SAPIENZA
UNIVERSITÀ DI ROMA

TEMPORAL FINE STRUCTURE PROCESSING, PITCH AND SPEECH PERCEPTION IN COCHLEAR IMPLANT RECIPIENTS

PhD thesis

Defended by

Hilal Dincer D'Alessandro

Curriculum: Pathologies of the Auditory System
School of PhD in Clinical-Experimental Neuroscience and Psychiatry
Sapienza University of Rome
Rome, 28 February 2017

Advisor
Prof. Patrizia Mancini, MD PhD

DOCTORAL COMMITTEE

Chairman:	**Prof. Dr. Alfredo Berardelli**
Advisor:	**Prof. Dr. Patrizia Mancini**
Members:	**Prof. Dr. Alberto Siracusano**
	Prof. Dr. Fabio Placidi
	Prof. Dr. Stefano Ferraina
External Members:	**Prof. Dr. Fabrizio Ottaviani**
	Prof. Dr. Marco Alessandrini

ACKNOWLEDGEMENTS

Firstly, I would like to thank to all the study participants and their families without whose contributions this research could not be conducted.

I wish to express my sincere gratitude to my advisor Prof. Dr. Patrizia Mancini for providing me an opportunity to join this PhD program and for her continuous support of my PhD study.

Besides my advisor, I would like to thank to Prof. Dr. Alfredo Berardelli, Prof. Dr. Elio De Seta and Prof. Dr. Roberto Filipo for providing me an opportunity to work in the context of this PhD program.

I would like to thank to Prof. Dr. Fabrizio Ottaviani and Prof. Dr. Marco Alessandrini for reviewing my thesis and for their time as well as to my thesis committee: Prof. Dr. Alberto Siracusano, Prof. Dr. Fabio Placidi, and Prof. Dr. Stefano Ferraina, for their time and interest.

I would especially like to thank to Dr. Deborah Ballantyne, Dr. Ersilia Bosco and Dr. Maria Nicastri from Cochlear Implant Center of Policlinico Umberto I of Sapienza University of Rome. All you have been there to support me for my PhD thesis. It is always a big pleasure to collaborate with you.

I am grateful to Dr. Patrick Boyle, for his great help at the preparation of the Italian STARR test material and his continuous support and advices for the Italian adaptation.

Lastly, I would like to thank to my mother Suna and my brother Ozan for all their love and support.

And to Flavio,
Thank you to exist and to stay with me!

Prior Publication of Content

Material covered in Chapter 2 has previously been published in Dincer D'Alessandro H, Filipo R, Ballantyne D, Attanasio G, Bosco E, Nicastri M, Mancini P (2015) Low-frequency pitch perception in children with cochlear implants in comparison to normal hearing peers. Eur Arch of Otorhinolaryngol 272:3115-3122.

Material covered in Chapter 3 has previously been published in Dincer D'Alessandro H, Ballantyne D, De Seta E, Musacchio A, Mancini P (2016) Adaptation of the STARR test for adult Italian population: A speech test for a realistic estimate in real-life listening conditions. Int J Audiol 55:262-267.

ABSTRACT

Cochlear implant (CI) recipients usually complain about poor speech understanding in the presence of noise. Indeed, they generally show ceiling effects for understanding sentences presented in quiet, but their scores decrease drastically when testing in the presence of competing noise. One important aspect that contributes to speech perception skills, especially when listening in a fluctuating background, has been described as Temporal Fine Structure (TFS) processing. TFS cues are more dominant in conveying Low Frequency (LF) signals linked in particular to Fundamental Frequency (F0), which is crucial for linguistic and musical perception. A§E Harmonic Intonation (HI) and Disharmonic Intonation (DI) are tests of pitch perception in the LF domain and their outcomes are believed to depend on the availability of TFS cues. Previous findings indicated that the DI test provided more differential LF pitch perception outcomes in that it reflected phase locking and TFS processing capacities of the ear, whereas the HI test provided information on its place coding capacity as well. Previous HI/DI studies were mainly done in adult population showing abnormal pitch perception outcomes in CI recipients and there was no or limited data in paediatric population as well as HI/DI outcomes in relation to speech perception outcomes in the presence of noise.

One of the primary objectives of this thesis has been to investigate LF pitch perception skills in a group of pediatric CI recipients in comparison to normal hearing (NH) children. Another objective was to introduce a new assessment tool, the Italian STARR test which was based on measurement of speech perception using a roving-level adaptive method where the presentation level of both speech and noise signals varied across sentences. The STARR test attempts to reflect a better representation of real world listening conditions where background noise is usually present and speech intensity varies according to vocal capacity as well as the distance of the speaker. The Italian STARR outcomes in NH adults were studied to produce normative data, as well as to evaluate interlist variability and learning effects. Finally, LF pitch perception outcomes linked to availability of TFS were investigated in a group of adult CI recipients including bimodal users in relation to speech perception, in particular Italian STARR outcomes.

Results were interesting: Although the majority of CI recipient children showed abnormal outcomes for A§E, their scores were considerably better than in the adult CI users. Age had a statistically significant effect on performance in both children and adults; younger children and older adults tended to show poorer performance. Similarly, CI recipient adults (even the better performers) showed abnormal STARR outcomes in comparison to NH subjects and group differences were statistically significant. The duration of profound deafness before implantation had a significant effect on STARR performance. On the other hand, the significant effect of CI thresholds re-emphasized the sensitivity of the test to lower level speech which a CI user can face very often during everyday life. Analysis revealed statistically significant correlations between HI/DI and STARR performance. Moreover, contralateral hearing aid users showed significant bimodal benefit for both HI/DI and STARR tests. Overall findings confirmed the usefulness of evaluating both LF pitch and speech perception in order to track changes in TFS sensitivity for CI recipients over time and across different listening conditions which might be provided by future technological advances as well as to study individual differences.

Key Words: Speech audiometry; speech perception in noise; speech reception threshold (SRT); cochlear implants; pitch perception; Temporal Fine Structure

RIASSUNTO

I portatori di impianto cocleare (CI), solitamente, si lamentano della scarsa comprensione del parlato in presenza di rumore. In effetti, molti di loro mostrano di avere un ‘Effetto soffitto’ nella comprensione di frasi presentate in silenzio e, tuttavia, i punteggi delle loro performance diminuiscono drasticamente quando si effettuano in concomitanza di rumore. Un aspetto importante che influisce sulla percezione vocale, soprattutto quando l'ascolto ha uno sfondo fluttuante, è l’elaborazione di struttura fine temporale (TFS). Le informazioni di TFS sono più dominanti nel trasmettere segnali a bassa frequenza legati, in particolare, alla frequenza fondamentale (F0), la quale è cruciale per la percezione linguistica e musicale. A§E Harmonic Intonation (HI) e Disharmonic Intonation (DI) sono test di percezione di pitch nel dominio a bassa frequenza e si ritiene che i risultati ottenuti usando questi test, dipendano dalla disponibilità delle informazioni di TFS. I risultati precedenti hanno mostrato che il DI test fornisce più informazioni differenziali (rispetto a HI) sulla percezione di pitch a bassa frequenza in quanto riflettono le capacità di ‘phase locking’ e di elaborazione di TFS; mentre il test HI fornisce informazioni anche sulla capacità di ‘place coding’. Gli studi condotti in precedenza hanno preso in esame, prevalentemente, la popolazione adulta, ed hanno mostrato risultati anomali di percezione di pitch a bassa frequenza nei portatori di CI; in tali studi, i dati nella popolazione pediatrica erano sostanzialmente assenti, cosi come i risultati dei test HI/DI in relazione ai risultati della percezione vocale in presenza di rumore.

Questa tesi si è posta, quale obiettivo primario, lo studio delle abilità di percezione di pitch a bassa frequenza in un gruppo di bambini con CI rispetto alla popolazione normoudente. Un secondo obiettivo è stato quello poi di introdurre in Italiano un nuovo strumento di valutazione, chiamato STARR test, che si basa sulla misura di percezione vocale, utilizzando un metodo adattivo in cui il livello di presentazione di entrambi i segnali, vocali e di rumore, varia tra frasi. Lo STARR test si prefigge quale scopo principale quello di rappresentare in modo efficiente le condizioni di ascolto del mondo reale. I risultati dello STARR test in adulti normoudenti sono stati studiati per produrre dati normativi, nonché per valutare la variabilità delle liste e l’effetto di apprendimento. Infine, i risultati della percezione di pitch a bassa frequenza, legati alla disponibilità di TFS, sono stati studiati e valutati in un gruppo più ampio di adulti con CI, inclusi gli utenti bimodali, in relazione ai risultati dello STARR italiano.

I risultati ottenuti sono di notevole interesse: nella percezione di pitch, i punteggi dei bambini sono notevolmente migliori rispetto a quelli degli adulti con CI, anche se, in entrambi i gruppi, la maggior parte dei risultati si colloca in un intervallo anomalo. Nello studio condotto in questa tesi, l’età ha avuto un effetto statisticamente significativo sulle performance sia per i bambini, sia per gli adulti. I bambini più piccoli e gli adulti più anziani hanno avuto la tendenza a mostrare performance più scadenti. Per lo STARR test, gli adulti con CI, anche includendo quelli con la performance migliore, hanno mostrato risultati scadenti rispetto ai soggetti normoudenti, e i gruppi hanno mostrato differenze statisticamente significative. La durata della sordità profonda, prima dell'impianto, ha avuto un effetto significativo sulla performance dello STARR test. D’altra parte, l'effetto di soglie con CI ha enfatizzato la sensibilità del test per il parlato di livello basso (che un utente con CI incontra spesso nella vita quotidiana). L'analisi ha rivelato correlazioni statisticamente significative tra le performance dei test HI/DI e dello STARR test. Inoltre, i portatori di protesi acustica controlaterale hanno mostrato un beneficio bimodale significativo per entrambi i test. Nel complesso, i risultati hanno confermato l’efficacia e l’utilità della valutazione, sia della percezione di pitch a bassa frequenza, sia della percezione vocale. Inoltre, entrambi i test consentono, da un lato, di tenere traccia delle modifiche della sensibilità di TFS per i portatori del CI, nel tempo o tra le diverse condizioni di ascolto; e, dall’altro, di studiare le differenze individuali.

Parole chiave: audiometria vocale; percezione vocale in rumore; soglie di riconoscimento vocale (SRT); impianto cocleare; percezione di pitch; struttura fine temporale

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AB	Advanced Bionics
ACE	Advanced Combination Encoder
A§E	Auditory Speech Sounds Evaluation
BM	Basilar Membrane
CI	Cochlear Implant
CIS	Continuous Interleaved Sampling
DI	Disharmonic Intonation
EAS	Electro-Acoustic Stimulation
F0	Fundamental Frequency
FSP	Fine Structure Processing
FS4	Fine Structure on 4 channels
FS4-p	Parallel stimulation FS4
HA	Hearing Aid
HF	High Frequency
HI	Harmonic Intonation
HINT	Hearing in Noise Test
HiRes	HiResolution
HiRes-S	Sequential stimulation HiResolution
HiRes-P	Parallel stimulation HiResolution
HiRes 120	HiResolution with the Fidelity 120
HS	Head Shadow
IHC	Inner Hair Cells
IHR	Institute of Hearing Research
ITD	Interaural Time Delay
JND	Just Noticeable Difference
LF	Low Frequency
NH	Normal Hearing
OHC	Outer Hair Cells
OLSA	Oldenburg Sentence Test

PTA	Pure Tone Average
SAS	Simultaneous Analogue Stimulation
SD	Standard Deviation
SM	Scala Media
SNR	Signal-to-Noise Ratio
SP	Speech Processor
SPSS	Statistical Package for Social Sciences
SRT	Speech Reception Threshold
SQ	binaural Squelch
ST	Scala Tympani
STARR	Sentence Test with Adaptive Randomized Roving Level
SU	binaural SUmmation
SV	Scala Vestibuli
TFS	Temporal Fine Structure
WRS	Word Recognition Score

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1. Introduction

Hearing loss, also known as hearing impairment, is considered as one of the most common human disorders and may arise from dysfunction of any part of the auditory pathway. Nowadays, several medical, surgical and technological tools are available for the treatment of hearing loss depending mainly on type and degree. In the case of cochlear dysfunction, options may include traditional amplifying systems with Hearing Aids (HA) or Cochlear Implants (CI) which bypass the severely impaired cochlea [1] and directly stimulate the acoustic nerve.

Deafness is often caused by the absence or degeneration of sensory hair cells in the cochlea and in the case of a pathological cochlea, the logical approach would be to bypass the damaged part and to stimulate spiral ganglion cells electrically with a cochlear implant. The fundamental criterion for cochlear implantation is a majority of missing or non-functioning cochlear hair cells where information from the acoustic environment transmitted through a normal or near-normal middle ear cannot be transduced into effective electrical signals that travel along the body's natural auditory system to the brain so that comprehension can take place. If there is a retrocochlear pathology, a cochlear implant will not be useful to restore deafness [2].

1.1 Normal Hearing

There are 3 major parts to the human peripheral auditory system: the outer ear, the middle ear and the inner ear (as shown in Figure 1). The outer ear consists of the pinna (also called the auricle) and the ear canal (also called external auditory canal) ending at the eardrum. The tympanic membrane is generally considered to be part of the middle ear system which lies between the outer ear and the inner ear, and includes an air-filled cavity called the tympanic cavity and the three ossicles which are small bones that function together to receive, amplify, and transmit the sound from the tympanic membrane to the inner ear (tympano-ossicular chain). The ossicles are the malleus, incus, and the stapes. The stapes is the smallest bone in the body. The middle ear also connects to the upper throat at the nasopharynx via the pharyngeal opening of the Eustachian tube which provides aeration and drainage of the middle ear system, and makes it possible for air pressure to be the same on both sides of the eardrum.

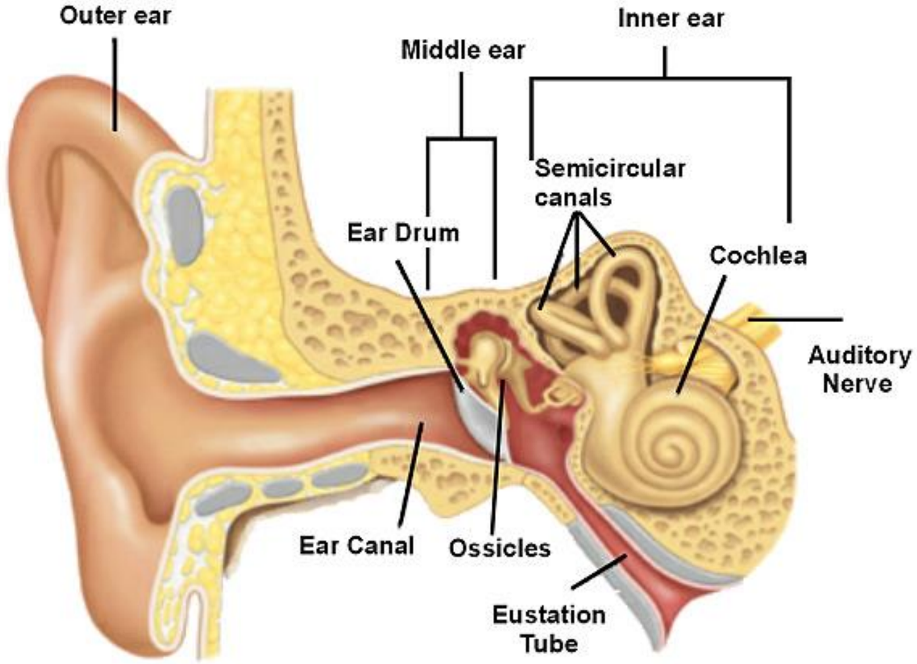

Figure 1 Major parts to the human peripheral auditory system: the outer ear, the middle ear and the inner ear. Source: http://www.erzetich-audio.com

The inner ear contains structures which are key to hearing and balance; it consists of the vestibule, which lies to the medial side of the oval window, the cochlea anteriorly, and the three semicircular canals posteriorly. The inner ear sits in the labyrinth, which passes through the temporal bone and contains a continuous membranous canal which can be considered to be a duct within a duct (as shown in Figure 2). The upper duct is called the Scala Vestibuli (SV) and the lower duct is called the Scala Tympani (ST); both contain a fluid called perilymph. They are separated by a middle duct, the Scala Media (SM), which contains endolymph. They meet at the far end of the tube at an opening called the helicotrema. The stapes at the oval window is at the base of SV and the round window is at the basal end of the ST. The SM is separated from the SV above it by Reissner's membrane, and from the ST below it by the Basilar Membrane (BM) [3].

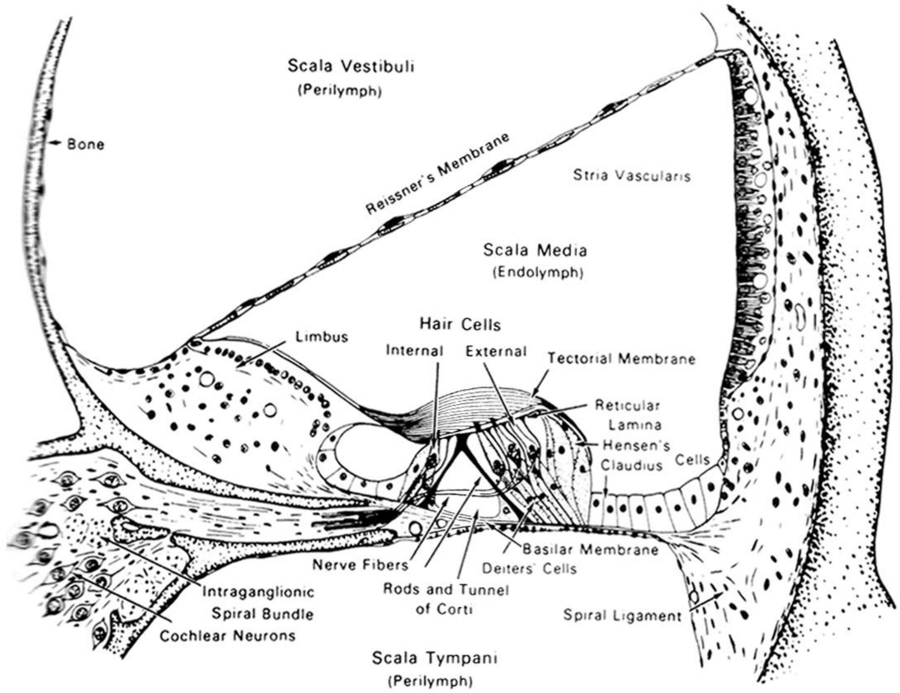

Figure 2 Cross-sectional view of the cochlear duct. Source: Adapted from Davis [4].

The part of the inner ear that is concerned with hearing is the cochlea which contains the organ of Corti that sits on the BM. The vestibular (balance) system consists of three semicircular canals and two further structures known as the utricle and saccule. The organ of Corti has hair cells that are the sensory receptors for hearing and these cells are in contact with the nerve cells of the VIII cranial nerve, which connects the peripheral ear to the central nervous system. The auditory branch of the eight nerve is generally called the auditory or cochlear nerve whilst the vestibular branches are often referred to as the vestibular nerve. The conductive system, which consists of the outer and middle ear structures, transduces sound and transmits the stimulus to the inner ear. Hence, the sensorineural system - cochlea and eight cranial nerve - induces the physiological response to the stimulus, activation of the nerve cells and the encoding of the sensory response into a neural signal. The aspects of the central nervous system that deal with this neurally encoded signal are generally called the central auditory nervous system [3].

Under conditions of normal hearing, sound waves which travel through the air reach the tympanic membrane via the ear canal, causing vibrations that move the three ossicles [5]. This produces a coordinated movement of the ossicular chain resulting in a piston-like movement of the stapes. It is the “footplate” of the stapes, which is attached to the oval window whose inward and outward movements induce pressure oscillations in the cochlear fluids, which in turn give rise to a traveling wave displacing fluids along the BM. This membrane has graded mechanical properties: At the base of the cochlea (near the stapes and oval window) it is narrow and stiff whilst at the other end (near the apex) it becomes progressively wider and less stiff. The resulting traveling wave, which propagates from the base to the apex of the cochlea, is characterized by points of maximal response based on the frequency or frequencies of the pressure oscillations within the cochlear fluids. For an oscillation with a single frequency, the magnitude of displacement increases up to a particular point along the membrane and then drops sharply thereafter. Low frequencies produce maxima near the apex whereas high frequencies produce maxima near the base of the cochlea. Motion of the BM is sensed by the sensory hair cells in the cochlea, which are attached to the top of the BM within the organ of Corti. The cells are arranged in four rows along the total length of the cochlea. The cells in the innermost row are called the Inner Hair Cells (IHC), and the cells in the remaining rows are called the Outer Hair Cells (OHC). The IHCs are closest to the modiolus or “center core” of the cochlea. Each hair cell has fine rods of protein, called stereocilia, emerging from one end. When the BM moves at the location of a hair cell, the rods are deflected as if they are hinged at their bases. Such deflections increase the release of a chemical transmitter substance at the base of the IHCs, whereas deflections in the opposite direction inhibits its release. In contrast, deflections of the stereocilia of the OHCs produce electromotile changes in the length of the cells, which in turn increase the sensitivity and sharpen the "tuning" of the BM to frequencies that correspond closely to the position(s) of the stimulated cells. Thus, the OHCs act as a highly selective biological amplifier. The increases in chemical transmitter substance at the bases of the IHCs increase discharge activity in the immediately adjacent auditory neurons, whereas reduction in the substance inhibits activity. Changes in neural activity thus reflect events at the BM. These changes are transmitted to the brain via the auditory nerve, which is effectively a collection of all neurons that innervate the cochlea [5].

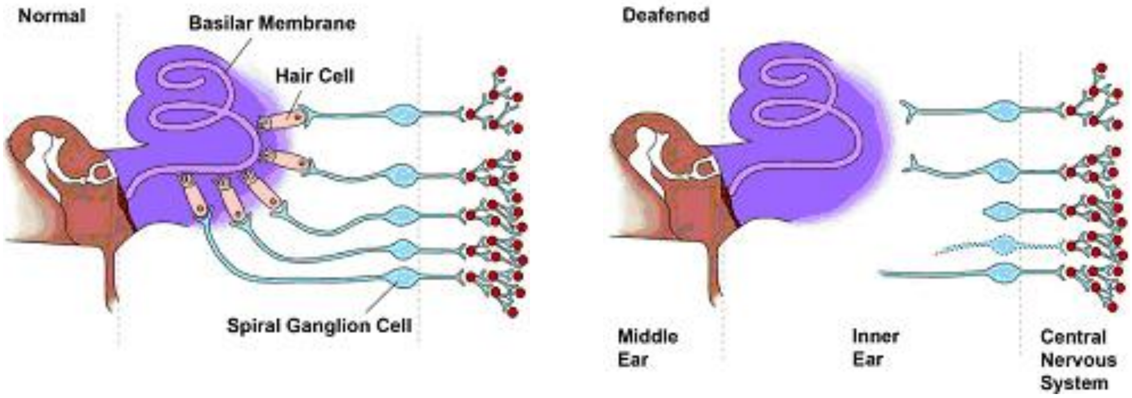

Figure 3 Anatomical structures in normal and deafened ears. Source: Wilson and Dorman [5].

1.2 Hearing Loss

Hearing loss is measured as the average elevation in pure tone thresholds in relation to normal hearing (-10 to 15 dBHL) and is classified by its degree. The most commonly accepted scheme is that proposed by Clark [6] as shown in Table 1.

HL Degree	**HL Range (dB HL)**
Slight	16 to 25
Mild	26 to 40
Moderate	41 to 55
Moderately severe	56 to 70
Severe	71 to 90
Profound	91+

Table 1 Classification of degree of hearing loss proposed by Clark [6].

There are two basic types of hearing loss: conductive and sensorineural hearing loss. Conductive hearing loss results from an abnormality before the cochlea. Any obstruction or malformation which impedes the transfer of sound from the environment through the outer and middle ear, thus attenuating it, will result in a conductive hearing loss. This type of loss is within the mild to moderate range, characteristically ranging from 20 dB HL to a maximum of 60 dB HL. The primary effect of a conductive hearing loss is a loss of intensity. Sensorineural hearing loss occurs as a result of damage to the inner ear (cochlear hearing loss) or to the auditory nerve pathways (retrocochlear hearing loss) [7]. The principal cause of hearing loss is partial or complete destruction of the sensory hair cells that are extremely fragile structures. The hair cells are subject to a wide variety of damage which include but are not limited to genetic defects, infectious diseases, overexposure to loud sounds, drugs and aging. Damage to the OHCs elevates hearing thresholds and degrades frequency resolution whilst damage to the IHCs produces difficulty in speech perception and is characterized by more profound losses including total deafness. The IHCs are largely or completely absent in the deaf or deafened cochlea, thus inhibiting the connection between the peripheral and central auditory systems.

The CI function is to bypass the pathological hair cells by directly stimulating the neurons in the auditory nerve. Figure 3 is the anatomical illustration of the deafened cochlea (a complete absence of hair cells- an anatomical situation which can be usually faced by CI specialists) in comparison to a normal auditory system including the tympanic membrane, the three ossicles, the oval window, the BM, the IHCs, and the adjacent neurons of the auditory nerve. A small number of cells may remain for some patients, usually in the apical (Low Frequency -LF) part of the cochlea. If not stimulated, the peripheral parts of the neurons undergo “degeneration” and cease to function. Fortunately, even after prolonged deafness or etiologies such as meningitis, some usually survive [5,8].

1.3 Cochlear Implants

Over the last years, cochlear implantation has become a common choice for the rehabilitation of bilateral, severe to profound, cochlear hearing loss and CI systems have proved to offer useful auditory information for the perception of environmental sounds, speech and music. The causes of deafness that have been associated with CI recipient individuals range from unknown, genetic or inherited pathology to unpredictable, accidental deafness due to trauma or infection [2].

1.3.1 Development of Cochlear Implants

Interest in the electrical stimulation of the auditory pathway extends to the late 18th century and begins with the experiment of the Italian physicist Alessandro Volta (1800) in which he stimulated his own ears with electrical current applied to metal rods at a voltage approximating 50 V. The outcome was reported as an unpleasant sound-like sensation which he described as “a shock within the head” (uno shock nella testa) followed by “a sound similar to that of boiling thick soup” (un rumore simile a una zuppa densa che ribolle) [1,9,10].

Considerably later, in 1957, Djourno and Eyries [11] carried out an experiment by directly stimulating the auditory nerve. They placed a wire on the auditory nerve of a patient that had surgery for facial nerve paralysis resulted from previous cholesteatoma surgery. The implantation procedure did not require surgical invasion of the ear due to prior opening of the cochlea. When current was applied to the wire, the patient reported some auditory sensation similar to a “roulette wheel of the casino” and a “noise of a cricket”. Eventually, the patient showed improved lipreading skills and developed limited recognition of common words [1,10].

Although there were some doubts about safety and reliability of electrical stimulation of auditory nerve cells at that stage, the experience of Djourno and Eyries [11] suggested that activation of the auditory periphery in humans through an electrical device could efficiently provide useful information to the central auditory pathway [10,11]. As a result, initiatives at stimulating the auditory nerve for clinical benefit began in the United States. In 1961, House and Doyle [12] reported data from two adults with profound deafness whose auditory nerve was stimulated electrically by an electrode placed on and then through the round window and into the ST of the inner ear. Both individuals reported auditory sensations. They noted that loudness changed with level of stimulation and the pitch of the stimulus changed with variation in the rate of stimulation. In 1964, Simmons placed an electrode through the promontory into the vestibule and directly onto the modiolus of the cochlea. These individuals could detect changes in duration and had the sensations of tonality. In 1972, the first commercially marketed system, became available [12,13,14]. This system, the 3M/House device consisted of a single electrode implant and a Speech Processor (SP) (as shown in Figure 4). From 1972 into the mid-1980s, over 1000 people were implanted with this device. In 1980, age criteria for use of this device was lowered from 18 to 2 years. By the mid-1980s, several hundred children had been implanted with the House 3M single-channel device [1,10].

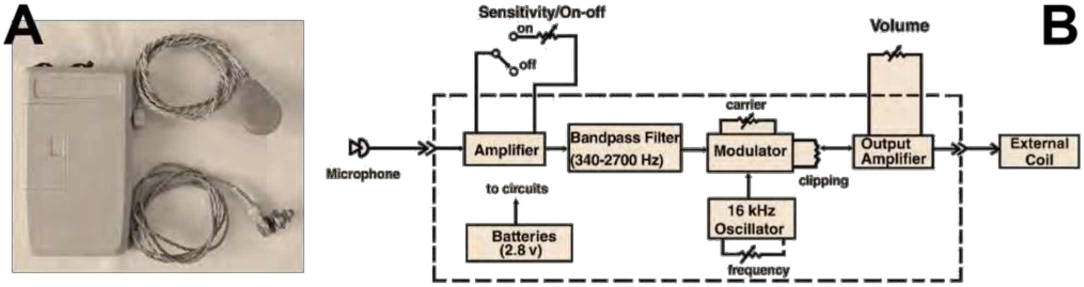

Figure 4 The House 3M single-channel cochlear implant. A: the body worn speech processor, B: signal processing diagram of the House 3M implant system. Source: Vaerenberg [15].

Multiple-channel devices were introduced in 1984, and the development of single- and multichannel systems moved hand in hand in the 1990s. Single channel implants deliver auditory information through a single electrode whereas multi-channel implants convey different parts of the signal via several distinct channels that stimulate different regions of the cochlea. The terms single and multi-channel are used to describe the number of active electrodes through which different information is sent; whereas single or multi-electrode refers to the number of electrodes in the implant [1]. The single-channel implants experienced successful use in terms of providing basic access to acoustic information where some patients were able to make successful use of the simple stimuli [16]. The development of multichannel systems required advanced technologies of digital signal processing (DSP) chip design, miniaturization, battery consumption and other engineered capabilities. They outweighed single-channel devices based on enhanced spectral perception and enhanced speech recognition capabilities, as shown in large adult clinical trials [2,10,17,18]. Most patients who have had single channel implants replaced with multi-channel implants have shown varying degrees of improvement in speech and environmental sound recognition [19,20]. CI centers nowadays choose to use multi-channel, intra-cochlear implants as they give better performance [1].

1.3.2 Basics of cochlear implants

Cochlear implantation is mainly based on the following principles: foreign, biocompatible materials can be placed within the human body without being rejected [2,21] and auditory nerve fibres respond to electrical stimulation [2,22].

To date, four major CI manufacturers exist in the market: Cochlear Ltd. (Australia), Advanced Bionics (Switzerland), MED-EL (Austria), Oticon Medical/Neurelec (France). Although the market offers a wide variety in technical and cosmetic features, CI systems basically consist of two parts: an internal part that is surgically implanted and an external part that is called speech processor. Figure 5 shows the essential components of a cochlear implant system.

Figure 5 The essential components of a conventional cochlear implant. (A) External parts: (1) Behind the Ear Speech Processor (2) transmitter coil. (B & C) Internal parts: (3) receiver-stimulator (4) electrode array. Source: Vaerenberg [15].

CI systems share a common working principle. The microphone which is located on the SP picks up environmental sounds and converts these analog sounds into digital information. This information is sent via radiofrequency transmission from an external coil to an internal receiver-stimulator implanted under the skin. The receiver-stimulator transforms the signals into patterns of electrical stimulation and delivers to the electrode array. The auditory nerve fibres in the cochlea pick up the signals and convey them through the natural auditory pathway to the brain, giving the sensation of the sound.

CI systems use a transcutaneous link. The link is bidirectional to allow transmission of data from the implanted components out to the external components (SP and coil) as well as transmission of data from the SP to the implanted receiver-stimulator. The data sent from the implanted parts to the external parts may include:

- information about the status of the receiver/stimulator;
- impedance of each single electrode;
- voltages at unstimulated electrodes;
- neural evoked potentials.

The electrodes and electrode carrier (together called electrode array) for CI systems are placed in the ST, which offers an accessible site that is close to the spiral ganglion. Figure 6 shows a cross-section of the implanted cochlea. It shows the three chambers and the partial insertion of an electrode array into the ST [5,23]. The electrodes should be biocompatible, mechanically stable and should facilitate atraumatic insertion. In general, flexible arrays facilitate insertion. The array is inserted through a drilled opening made by the surgeon in the bony shell of the cochlea overlying the ST and close to the base of the cochlea (called a “cochleostomy”). Alternatively, the array may be inserted through the round window membrane, which also is close to the basal end of the cochlea and ST. The cochleostomy provides a “straighter shot” into the ST than the round window approach [5]. The number of intracochlear electrode contacts available ranges from 12 to 22. In addition to the intracochlear electrodes, 1 or 2 electrodes are positioned outside of the cochlea. These electrodes serve as reference or ground electrodes [24].

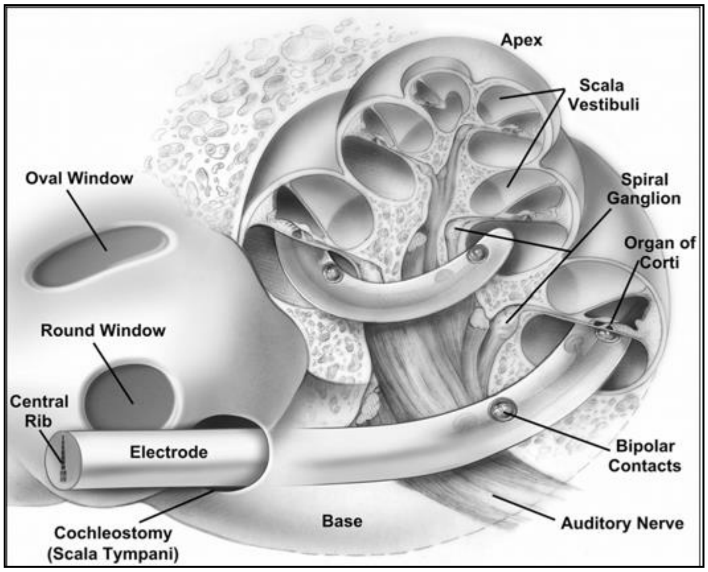

Figure 6 Cross section of implanted cochlea. Source: Wilson and Dorman [5].

The depth of insertion can be limited by the decreasing lumen of the ST from base to apex, the curvature of the cochlear spiral, and an uneven and unsmooth lumen, particularly in the apical region. Insertions are typically 18 to 26 mm and no array has been inserted farther than about 30 mm (The length of the human cochlea is typically about 35 mm). In some cases, only shallow insertions are possible, such as when bony obstructions in the lumen impede further insertion. Different electrodes in the implanted array may stimulate different populations of neurons. Neurons at different positions along the length of the cochlea respond to different frequencies of acoustic stimulation in normal hearing subjects. CI systems attempt to mimic or reproduce this “tonotopic” encoding by stimulating basally situated electrodes (first turn of the cochlea and lower part of Figure 6) to indicate the presence of High Frequency (HF) sounds and by stimulating electrodes at more apical positions (deeper into the ST and ascending along the first and second turns in Figure 6) to indicate the presence of sounds with lower frequencies [5,23].

Intracochlear electrodes can be stimulated in a monopolar or bipolar configuration. In the monopolar configuration, each intracochlear electrode is stimulated with reference to a remote electrode, usually in the temporalis muscle or outside of the case of the implanted receiver-stimulator. In the bipolar configuration, one intracochlear electrode is stimulated with reference to another (adjacent) intracochlear electrode. Different pairs of electrodes are used to stimulate different sites along the electrode array [24]. In Figure 6, closely spaced pairs of bipolar electrodes are shown, but all present-day implant systems use the monopolar coupling configuration since it supports performance that is at least as good as bipolar coupling and requires less current and battery power to produce auditory percepts. The spatial specificity of stimulation with an ST electrode most likely depends on multiple factors, including the orientation and geometric arrangement of the electrodes, the proximity of the electrodes to the target neural structures, and the condition of the implanted cochlea in terms of nerve survival and ossification. An important goal of electrode design is to maximize the number of largely nonoverlapping populations of neurons that can be addressed within the electrode array. However, evidence suggests that no more than 4 to 8 independent sites are available with current designs, even for arrays with as many as 22 electrodes [25,26,27]. Most likely, the number of independent sites is limited by substantial overlaps in the electric fields from adjacent (and more distant) electrodes. The overlaps are unavoidable for electrode placements in the ST because the electrodes are "sitting" in the highly conductive fluid of the perilymph and, additionally, are relatively far away from the target neural tissue in the spiral ganglion. A closer apposition of the electrodes to the inner wall of the ST would move them a bit closer to the target cells, and such placements have been shown in some cases to produce an improvement in the spatial specificity of stimulation [28]. However, a large gain in the number of independent sites may well require a fundamentally new type of electrode or a fundamentally different placement of electrodes. Figure 6 shows a complete presence of hair cells (in the labeled organ of Corti) and a pristine survival of cochlear neurons. However, the number of hair cells is zero or close to in cases of total deafness. In addition, survival of neural processes peripheral to the ganglion cells (the "dendrites") is rare in the deafened cochlea. Survival of the ganglion cells and central processes (the axons) ranges from scarce to substantial. The pattern of survival is usually not uniform, with reduced or sharply reduced cell counts in certain regions of the cochlea and the neural substrate or target for a cochlear implant can be quite different between patients [5].

Despite above mentioned limitations, CI recipients over the last years have benefited from advances in battery, integrated circuit and DSP chip technologies, in that the developments have allowed smaller and more capable SPs and implanted receiver-stimulators. The SPs are available in different models and are usually named according to their wearing styles: body worn processors (worn on the belt or in a pocket usually to offer a more robust solution for young children) and behind the ear processors (usually preferred by adults). The trend in cochlear implantation has been towards achieving better patient performance as well as attempting to improve cosmetic features.

In the 1990s, clinical and basic science investigations produced changes in implant technology and in clinical approaches to cochlear implantation. Electrode and SP designs have evolved to produce encoding strategies that are associated with higher performance levels. Simultaneously, along with device development and observations of safety and reliability there has been emphasis on earlier implantation in children. There is now recognition of the required services for children to optimize implant performance and the structure of the interaction needed among the implanted child, family members, school staff, and implant team professionals. There is now substantially greater potential for open-set speech understanding in children and adults. Technologic advances of the past decade have refined speech encoding strategies and have expanded implant candidacy [10,29,30].

1.3.3 Cochlear Implant Fitting and Processing Strategies

A normally hearing ear can discriminate speech by detecting changes in the frequency and intensity (or pitch and loudness referred as their subjective percepts) with time. Thus, the CI aims to represent the sound input so that the recipient can detect pitch and loudness variations and have the ability to understand speech [1]. However, the amount of information that can be presented and perceived with a cochlear implant is much less than that for a normal hearing (NH) person who is listening to an unprocessed acoustic signal, e.g. CI’s capacity is restricted by the limited number of electrodes (max. 22 intracochlear electrodes currently) and by the stimulus rate. Additionally, the dynamic range of stimulus amplitudes from auditory threshold to loud percepts is in the order of 10 to 20 dB for electrical pulses in comparison to the order of 100 dB for acoustical stimuli [9].

The SP is usually activated in CI recipients at 3 to 4 weeks after surgery. Initially CIs should be programmed to activate the system and to make the recipient hear sounds. Subsequently, fitting sessions should be done regularly (but less over time), the goal being to achieve the most appropriate configuration for each recipient.

CI systems offer a choice of different speech coding strategies and variables that can be adjusted during fitting sessions. The parameters that can usually be changed are as follows: thresholds, comfortable levels, active channels or electrodes, stimulation rate, stimulation mode, frequency boundaries, stimulation cycle, sampling rate, pulse width, dynamic ranges, smoothing cut-off filter, automatic channel selection, noise suppression, input dynamic range, compression, rectification mode, pulse rate per channel and more [2].

Initially sound information was transmitted via the use of an analogue waveform with continuous and simultaneous stimulation of the electrodes. Analogue waveform was first used in the Compressed Analogue (CA) strategy of the Ineraid cochlear implant system and later in the Simultaneous Analogue Stimulation (SAS) strategy of the Clarion (Advanced Bionics- AB). The SAS strategy consisted of 16 electrodes used for bipolar stimulation (eight pairs of one active and one reference electrode). The aim of bipolar stimulation was to reduce current spread and to minimize channel interaction.

Current CI models make use of pulsatile waveforms that consist of series of pulses extracted from the incoming signal and delivered to different channels based on their frequency. Each pulse is presented to each channel sequentially in order to minimize channel interactions and to maximize spectral information. Pulsatile strategies do not convey the whole waveform but rapidly updated samples of the sound signal. To represent adequately changes in the signal with time, rapid updating of the incoming signal is required. Most pulsatile strategies use sequential stimulation and fall into two broad categories: Continuous Interleaved Sampling (CIS) strategy and the n of m or spectral maxima type of strategy [1].

CIS Strategy

One of the most effective approaches for representing speech with current CI technology is the CIS strategy [5]. This strategy was the first, fast rate strategy developed by Wilson et al. [31]. CIS gets its name from the continuous sampling of the (compressed) envelope signals by rapidly presented pulses that are interleaved across electrodes. It filters sound information into bands of frequencies with a bank of bandpass filters. Envelope variations in the different bands are represented at corresponding electrodes in the cochlea with modulated trains of biphasic electrical pulses. The envelope signals extracted from the bandpass filters are compressed with a nonlinear mapping function prior to the modulation in order to map the wide dynamic range of sound in the environment (up to about 100 dB) into the narrow dynamic range of electrically evoked hearing (about 10 dB). The output of each bandpass channel is directed to a single electrode, with low-to-high frequency channels assigned respectively to apical-to-basal electrodes, to mimic the frequency mapping in the normal cochlea. The pulse trains for the different channels and corresponding electrodes are interleaved in time so that the pulses across channels and electrodes are non-simultaneous. This eliminates a principal component of electrode interaction, which otherwise would be produced by direct vector summation of the electric fields from different (simultaneously stimulated) electrodes. The corner or “cutoff” frequency of the low-pass filter in each envelope detector is usually set at 200 Hz or higher so that the Fundamental Frequencies (F0s) of speech sounds are represented in the modulation waveforms. All the currently available cochlear implants can be programmed with the CIS strategy but the implementation of the strategy may vary in different implants for parameters such as filtering, envelope extraction, the number of channels, the pulse rate and the update rate. CIS implementations use up to 22 channels and corresponding stimulus sites [5].

The *n*-of-*m*, SPEAK and Advanced Combination Encoder (ACE) Strategies

The spectral maxima strategies use a channel-selection scheme in which the envelope signals for the different channels are scanned prior to each frame of stimulation across the intracochlear electrodes in order to identify the signals with the *n*-highest amplitudes from among *m* processing channels. Stimulus pulses are delivered only to the electrodes that correspond to the channels with the highest amplitudes. Examples of spectral maxima type strategies are the Nucleus SPEAK and ACE strategies and the Medel n of m strategy. The parameter *n* is fixed in the *n*-of-*m* and ACE strategies and it can vary from frame to frame in the SPEAK strategy, depending on the level and spectral composition of the input signal from the microphone. Stimulus rates typically approximate or exceed 1,000 pulses/sec/selected electrode in the *n*-of-*m* and ACE strategies and approximate 250 pulses/sec/selected electrode in the SPEAK strategy. The designs of the *n*-of-*m* and ACE strategies are essentially identical and are similar to CIS except for the channel-selection feature. The SPEAK strategy uses much lower rates of stimulation and an adaptive *n*. The channel selection or "spectral peak picking" scheme used in the *n*-of-*m*, ACE, and SPEAK strategies is designed in part to reduce the density of stimulation whilst representing the most important aspects of the acoustic environment. The deletion of low-amplitude channels for each frame of stimulation can reduce the overall level of masking or interference across electrodes and stimulus regions within the cochlea. To the extent that the omitted channels do not contain significant information, such "unmasking" may improve the perception of the input signal by the patient. Furthermore, for positive speech-to-noise ratios, selection of the channels with the greatest amplitudes in each frame may emphasize the primary speech signal with respect to noise [5].

SPEAK is a slow-rate (180-300 pps) spectral maxima strategy. The SP extracts up to nine maxima from the incoming signal and presents these maxima to different electrodes of 20 active electrodes inside the cochlea according to their frequency. The average number of maxima is six but may vary according to the incoming signal. Frequency bands are typically allocated within the range 187-7937 Hz but alternative frequency allocations can be set.

The ACE is a fast rate, flexible spectral maxima strategy. Up to 22 channels can be used and up to 20 maxima, although 8 to 12 maxima are most widely used. The frequency band allocation is variable but usually 187-7937 Hz. Stimulation rates up to 2400 pps per channel are available with a maximum overall rate of 14400 pps. Medel n of m strategy extracts up to 11 spectral peaks from the signal to be delivered to a maximum of 12 electrodes. Fast rates up to a maximum of overall rate of 18000 pps can be used. The spectral peaks are extracted up to 7.5 kHz [1].

HiResolution (HiRes) Strategy and HiRes with Fidelity 120

In 1912, David Hilbert showed that signals can be decomposed into temporal envelope (the relatively slow variations in amplitude over time) and Temporal Fine Structure (TFS, a frequency modulated carrier, rapid oscillations with rate close to the center frequency of the signal). Figure 7 illustrates an example of such a decomposition.

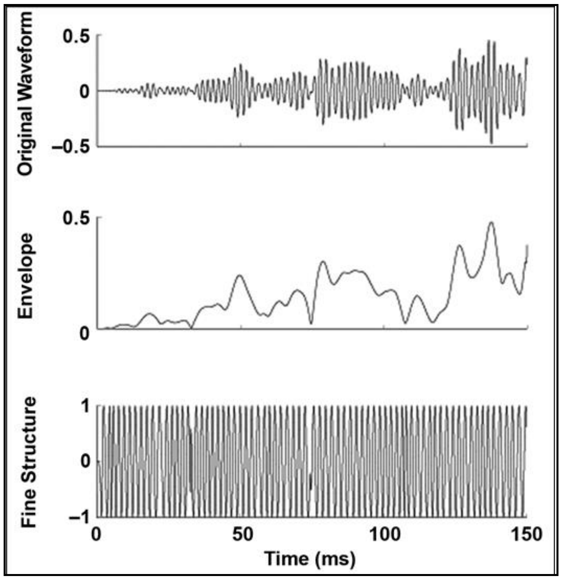

Figure 7 Decomposition of signal using Hilbert transformation. Source: Wilson and Dorman [5].

In 2002, Smith et al. [32] investigated the importance of envelope and TFS information for auditory perception in NH subjects. They synthesized novel stimuli called “auditory chimeras” which had the envelope of one sound and the TFS of another sound. Hence, the chimeras resulted in conflicting cues—the envelope variations in a given number of bands for one sound versus the TFS variations in the same bands for another sound. Pairings consisted of sentences versus different sentences, sentences versus noise, melodies versus different melodies, and sentences with an Interaural Time Delay (ITD) which corresponded to a sound image to the left side of the subject versus the same or different signal but with an ITD to the right.

Performance depended on the type of sounds in each pairing and on the number of processing channels. Speech was identified by its envelope information for a minimum of eight channels whereas the TFS information was more important for one or two channels. For intermediate numbers of channels, both envelope and TFS information contributed to sentence recognition. On the other hand, melodies – which required greater spectral resolution - were recognized almost exclusively by their TFS information for a minimum of 32 channels and envelope cues became dominant at 48 and 64 channels. Lateralization of sentences was difficult for subjects with a small number of channels, but performance improved with increasing numbers up to the test limit of 32. Lateralization was cued by the TFS information in all cases. These findings indicated the importance of TFS information for speech perception using less than about 8 processing channels and for music perception using less than about 40 channels. In addition, the findings indicated that ITD cues may be represented by TFS but not envelope information for any number of channels up to (at least) 32. Present-day electrode arrays and processors do not support more than 4 to 8 channels of perceptually separable information. In this 4 to 8 range, both envelope and TFS information contribute to speech perception, whereas music information is conveyed almost uniquely by TFS cues [5,32].

In recent years, an important aim of CI research has been to develop strategies for enhancing the representation of TFS information so that implanted patients could benefit better pitch perception and sensitivity to ITD. Improved pitch perception should help music appreciation and convey prosody cues in speech. Thus, it may improve speech perception in particular among speakers of tonal languages, where pitch is crucial to distinguish different words. Better ITD sensitivity may help bilateral CI recipients in taking advantage of binaural cues that NH listeners use to distinguish speech among competing sound sources [32].

The HiRes strategy has been the first approach among these strategies. It uses relatively high rates of stimulation and high envelope cutoff frequencies to improve TFS information. Although only envelope information is presented with the processing strategies, frequencies included in the envelopes generally range up to 200 to 400 Hz or even higher in the HiRes strategy. Thus, substantial TFS information is represented and may be at least partially perceived within LF range [5].

An alternative approach has been to represent the TFS information within bands using multiple sites of stimulation for each band instead of a single site for each band. This approach is called the HiRes with the Fidelity 120 (HiRes 120) and is a variation of HiRes strategy. It makes use of "virtual channels" in order to increase the number of discriminable sites beyond the number of physical electrodes. The term "current steering" is also used to refer "virtual channels". This concept was first introduced by Wilson et al. in the early 1990s [33,34,35]. With virtual channels or current steering, adjacent electrodes can be stimulated simultaneously. In this way, the perceived pitch can be shifted in any direction with respect to the percepts elicited with stimulation of either of the electrodes alone. Studies with CI recipients revealed that pitch could be manipulated through various choices of simultaneous and single-electrode conditions [36]. For example, if the most apical electrode was stimulated alone, subjects reported a low pitch. If the next electrode in the array was stimulated alone, a higher pitch was reported. For the majority of subjects, an intermediate pitch was perceived by stimulating the two electrodes together with identical in-phase pulses.

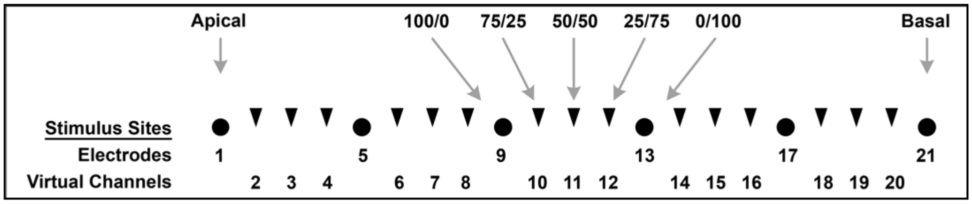

Figure 8 Diagram of stimulus sites used in virtual channel interleaved sampling processors and other similar processors. Source: Wilson et al. [36].

The concept of virtual channels can be extended to include a quite high number of sites and corresponding pitches by using varying ratios of the currents delivered between electrodes that are simultaneously stimulated. As shown in Figure 8, stimulus site 1 is produced by stimulation of electrode 1 alone, stimulus site 2 by simultaneous stimulation of electrodes 1 and 2 with a pulse amplitude of 75 percent for electrode 1 and 25 percent for electrode 2, and so on. The total number of sites and corresponding pitches that might be produced for a good subject in the illustrated case is 21, with 6 intracochlear electrodes [36].

In the HiRes 120 strategy, 8 sites are allocated to each of 15 bandpass ranges in order to form 120 sites. The different sites for each channel are created with eight different ratios of currents delivered to the two adjacent electrodes assigned to that bandpass range. One of each of the eight ratios is used at a time, and the stimuli for the different channels are presented in a non-overlapping sequence, as in the CIS strategy. However, unlike the CIS strategy, two electrodes are stimulated simultaneously (with the selected amplitude ratio) at each update, rather than stimulation of a single electrode at each update.

The HiRes 120 strategy suggests that a higher number of available pitches may result in greater spectral resolution hence giving patients access to relatively small frequency differences thus enhancing speech perception particularly under adverse conditions, and music perception, which is generally quite poor with the CIS and other related strategies, as might be expected from the findings of Smith et al. [32]. However, a high number of available pitches or discriminable sites does not guarantee a high number of effective channels in CI recipients and furthermore “virtual pitches” may well be inherent in standard CIS and related strategies using sequential stimulation, in that intermediate pitches can also be created with nonsimultaneous stimulation of adjacent (or more distant) electrodes so long as the pulses are relatively close in time [5,37,38,39].

Fine Structure Processing (FSP) Strategy

More recently, new processing options have been introduced. In one approach, the FSP strategy, the timing of positive zero crossings in the output of the bandpass filter with the lowest center frequency, or in the outputs of up to four bandpass filters with the lowest center frequencies, is marked with the presentation of a short group of pulses for the corresponding channel(s) and site(s) of stimulation instead of the continuous presentation of pulses for CIS channels. The overall amplitude of the pulse bursts for these processing channels is determined by the magnitude of energy in the band for each channel, like in CIS. The remaining higher frequency channels make use of CIS processing. The pulses for the lower frequency channels are also interleaved across electrodes, including the electrodes presenting the CIS stimuli [5,40,41].

There is some evidence that the FSP and related approaches provide an advantage compared to CIS and other envelope-based strategies to the extent that single pulses or short groups of pulses represent temporal events in the LF channel(s) better than the continuous (and time varying) modulations for the same channels in envelope-based strategies [5,40,41,42,43].

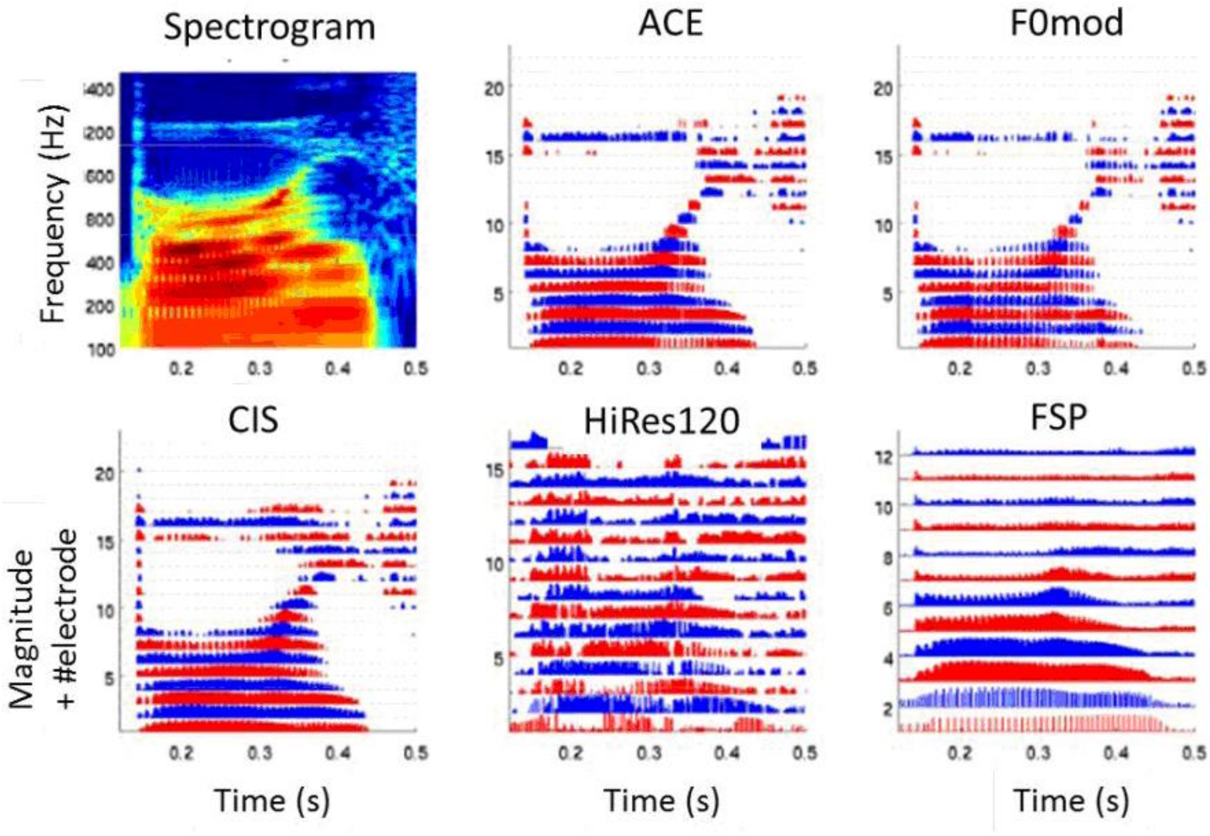

Figure 9 Spectrogram and electrodograms of the word "boy". The signal was presented at an average RMS level of 60 dB SPL. For the electrodograms, the vertical axis indicates the channel, and the height of each vertical line represents the magnitude of the pulse. The magnitude is expressed in different units for different strategies. The red and blue colors visually distinguish adjacent channels. Source: Wouters et al. [44].

1.3.4 Performance with cochlear implants

Postoperative outcomes for the perception of meaningful speech information are very satisfactory in CI recipient adults and children. Current CI technology offers good opportunities for formal and informal language acquisition in deaf children [45,46,47,48] as well as very good speech understanding in quiet environments in postlingually deafened adults [49]. Indeed, a principal conclusion of the 1995 National Institutes of Health (NIH) Consensus Conference on Cochlear Implants in Adults and Children [29] was that “A majority of those individuals with the latest speech processors for their implants will score above 80 percent correct on high context sentences, even without visual cues.” [5].

However, CI recipients usually complain about speech understanding under noisy listening conditions which are very common in everyday communication environment. Findings indicate that though many of them show ceiling effects for understanding speech in quiet, their performance decreases considerably when testing takes place in the presence of background noise [50]. Another common complaint of CI recipient people is the music perception and appreciation. Although rhythm perception usually is good, they generally have difficulty hearing pitch or melody [51,52]. On the other hand, adult CI users commonly complain that they do not enjoy listening to music [53,54,55,56] whilst parents of implanted children report that their children enjoy listening to music and singing with the CI [57]. It may be the case that children who acquire hearing with a cochlear implant may encode sounds differently than adults. Consequently, children may be able to use acoustic cues that are not perceptible or available to implanted adults to hear and appreciate music.

The trend in research on cochlear implantation over the past years has been towards improvements in speech perception in the presence of noise by implementing developments such as directional microphones and noise reduction algorithms as well as by introducing advanced settings in technical parameters such as Automatic Gain Control and Input Dynamic Range [58,59,60,61]. An important technological improvement in most of the SPs that are available on the market regards the use of multiple microphones in order to increase the selectivity of the directional pattern. With two microphones, sounds originating between and in front of the microphones produce microphone outputs that are in phase with each other, whereas sounds originating at other locations produce microphone outputs that are not. Summation of such microphone outputs produces larger signals for the in-phase conditions, emphasizing sounds in front of the microphones and suppressing sounds from other locations. It has been shown that the addition of a second microphone to a CI system improves speech reception performance under difficult listening conditions with reduced Signal-to-Noise Ratio SNR [5].

Binaural hearing provides important benefits in comparison to monaural hearing especially under challenging listening conditions [62,63,64]. One main advantage of binaural hearing is defined as the improvement in speech perception in the presence of noise. Three specific binaural effects are believed to benefit NH listeners and those with a hearing loss when listening to speech in noise: Head Shadow (HS), binaural SQuelch (SQ), and binaural SUmmation (SU). Although initially cochlear implantation used to be monolateral, over the past years the trend has been towards implanting patients bilaterally to make use of the binaural hearing advantages provided by the capacity of the central auditory system to process stimuli received from each ear and to reproduce it with a higher SNR by comparing interaural time and intensity differences or by the physical placement of the head which acts as an acoustic barrier and leads to an increase in SNR in the ear far from the noise when signal and noise are spatially separate. Research in normal hearing subjects indicated a 3 dB improvement in SQ for the binaural Speech Reception Threshold (SRT) and an average increase of 3 dB SNR for HS which is more dominant for attenuation of high frequencies and can cause even 8 to 10 dB of improvement [65]. Several studies indicate that these effects may improve speech recognition in bilateral CI recipients [66]. Schleich et al. [67] evaluated SRTs in adults with bilateral CIs under three listening conditions and found an average improvement of 6.8 dB from HS, of 0.9 dB improvement for SQ and a 2.1 dB improvement from SU. Similarly, Livotsky et al. [68] and Buss et al. [69] showed that HS was resulting in the greatest effect. An important factor that affected differences in performance was the timing of implantation: bilateral CI use was found to be more effective when implantation was done simultaneously or sequentially with the shortest possible time interval [70].

On the other hand, the extension of indications for candidature to severe as opposed to uniquely profound hearing losses has led to an increasing number of CI recipients wearing an HA to make use of LF residual hearing on the contralateral side. Hence, contralateral HA use offers an alternative to bilateral cochlear implantation in that unilateral CI recipients benefit from the LF cues provided through acoustic signals from contralateral HA in addition to electrical signals from the CI. This has been named as "bimodal benefit" [71]. Bimodal benefit in CI recipients has recently received much attention and previous studies have shown a significant positive effect on speech recognition in noise and on functional performance in daily life as well as on the improvement of localization, pitch and music perception skills [65,72,73,74].

Previous studies showed discrepancies for correlations between bimodal benefit and audiological outcomes such as unaided pure tone average and aided free field audiometry as well as the duration of CI experience and the duration of HA experience prior to cochlear implantation. There were also studies showing the positive effect of degree of LF residual hearing or longer duration of HA experience prior to implantation on bimodal benefit [65]. Some studies even found an adverse effect of better hearing thresholds at mid-to-high frequencies [71,75]. Bimodal findings were promising for unilateral CI recipients with profound hearing loss and with no LF residual hearing especially in countries where bilateral implantation is still not reimbursed. On the other hand, in countries without any financial restrictions, the decision depends more on evaluation of the amount of benefit that a second CI or a contralateral HA can provide for individual subjects by taking into consideration better time-based cues that HA can convey to an ear with LF residual hearing in comparison to CI [76,77].

A wide range of outcomes has been found for the various multichannel implants currently on the market. Different patients using identical implant devices may show quite different speech perception scores. This indicated the importance of patient variables in the design and performance of implant systems. Such variables may include differences among patients in the survival of neural elements in the implanted cochlea, proximity of the electrodes to the target neurons, depth of insertion for the electrode array, integrity of the central auditory pathways, and pre-existing cognitive and language skills [9,10].

Objectives of the thesis

The aim of this thesis were as follows:

- To demonstrate outcomes for LF pitch perception in CI recipient children in comparison to their NH peers, for which very limited results exist. This was due to the lack of standard pitch perception tests focused on the auditory capacities within the LF domain, as well as tests applicable to children. For this purpose, the clinical applicability of two new LF pitch perception tests (Harmonic Intonation-HI / Disharmonic Intonation -DI) and the effect of chronological age on HI/DI performance were investigated in a group of children with cochlear implants and normal hearing. HI/DI outcomes are believed to depend on the availability of TFS cues which are crucial for speech perception, especially in the presence of fluctuating background noise which is a common listening condition in everyday life and a big challenge for CI users.
- To introduce the Italian adaptation of the STARR test, a speech assessment tool based on adaptive randomized roving levels across sentences which mimics challenging real world listening conditions. The STARR test could be a supplement to the Italian speech assessment battery for use with hearing-impaired populations with auditory prostheses and could contribute to cross-language studies.
- To study in adult CI recipients:
    - The LF pitch perception outcomes linked to TFS processing capacities, in a large group
    - The Italian STARR outcomes in relation to other speech perception tests
    - The interaction between TFS processing and speech perception outcomes in particular those of the Italian STARR.
    - The effects of variables such as age, the duration of profound deafness before implantation, the duration of CI experience and hearing thresholds.
    - The correlations between TFS processing and speech perception outcomes in relation to the amount of bimodal benefit.

CHAPTER 2: LOW-FREQUENCY PITCH PERCEPTION IN CHILDREN WITH COCHLEAR IMPLANTS IN COMPARISON TO NORMAL HEARING PEERS[i]

Abstract

The aim of the present study was to investigate the application of two new pitch perception tests in children with cochlear implants (CI) and to compare CI outcomes to normal hearing (NH) children, as well as investigating the effect of chronological age on performance. The tests were believed to be linked to the availability of Temporal Fine Structure (TFS) cues. 20 profoundly deaf children with CI (5–17 years) and 31 NH peers participated in the study. Harmonic Intonation (HI) and Disharmonic Intonation (DI) tests were used to measure low-frequency pitch perception. HI/DI outcomes were found poorer in children with CI. CI and NH groups showed a statistically significant difference ($p<0.001$). HI scores were better than those of DI test ($p<0.001$). Chronological age had a significant effect on DI performance in NH group ($p<0.05$); children under the age of 8.5 years showed larger inter-subject-variability; however, the majority of NH children showed outcomes that were considered normal at adult-level. For the DI test, bimodal listeners had better performance than when listening with CI alone. HI/DI tests were applicable as clinical tools in the pediatric population. The majority of CI users showed abnormal outcomes on both tests confirming poor TFS processing in the hearing-impaired population. Findings indicated that the DI test provided more differential low frequency pitch perception outcomes in that it reflected phase locking and TFS processing capacities of the ear, whereas HI test provided information of its place coding capacity as well.

2.1 Introduction

Cochlear implantation has become a standard procedure for the (re)habilitation of the profoundly deaf. In fact, several prelingually deafened children with cochlear implants (CI) are able to acquire age-appropriate competence in formal and informal aspects of language [45,46] whilst postlingually deafened adult CI recipients show considerable benefit for speech understanding under quiet listening conditions [49]. However, speech understanding in noisy situations and music perception/appreciation still continue to be an important issue for CI users with more frustration in the adult population [56,78]. Even though studies reveal some positive effects of bilateral implantation and bimodal listening through contralateral acoustic amplification aids, the results are still not comparable with normal hearing (NH) population and the amount of benefit reflects inter-subject variability [73,79].

A key element for music and speech perception has proved to be pitch [80,81]. Pitch is known to be coded by two peripheral auditory mechanisms: place coding and phase locking. Place coding is based on tonotopic excitation in which pitch cues are conveyed through spatial alteration of nerve fibers and it is believed that high frequency (HF) coding is dominated by this physiological mechanism. On the other hand, phase locking is more dominant in conveying pitch for low

---

[i] **Material covered in Chapter 2 has previously been published in Dincer D'Alessandro H, Filipo R, Ballantyne D, Attanasio G, Bosco E, Nicastri M, Mancini P (2015) Low-frequency pitch perception in children with cochlear implants in comparison to normal hearing peers. Eur Arch of Otorhinolaryngol 272:3115-3122.**

frequency (LF) signals and this time-based mechanism locks onto the Temporal Fine Structure (TFS) of the signal and conveys intonation by keeping the auditory nerve fibers’ firing rate at the same frequency as the signal [82].

An important aspect that contributes to pitch perception skills has been described as TFS processing which is defined as rapid oscillations with average rate close to the center frequency of the signal [79,81] and is assumed to be reflected in the phase locking patterns within the auditory nerve [83]. TFS information is proved to have an effect on speech perception in fluctuating background sounds, in speaker identification and in the understanding of tonal languages where differences in Fundamental Frequency (F0) over time may lead to semantic changes [81]. Listeners with hearing impairment have a reduced capacity to use TFS information and the evidence comes from studies in which performance using unprocessed speech was compared to those with only TFS or envelope cues (the relatively slow modulations in amplitude over time). It was found that hearing-impaired subjects performed considerably poorer in speech tests only with TFS cues in comparison to NH ones while both groups had good scores in speech tests with unprocessed or envelope cues only [81,84]. Both TFS and envelope information of sounds are represented in the timing of neural discharges. However, recent CI technology is based on conveying mainly envelope information in different frequency bands whilst TFS information is mostly discarded. Envelope information is sufficient for speech understanding in quiet but fine structure is needed for more difficult listening tasks as in the presence of a fluctuating noise and envelope cues alone are insufficient to permit the perceptual segregation of mixtures of sounds [79,81].

Pitch perception skills have been recently studied in adults via two new tests called Harmonic Intonation (HI) and Disharmonic Intonation (DI) that focus on the spectral discrimination of the auditory system in the LF domain. HI/DI outcomes are expected to be indicative of the capability of the inner ear to use its phase locking mechanism and therefore, are believed to depend on the availability of TFS cues [82,85]. HI/DI outcomes in adults showed abnormal pitch perception in CI recipients in comparison to NH subjects. Further studies have been done in adults with LF or HF hearing losses and in adult CI users with electric stimulation only or electro-acoustic stimulation (EAS). Results indicated worse performance in the group with LF hearing loss for both tests, with even poorer scores on DI test [76,82] and better DI outcomes in CI users with an EAS processor rather than in those with electrical stimulation only [86].

However, so far outcomes in the pediatric population are missing and this, together with the following motivations, lead us to conduct the present study:

- Previous pitch perception studies have been done mainly in adult CI recipients and there are limited findings in CI children as compared to NH peers owing to difficulties in testing them under difficult conditions and due to the lack of standard pitch perception tests [87], e.g. two recent studies used acoustic musical instruments to test pitch discrimination in children but their big challenge has been to avoid the potential use of intensity cues [88,89];

- HI/DI tests are focused on the auditory capacities in the LF domain that are linked to TFS processing which plays an important role for speech understanding in fluctuating noise, known as a big challenge for CI users and a common listening condition in real life environments [81];

- Both are standard pitch perception tests which can permit replication/comparison by different centers, clinicians and subjects following standard criteria [82],

- Pitch perception outcomes in the pediatric population may differ from adults due to cognitive bias depending on test difficulty. However, HI/DI tasks are based on a same/different discrimination paradigm, which is an easy cognitive task and a primary skill developed by the deaf child after cochlear implantation [90];

- Intensity roving is applied to avoid the use of loudness cues by the listeners [82];

- Test duration is short and this may facilitate the application in children who have limited duration of concentration/attention [85].

Hence, our study aimed to evaluate whether HI/DI tests were clinically applicable in children with CI, to assess their LF pitch perception skills using HI/DI tests and to compare their outcomes to NH peers as well as to investigate the effect of chronological age on HI/DI performance.

2.2 Materials and Methods

2.2.1 Participants

20 profoundly deaf children who had been consistent CI users for at least 6 months and who did not have any additional disabilities participated in the study group [91]. Their ages varied from 5 to 17 years (mean age=12yrs, SD=3.1). Their CI model and sound coding strategies were as follows: 16 Advanced Bionics users (implanted with 90K, CII or C1 device with HiFocus1J electrode and fitted with HiRes-S, HiRes-S Fidelity 120, SAS or CIS); 3 Med-El users (implanted with Concerto device with Flex28 electrode and fitted either with FS4 or FS4-p); 1 Cochlear Freedom user (implanted with CI24RE with straight electrode and fitted with ACE). Mean duration of CI experience was 94 months (SD=42.7). The group consisted of 12 unilateral, 5 bilateral and 3 bimodal listeners. All bilateral CI recipients were implanted simultaneously except P15 who had an early implantation on one ear but a late implantation sequentially on the contralateral ear. Unilateral and bilateral CI listeners showed no degree of LF residual hearing in both ears that may have interfered with pitch perception outcomes (hearing thresholds ≥ 85 dB HL for frequencies 125-1000 Hz), whilst bimodal listeners did have residual hearing on side of contralateral hearing aid (HA). Demographic information of individual CI recipients in relation to their listening mode is given in Table 2.

The control group consisted of 31 NH peers (16 male and 15 female). They had hearing thresholds ≤ 20 dB HL for frequencies between 250-6000 Hz on both ears (mean threshold= 6.8 dB HL, SD= 6.7) and did not have any otologic history. Mean age in this group was 10.5 years (range= 5-17yrs, SD=3.2). All CI and NH subjects were placed in mainstream schools.

This study was approved by the local Ethical Committee and parents’ consent was given freely.

2.2.2 Procedures

CI programs for individual recipients were controlled prior to testing. Unaided hearing thresholds were measured via an Aurical audiometer and TDH39 headphones in a sound treated room at frequencies between 125-6000 Hz using a warble tone as were aided thresholds in Free Field through a loudspeaker placed at 0º azimuth at 1m distance from the subject’s head. Speech recognition in quiet was tested at 65 dB SPL in the daily listening mode of individual CI recipients, using standard Italian phonetically balanced bi-syllabic words for pediatric population [92].

HI/DI tests which were part of the Auditory Speech Sounds Evaluation (A§E) suite were used to evaluate LF pitch perception skills in CI users under their daily listening mode whilst performances were measured additionally in the CI only condition for bimodal listeners. CI only condition in bimodal listeners was performed by occluding the contralateral ear with an ear foam plug plus a circumaural headphone to avoid a potential contribution of the non-implanted ear [73].

HI/DI tests were based on a same/different discrimination task between two consecutive complex tones: one intonating and one non-intonating. The non-intonating stimulus represented a harmonic complex signal of a F0 of 200 Hz and 3 higher harmonics presented at lower intensities than F0 (-6 dB at 2F0, -12 dB at 3F0 and -18 dB at 4F0). In both tests, the non-intonating sound was contrasted to an intonating sound. In HI test the intonating sound was characterized with a frequency sweep of F0 together with all harmonics [from NF0 to N(F0+ΔF), N ranging from 1 to 4] whilst in the DI test a sweep of only F0 [F0 to F0+ΔF] leading to a disharmonic intonation was concerned. For both tests, the sweep was linear and introduced at 330ms after the start of the signal and lasted 120ms. Total duration of each stimulus was 600ms and the two consecutive stimuli were separated with a 500ms inter-stimulus interval. White noise was added to the stimuli (SNR +10.9 dB) so that they sounded more natural and intensity roving (±2 dB) was applied to avoid the use of loudness cues by the listeners [82].

Test orders were counterbalanced across subjects in order to minimize the effects of learning and attention factors. Subjects were asked to listen to two consecutive stimuli at 70 dB SPL and to discriminate whether they were the same or different. Both tests were carried out after familiarizing and training the subject with the stimuli and test task.

For each subject, the Just Noticeable Difference (JND) - the smallest ΔF that the subject could discriminate - was calculated by the software using an adaptive staircase procedure where ΔF (41 Hz at test start, range 0 to 214 Hz, 0 Hz means no change between two signals) was increased for an incorrect response and decreased for a correct one until estimating the 50%-point on a subject's psychometric curve. If JND could not be found within 100 trials, it was set to 220 Hz which was above the maximum ΔF value [76,82]. The details of this adaptive staircase procedure are described in a publication by Vaerenberg et al. [93].

2.2.3 Data Analysis

Data analysis was carried out using the Statistical Package for Social Sciences (SPSS) version 19.0 (Chicago, IL, USA). Non-parametric statistical tests were used due to small CI sample size. Differences between CI and NH groups were investigated by Mann-Whitney U test and Wilcoxon Signed Ranks test was used to compare HI/DI results within subjects. Percentage of outcomes that fell within the normal clinical zone (JNDs ≤ 4 Hz for HI and ≤ 10 Hz for DI as per normative data collected by Vaerenberg et al. in adult population) was computed for each group in order to compare pediatric outcomes to adult ones [82]. Spearman rank-order correlations were analyzed to investigate the correlations between HI/DI tests, chronological age, duration of CI experience, age at implantation, listening mode and speech recognition score in CI group whilst correlations were analyzed between HI/DI tests and chronological age in NH group. The cut-off level for statistical significance was set at 0.05.

ID	Age	Gender	Onset / Etiology	CI Ear	CI Age	CI Experience	Sound Coding Strategy	Channel's Number & Bandwidth for F0	Listening Mode	Audiometry (dB HL)	WRS (%)	HI JND (Hz)	DI JND (Hz)
P1	8;5	M	Congenital/Unknown	R	1;7	6;9	HiRes-S Fid.120	1 / 238 – 442 Hz	Unilateral	20,8	90	36.0	54.0
P2	8;7	M	Congenital/Unknown	L	2;0	6;6	HiRes-S Fid.120	1 / 238 – 442 Hz	Unilateral	21,4	100	4.0	49.0
P3	9;7	F	Congenital/Unknown	R	2;5	7;1	HiRes-S Fid.120	1 / 238 - 442 Hz	Unilateral	32,5	90	2.0	43.0
P4	9;7	F	Congenital/Connexin26	L	2;5	7;1	HiRes-S Fid.120	1 / 238 - 442 Hz	Unilateral	36,7	90	9.0	21.0
P5	9;11	F	Congenital/Unknown	R	4;0	5;10	HiRes-S Fid.120	1 / 238 - 442 Hz	Unilateral	22,1	100	7.0	49.0
P6	10;6	F	Congenital/Unknown	R	1;8	8;9	HiRes-S Fid.120	1 / 238 - 442 Hz	Unilateral	23,6	90	45.0	77.0
P7	11;3	M	Congenital/Unknown	R	3;10	7;4	HiRes-S	1 / 250 - 416 Hz	Unilateral	26,4	90	7.0	33.0
P8	11;10	M	Congenital/Connexin26	R	3;2	8;7	HiRes-S	1 / 250 - 416 Hz	Unilateral	20	90	3.0	27.0
P9	13;5	F	Congenital/Unknown	R	1;9	11;7	HiRes-S	1 / 250 - 416 Hz	Unilateral	26,7	95	21.0	172.0
P10	15;0	F	Congenital/Unknown	R	2;7	12;4	SAS	1 / 350 - 494 Hz	Unilateral	22,9	80	3.0	3.0
P11	16;11	F	Congenital/Unknown	L	4;3	12;7	CIS	1 / 350 - 494 Hz	Unilateral	22,1	60	4.0	36.0
P12	17;7	M	Congenital/Unknown	R	4;0	13;6	CIS	1 / 350 - 494 Hz	Unilateral	22,1	85	10.0	24.0
P13	5;9	M	Congenital/Connexin26	B	1;11	3;9	HiRes-S Fid.120	1 / 238 - 442 Hz	Bilateral	R=26,4 / L=27,1	100	1.0	47.0
P14	10;10	M	Congenital/Unknown	B	1;11	8;10	HiRes-S	1 / 250 - 416 Hz	Bilateral	R=21,4 / L=22,9	70	7.0	49.0
P15	12;10	M	Congenital/Unknown	B	2;0	10;9	FS4-p	2 / 198 - 325 Hz	Bilateral	R=23,6 / L=25	80	4.0	2.0
P16	14;7	M	Congenital/Unknown	B	5;5	9;1	HiRes-S	1 / 250 - 416 Hz	Bilateral	R=27,1 / L=26,1	60	21.0	16.0
P17	15;2	M	Progressive/Unknown	B	14;6	0;7	FS4-p	2 / 198 - 325 Hz	Bilateral	R=29,3 / L=30	100	3.0	3.0
P18	9;7	M	Congenital/Unknown	R	4;4	5;2	ACE	22 / 188 - 313 Hz	Bimodal	CI=29,3 / HA=38,8	80	2.0 14.0*	22.0 22.0*
P19	13;2	M	Congenital/Unknown	R	3;3	9;10	HiRes-S Fid.120	1 / 238 - 442 Hz	Bimodal	CI=26,7 / HA=51,3	100	6.0 6.0*	45.0 19.0*
P20	15;0	F	Congenital/Premature	R	14;2	0;9	FS4	1 / 198 - 325 Hz	Bimodal	CI=21,4 / HA=36,3	70	67.0 27.0*	220.0 12.0*

Table 2 Demographic information and audiological outcomes for CI users. Average aided audiometric thresholds were calculated between 125-6000 Hz for CI and between 125-1000 Hz for HA only conditions. CI+HA outcomes for bimodal listeners are marked with *. (WRS= Word Recognition Score).

2.3 Results

All CI recipients and NH children that participated in this study were able to perform both HI/DI test tasks. Individual HI/DI outcomes as well as aided audiometric thresholds and speech recognition scores for CI recipients are shown in Table 2. Figure 10 represents HI/DI outcomes of the CI group in comparison to the NH group.

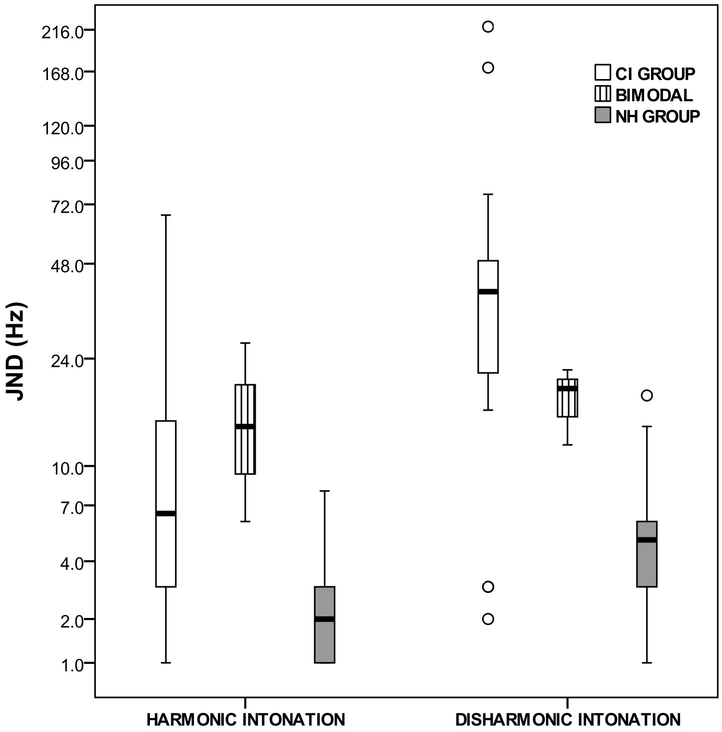

Figure 10 Pitch perception performance in CI subjects in comparison to NH peers.

The statistical comparison treated outcomes of unilateral/bilateral CI users plus CI only condition in bimodal listeners versus NH group (N=51). Analysis revealed a significant difference between NH and CI groups for both HI/DI tests ($p < 0.001$). Pitch perception results between HI and DI tests differed significantly within both groups ($p < 0.001$): outcomes were better for HI than for DI [median JND= 6.5 Hz (range 1.0 to 67.0 Hz) and 39.5 Hz (range 2.0 to 220.0 Hz) in CI group versus median JND= 2.0 Hz (range 1.0 to 8.0 Hz) and 5.0 Hz (range 1.0 to 18.0 Hz) in NH group for HI and DI respectively]. In the CI group, 40% of HI outcomes fell within the adult normal zone [10] versus 15% of DI outcomes. In the NH group, the majority of children, 93.5% for HI versus 90.3% for DI, showed outcomes that were considered normal in adults. However, some children under the age of 8.5yrs had higher JND scores: 18% (2 out of 11 children) on HI versus

27% (3 out of 11 children) on DI. The analysis of correlation revealed a significant negative correlation between chronological age and JND score for DI ($p < 0.05$) but not for HI one ($p > 0.05$).

On the other hand, in CI group, the correlations between HI and DI outcomes were statistically significant ($p < 0.05$). Listening mode correlated significantly with DI performance ($p < 0.05$): bimodal users (2 out of 3 children) showed better DI outcomes in bimodal listening than CI only condition [median HI JND= 6.0 Hz (range 2.0 to 67.0 Hz) versus 14.0 Hz (range 6.0 to 27.0 Hz) and median DI JND= 45.0 Hz (range 22.0 to 220.0 Hz) versus 19.0 Hz (range 12.0 to 22.0 Hz) in CI only and bimodal listening conditions respectively, N=3].

2.4 Discussion

The results of the present study indicated that HI/DI tests evaluating LF pitch perception were clinically applicable in both NH/CI recipient children of 5 years and older. Overall, LF pitch perception was found considerably poorer in the majority of children with CI than in their NH peers. HI/DI outcomes showed a significant positive correlation in CI recipient children: children with higher JNDs on HI tended to have higher JNDs on DI test as well. Furthermore, both groups showed better HI scores than DI. Better performance on HI than DI could be mainly due to the fact that HI test makes use of both place and time-based codes because of the sweep of the harmonics together with F0. This could lead one or more of the harmonics to move to adjacent channels providing a place cue for lower JNDs. However, DI test is dominated by time-based codes because of the sweep of F0 alone whilst its harmonics were kept fixed at their initial frequency and therefore, it is not likely that a small change in F0 could lead to a different electrode being stimulated. Moreover, temporal beatings that may serve as cues to CI users can be induced only for very high stimulus levels ($\Delta F > 150$ Hz) in the DI signal. Therefore, it seems likely that HI test evaluates the availability of both TFS and place cues whereas DI test provides more differential outcomes on phase locking and TFS processing capacities [82,86].

The present findings in the majority of NH children were similar to those for NH adults studied previously by Vaerenberg et al. [82] [median JNDs 2.0 (HI) and 5.0 Hz (DI) in children versus median JNDs 2.0 (HI) and 3.0 Hz (DI) in adults]. However, surprisingly the outcomes in CI recipient children revealed better scores in comparison to the previous study by Schauwers et al. [76] carried out on 21 postlingually deafened adults with unilateral CI where median JNDs were found to be 16.0 (HI) and 139.0 Hz (DI). Such discrepancies could be attributed to the small CI sample size in both studies. On the other hand, a similar finding was observed by Looi and Radford [78] in pitch ranking skills for CI recipient children performing more accurately than adult counterparts. Even though it is known that both frequency and time domain information are present in the peripheral auditory system and the frequency map is maintained to some degree throughout the auditory system up to the primary auditory cortex, there is recent evidence for the existence of cortical neurons beyond the primary auditory cortex that are tuned to pitch [80]. These findings can lead to questions as to whether a congenitally or early impaired auditory system, owing to higher cortical plasticity [94], might be able to develop strategies to compensate for poor LF pitch information provided by the CI. Another question would be if children with CI might have benefited from a larger spiral ganglion neuron population as demonstrated by Miura et al. [95]. In the present study as well as in Schauwers et al. findings [76], the majority of CI recipients showed abnormal results confirming poor TFS processing of the CI. However, CI recipient children seemed to perform considerably better on DI test than most of the adults with CI [76] and actually, this aspect would need to be studied in larger CI populations to investigate the possible effects of age.

Previous studies in children differed in their findings for predictors of outcome of pitch perception performance. Some indicated age at implantation and/or duration of CI experience as predictors in CI users whilst others noted a positive effect of increasing chronological age but not age at implantation [52,89]. The analysis of correlations in the present study revealed no significant correlations between pitch perception performance, chronological age and age at implantation in the CI group. However, in the NH group the chronological age had a significant effect on DI performance; a small number of NH children under the 8.5 years of age showed considerably higher JNDs and some of them had JNDs that are considered abnormal in the adult NH population [82]. Although poorer outcomes in some younger NH children may reflect cognitive aspects in development and/or limited duration of attention/concentration rather than the intrinsic pitch perception capacity itself, the outcomes in the majority of NH children confirm that HI/DI tests are not influenced substantially by cognitive bias.

It has been thought that one way to improve the representation of TFS for CI users could be to add acoustic hearing through a contralateral HA or through EAS stimulation on the same ear [79]. Indeed, HI/DI tests have been previously studied in EAS processor users by comparing their performance in electric stimulation only and EAS conditions. The results confirmed an improvement in DI score whereas HI score was the same for both listening conditions (median HI JND 7.0 Hz in both conditions; median DI JND 44.0 Hz in electric stimulation versus 12 Hz in EAS) [86]. Our findings in bimodal listeners showed a similar tendency between CI only and bimodal listening conditions. 2 out of 3 children had remarkable bimodal benefit on DI test which could be explained by processing of additional LF pitch cues that were provided via the HA by remaining phase-locking capacities of residual hearing. However, the outcomes in bimodal listeners were still worse than those of NH counterparts.

An important limitation of this study has been the small CI sample size especially for bimodal listeners. Moreover, CI recipient children were evaluated in their daily listening mode mainly in order to avoid an acclimatization effect and longer test duration. Bimodal listeners were believed to benefit from the additional LF cues provided by the contralateral HA and therefore bimodal listening was compared to CI only listening mode whereas bilateral CI advantage was not expected for this kind of task [79]. On the other hand, similarly with findings of Schauwers et al. [76], the small CI sample size in this study did not allow us to investigate, especially for DI test, the reason why some CI recipient children had very low JNDs that were considered normal even at an adult level (40% of HI versus 15% of DI outcomes), although they did not have LF residual hearing that could interfere with outcomes (hearing thresholds ≥ 85 dB HL for frequencies ≥ 125 Hz). A possible explanation for low JNDs could be the spectral leakage into adjacent channels. Hence, fluctuations in the amount of cross-channel leakage during frequency shifts might have served as cues for some CI listeners [86]. Moreover, the device type, in particular the sound coding strategy, may have an impact, i.e. coding strategies such as HiRes and FS4 that are based on channel-specific real-time filters may provide better temporal cues rather than those such as Fidelity 120 and ACE that are based on Fast Fourier Transform analysis which leads to temporal smearing [96,97]. On the other hand, FS4, where F0 (200 Hz) is represented by the 2nd apical channel - has a greater extension to the LF range if compared to extended LF filters of HiRes. In strategies such as HiRes and Fidelity 120, F0 is represented by the first channel, although the band-pass filter shape induces an intensity reduction [5] and reduced loudness might have contributed to higher JNDs. It would be interesting to study aspects such as the effect of the spectral range, the number of stimulating electrodes, stimulation rate and temporal envelope extraction on HI/DI outcomes in large populations.

Furthermore, findings of this study indicated that speech recognition scores in quiet were not significantly correlated with HI/DI outcomes. This result was consistent with the previous study in EAS processor users in which Vaerenberg et al. [86] found significant differences between electric only and EAS listening modes in speech in noise results as well as a remarkable improvement in DI performance. Moreover, the role of TFS cues is emphasized in difficult listening tasks such as listening in the dips which is defined as detecting a signal in a fluctuating background [81]. Actually, it would be interesting to investigate the correlation between DI test results and speech recognition tests in fluctuating signal/noise levels in particular as they may represent real life situations more adequately [50].

A§E phonemic discrimination tasks are used in our clinical routine as they provide useful information in selection and follow-up of CI patients [82,98]. HI/DI tests can be a supplement of this assessment battery for analysis in the low-frequency domain. Both tests seem to be applicable clinical tools for the evaluation and monitoring of LF pitch perception skills in children of 5 years and older, but it should be considered that inter-subject variability is larger in children right up to 8.5 years of age. The DI test in particular seems to provide more differential outcomes in the evaluation of LF processing capacities whereas HI test offers information about the availability of both place and time-based codes. HI/DI tests, in further studies with larger populations, may provide better information about bimodal benefit, as well as about CI models and technology apropos the improvement of pitch perception. In particular, it would be useful to investigate the role of HI/DI outcomes in the decision process between bimodal versus bilateral CI use.

2.5 Conclusion

HI/DI tests- which evaluate LF pitch perception- were found to be applicable both to CI recipient children and to their NH peers. Outcomes in NH children were similar to those found for the adult population; nevertheless, children under the age of 8.5yrs showed larger inter-subject variability and a small number of them had abnormal JNDs. CI recipients performed less accurately than their NH peers, especially on DI test. Contralateral HA users tended to have better performance for the DI test, reflecting a bimodal benefit in the low-frequency domain. Findings supported that DI is dominated by phase locking capacity versus HI test benefiting from place cues as well. It is better to use both tests together in order to gather useful information for comparative analysis of TFS versus place cues availabilities.

CHAPTER 3: ADAPTATION OF THE STARR TEST FOR ADULT ITALIAN POPULATION: A SPEECH TEST FOR A REALISTIC ESTIMATE IN REAL LIFE LISTENING CONDITIONS[ii]

Abstract
Objectives: To introduce the Italian adaptation of the STARR test based on a roving-level adaptive method to mimic challenging real life listening conditions that can be used in people with auditory prostheses.
Design: Normative data were collected and interlist-variability as well as learning effects were investigated using a within-subject design with repeated measures.
Study sample: A group of 32 normal hearing (NH) adults participated in the study.
Results: The average speech reception threshold (SRT) for NH subjects was -8.4 dB SNR. The variability of mean SRTs across test lists was relatively small (≤ 1 dB for all test lists). The statistically significant differences between lists was eliminated after applying correction factors. On the basis of variability for the corrected SRTs within each subject, a difference of 2.8 dB in SRT was meaningful for outcome comparisons using one test list per condition and 2 dB using two lists per condition. Statistical analysis did not show any significant learning effects.
Conclusions: Findings in NH listeners suggested that the Italian STARR test could be a promising supplement to existing speech assessment tools. Further studies in populations with hearing impairment could contribute to cross-language studies.

3.1 Introduction

Everyday communication environment due to changes in both speech and noise levels usually brings challenges to listeners, in particular to people with auditory prostheses in situations such as group conversation where the recipient may have to deal with someone who uses quite a large vocal effort and is placed close to the recipient's microphone and someone who uses much less vocal effort and is located further from the recipient's microphone [50,99]. In the attempt to develop a speech perception test that was representative of adverse listening situations and sensitive to differences in performance between various settings in auditory prostheses, Boyle et al. [100] and Haumann et al. [99] introduced a new test approach which was based on measurement of speech perception using a roving-level adaptive method where the presentation level of both speech and noise signals varied across sentences. The two factors that resulted in the greatest variation to listeners were change in presentation level and modification of signal-to-noise ratio (SNR). Whilst the sentences were presented at a roving level, the noise was adapted automatically to obtain SNR at which the subject reaches the 50% correct level referred to as the

---

[ii] **Material covered in Chapter 3 has previously been published in Dincer D'Alessandro H, Ballantyne D, De Seta E, Musacchio A, Mancini P (2016) Adaptation of the STARR test for adult Italian population: A speech test for a realistic estimate in real-life listening conditions. Int J Audiol 55:262-267.**

Speech Reception Threshold (SRT). The test required the listener to understand speech in the presence of competing noise and to do this for an unpredictable presentation level that should cover the constantly changing range of levels that might be encountered in most everyday life situations. At the initial application of this test, Boyle et al. [100] studied difference in performance between two types of signal processing in 6 cochlear implant (CI) recipients using a crossover design and they found significant differences after one month of experience. Subsequently, Haumann et al. [99] investigated the effect of processor models in 55 CI users using a German test material based on HSM sentences [101]. The participants were divided into five groups according to their CI processor model and all groups were matched for demographic factors and traditional speech perception test scores. The groups showed significant outcome differences when tested with roving-levels across sentences (the mean SRTs ranged from -1 to +6.4 dB), although they performed similarly on HSM sentence test using a fixed speech presentation level (at 65 dB SPL) and a fixed CCITT noise (+10 dB SNR). These findings suggested that a test using roving levels could reveal differences between processor designs that were not shown when a fixed presentation level and fixed SNR were used. Later on, Boyle et al. [50] created the Sentence Test with Adaptive Randomized Roving level (STARR) test which made use of the IEEE sentences [102] by adapting them into British-English. The recordings were done by two native speakers, one male and one female, where three consecutive lists of IEEE sentences were combined to one STARR test list which resulted in a total of 25 lists, each with 30 items. The STARR test was applied to adults with normal hearing (NH) as well as to a group of adult CI recipients and the outcomes supported in particular a noticeable difficulty of CI users under these challenging test conditions, showing that for normal hearing subjects the effect of roving was minimal whereas for CI users it was much greater [50].

The present study aimed to introduce an Italian version of the roving-level adaptive test method. For this purpose, the STARR test was adapted into Italian, normative data were collected for an adult population, and interlist-variability as well as learning effects were investigated.

3.2 Materials and Methods

3.2.1 Participants

A group of 32 NH adults (15 female and 17 male) reporting no otologic history and no hearing complaints participated in the study. Their ages varied between 18 and 53 years (mean=32yrs, SD=11). All had hearing thresholds ≤ 20 dB HL for frequencies between 250-8000 Hz on both ears (mean=10 dB HL, SD=6). This study was approved by the Local Ethical Committee and subjects' consent was given freely.

3.2.2 Italian STARR Test Design

The Italian STARR test made use of sentences from the standard Italian speech recognition test developed by Cutugno et al. [92]. The original corpus consisted of 200 meaningful sentences which were organized into 10 test lists, each of 20 sentences. The sentences were selected based on lexical and morpho-syntactic characteristics in order to make them more easily accessible to a heterogeneous group, coming from different regions of Italy [103]. Independent, short, simple everyday sentences were used. 100 phonemically balanced bisyllabic words [104] were included within the sentences; criteria regarding the frequency of use in contemporary Italian were followed [105,106].

The recordings were done in a double-walled sound room with additional sound absorption material on the walls and ceiling (complied with ANSI S3.1-1999) [107]. The speaker was a male native Italian phonetician using clear conversational speech with natural rate and natural vocal effort without emphasizing any key words. During recordings the speaker was monitored by another phonetician in order to control pronunciation errors. Sentences misread or pronounced incorrectly were repeated. Recordings were made directly to digital format with a Neumann condenser microphone (TLM 193) which was directly connected to a Tascam DAT recorder with 16-bit resolution at 44100 Hz. Following the recording session, the contents of the DAT were transferred digitally to a computer. All signals were analysed using KAY 4300 Computerized Speech Laboratory (CSL) system. Equalization of the levels, when necessary, was done using a software for digital processing of audio signals (Sound Forge 4.0, Sonic Foundry). For each sentence, the integral value of RMS expressed in dBV was calculated and subsequently all sentences were equalized considering ±1 dB as normal range in respect to the average value. Sentences were separated by seven-second intervals.

In the Italian STARR test, sentence material recorded by Cutugno et al. [92] was transferred digitally to a PC and edited using the CoolEdit 2000 sound-editing software. Each sentence was isolated and adjusted in level such that the total root-mean-square power was 20 dB below full scale for a.wav file, i.e. -20 dBV. In pilot testing of STARR test, it was found that a set of 15 sentences provided a practical balance for a clinical test: sufficient sentences were delivered to allow convergence around an SNR range that represented the abilities of a subject to understand speech in the presence of competing noise, while avoiding an overly long test that would induce fatigue [108]. Therefore, 150 sentences were selected from the original corpus of 200 sentences and each test list consisted of 15 sentences (5 sentences were removed from each test list of Cutugno et al. [92] maintaining their list boundaries). The sentences typically consisted of 5 words with 9 to 13 syllables (median = 5 words across all sentence lists as well as for individual lists, range 3-7). Each sentence was used only once during the whole procedure in order to avoid repetitions in case of outcome comparisons using different test lists during same sessions. Three key words (noun, verb, adverb or adjective) were allocated for each sentence and speech-shaped noise was used as competition. Figure 11 illustrates both the spectrum of the first 10 sentences from list 1 and that of the noise. The speech spectrum showed maximum energy between 100 and 500 Hz followed by a drop with an average slope of 10 dB per octave up to 8000 Hz. The noise had a spectrum resembling long-term spectrum of the speech test material and was based on recommendations of the Institute of Hearing Research (IHR) in the UK. The spectrum of the noise used by the IHR - produced by shaping a white noise and used for various IHR speech tests - was flat from 100 Hz to 1000 Hz and then dropped with an average slope of 11 dB per octave to 8000 Hz.

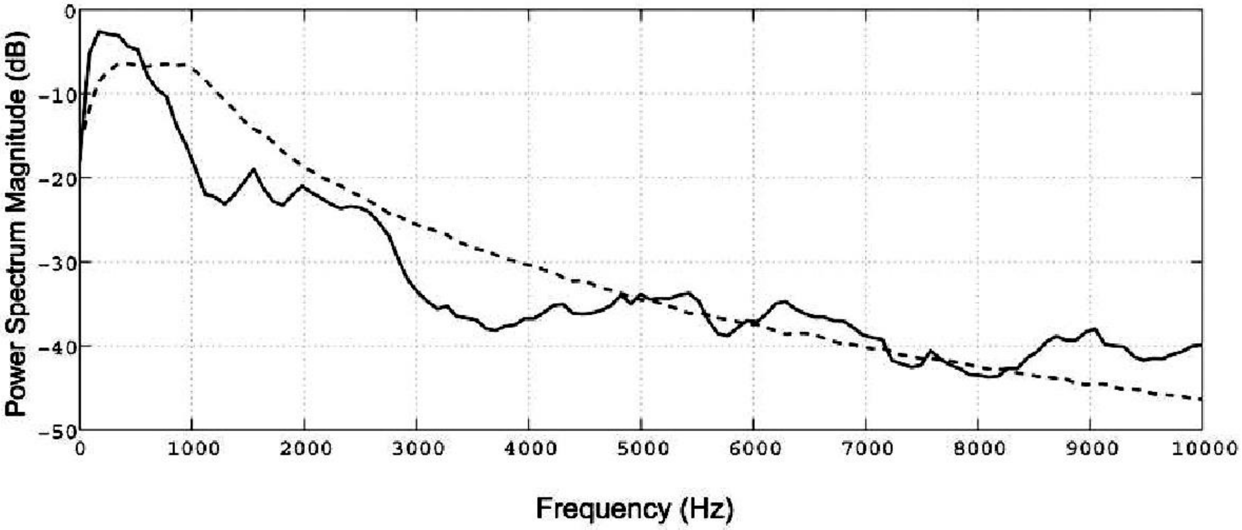

Figure 11 The spectrum of the speech (continuous line) and that of the noise (dotted line). The speech spectrum showed maximum energy between 100 and 500 Hz and then dropped with an average slope of 10 dB per octave up to 8000 Hz. The noise spectrum was approximately flat from 100 Hz to 1000 Hz and then dropped with an average slope of 11 dB per octave up to 8000 Hz.

The STARR software was written in Visual Basic and provided the clinician with a graphical user interface to both deliver and score the test. Three presentation levels were used within each test list: 50, 65 and 80 dB SPL. This range was selected in line with the levels that were typically explored in hearing instrument research. Once the clinician selected a test list to be presented, the software randomly selected both a sentence and a presentation level at which to begin the procedure. This process continued with any sentence from the list being presented only once and an equal number of presentations being made at each presentation level, i.e. 5 at 50, 5 at 65 and 5 at 80 dB SPL. Each time any test list was presented to a new subject, a different presentation order of the sentences was likely to occur and any given sentence was likely to be presented at a different presentation level.

3.2.3 Procedure

All participants underwent testing in a sound-proofed booth of 2x2 metres. The stimulus was presented via the PC and a preamplifier connected directly to a single loudspeaker. Both sentence material and noise came from 0° azimuth with loudspeaker at 1m from participant's head. Before each test session, calibration was performed. Each sentence was matched in terms of RMS energy. All gains needed to apply the desired sentence presentation level and the SNR were automatically calculated by the program software. This meant that only a single point calibration was necessary. A 10 second duration noise burst, centred around 1 kHz, was used for calibration. The burst was arranged to be 11 dB higher than the sentence RMS level. By delivering the calibration noise burst such that a level of 91 dB SPL was observed on a sound level meter, placed where the subject's head would be during the test, the correct sentence levels would be automatically produced by the software. For example, a gain of -11, -26 or -41 dB was applied to a sentence to produce maximum, middle or minimum presentation levels of 80, 65 and 50 dB SPL respectively.

In order to verify inter-list variability, all participants were tested using all 10 test lists. List order was counterbalanced by staggering the allocation of test lists to participants. No practice list was used in order to investigate any learning effects. The test was started after explaining the task to participants and warning them about the possibility of sometimes hearing quite loud stimuli. The upper sound pressure limit was set to 91 dB SPL resulting in a -10 dB SNR for the highest sentence presentation level of 80 dB SPL. Beyond this limit, the program automatically decreased both the sentence and noise levels, maintaining the required SNR while avoiding uncomfortable stimuli. However, this happened very rarely in practice as it was extremely difficult to understand the speech at the -10 dB SNR, especially at such high presentation levels. Participants were told that they could ask for a break whenever they needed but none of them requested it, since a test session including hearing assessment never exceeded 45 minutes. After presentation of a sentence, participants were asked to repeat it as accurately as possible, while explaining that not every word needed to be correct and encouraging them to guess when not sure.

For scoring, single key words were highlighted in white on the screen becoming green to indicate words repeated correctly. The clinician could either click on single key words or, alternatively, select the “all” or “none” options to indicate participant’s response. The SNR was +20 dB initially and varied adaptively according to the participant’s response. Where the minimum specified number of key words was correctly identified, typically 2 out of 3, the sentence was considered correct and the SNR used for the next sentence was made more adverse. Where insufficient key words were correctly identified, the sentence was considered incorrect and the SNR for the next sentence to be presented would become more favourable. The initial step size was 10 dB, dropping to 5 dB after the first reversal of the adaptive track and dropping again to the final step size of 2.5 dB after a further reversal. The SNR was varied by adjusting the noise level while keeping the speech level at 50, 65 and 80 dB SPL and using the same SNR for all three levels. The SRT was computed by averaging the SNRs for the last nine (7th to 15th) sentences, along with the SNR at which a next (16th) sentence would have been presented. Clicking on the OK button finalized scoring for individual sentences and the next sentence was presented by the software until the end of the test list where the SRT value was displayed automatically.

3.3 Results

3.3.1 Normative data

A mean SRT was calculated averaging the SRTs for each participant across all test lists. The mean SRTs for individual participants ranged between -5.8 to -9.7 dB SNR. Subsequently, an overall mean SRT was produced averaging the SRTs across all participants (Mean SRT= -8.4 dB SNR, SD=0.9).

3.3.2 Inter-list variability

To assess the variation in difficulty between the test lists and to calculate deviations from the participant's mean, the SRTs across all test lists were averaged for each participant and the SRT score for each test list was subtracted from the average SRT for that participant. The deviations for each test list were averaged across all participants in order to calculate correction factors for each list. Figure 12 represents the deviation of SRT values from the overall mean for each test list. The deviations for lists 1, 3, 4, 6, 8 were reasonably small (<0.5 dB). The largest deviations were 0.9 dB for list 7 and 1 dB for list 10. The SD of the deviations was 0.7 dB.

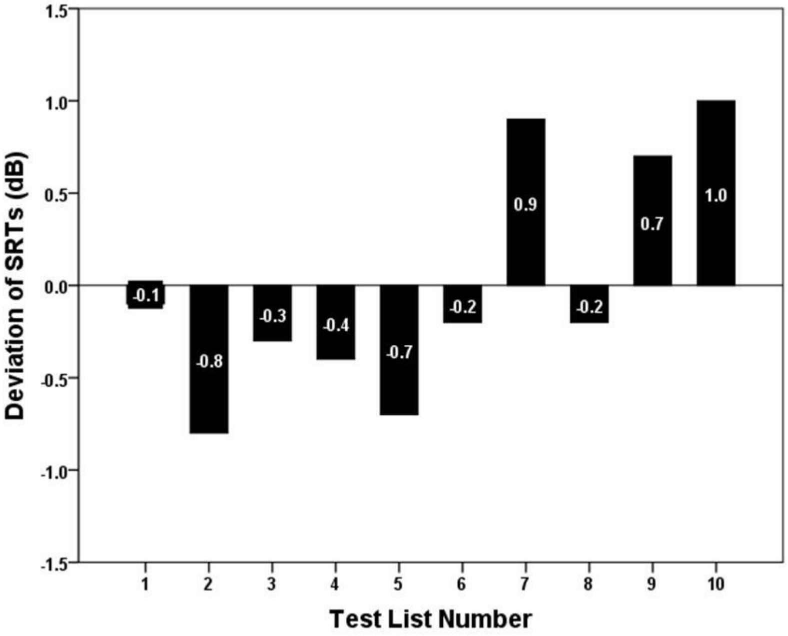

Figure 12 Deviation of SRT values from overall mean for each test list. SRT, speech reception threshold.

The SRT obtained using a given test list was corrected by subtracting the deviation for that list in order to compensate for differences in list difficulty. The reliability of SRT estimates was assessed by calculating the SD across all corrected SRTs (of all test lists) for each participant and this SD was averaged across all participants. This value was compared to the average SD without applying correction factors in order to evaluate the effect of the application of correction for each test list. The average SD across all corrected SRTs was 1 dB in comparison to the average SD of 1.2 dB without applying correction factors (uncorrected SRTs).

To determine the meaningful difference in SRT value for an outcome comparison (e.g. when comparing two different CI processors) by using one single test list under each listening condition, the SD of the difference in SRT between two lists was $\sqrt{2}$ times SD. A difference was considered meaningful if it was more than 2 SD; therefore, the difference in SRT was calculated as $2\sqrt{2}$ times SD. Therefore, if an outcome comparison is done by using one single test list under two listening conditions, a difference of 2.8 dB in SRTs would be considered meaningful and this value would be 3.4 dB in case of uncorrected SRTs. If two test lists were to be used per condition, the corresponding values would be 2 dB and 2.4 dB for corrected and uncorrected SRTs respectively.

Statistical analysis was carried out using Statistical Package for Social Sciences (SPSS) version 19.0 (Chicago, IL, USA). One-way repeated measures analysis of variance (ANOVA), with Greenhouse-Geisser correction, was conducted by using paired t-test procedures with Bonferroni correction for multiple comparisons in order to investigate the statistical differences between the test lists with uncorrected and corrected SRTs. A significant main effect of the test lists was found for uncorrected SRTs [$F(6.16, 191.19)=12.17, p=0.000$] where post hoc tests indicated significant differences of lists 7, 9 and 10 to the rest of the test lists ($p<0.005$). However, no significant main effect of test list was found with corrected SRTs [$F(6.16, 191.19)= 0.13, p=0.993$].

3.3.3 Learning effects

To evaluate learning effects, the SRTs were averaged according to the order of testing without considering the list number. The SRTs averaged across test lists that are administered in the same serial position (first, second, etc.) are illustrated in Figure 13. The mean SRTs for test lists 1 to 10 changed within a range of 0.7 dB. A one-way repeated-measures ANOVA did not show any significant main effect of the list presentation order [F(9, 279)=1.60,p=0.114].

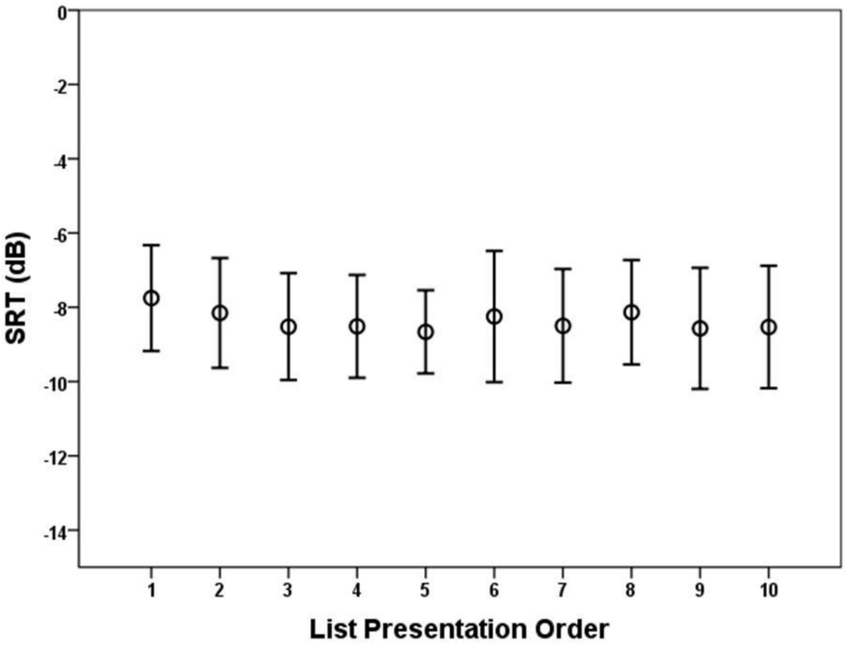

Figure 13 The SRTs averaged across test lists considering list presentation order. Error bars show ±1 SD. SRT, speech reception threshold.

3.4 Discussion

Improvements in prostheses such as hearing aids and cochlear implants are very crucial to enabling the patient to reach her/his best potential. In parallel to such progress, there is an increasing need for new speech assessment tools that mimic challenging real life listening conditions where background noise is usually present and speech level varies according to vocal capacity and distance of the speaker. A recent attempt to meet this need has been done by Boyle et al. [50] by introducing the STARR test which is based on adaptive randomized roving levels across sentences and the present study is the Italian adaptation of the original STARR test. This test was found particularly sensitive to lower level speech and is believed to provide a better estimate of improvements in technical settings of auditory prostheses [50].

Outcomes for STARR in a British-English NH population [50] were based on two speakers-male/female- whereas the Italian version had a male speaker only as in Cutugno et al. [92]. The mean SRT in Italian NH population (-8.4 dB, range -5.8 to -9.7 dB) was similar but slightly lower than the mean SRTs for the male (-6.1 dB, range -2.8 to –9.3 dB) and the female (-5.7 dB, range -0.8 to -10.3 dB) speakers as well as the mean across them (-5.9 dB, range -2.1 to -9.6 dB) studied with the original version of the STARR test. Such outcome differences have been found in other multilingual speech perception tests such as the Oldenburg Sentence Test (OLSA)/Matrix Sentence Test [109] and the Hearing in Noise Test (HINT) [110] and can be attributed to language and speaker dependent factors [111].

A key aim of this work was to verify and strengthen the clinical reliability of the Italian STARR test. For this purpose, the variations in difficulty between the test lists were assessed. The variability of mean SRTs across test lists for NH subjects was ≤ 1 dB for all sentence lists. The findings indicated that the SD of the deviations was 0.7 dB. This value was 0.6 dB in the case of Boyle et al. [50] and both outcomes were highly consistent. Although the deviations for the majority of the Italian STARR test lists were reasonably small (<0.5 dB), they were higher than in multilingual matrix tests (range 0.13 to 0.2 dB) [112]. Some test lists (7, 9 and 10) had considerably larger deviations and showed statistically significant differences with the rest of the test lists. However, the SRT obtained through a certain list was corrected by subtracting the deviation for that list from the raw SRT and the statistically significant differences between test lists were eliminated after this procedure.

On the other hand, the average SD across all corrected SRTs was 1 dB in comparison to the average SD of 1.2 dB without applying correction factors. The benefit of applying correction factors was remarkable. Outcome comparisons using one single test list under two different listening conditions in a within subject design considered a difference of 2.8 dB to be meaningful in corrected SRTs versus 3.4 dB in case of uncorrected SRTs. The corresponding values became 2.0 dB and 2.4 dB for corrected and uncorrected SRTs respectively when two test lists were used to assess each condition. In the light of these findings, the authors suggest using two test lists per condition and applying correction factors in order to ensure more reliable outcomes. On the other hand, in order to avoid a reduction in the number of available test lists while maintaining reasonably low interlist-variability, it would seem worthwhile to increase the number of items for each test list as in the original version of the STARR test by adding 15 sentences recorded by the female voice to each list. This would lead the SRT to be computed by averaging the SNRs for a total of 20 sentences instead of 10 in the present Italian version and would help the variability to decrease as the number of independent items scored in the task would be increased [113]. Alternatively, equalisation of sentence intelligibility, which is a standard reliability optimization procedure, could be applied [109,110,111,114]. In the Italian STARR test, the question of accounting for differences afterwards was preferred as in the original version in order to maintain low inter-list variability; applying correction factors was required to compensate for differences in difficulty across test lists so that they effectively provide consistent outcomes and can be used interchangeably.

Learning effects were investigated in that no practice list was used prior to testing and the SRTs were compared only considering order of list presentation. Similarly with Boyle et al. [50], the mean SRT for the first list presented was slightly higher than the next lists and the mean SRTs (lists 1-10) covered a range of only 0.7 dB. Moreover, statistical analysis confirmed no significant main effect of order of list presentation. Nevertheless, it should be considered that the SRT calculation of the STARR test is based on averaging the SNRs for the last nine sentences along with the SNR at which a next sentence would have been presented and the experience through initial sentences may have contributed to learning of the NH participants. However, it seems reasonable to expect that people with auditory prostheses may not compensate as quickly as NH people. Therefore, it would be better to administer one practice list before a test session with STARR in order to minimise learning effects in populations with hearing impairment. This recommendation can also be supported with the findings of other speech recognition tests such as the Canadian Francophone HINT [115] and the Italian OLSA Test [114] that have found significant differences between testing trials suggesting a possible practice effect in NH populations.

It has been shown that speech intelligibility can be affected by phonemic content, word familiarity, sentence length, RMS levels and sentence difficulty [110,116]. Hence, the creation of Italian STARR lists, based on the original material provided by Cutugno et al. [92], took into consideration aspects such as word familiarity, sentence length, syntactic structure, equalization of RMS level and minimum number of items per test list. However, an important limitation of the Italian STARR test was that the phonemic distribution of the entire original sentence set was not matched although phonemically balanced bisyllabic words [104] were equally distributed throughout the test lists. Moreover, further studies will be required to investigate use in people with auditory prostheses who have greater difficulty in understanding speech, especially at lower levels and in the presence of noise. Such outcomes may provide useful information for optimizing fitting procedures, patient performance and contributing to cross-language studies.

3.5 Conclusion

The present study showed findings for the Italian STARR test which was based on a roving-level adaptive method. Outcomes were in line with previous research in NH population and the variability of mean SRTs across lists was relatively small (≤ 1 dB for all test lists). The benefit of applying correction factors was basic to improval of reliability. Statistical analysis showed no significant learning effects.

CHAPTER 4. TEMPORAL FINE STRUCTURE PROCESSING, PITCH AND SPEECH PERCEPTION IN COCHLEAR IMPLANT RECIPIENTS

4.1 Introduction

The envelope and Temporal Fine Structure (TFS) are known to be two important acoustic cues for speech intelligibility [117]. The envelope is an amplitude modulated carrier with relatively slow modulations over time that are superimposed on a more rapidly varying TFS which is a frequency modulated carrier [81]. The envelope traditionally has been considered as the most important carrier of information for speech signals. However, envelope cues alone are insufficient to permit the spatial separation of multiple voices and therefore, it is believed that TFS information may be crucial for speech perception in the presence of noise, especially for the ability to listen in fluctuating background sounds [79,81]. Present-day cochlear implant (CI) design allows conveying mainly envelope information in different frequency bands whilst TFS information is mostly discarded and this fact is thought to contribute to the big difficulty of CI recipients when listening in the presence of noise while most of them show high level of speech understanding in quiet [50,81,118].

TFS processing has proved to contribute to pitch perception which is important for speech perception [80,81]. Neurophysiological studies in animals have shown that pitch can be encoded by two auditory mechanisms: place coding that is based on tonotopic excitation where pitch cues are conveyed through spatial alteration of nerve fibers and phase locking which is a time-based mechanism that locks onto the TFS of the signal [119]. Place coding is more dominant in conveying pitch for high frequency (HF) signals whereas low frequency (LF) coding is dominated by phase locking and TFS processing is assumed to be reflected in the phase locking patterns within the auditory nerve [83].

Two recent pitch perception tests, A§E Harmonic Intonation (HI) and Disharmonic Intonation (DI), are expected to be indicative of the capability of the ear to use its phase locking mechanism and therefore, are believed to depend on the availability of TFS cues since both tests focus on the spectral discrimination of the auditory system in the LF domain linked in particular to Fundamental Frequency (F0) [82,85]. Previous HI/DI outcomes in CI recipients revealed abnormal pitch perception in comparison to normal hearing (NH) subjects [76,82,120]. Furthermore, Vaerenberg et al. [86] studied in 6 adult CI users outcomes for pitch and speech perception comparing electric stimulation only to electro-acoustic stimulation (EAS). Differences for speech perception performance in quiet were not statistically significant. However, results indicated a remarkable improvement in DI with an EAS processor compared to electrical stimulation only, just as speech perception outcomes at +10 dB SNR showed statistically significant differences. When testing speech perception in noise, their study has used a method based on a fixed Signal-to-Noise Ratio (SNR) where speech-shaped noise fixed at 65 dB SPL was presented with varying speech signal (SNRs at +10, +5, 0, -5 dB) using monosyllabic CVC words and performance was measured as a percent-correct score at that SNR.

However, in real world listening conditions, the presence of noise in a fluctuating background together with the variability of speech levels according to vocal capacity as well as to the distance of the speaker brings further challenges to listeners. A new assessment tool, the STARR test [50,99], is based on measurement of speech perception using a roving-level adaptive method. Whilst the sentences are presented at a roving level, the noise is adapted automatically to obtain SNR at which the subject reaches the 50% correct level referred to as the Speech Reception Threshold (SRT). The STARR test in British English language was applied to adults with normal hearing as well as to a group of adult CI recipients and the outcomes supported the big difficulty of CI users under these challenging test conditions [50]. The STARR test has been also adapted into Italian language and normative data as well as interlist-variability and learning effects were studied previously [121]. However, outcomes for Italian speaking CI users are so far missing.

The above mentioned background as well as lack of outcomes regarding TFS processing in relation to speech perception capacities, especially when both speech and noise signals are fluctuating just as in realistic listening conditions lead us to conduct the present study with the aim to investigate in adult CI recipients:

- The LF pitch perception outcomes linked to TFS processing capacities, in a large group;
- The Italian STARR outcomes in relation to other speech perception tests;
- The interaction between TFS processing and speech perception outcomes in particular those of the Italian STARR;
- The effects of variables such as age, the duration of profound deafness before implantation, the duration of CI experience and hearing thresholds;
- The correlations between TFS processing and speech perception outcomes in relation to the amount of bimodal benefit.

4.2 Materials and Methods

4.2.1 Participants

The participants in the study were 43 postlingually deafened adult CI recipients (17F, 26M), aged 18-83 years (mean=57yrs, SD=17.2) at the time of testing. All were consistent CI users with at least 3 months of CI experience (range 3 to 208mths, mean=52.7mths, SD=50.9). The mean duration of profound deafness before implantation was 75 months (range 3 to 360mths, SD=98.6) and this value was longer than 10yrs for 37% of the group. The study group consisted of 23 unilateral, 6 bilateral and 14 bimodal listeners. All bilateral CI recipients were implanted simultaneously with the same CI model on both ears. There was one exception who had a different model on the contralateral ear 3,5yrs after the first implantation (a Cochlear Freedom device on one ear and a Medel Concerto device on the contralateral ear).

Assessment regarded a total of 49 CI ears where 15 were implanted with Advanced Bionics devices [14 ears implanted with 90K device and fitted either with HiRes-S (n=8), HiRes-S with Fidelity 120 (n=5) or HiRes-P with Fidelity 120 (n=1) plus 1 recipient implanted with C1 device and fitted with CIS]; 30 were implanted with Med-El devices [all implanted with the Concerto and fitted either with FS4 (n=20) or FS4-p (n=10)]; 4 were implanted with Cochlear Freedom devices [all implanted with CI24RE and fitted with ACE]. The number of active electrodes all ranged from 12 to 22 in the group. The bandwidth for the most apical electrode varied between 250 to 416 Hz for HiRes users, from 238 to 442 Hz for Fidelity 120 users, 198 to 325 Hz for FS4 users, 188 to 313 Hz for ACE users and 350 to 494 Hz for the CIS user.

All unilateral and bilateral CI listeners showed no degree of LF residual hearing in both ears that may have interfered with pitch perception outcomes (hearing thresholds ≥ 85 dB HL for frequencies 125-1000 Hz), whilst bimodal listeners did have residual hearing on side with contralateral hearing aid (HA). All participants had CI thresholds that were ≤ 40 dB HL (mean= 34.5 dB HL, SD= 5.7) for octave frequencies between 125-8000Hz. For bimodal listeners, mean unaided threshold on contralateral ear (125-4000 Hz) was 83.8 dB HL (range 59.2 to 112.5 dB HL, SD=13.8) and mean aided threshold was 63.8 dB HL (range 37.5 to 81.7 dB HL, SD=13.3).

This study was approved by the Local Ethical Committee and subjects' consent was given freely.

4.2.2 Procedure

Fitting

CI programs for individual recipients were controlled prior to testing. Since existing HA programs were to be used during testing, all bimodal listeners were asked to visit their HA providers shortly before their appointment in our center. Furthermore, following a regular CI fitting session, most comfortable levels were verified in live-speech when listening together with contralateral HA in order to avoid any discomfort due to a loudness summation effect.

Hearing thresholds

Unaided pure tone thresholds were recorded via an Aurical audiometer and TDH39 headphones in a standard sound-proofed booth at octave frequencies between 125-8000 Hz using a warble tone as were aided thresholds in Free Field through a loudspeaker placed at 0º azimuth at 1m distance from the subject's head.

HI/DI performance

A§E - HI/DI tests were used separately in each ear to evaluate LF pitch perception skills that are linked to TFS processing in CI users whilst performances were measured additionally in the CI plus HA listening condition for bimodal users. CI only condition in bimodal listeners was performed by occluding the contralateral ear with an ear foam plug plus a circumaural headphone to avoid a potential contribution of the non-implanted ear [73].

HI/DI tests were based on a discrimination task between two consecutive complex tones: one intonating and one non-intonating. The non-intonating stimulus represented a harmonic complex signal of a F0 at 200 Hz and its 3 higher harmonics presented at levels lower than F0 (-6 dB at 2F0, -12 dB at 3F0 and -18 dB at 4F0). In both tests, the non-intonating sound was contrasted by an intonating sound. In the HI test the intonating sound was characterized with a frequency sweep of F0 together with all harmonics [from NF0 to N(F0+ΔF), N ranging from 1 to 4] whereas in the DI test a sweep of only F0 [F0 to F0+ΔF] was used resulting in a disharmonic intonation. For both tests, the sweep, that was linear, lasted 120ms and was introduced at 330ms after the start of the signal. Total duration of each stimulus was 600ms and the two consecutive stimuli were separated with a 500ms inter-stimulus interval. White noise was added to the stimuli (SNR +10.9 dB) so that they sounded more natural and intensity roving (±2 dB) was applied in order to avoid the use of loudness cues [82].

HI/DI test orders were counterbalanced across subjects in order to minimize the learning effects. Two consecutive stimuli were presented to subjects at 70 dB SPL who were asked to discriminate between same or different. Testing was carried out after training for both the stimuli and the test task.

For each subject, the Just Noticeable Difference (JND) - the smallest ΔF that the subject could discriminate - was calculated by the software using an adaptive staircase procedure where ΔF (41 Hz at test start, range 0 to 214 Hz, 0 Hz means no change between two signals) was increased for an incorrect response and decreased for a correct one until reaching an estimate of the 50%-point on a subject’s psychometric curve. If it could not be found within 100 trials, JND was set to 220 Hz which was above the maximum ΔF value [76,82]. This adaptive staircase procedure is described in detail in a study by Vaerenberg et al. [93].

Speech perception

Similarly, speech perception performance was evaluated ear by ear for all participants and additionally in bimodal listening condition for contralateral HA users. The stimulus was presented via a computer and a preamplifier connected directly to a single loudspeaker. Both sentence material and noise were presented from a loudspeaker at 0° azimuth at 1m from participant’s head.

The test battery consisted of standard Italian phonetically balanced bi-syllabic words for an adult population [92]. Speech recognition in quiet (Word Recognition Score- WRS) was tested at 65 dB SPL whilst the performance in noise was evaluated presenting words at +10 and +5 dB SNR ((SNR+10 and SNR+5) with speech-shaped noise at 65 dB SPL.

Italian STARR

The Italian STARR test made use of sentences from the standard Italian test battery [92]. The material consisted of 10 test lists each containing 15 sentences, all recorded with male voice. For competition CCIT noise was used resembling long-term spectrum of the speech test material. Three presentation levels (50, 65 and 80 dB SPL) were used for sentences and there were 5 presentations at each of these levels within each test list. The details for Italian STARR test design are described in a previous publication [121].

The test was started after explaining the task to participants and warning them about the possibility of sometimes hearing quite loud stimuli. The upper sound pressure limit was set to 91 dB SPL resulting in a -10 dB SNR for the highest sentence presentation level of 80 dB SPL. Beyond this limit, the program automatically decreased both the sentence and noise levels, maintaining the required SNR while avoiding uncomfortable stimuli. After presentation of a sentence, participants were asked to repeat it as accurately as possible, while explaining that not every word needed to be correct and encouraging them to guess when not sure. For scoring, and for the benefit of the operator, single key words were highlighted in white on the screen becoming green to indicate words repeated correctly.

The clinician could either click on single key words or, alternatively, select the “all” or “none” options to indicate participant’s response. The SNR was +20 dB initially and varied adaptively according to the participant’s response. Where the minimum specified number of key words was correctly identified, typically 2 out of 3, the sentence was considered correct and the SNR used for the next sentence was made more adverse. Where insufficient key words were correctly identified, the sentence was considered incorrect and the SNR for the next sentence to be presented would become more favourable. The initial step size was 10 dB, dropping to 5 dB after the first reversal of the adaptive track and dropping again to the final step size of 2.5 dB after a further reversal. The SNR was varied by adjusting the noise level while keeping the speech level at 50, 65 and 80 dB SPL and using the same SNR for all three levels. The SRT was computed by averaging the SNRs for the last nine sentences, along with the SNR at which a next sentence would have been presented. Clicking on the OK button finalized scoring for individual sentences and the next sentence was presented by the software until the end of the test list where the SRT value was displayed automatically. All participants were tested using two test lists following a practice list which was administered to minimize any learning effects. All the tests were administered on the same day.

4.2.3 Statistical Analysis

Data analysis was carried out using the Statistical Package for Social Sciences (SPSS) version 19.0 (Chicago, IL, USA). Non-parametric statistical tests were used since one-sample Kolmogorov-Smirnov tests revealed that outcomes were not normally distributed ($p<0.001$). Differences between tests were investigated by Wilcoxon Signed Ranks test as well as differences between CI only and bimodal listening conditions within subjects. Percentage of outcomes that fell within the normal clinical zone (JNDs $\leq$ 4 Hz for HI and $\leq$ 10 Hz for DI as per normative data collected by Vaerenberg et al. in adult population) was computed in order to compare outcomes to those of people with normal hearing [82]. For STARR test, the scores obtained from two lists were averaged for each ear/condition and percentage of meaningful STARR score was calculated considering scores that were less than 20 dB SNR. Spearman rank-order correlations were used to investigate the correlations between HI/DI tests and speech perception outcomes as well as to evaluate the effect of variables such as age, duration of CI experience, duration of deafness, unaided/aided PTA thresholds (PTA <1000 Hz, PTA≥1000 Hz, PTA ≤8000 Hz with CI and ≤4000 Hz with HA). The cut-off level for statistical significance was set to 0.05.

4.3 Results

4.3.1 HI/DI performance

The analysis regarded outcomes of unilateral/bilateral CI users (based on individual ears) plus CI only condition in bimodal listeners (N=49). Pitch perception results between HI and DI tests indicated statistically significant differences ($Z= -5.5$, $p<0.001$): outcomes were better for HI than for DI [median JND= 27.0 Hz (range 1.0 to 220.0 Hz) and 147.0 Hz (range 7.0 to 220.0 Hz) for HI and DI respectively]. 12% of HI outcomes were within the normal zone versus 8% of DI outcomes.

The correlations between HI and DI outcomes were statistically significant ($r_s=0.68$, $p<0.001$). There was a significant positive correlation between age and JND score for both HI and DI ($r_s=0.51$ and 0.40 respectively, $p<0.005$). Other factors such as the duration of CI experience and the duration of deafness were not significantly correlated with pitch perception outcomes ($p>0.05$). Figure 14 shows individual HI/DI outcomes relative to age of CI recipients.

Bimodal users showed better HI and DI outcomes in bimodal listening than CI only condition [median HI JND=39.5 Hz (range 1.0 to 220.0 Hz) versus 6.5 Hz (range 3.0 to 133.0 Hz) and median DI JND=133.5 Hz (range 9.0 to 220.0 Hz) versus 38.5 Hz (range 7.0 to 220.0 Hz) in CI only and bimodal listening conditions respectively, N=14]. Differences were statistically significant for both HI ($Z = -2.6$, $p\leq0.001$) and DI ($Z = -2.9$, $p<0.005$). DI outcomes in bimodal listening condition showed a significant correlation with unaided PTA thresholds for octave frequencies lower than 1000 Hz ($r_s=0.68$, $p=0.008$). HI/DI outcomes at CI only in comparison to bimodal listening condition are given in Figure 15.

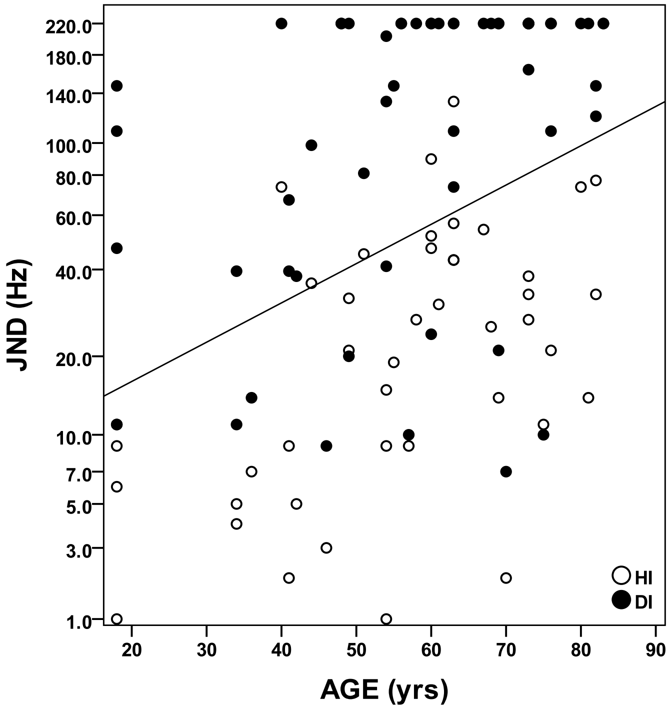

Figure 14 Individual HI/DI outcomes relative to age of CI recipients.

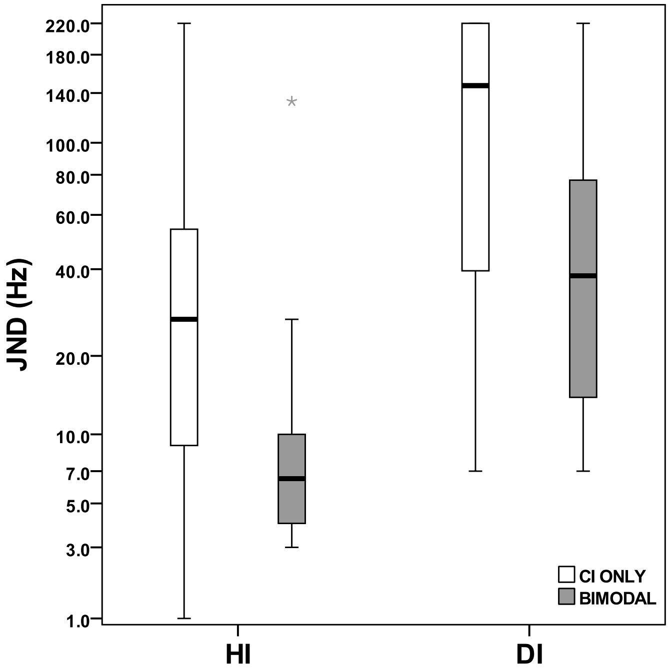

Figure 15 HI/DI outcomes at CI only in comparison to bimodal listening.

4.3.2 Italian STARR outcomes in relation to other speech perception tests

All participants were able to complete the test. Median STARR outcome in CI only listeners was 14.8 dB SNR (range -1.8 to 125.0 dB, N=49). 67% of the group had a STARR score that was less than 20 dB SNR. In this subgroup, median STARR score was 8.6 dB SNR (range -1.8 to 19.5, N=33).

Analysis showed statistically significant negative correlations between STARR outcomes and other speech recognition measures ($p<0.001$). Table 3 illustrates outcomes for individual tests as well as Spearman's rho ($r_{s)}$ and p values regarding the correlations with STARR.

	Median	**Range**	**r_s**	**p**
WRS in quiet	82 %	22 to 100 %	**-0.65**	**0.000**
SNR+10	35 %	0 to 83 %	**-0.52**	**0.000**
SNR+5	12 %	0 to 75 %	**-0.61**	**0.000**

Table 3 Outcomes for individual speech perception tests as well as Spearman's rho and p values relative to Italian STARR test.

The duration of deafness and CI thresholds had a significant effect on the STARR performance ($r_s=0.36$ and $r_s=0.30$ respectively, $p<0.05$) whereas other factors such as age and CI experience were not significantly correlated ($p>0.05$).

Bimodal users showed better speech recognition outcomes in bimodal listening than CI only condition [median WRS in quiet= 76% versus 85%, median SNR+10=35 versus 40%, median SNR+5= 5 versus 22% and median STARR=17.3 versus 8.1 dB SNR in CI only and bimodal listening conditions respectively, N=14]. Results are shown in Figure 16.

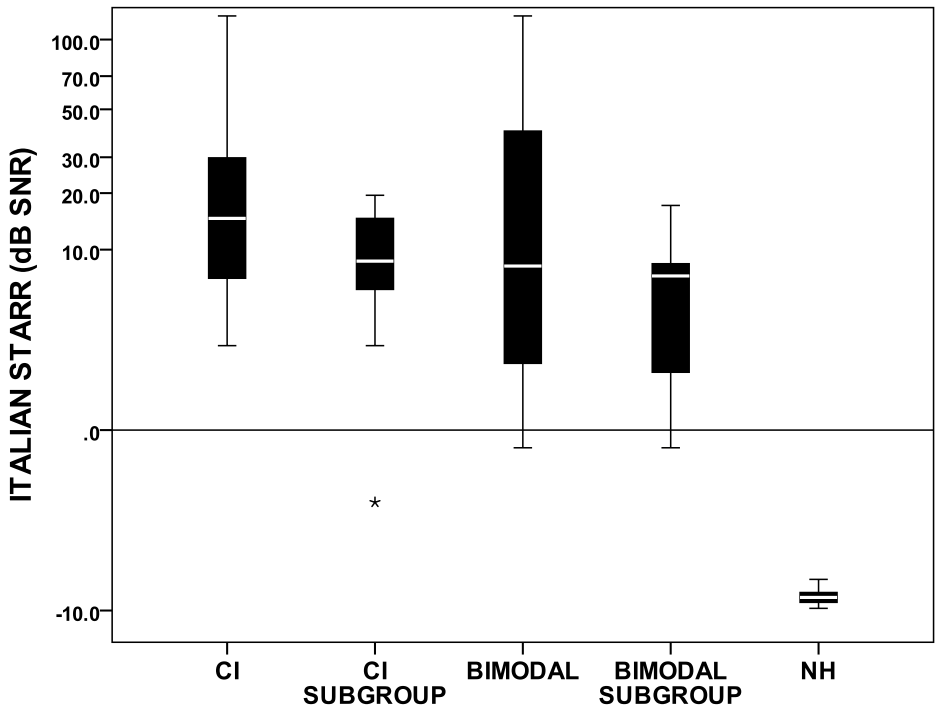

Figure 16 Italian STARR outcomes. Data for normal hearing (NH) people is received from Dincer D'Alessandro et al. [121].

Differences were statistically significant for speech perception in noise whilst they were not significant for speech understanding in quiet (Z= -2.9, p<0.005 in STARR). The details of analysis are given in Table 4. When listening with CI only, 9 out of 14 listeners (65%) had a STARR score that was less than 20 dB SNR instead of 10 (71%) when listening bimodally. 7 out of these 10 subjects had a meaningful amount of bimodal benefit [greater than 2.4 dB SNR improvement as indicated by Dincer D'Alessandro et. al. [121] when using two test lists per condition]. On the other hand, one subject had an improvement of 2.2 dB SNR whilst the other two had a deterioration of 1.5 and 3.0 dB SNR when listening bimodally. STARR outcomes did not show any significant correlations with unaided nor aided PTA thresholds for HA ear (p>0.05).

	CI only listening		**Bimodal Listening**		**Differences**	
	Median	**Range**	**Median**	**Range**	**z**	**p**
HI (Hz)	39.5	1.0 to 220.0	6.5	3.0 to 133.0	**-2.6**	**0.001**
DI (Hz)	133.5	9.0 to 220.0	38.5	7.0 to 220.0	**-2.9**	**0.003**
STARR (dB SNR)	17.3	-1.8 to 125.0	8.1	-0.3 to 125.0	**-2.9**	**0.004**
WRS (%)	76	30 to 100	85	50 to 100	-1.5	0.130
SNR+10 (%)	35	0 to 60	40	0 to 98	**-2.0**	**0.050**
SNR+5 (%)	5	0 to 40	22	0 to 75	**-2.0**	**0.050**

Table 4 Within-subject comparisons for CI only versus bimodal listening conditions. Statistically significant differences are in bold.

4.3.3 HI/DI in relation to STARR and other speech perception tests

For all participants including CI plus HA listening condition in bimodal users (N=63), analysis of correlations between HI/DI outcomes and Italian STARR showed statistically significant positive correlations (r_s=0.40, $p<0.005$ for HI and r_s=0.30, $p<0.05$ for DI). Figure 17 shows the correlations between HI/DI and Italian STARR outcomes.

On the other hand, HI was significantly correlated with WRS in quiet (r_s= -0.46, $p<0.001$), with WRS+10 (r_s= -0.38, $p<0.005$) whilst DI was not significantly correlated with other speech perception tests. The details are given in Table 5.

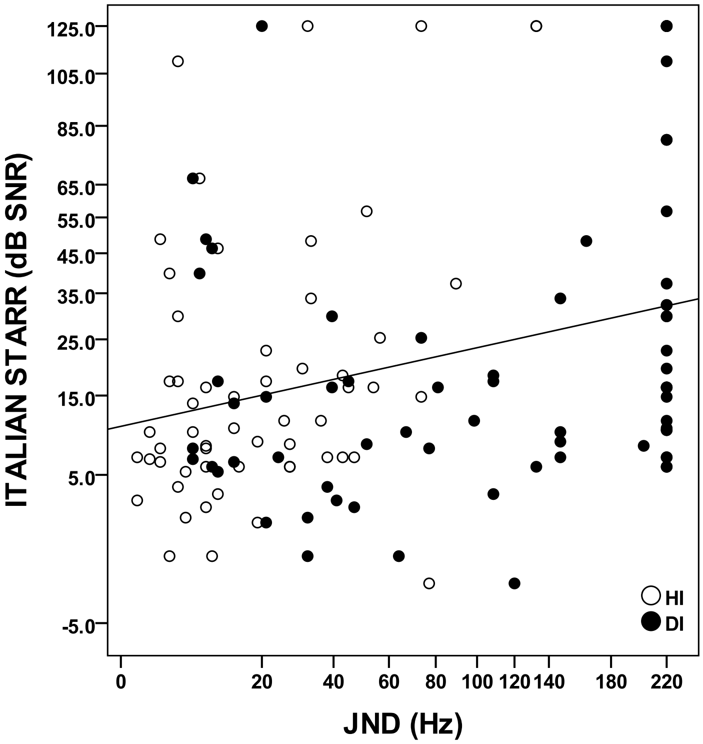

Figure 17 The correlations between HI/DI and Italian STARR outcomes.

	HI		DI	
	r_s	p	r_s	p
ITALIAN STARR	**0.40**	**0.002**	**0.30**	**0.036**
WRS	**-0.46**	**0.000**	-0.22	0.077
SNR+10	**-0.38**	**0.003**	-0.18	0.186
SNR+5	-0.26	0.067	-0.16	0.239

Table 5 HI/DI outcomes in relation to Italian STARR and other speech perception tests. Statistically significant correlations are in bold.

4.4 Discussion

One of the primary objectives of the present study was to investigate TFS processing that is assumed to be reflected in the LF pitch perception capacities for a larger CI recipient group. For this purpose, A§E Harmonic Intonation and Disharmonic Intonation tests were used. Findings were similar to outcomes studied previously on a smaller group of adult CI recipients [76] where median JNDs were found to be 16.0 and 139.0 Hz for HI and DI respectively. The corresponding values were 27.0 Hz and 147.0 Hz in the present study. Although CI recipients had significantly better scores for HI than for DI, the majority showed abnormal outcomes for both HI and DI tests. Significant correlations between HI/DI outcomes confirmed that both tests targeted LF processing capacities. However, relative performance deterioration for the DI test was re-indicating DI with more differential outcomes on phase locking and TFS processing capacities whereas the HI test, due to the sweep of F0 together with its harmonics, could provide some HF cues in the complex signal and therefore, additional place cues for lower JNDs [82,86]. Furthermore, there was a strong effect of age on both HI/DI outcomes indicating that TFS sensitivity worsened with increasing age. Indeed, it is well-known that aging is accompanied by changes in physiological, psychophysical and psychological domains. As a result of increase in longevity, the effects of aging on the auditory system have received much attention recently [122,123]. Previous studies on TFS sensitivity were carried out in NH populations and results indicated a significant effect of aging both for binaural, and monaural listening tasks like in our study. However, more studies are required to give a precise definition of the role of changes in peripheral sensory processing or in the central auditory system or in cognition and attention associated with aging [83,124].

Another objective was to evaluate the Italian STARR outcomes in adult CI recipients as well as to study the correlations with other speech perception tests that are commonly used in Italian. The STARR test originally existed in British-English and was based on both male and female speakers instead of a male speaker only in the Italian version. Present findings indicated that median Italian STARR score (14.8 dB SNR) was lower than the British-English version (22 dB, 34 dB and 28 dB SNR for the male, female speakers and their overall mean respectively) studied by Boyle et al. [50]. However, this difference certainly was due to the difference in the measure of central tendency (in the present study, median value had to be used since the data was not normally distributed).

On the other hand, Boyle et al. [50] reported that for SRTs lower than 20 dB SNR, performance during the adaptive track was related in a more orderly way to the SNR whereas higher SNRs did not materially change the performance and the score was determined basically by the ability to understand speech in quiet rather than by the SNR. Therefore, a subgroup of participants with SRTs better than 20 dB was further analysed to investigate performances that were obtained in a meaningful way from the test. Although all participants were able to complete the test, 67% of the study group had an SRT that was considered meaningful. However, it should be considered that 37% of the overall group had a duration of profound deafness longer than 10yrs and its significant effect on STARR performance was among the remarkable findings of the present study. In the subgroup, the median STARR score was 8.6 dB SNR. In the study of Boyle et al. [50] for British-English population, 88% of the CI users achieved to complete the test whilst the subgroup consisted of only 40% of the overall group. The mean SRT for the male speaker was a bit lower (5.9 dB) in their case whilst the overall mean for both speakers (9.4 dB) was very similar to Italian STARR outcomes. Such differences in multilingual speech perception tests are to be expected and can be attributed to language as well as speaker dependent factors [111]. On the other hand, both the overall success rate and the percentage of subgroup were smaller in the Boyle et al. study [50]. But their group concerned CI recipients who performed more poorly than typical and the duration of profound deafness before implantation was even longer. Moreover, their subjects consisted of CI users with older-generation technology and their processors were not optimally adjusted. On the contrary, in another similar study, Haumann et al. [99] tested 55 German-speaking CI users using roving levels and all their subjects were able to achieve meaningful SRTs. Nevertheless, all these studies including the present one showed abnormal SRTs in CI recipients confirming their difficulty for listening in the presence of noise. Even the better performers (with SRTs lower than 20 dB) showed a vast difference in comparison to NH people (17 dB poorer than NH group in Italian STARR). Furthermore, the significant effect of hearing thresholds on Italian STARR performance, even in a group that achieved typical target CI thresholds ≤ 40 dB HL, was re-emphasizing the sensitivity of the test to lower level speech which a CI user can face very often during everyday life.

A further objective of this study was to investigate outcomes in contralateral HA users. The previous results by Vaerenberg et al. [86] and Dincer D'Alessandro et al. [120] indicated a bimodal benefit only in DI scores whereas HI scores tended not to change between CI only and bimodal listening conditions. However, present findings showed considerable bimodal benefit for both HI and DI outcomes which could be explained by the small number of bimodal listeners in all these studies; nevertheless, the present number is bigger and it is reasonable to expect bimodal benefit for the kind of tasks in both tests. On the other hand, bimodal benefit was remarkable in Italian STARR performance as well; bimodal users showed better STARR outcomes in bimodal listening than CI only condition [median=17.3 versus 8.1 dB SNR in CI only and bimodal listening conditions respectively]. Differences were statistically significant. When listening with CI only, 9 out of 14 listeners (65%) had a STARR score that was less than 20 dB SNR instead of 10 (71%) when listening bimodally. 7 out of these 10 subjects had a meaningful amount of bimodal benefit (greater than 2.4 dB SNR improvement as indicated by Dincer D'Alessandro et al. [121] when using two test lists per condition). Similarly, bimodal benefit had a significant effect on word recognition in noise whereas scores for listening in quiet did not show statistically significant differences. However, overall outcomes in bimodal listeners were still worse than those of NH counterparts. Although speech perception outcomes including those for STARR did not show any significant correlations with unaided or aided PTA thresholds for the HA ear, DI outcomes in bimodal listening condition showed a strong effect of unaided PTA thresholds for octave frequencies lower than 1000 Hz suggesting that bimodal listeners do benefit from the additional LF cues and by remaining phase-locking capacities of residual hearing and in that way the representation of TFS for CI users was improved [79].

Finally, this study aimed to analyze the outcome comparisons between pitch and speech perception. The idea was to get a better understanding of the link between LF pitch perception and the processing of TFS in CI users as well as implications for speech perception performance in particular when using a test that attempts to better represent everyday listening situations. It was previously shown that speech recognition scores in quiet were not significantly correlated with HI/DI outcomes in pediatric population [120]. Furthermore, as mentioned before, outcomes in EAS processor users revealed significant differences between electric only and EAS listening modes in speech in noise results as well as a remarkable improvement in DI performance [86]. Although the strength was moderate, present findings showed that HI was significantly correlated with WRS in quiet, with SNR+10 and with Italian STARR outcomes, whilst DI was significantly correlated only with Italian STARR scores indicating again DI to provide more differential outcomes on phase locking and TFS processing capacities since TFS cues are emphasized in difficult listening tasks such as listening in the dips which is defined as detecting a signal in a fluctuating background [81].

The small sample size for bimodal and bilateral listeners was an important limitation of the present study. Bilateral CI users were tested one ear at a time (monolateral listening only) in order to avoid longer testing during the same day but especially due to the limited number of test lists that were available for Italian STARR. Actually, it would be interesting to study outcome comparisons in relation to bilateral benefit as well. Moreover, the group size did not allow us to study any possible impact of the device type, in particular the sound coding strategy, on the performance. Furthermore, it would be useful to study clinical usefulness of HI/DI tests as a predictor of music perception and appreciation in CI users since the availability of LF pitch and TFS cues are even more dominant for music perception [125].

CHAPTER 5. CONCLUSION

CI recipients usually complain about their difficulties for speech understanding in noisy environments. One important aspect that contributes to this fact is recognized as poor TFS processing of cochlear implants. This thesis aimed to investigate the following three topics. First, the LF pitch perception skills that are believed to be linked to TFS processing, were studied in a group of pediatric and adult CI recipients. Secondly, the STARR test which attempts to mimic challenging real life listening conditions was adapted in order to introduce the test into Italian. Finally, the interactions between TFS processing, pitch and speech perception outcomes as well as the effects of demographic factors and the amount of bimodal benefit were investigated.

Findings suggested the following:

- HI/DI outcomes evaluating LF pitch perception were found to be abnormal in the majority of both pediatric and adult CI recipients, confirming poor TFS processing capacities of cochlear implants. However, DI performance was considerably better in children than in adult CI recipients.

- HI/DI outcomes showed a significant positive correlation: subjects with higher JNDs on HI tended to have higher JNDs on DI test as well. However, performance was significantly worse for the DI test. This finding indicated the DI test as providing more differential LF pitch perception outcomes in that it reflected phase locking and TFS processing capacities of the ear, whereas HI test provided information of its place coding capacity as well.

- HI/DI tests were clinically applicable in children of 5 years and older. Chronological age had a significant effect on DI performance. NH children under the age of 8.5 years showed larger inter-subject-variability; however, the majority of them showed outcomes that were considered normal at adult-level.

- Similarly, age in adult CI recipients, showed a strong effect on both HI/DI outcomes indicating that TFS sensitivity worsened with increasing age.

- Contralateral HA users had remarkable bimodal benefit on both HI/DI tests. Moreover, DI outcomes in bimodal listening condition showed a strong effect of unaided PTA thresholds for octave frequencies lower than 1000 Hz suggesting bimodal listeners to benefit from the additional LF cues and by remaining phase-locking capacities of residual hearing. However, the outcomes in bimodal listeners were still worse than those of NH counterparts.

- Findings in NH listeners suggested that the Italian STARR test could be a promising supplement to existing speech assessment tools. The average SRT for NH and CI recipient subjects was consistent with SRTs reported for sentence testing by other researchers. The variability of mean SRTs across test lists was relatively small. Statistical analysis showed no significant learning effects. The outcomes for Italian STARR test showed statistically significant correlations with those for standard word recognition test in Italian.

- STARR outcomes showed abnormal SRTs in CI recipients confirming their difficulty of listening in the presence of noise. Even the better performers had a vast difference in comparison to NH people.

- The success rate for Italian STARR test was excellent; all subjects managed to complete the test. But only 67% of the study group had an SRT that was considered meaningful.

- The duration of profound deafness had a significant effect on STARR performance.

- The significant effect of CI thresholds on Italian STARR performance re-emphasized the sensitivity of the test to lower level speech which a CI user can face very often during everyday life.

- Similarly with LF pitch perception outcomes, bimodal users had a significant benefit on speech perception in noise whereas scores for listening in quiet did not show statistically significant differences. However, outcomes in bimodal listeners were still worse than those of NH counterparts.

- HI was significantly correlated with WRS in quiet, with SNR+10 and Italian STARR outcomes whilst DI was significantly correlated only with Italian STARR scores indicating again DI to provide more differential outcomes on phase locking and TFS processing capacities since TFS cues are emphasized in difficult listening tasks such as listening in the dips which is defined as detecting a signal in a fluctuating background.

REFERENCES

[1] Flynn S (2003) Cochlear implant systems. In McCormick B & Archbold S (Eds) Cochlear implants for young children. London: Whurr Publishers 11-48.

[2] Allum DJ (1996) Basics of cochlear implant systems. In Allum DJ (Ed) Cochlear implant rehabilitation in children and adults. London: Whurr Publishers 1-21.

[3] Gelfand SA (2001) Essentials of Audiology second edition. Thieme, New York, USA.

[4] Davis H (1962) Advances in the neurophysiology and neuroanatomy of the cochlea. J Acoust Soc Am 34:1377-1385.

[5] Wilson BS, Dorman MF (2008) Cochlear implants: current designs and future possibilities. J Rehabil Res Dev 45:695-730.

[6] Clark JG (1981) Uses and abuses of hearing loss classification. Asha 23:493.

[7] Stewart L, Adams D (1997) Deafness: Its Implications. In McCracken W., Laoide-Kemp S. (Eds) Audiology in Education. London: Whurr Publishers 79-106.

[8] Hinojosa R, Marion M (1983) Histopathology of profound sensorineural deafness. Ann N Y Acad Sci 405:459–84.

[9] Niparko JK, Iler Kirk K, Mellon NK, Robbins AM, Tucci DL, Wilson BS (2000) Cochlear Implants Principles and Practices. Lippincott Williams & Wilkins, Philadelphia, USA.

[10] Niparko JK, Wilson BS (2000) History of Cochlear Implants. In Niparko JK, Iler Kirk K, Mellon NK, Robbins AM, Tucci DL, Wilson BS (Eds). Cochlear Implants Principles and Practices. Lippincott Williams & Wilkins, Philadelphia, 103-107.

[11] Djourno A, Eyres C (1957) Prosthèse auditive par excitation électrique à distance du nerf sensoriel àl'aid d'un bobinage inclus á demeure. [Auditory prosthesis by means of a distant electrical stimulation of the sensory nerve with the use of an indwelling coil.] Presse Medicale 35:14-17.

[12] House WF, Berliner K, Crary W, Graham M, Luckey R, Norton N, Selters W, Tobin H, Urban J, Wexler M (1976) Cochlear Implants. Ann Otol Rhinol Laryngol Supp 85:1-93.

[13] House WF, Berliner KI (1986) Safety and efficacy of the House/3M cochlear implant in profoundly deaf adults. Otolaryngol Clin North Am 19:275-286.

[14] House WF, Berliner KI (1991) Cochlear implants: from idea to clinical practice. In Cooper H (Ed.) Cochlear Implants: A Practical Guide. London: Whurr 9-33.

[15] Vaerenberg B (2014) Programming Cochlear Implants for Auditory Performance. PhD Dissertation. The University of Antwerp, Antwerp, Belgium.

[16] Hochmair-Desoyer I, Hochmair E (1996) Implementation of coding strategies in cochlear implant systems and comparative results. Central East Eur J Otorhinolar Head Neck 1:33-41.

[17] Gantz BJ, Tyler RS, Knutson JF, et al. (1988) Evaluation of five different cochlear implant designs: audiologic assessment and predictors of performance. Laryngoscope 10:1100-1106.

[18] Cohen NL, Waltzman SB, Fisher SG (1993) the VAH Cochlear Implant Cooperative Study Group: Prospective, randomized study of cochlear implants. The Department of Veterans Affairs Cochlear Implant Study Group. N Engl J Med 328:233-237.

[19] Doyle PJ, Pijl S, Noel FJ (1991) The cochlear implant: a comparison of single and multi-channel results. J Otolaryngol 20:204-208.

[20] Eyles JE, Aleksy WL, Boyle PJ (1995) Performance changes in University College Hospital/ Royal National Institute for the Deaf single-channel cochlear implant users upgraded to the Nucleus 22 channel cochlear implant system. Ann Otol Rhinol Laryngol Supp 104:263-265.

[21] Shepherd RK, Franz B, Clark GM (1990) The biocompatibility and safety of cochlear prostheses. In Clark GM, Tong YT, Patrick JF (Eds) Cochlear Prosthesis. Melbourne: Churchill Livingstone 69-98.

[22] Luxford WM, Brackman DE (1985) The history of cochlear implants. In Gray R (Ed) Cochlear Implants. San Diego: CA: College Hill Press 1-26.

[23] Baumann U, Nobbe A (2004) Pulse rate discrimination with deeply inserted electrode arrays. Hear Res 196:49–57.

[24] Wilson BS (2000) Cochlear Implant Technology. In Niparko JK, Iler Kirk K, Mellon NK, Robbins AM, Tucci DL, Wilson BS (Eds). Cochlear Implants Principles and Practices. Lippincott Williams & Wilkins, Philadelphia 103-107.

[25] Wilson BS (1997) The future of cochlear implants. Br J Audiol 31:205–25.

[26] Kiefer J, Von Ilberg C, Rupprecht V, Hubner-Egener J, Knecht R (2000) Optimized speech understanding with the continuous interleaved sampling speech coding strategy in patients with cochlear implants: Effect of variations in stimulation rate and number of channels. Ann Otol Rhinol Laryngol 109:1009–20.

[27] Friesen LM, Shannon RV, Baskent D, Wang X (2001) Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. J Acoust Soc Am 110:1150–63.

[28] Cohen LT, Saunders E, Knight MR, Cowan RS (2006) Psychophysical measures in patients fitted with Contour and straight Nucleus electrode arrays. Hear Res 212:160–75.

[29] National Institutes of Health Consensus Conference (1995) Cochlear implants in adults and children. JAMA 274:1955–61.

[30] Wilson BS, Schatzer R, Lopez-Poveda EA, Sun X, Lawson DT, Wolford RD (2005) Two new directions in speech processor design for cochlear implants. Ear Hear 26:73S–81S.

[31] Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM (1991) Better speech recognition with cochlear implants. Nature 352:236–238.

[32] Smith ZM, Delgutte B, Oxenham AJ (2002) Chimaeric sounds reveal dichotomies in auditory perception. Nature 416:87–90.

[33] Donaldson GS, Kreft HA, Litvak L (2005) Place-pitch discrimination of single- versus dual-electrode stimuli by cochlear implant users (L). J Acoust Soc Am 118:623–26.

[34] Firszt JB, Koch DB, Downing M, Litvak L (2007) Current steering creates additional pitch percepts in adult cochlear implant recipients. Otol Neurotol 28:629–36.

[35] Koch DB, Downing M, Osberger MJ, Litvak L (2007) Using current steering to increase spectral resolution in CII and HiRes 90K users. Ear Hear 28:39S–41S.

[36] Wilson BS, Schatzer R, Lopez-Poveda EA (2006) Possibilities for a closer mimicking of normal auditory functions with cochlear implants. In: Waltzman SB, Roland JT Jr, editors. Cochlear implants. 2nd ed. New York (NY): Thieme 48–56.

[37] McDermott HJ, McKay CM (1994) Pitch ranking with nonsimultaneous dual electrode electrical stimulation of the cochlea. J Acoust Soc Am 96:155–62.

[38] Kwon BJ, Van den Honert C (2006) Dual-electrode pitch discrimination with sequential interleaved stimulation by cochlear implant users. J Acoust Soc Am 120:EL1–6.

[39] Nobbe A, Schleich P, Zierhofer C, Nopp P (2007) Frequency discrimination with sequential or simultaneous stimulation in MED-EL cochlear implants. Acta Otolaryngol 127:1266–72.

[40] Hochmair I, Nopp P, Jolly C, Schmidt M, Schösser H, Garnham C, Anderson I (2006) MED-EL Cochlear implants: State of the art and a glimpse into the future. Trends Amplif 10:201–19.

[41] Arnoldner C, Riss D, Brunner M, Durisin M, Baumgartner WD, Hamzavi JS (2007) Speech and music perception with the new fine structure speech coding strategy: Preliminary results. Acta Otolaryngol 127:1298–1303.

[42] Green T, Faulkner A, Rosen S (2004) Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. J Acoust Soc Am 116:2298–2310.

[43] Riss D, Hamzavi JS, Blineder M, Flak S, Baumgartner WD, Kaider A, Arnoldner C (2016) Effects of Stimulation Rate With the FS4 and HDCIS Coding Strategies in Cochlear Implant Recipients. Otol Neurotol 37:882-888.

[44] Wouters J, McDermott HJ, Francart T (2015) Sound coding in cochlear implants: from electric pulses to hearing. IEEE Signal Process Mag 32:67-80.

[45] Lederberg AR, Schick B, Spencer PE (2013) Language and literacy development of deaf and hard-of-hearing children: successes and challenges. Dev Psychol 49:15-30.

[46] Nicastri M, Filipo R, Ruoppolo G et al. (2014) Inferences and metaphoric comprehension in unilaterally implanted children with adequate formal oral language performance. Int J Pediatr Otorhinolaryngol 78:821-827.

[47] Mancini P, Dincer D’Alessandro H, Guerzoni L et al. (2015) Adequate formal language performance in unilateral cochlear implanted children: Is it indicative of complete recovery in all linguistic domains? Insights from referential communication. Int J Pediatr Otorhinolaryngol 79:598-604.

[48] Eisenberg LS, Fisher LM, Johnson KC, Ganguly DH, Grace T, Niparko JK; CDaCI Investigative Team (2016) Sentence Recognition in Quiet and Noise by Pediatric Cochlear Implant Users: Relationships to Spoken Language. Otol Neurotol 37:e75-81.

[49] Park E, Shipp DB, JM Chen, Nedzelski JM, Lin VY (2011) Postlingually deaf adults of all ages derive equal benefits from unilateral multichannel cochlear implant. J Am Acad Audiol 22:637-643.

[50] Boyle PJ, Nunn TB, O'Connor AF, Moore BCJ (2013) STARR: A speech test for evaluation of the effectiveness of auditory prostheses under realistic conditions. Ear Hear 34:203-212.

[51] Kong Y-Y, Cruz R, Jones A, Zeng FG (2004) Music perception with temporal cues in acoustic and electric hearing. Ear Hear 25:173-185.

[52] Hsiao F, Gfeller K (2012) Music perception of cochlear implant recipients with implications for music instruction: a review of literature. Update Univ S C Dep Music 30:5–10.

[53] Fujita S, Ito J (1999) Ability of Nucleus cochlear implantees to recognize music. Ann Otol Rhinol Laryngol 108:634-640.

[54] Leal MC, Shin YJ, Laborde M-L, Lugardon S, Andrieu S, Deguine O, Fraysse B (2003) Music perception in adult cochlear implant recipients. Acta Otolaryngol 123:826-835.

[55] Mirza S, Douglas S, Lindsey P, Hildreth T, Hawthorne M (2003) Appreciation of music in adult patients with cochlear implants: a patient questionnaire. Coch Impl Intl 4:85-95.

[56] McDermott HJ (2004) Music perception with cochlear implants: a review. Trends Amplif 8:49-82.

[57] Dincer H, Vickers D, Boyle P, Kırkım G, Olgun L, Incesulu A, Yücel E, Önder S, Akın I, Korkmaz H (2009) Development of musical behaviour in children with cochlear implants. In Ozgirgin ON (Ed) Surgery of the Ear: Current Topics**.** Rekmay Publishing, Ankara, 453-501.

[58] Spahr AJ, Dorman MF (2004) Performance of subjects fit with the Advanced Bionics CII and Nucleus 3G cochlear implant devices. Arch Otolaryngol Head Neck Surg 130:624-628.

[59] Spahr AJ, Dorman MF, Loiselle LH (2007) Performance of patients using different cochlear implant systems: effects of input dynamic range. Ear Hear 28:260-275.

[60] Baudhuin J, Cadieux J, Firszt JB, Reeder RM, Maxson JL (2012) Optimization of programming parameters in children with the Advanced Bionics cochlear implant. J Am Acad Audiol 23:02-312.

[61] McCreery RW, Venediktov RA, Coleman JJ, Leech HM (2012) An evidence-based systematic review of directional microphones and digital noise reduction hearing aids in school-age children with hearing loss. Am J Audiol 21:295-312.

[62] Cherry EC (1953) Some experiments upon the recognition of speech, with one and with two ears. J Acoust Soc Am 25:975-979.

[63] Byrne D (1981) Clinical issues and options in binaural hearing aid fitting. Ear Hear 2:187-193.

[64] Veugen LC, Hendrikse MM, van Wanrooij MM, Agterberg MJ, Chalupper J, Mens LH, Snik AF, John van Opstal A (2016) Horizontal sound localization in cochlear implant users with a contralateral hearing aid. Hear Res 336:72-82.

[65] Firszt JB, Reeder RM, Skinner MW (2008) Restoring hearing symmetry with two cochlear implants or one cochlear implant and a contralateral hearing aid. J Rehabil Res Dev 45:749-768.

[66] Kokkinakis K, Pak N (2014) Binaural advantages in users of bimodal and bilateral cochlear implant devices. J Acoust Soc Am 135:EL47-53.

[67] Schleich P, Nopp P, and D'Haese P (2004) Head shadow, squelch and summation effects in bilateral users of the MED-EL COMBI 40/40+ cochlear implant. Ear Hear 25:197-204.

[68] Livotsky R, Parkinson A, Arcaroli J, Sammeth C (2006) Simultaneous bilateral cochlear implantation in adults: A multicenter clinical study. Ear Hear 27:714-731.

[69] Buss E, Pillsbury HC, Buchman CA, Pillsbury CH et al. (2008) Multicenter US bilateral MED-EL cochlear implantation study: Speech perception over the first year of use. Ear Hear 29:20-32.

[70] Sharma A, Dorman MF, Kral A (2005) The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. Hear Res 203:134-43.

[71] Mok M, Grayden D, Dowell RC et al. (2006) Speech perception for adults who use hearing aids in conjunction with cochlear implants in opposite ears. J Speech Lang Hear Res 49:338-351.

[72] Ching TYC, Incerti P, Hill M (2004) Binaural benefits for adults who use hearing aids and cochlear implants in opposite ears. Ear Hear 25:9-21.

[73] Dorman MF, Gifford RH, Spahr AJ, McKarns SA (2008) The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies. Audiol Neurootol 13:105-12.

[74] Van Hoesel RJM (2012) Contrasting benefits from cochlear implants and hearing aids in cochlear implant users. Hear Res 288:100-113.

[75] Mok M, Galvin KL, Dowell RC et al. (2010) Speech perception benefit for children with a cochlear implant and a hearing aid in opposite ears and children with bilateral cochlear implants. Audiol Neurotol 15:44-56.

[76] Schauwers K, Coene M, Heeren W, Del Bo L, Pascu A, Vaerenberg B, Govaerts PJ (2012) Perception of pitch changes in hearing impaired adults with aided and unaided hearing loss. J Hear Sci 2:OA25-34.

[77] Dincer D'Alessandro H, Sennaroğlu G, Yücel E, Belgin E, Mancini P (2015) Binaural squelch and head shadow effects in children with unilateral cochlear implants and contralateral hearing aids. Acta Otorhinolaryngol Ital 35:343-349.

[78] Looi V, Radford CJ (2011) A comparison of the speech recognition and pitch ranking abilities of children using a unilateral cochlear implant, bimodal stimulation or bilateral hearing aids. Int J Pediatr Otorhinolaryngol 75:472-482.

[79] Cullington HE, Zeng FG (2011) Comparison of bimodal and bilateral cochlear implant users on speech recognition with competing talker, music perception, affective prosody discrimination, and talker identification. Ear Hear 32:16-30.

[80] McDermott HJ, Oxenham AJ (2008) Music perception, pitch and the auditory system. Curr Opin Neurobiol 18:452-463.

[81] Moore BJC (2008) The role of temporal fine structure processing in pitch perception, masking and speech perception for normal-hearing and hearing-impaired people. J Assoc Res Otolaryngol 9:399-406.

[82] Vaerenberg B, Pascu A, Del Bo L, Schauwers K, De Ceulaer G, Daemers K, Coene M, Govaerts PJ (2011) Clinical assessment of pitch perception. Otol Neurotol 32:736-741.

[83] Moore BJC, Vickers DA, Mehta A (2012) The effects of age on temporal fine structure sensitivity in monaural and binaural conditions. Int J Audiol 51:715-721.

[84] Lorenzi C, Gilbert G, Carn H, Garnier S, Moore BCJ (2006) Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. Proc Natl Acad Sci 103:18866-18869.

[85] Heeren W, Coene M, Vaerenberg B, Avram A, Cardinaletti A, Del Bo L, Pascu A, Volpato F, Govaerts PJ (2012) Development of A§E test battery for assessment of pitch perception in speech. Cochlear Implants Intl 13:206-219.

[86] Vaerenberg B, Péan V, Lesbros G, De Ceulaer G, Schauwers K, Daemers K, Gnansia D, Govaerts PJ (2013) Combined electric and acoustic performance with Zebra® speech processor: speech reception, place and temporal coding evaluation. Cochlear Implants Intl 14:150-157.

[87] See RL, Driscoll VD, Gfeller K, Kliethermes S, Oleson J (2013) Speech intonation and melodic contour recognition in children with cochlear implants and with normal hearing. Otol Neurotol 34:490–498.

[88] Yucel E, Sennaroglu G, Belgin E (2009) The family oriented musical training for children with cochlear implants: speech and musical perception results of two year follow-up. Int J Pediatr Otorhinolaryngol 73:1043-1052.

[89] Chen JK, Chuang AY, McMahon C, Hsieh JC, Tung TH, Li LP (2010) Music training improves pitch perception in prelingually deafened children with cochlear implants. Pediatrics 125:e793-800.

[90] Hillenbrand J (1983) Perceptual organization of speech sounds by infants. J Speech Hear Res 26:268–282.

[91] Filipo R, Bosco E, Mancini P, Ballantyne D (2004) Cochlear implants in special cases: deafness in the presence of disabilities and/or associated problems. Acta Otolaryngol Suppl 552:74-80.

[92] Cutugno F, Prosser S, Turrini M (2000) Audiometria Vocale, vol. IV. Padova, GN Resound Italia.

[93] Vaerenberg B, Heeren W, Govaerts PJ (2013) Managed estimation of psychophysical thresholds. J Hear Sci 3:19-31.

[94] Kral A, Sharma A (2012) Developmental neuroplasticity after cochlear implantation. Trends Neurosci 35:111-122.

[95] Miura M, Sando I, Hirsch BE, Orita Y (2002) Analysis of spiral ganglion cell populations in children with normal and pathological ears. Ann Otol Rhinol Laryngol 111:1059-1065.

[96] Nie K, Barco A, Zeng FG (2006) Spectral and temporal cues in cochlear implant speech perception. Ear Hear 27:208-217.

[97] Drennan WR, Won JH, Nie K, Jameyson E, Rubinstein JT (2010) Sensitivity of psychophysical measures to signal processor modifications in cochlear implant users. Hear Res 262:1-8.

[98] Mancini P, Bosco E, D’Agosta L, Traisci G, Nicastri M, Giusti L, Musacchio A (2010) Testing auditory skills in children CI users: is phonemic discrimination related to acoustic variables only?. Cochlear Implants Intl 11:332–335.

[99] Haumann S, Lenarz T, Büchner A (2010) Speech perception with cochlear implants as measured using a roving-level adaptive test method. ORL J Otorhinolaryngol Relat Spec 72:312-318.

[100] Boyle PJ, Büchner A, Stone MA, Lenarz T, Moore BC (2009) Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants. Int J Audiol 48:211-221.

[101] Hochmair-Desoyer I, Schulz E, Moser L, Schmidt M (1997) The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users. Am J Otol 18:S83.

[102] O’Neill JJ (1957) Recognition of intelligibility test materials in context and isolation. J Speech Hear Disord 22:87-90.

[103] Voghera M (1992) Sintassi e intonazione nell’italiano parlato. Il Mulino, Bologna.

[104] Turrini M, Cutugno F, Maturi P, Prosser S, Albano Leoni F, Arslan E (1993) Nuove parole bisillabiche per audiometria vocale in lingua italiana. Acta Otorhinol Ital 13:63-67.

[105] De Mauro T (1989) Guida all’uso delle parole. Editori Riuniti, Rome.

[106] De Mauro T, Mancini F, Vedovelli M, Voghera M (1993) Lessico di frequenza dell’italiano parlato. ETAS Libri, Milan.

[107] ANSI/ASA S3.1-1999: Maximum permissible ambient noise levels for audiometric test rooms. Washington, USA: American National Standards Institute.

[108] Joffo LM, Boyle P (2010) Star^2 Validation Working Group: Cochlear Implant Users' Everyday-life performance assessed in clinical practice. Cochlear Implants Int 11:52-56.

[109] Wagener K, Kühnel V, Kollmeier B (1999) Entwicklung und Evaluation eines Satztests für die deutsche Sprache. 1. Design des Oldenburger Satztests. Z Audiol/Audiol Acoust, 38:4-15.

[110] Nilsson M, Soli SD, Sullivan JA (1994) Development of the hearing in noise test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am 95: 1085-1099.

[111] Zokoll MA, Hochmuth S, Warzybok A et al. (2013) Speech-in-noise tests for multilingual hearing screening and diagnostics1. Am J Audiol 22:175-178.

[112] Hochmuth S, Brand T, Zokoll MA, Castro FZ, Wardenga N, Kollmeier B (2012) A Spanish matrix sentence test for assessing speech reception thresholds in noise. Int J Audiol 51:536-544.

[113] Thornton AR, Raffin MJ (1978) Speech-discrimination scores modeled as a binomial variable. J Speech Hear Res 21:507–518.

[114] Puglisi GE, Warzybok A, Hochmuth S et al. (2015) An Italian matrix sentence test for the evaluation of speech intelligibility in noise. Int J Audiol 15:1-7.

[115] Vaillancourt V, Laroche C, Mayer C et al. (2005) Adaptation of the HINT (hearing in noise test) for adult Canadian Francophone populations. Int J Audiol 44:358-369.

[116] ISO 8253-3:2012 Acoustics – Audiometric test methods – Part 3: Speech audiometry.

[117] Chen F, Zhang YT (2008) A novel temporal fine structure-based speech synthesis model for cochlear implant. Signal Processing 88:2693-2699.

[118] Gifford R, Shallop J, Peterson A (2008) Speech recognition materials and ceiling effects: considerations for cochlear implant programs. Audiol Neurootol 13:193-205.

[119] Zeng FG (2002) Temporal pitch in electric hearing. Hear Res 174:101–6.

[120] Dincer D'Alessandro H, Filipo R, Ballantyne D, Attanasio G, Bosco E, Nicastri M, Mancini P (2015) Low-frequency pitch perception in children with cochlear implants in comparison to normal hearing peers. Eur Arch of Otorhinolaryngol 272:3115-3122.

[121] Dincer D'Alessandro H, Ballantyne D, De Seta E, Musacchio A, Mancini P (2016) Adaptation of the STARR test for adult Italian population: A speech test for a realistic estimate in real-life listening conditions. Int J Audiol 55:262-267.

[122] Pichora-Fuller M (2003) Processing speed and timing in aging adults: psychoacoustics, speech perception, and comprehension. Int J Audiol 42:S59–67.

[123] Dorman M, Spahr A, Gifford RH, Cook S, Zhanga T, Loiselle L, Whittingham J, Schramm D (2012) Bilateral and bimodal benefits as a function of age for adults fitted with a cochlear implant. J Hear Sci 2:EA37–EA39.

[124] Hopkins K, Moore BCJ (2011) The effects of age and cochlear hearing loss on temporal fine structure sensitivity, frequency selectivity, and speech reception in noise. J Acoust Soc Am 130:334-349.

[125] Heng J, Cantarero G, Elhilali M, Limb CJ (2011) Impaired perception of temporal fine structure and musical timbre in cochlear implant users. Hear Res 280:192-200.