR Am J Neuroradiol* 1995;16:1517-1524.
358. Verbist BM, Frijns JH, Geleijns J et al. Multisection CT as a valuable tool in the postoperative assessment of cochlear implant patients. *AJNR Am J Neuroradiol* 2005;26:424-429.

Acknowledgement

It is with immense gratitude that I acknowledge the support of my supervisors, Professor Dhooge and Professor Leman. It has been a great pleasure to work with them and to learn from their experience.

This thesis would have remained a dream had it not been for Prof. Limb. It was a great pleasure to work with him and his lab and his ideas and feedback were very important.

I would also like to thank Dr. Vermeire. She helped me a lot at the beginning of my project and without her many of the studies would have never been conducted.

It gives me great pleasure in acknowledging the support and help of Professor Della-Santina. He provided great feedback at the beginning of my thesis and it has always been a pleasure to work with him.

My deepest thanks goes to the European Commission who provided me with Marie Curie funding. The benefit of being in an initial training network was remarkable. Not only all the friends that I made during the numerous meetings, but also to get to know research centers all over Europe was a great experience.

Furthermore I would like to thank Prof. Büchner. He enabled me to do my secondment at the University of Hannover. Working at MHH was as remarkable experience. This increased the number of my potential CI subjects by eight and helped me to finish my experiments a lot faster.

Lastly and most importantly I would like to thank my family and my fiancée Dr. Daválos Bichara. Without their support I would have never been able to do a PhD program.

Acknowledging all the other friends and coworkers would require too many extra pages and of course I would like to thank all the participants of my experiments.