Universiteit Antwerpen

PROGRAMMING COCHLEAR IMPLANTS FOR AUDITORY PERFORMANCE

DISSERTATION SUBMITTED IN FULFILLMENT OF THE DEGREE

DOCTOR OF SCIENCE

AT THE UNIVERSITY OF ANTWERP, DEFENDED BY

BART VAERENBERG

Industrial supervisor	**Antwerp**	**Academic supervisor**
Prof. Dr. Paul J. Govaerts	June 2014	Prof. Dr. Joris J.J. Dirckx

FACULTY OF SCIENCE | DEPARTMENT OF PHYSICS | LAB OF BIOMEDICAL PHYSICS

Universiteit Antwerpen

RESULTAATGESTUURD PROGRAMMEREN VAN COCHLEAIRE IMPLANTATEN

PROEFSCHRIFT VOORGELEGD TOT HET BEHALEN VAN DE GRAAD VAN

DOCTOR IN DE WETENSCHAPPEN

AAN DE UNIVERSITEIT ANTWERPEN, TE VERDEDIGEN DOOR

BART VAERENBERG

Industriële promotor
Prof. Dr. Paul J. Govaerts

Wetenschappelijke promotor
Prof. Dr. Joris J.J. Dirckx

Antwerpen
Juni 2014

FACULTEIT WETENSCHAPPEN | DEPARTEMENT FYSICA | LABO BIOMEDISCHE FYSICA

Doctoral committee

Supervisors
Prof. Dr. Paul Govaerts
Prof. Dr. Joris Dirckx

Chairman
Prof. Dr. Jan Sijbers

Members
Prof. Dr. Sandra Van Aert
Prof. Dr. Luc Pieters

External members
Prof. Dr. Martine Coene
Prof. Dr. Filiep Vanpoucke

Contact information

Bart Vaerenberg
Otoconsult NV
Herentalsebaan 71
2100 Antwerp, Belgium

vaerenberg@otoconsult.com

+32 3 609 51 68

+32 485 628 718

www.otoconsult.com

@vaerenberg

Private Ph.D. defense

Tuesday, April 29th 2014 at 10h00
in meeting room De Rede, U.338
Campus Groenenborger
University of Antwerp

Public Ph.D. defense

Friday, June 13th 2014 at 16h00
in auditorium Jan Fabre, G0.10
Campus Middelheim
University of Antwerp

FACULTEIT WETENSCHAPPEN | DEPARTEMENT FYSICA | LABO BIOMEDISCHE FYSICA

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

A/D	analog to digital
A§E	auditory speech sounds evaluatiuon
AB	Advanced Bionics
ABR	auditory brainstem response
ACE	advanced combination encoder
ADG	acyclic directed graph
ADRO	adaptive dynamic range optimization
AE	additional effect
AEV	averaged electrode voltage
AGC	automatic gain control
AI	artificial intelligence
AIMD	active implantable medical device
AMD	active medical device
ANSI	American national standards institute
APW	automatic pulse width algorithm
ART	auditory nerve response telemetry
ASC	Autosensitivity
ASR	automatic speech recognition
ASW	adaptive sound window
BAHA	bone anchored hearing aid
BKB	Bench-Kowal-Bamford sentence test
C level	comfort level
CB	critical band
CF	centre frequency
CG	common ground
CI	cochlear implant
CIS	continuous interleaved sampling
CL	current level
CMVN	cepstral mean and variance normalization
CPT	conditional probability table
CSSS	channel specific sampling sequences
CU	current unit (MED-EL) or clinical unit (AB)
CVC	consonant-vowel-consonant
dB FS	deciBell full scale
dB HL	deciBell hearing level
dB SL	deciBell sensation level
DI	disharmonic intonation
DLI	difference limen of intensity
DTW	dynamic time warping
EABR	electrically evoked ABR
EART	electrically evoked acoustic reflex threshold
EAS	electroacoustic stimulation
ECAP	electrically evoked compound action potential
EDR	electrical dynamic range
EER	equal error rate
EFP	electrical field potential
EM	expectation maximization
EMR	electronical medical record
ENT	ear-nose-throat
ERB	equivalent rectangular bandwidth
ES	electric stimulation
ESR	electrically evoked stapedius reflex
ESRT	electrically evoked SRT
F0	fundamental frequency

FAT	frequency allocation table
FDA	U.S. food and drug administration
FFT	fast Fourier transform
FIR	finite impuls response
FOX	fitting to outcome expert
FS4	fine structure on 4 channels
FS4-p	parallel stimulation FS4
FSP	fine structure processing
GCP	good clinical practice
HDCIS	high definition CIS
HF	high frequency
HI	harmonic intonation
HI	hearing impaired
HiRes	AB's high resolution strategy
HiRes-P	parallel stimulation HiRes
HiRes-S	sequential stimulation HiRes
HIS	hospital infornation system
HSM	Hochmair-Schulz-Moser sentence test
IBK	Innsbruck
ICCI	intensity coding in cochlear implants
ICF	intensity coding function
ICI	independence of causal interaction
IDE	integrated development environment
IDR	input dynamic range
IEC	international electrotechnical commission
IHC	inner hair cell
IIDR	instantaneous input dynamic range
IMR	instantaneous mapping range
IPI	inter pulse interval
ISO	international organization for standardization
IWT	agency for innovation by science and technology
JND	just noticeable difference
k-NN	k-nearest neighbors
LAURA	Leuven and Antwerpen universities research auditory prosthesis
LF	low frequency
LS	loudness scaling
M level	most comfortable level
MAP	configured values for the tuneable parameters of the CI speech processor
MC	map condition
MCL level	most comfortable level
ME	map effect
MFCC	mel-frequency cepstral coefficients
MLE	maximum likelihood estimation
MP	monopolar
MPIS	main peaks interleaved sampling
NH	nornal hearing
NRI	neural response imaging
NRT	neural response telemetry
OAE	otoacoustic emissions
OC	outcome condition
OHC	outer hair cell
OOPN	object-oriented probabilistic network
PCA	principal-components analysis
PEST	parameter estimation by sequential testing
PGM	probabilistic graphical model
POMDP	partially-observable Markov decision process
POR	pending outcome request

pps	pulses per second
PSOLA	pitch synchronous overlap add
Q	loudness growth
QOL	quality of life
QR	quartile range
QU	charge unit
RF	radio frequency
RMS	root mean square
RMSE	RMS error
S/N	signal-to-noise ratio
SAT	speech audiometric test
SD	standard deviation
SHR	subject hit rate
SI	sentence intonation
SIR	speech intelligibility rating
SNHL	sensorineural hearing loiss
SNR	signal-to-noise ration
SP	speech processor
SPL	sound pressure level
SRT	stapedius reflex threshold
SRT	speech reception threshold
STARR	sentence test with adaptive randomized roving levels
T level	threshold
TE	temporal envelope
TEMA	threshold estimation by managed algorithm
TFS	temporal fine structure
THR	target hit rate
THR level	threshold level
tSHR	tolerant subject hit rate
tTHR	tolerant target hit rate
WSP	word stress pattern

PREFACE

Fox's victory mood, when all targets of auditory performance have been reached by a CI recipient.

MOTIVATION

The ear, when put into competition with our other senses, is probably the most important sensory organ. Hearing is essential to man's existence. The handicap of being hearing impaired carves deeply into all segments of society and all facets of man in his social context: in the elderly induced by age, in the unfortunate and reckless by trauma, but a fortiori in innocent children struck by fate and sentenced to not only deafness but also verbal muteness and intellectual deprivation.

The cochlear implant is the first technological development in human kind to bring a radical change to this predicament of severe sensory disability. To effectively replace the human sensory organ of hearing with a machine is an accomplishment for which only the humble word miracle is in place. As rudimentary and subpar to the healthy human ear as the cochlear implant may be today, so essential it is already for the social integration, the communicative abilities and basically the comfort and happiness of individual people.

To be working in the field of cochlear implants therefore is rewarding for 2 reasons: (1) being able to contribute to something that helps hearing impaired people to regain the ability to communicate and (2) the excitement of being at the forefront of an evolution towards the fusion of man and machine. It is amazing to see that machines no longer need to be separate entities; rather they can be integrated into our physical bodies. This mixture of biology and technology is fascinating. The present project and its prerequisite appetite for knowledge live by these incentives.

Almost 30 years have passed since the introduction of the multichannel cochlear implant (CI). Briefly, it provides direct electrical stimulation of the auditory nerve, bypassing damaged parts of the ear. A series of electrodes is surgically inserted into the inner ear. An external speech processor, usually worn behind the ear, picks up sounds from the environment and converts them into a digital signal. That signal is sent to the implant and delivered to the electrodes for stimulation of auditory nerve fibres. Hardware capabilities, speech processing strategies, ease of use and aesthetic design of cochlear implants have all evolved significantly over the years. Still, it is astonishing to see that a technology as bold and revolutionary as the cochlear implant has grown to become a routine intervention, taken for granted as the solution to profound deafness.

Huge challenges concerning the further development and introduction of cochlear implants worldwide, however, remain even to this day. One such a challenge is the process of "fitting": the tuning of the speech processor to the individual recipient, in order to provide an optimal auditory percept and aiming at maximal speech understanding for every patient. The many tuneable parameters available to customize the speech processor and the complexity of signal processing and stimulation strategies of current generation cochlear implant systems make that fitting is a non-trivial task, to say the least. Also the various pathological conditions that are encountered across recipients, of which the causes are mostly unknown and of which the origin may well lie in the peripheral as well as in the central auditory system, add to the difficulty and uncertainty of achieving an optimal tuning for any given individual.

Over the course of 30 years, knowledgeable experts have been perfecting their own approaches to the act of fitting. It has become craftsmanship; more of an art than a science. Not to say that the results of fitting at present are unsatisfactory. They are good to excellent in many cases, but the current practice leaves obvious room for improvement. No universally accepted methodology exists and the variation in procedures adopted by different CI clinics is large. Time has come to reflect upon this matter and take the next step. The world's expertise needs to be distilled into common knowledge, such that it can be validated, optimized and exported to those places where expert fitters or resources are scarce. In fact, the ever growing population of CI recipients worldwide will result in the necessity for an optimization over the current practice of fitting, even in the most experienced clinics and resource abundant settings, if one is to keep the quality and manageability of CI fitting under control.

Three components are essential for this attempt to process optimization: (1) defining targets for the system, (2) measuring the state of the system and (3) algorithms effective in moving the system's state towards target. The application of such a process optimization to the act of CI fitting is the subject of the present project. Its fundamental approach consists of measuring a CI recipient's auditory performance to make targeted adjustments to the speech processor, in order to improve hearing performance. Major investments have therefore been directed towards:

- the definition of targets for auditory performance;
- the development of measurement instruments for hearing assessment;
- modelling the impact of processor adjustments on hearing performance as assessed by these measurements and
- the clinical validation of the fitting paradigm incorporating the model

The present project is motivated by the belief that the systematized measuring of hearing performance, and modelling the effects of processor changes on outcome, is essential to improving the quality of CI fitting. This dissertation reports on the efforts of developing such a paradigm. It was given form by combining a number of manuscripts which have been published or submitted for publication in a scientific journal. The required introductory information, "educated commentary" and context explanation is given in between the manuscripts, such that the reader is guided through the project in a more fluent manner. It is hoped that the reader, at the end of this dissertation, is convinced that the present project, already now, yields real added value to the hearing impaired patient, but more importantly that it is a first step in a valid direction, which has to be taken in order to improve the art of CI fitting, to convert it into the science of CI fitting.

CONTEXT

This PhD research project was conducted in a consortium involving the University of Antwerp and Otoconsult, a privately owned company. Otoconsult was founded in 2007 as a spin-off company from the Eargroup, an ENT clinic specialized in otology and audiology. Since many years the members of the Eargroup have contributed to the development of new diagnostic and therapeutic techniques, such as middle ear transplantation surgery, cochlear implant surgery, skull base surgery, mechanical measurements of the middle ear, different tests of auditory performance, etc. Otoconsult now accommodates the research and development activities that have evolved from the Eargroup's years of experience in the field of audiology and otology. The results of these activities are targeted towards commercial exploitation. This unique mixture of academic context, clinical setting and economic finality requires high attention for theoretical correctness as well as for clinical pragmatics and final applicability.

The present PhD project being no exception, originated from an economical necessity to optimize CI fitting, but required fundamental research into a model of the relationship between CI processor parameters and hearing performance. Such a model needs to bridge between biology and technology. It needs to incorporate efficient algorithms while respecting the uncertainty inherent to behavioural methods. It needs to consider both quantitative and qualitative measures. It needs to encompass the many output variables expressed by outcome measures and input variables available as tuneable CI processor parameters. Therefore the model was expected to be highly complex and the endeavour of developing it high risk.

But only by such an in depth research approach it was estimated feasible to bring forth a viable application that could be commercially marketed. This direct commercial finality is different from typical PhD research, conducted in purely academic settings, and made the project a perfect fit for being funded through the Baekeland initiative of the Flemish government's agency for Innovation by Science and Technology (IWT).

Universiteit Antwerpen | de OOR groep | OTO CONSULT | IWT

The purpose of the Baekeland mandates is to support basic research that – if successful – has clear economic objectives and offers added value to the company involved in the project. However, the research should be directed towards achieving a doctorate (PhD) diploma and meet the accepted criteria for doctoral research. In other words, the project should fit within strategic basic research with an economic finality, defined as high quality research that is innovative and provides the PhD student with ample intellectual properties. It aims to build up scientific or technological knowledge as a basis for economic applications.

This also means that the results of the research cannot always be made publicly available. For reasons of intellectual property and the possibility for commercial exploitation, parts of this dissertation are only made available to the private doctoral committee, after non-disclosure agreement. Also during the course of the project, the delicate balance between contribution to science and industrial valorisation had to be maintained. This resulted in the fact that some of the information presented in this dissertation could not be submitted for publication in a scientific journal. Nonetheless it strives for the same scientific validity as can be expected from peer reviewed manuscripts.

AKNOWLEDGEMENTS

This PhD dissertation is the result of an entire team's aspiration, dedication and years of work. It is the spirit of the Eargroup that made this mission possible and that amounted to this dissertation. It is of great fortune to have landed in such an extraordinary setting, amongst people as generous and challenging as my colleagues at the Eargroup. I wish to express my gratitude towards them, for sharing their knowledge and competences, for their enthusiasm and devotion to the project, for their sympathy and kindness, for the marvellous times we spent together:

the Eargroup's inspirer and driving force, guru in otology and audiology: Prof. **Paul** Govaerts;

the Eargroup's seasoned and ever-committed audiologists: Ir. **Geert** De Ceulaer, **Kristin** Daemers, **Carina** De Beukelaer and Dr. **Karen** Schauwers;

the Eargroup's next-generation audiologists: **Karolien** Weekers, **Katiba** Tabaouni, **Magalie** Teunen, **David** Pascoal, **Lenne** Kerremans and **Jasmijn** Stuer;

the Eargroup's ENT physicians: Prof. **Olivier** Vanderveken and Dr. **Gerd** Claes;

the Eargroup's software engineers and computer scientists: **Jess** Van Thillo, **Claudio** Redavid, **Denis** Boigelot, **Joris** Schroeven and **Jasper** Vermorgen and

the Eargroup's fundamental linguist and recreational fund raiser: Prof. **Martine** Coene.

During my research I had the pleasure of meeting a number of wonderfully interesting people. I would like to say thanks to them, for sharing their expertise and for their contributions to the project, as well as for the good times we had:

conspiring researchers: Prof. **Wojtek** Kowalczyk and Dr. **Zoltan** Szlavik, mathematicians in computational intelligence; Prof. **Louis** ten Bosch, language and speech technologist; Drs. **Iñigo** Bermejo and Prof. **Javier** Díez, experts in artificial intelligence; Dr. **Patty** Hendrix, biologist; **Eric** Schipperen, **Ad** Vreeswijk and **Hans** Kleijnendorst, acousticians; **Herwig** De Smet, computer scientist, neurolinguist and ad hoc guinea pig; Dr. **Annemiek** Hammer and Dr. **Willemijn** Heeren, linguists; Prof. Christian **Lorenzi**, psychoacoustician; Dr. Luca **Del Bo**, bio-engineer and Dr. Alexandru **Pascu**, ENT physician;

conspiring engineers and clinical specialists of the 4 CI manufacturing companies: **Filiep** Vanpoucke, **Alexandre** Gault, **Vincent** Jourdes, **Vincent** Péan, **Kristof** De Ceulaer, **Dzemal** Gazibegovic, **Peter** Nopp, **Mathias** Kals, **Thomas** Stainsby, **Stefan** Lievens , **Guillaume** Lesbros, **Dan** Gnansia and **Philippe** Dykmans and

my fellow PhD students/co-workers/professors at the Laboratory of Biomedical Physics: Dr. **Joris** Soons, Dr. **Jef** Aernouts, Dr. **Jan** Buytaert, Dr. **Johan** Aerts and also **Daniël**, **Sam**, **John**, **Adriaan**, **Jana**, **Fred**, **Jan**, **William**, **Luc**, **Ilse** and **Greet** and Profs. **Decraemer** and **Wuyts**.

I also want to say thanks to those that had no direct interaction with my research, but without whom this dissertation would not have existed: to my brother **Jan** and sisters **Sofie** and **Carolien** for being family, to my parents Dr. **Brigitte** Dompas and Dr. **Marc** Vaerenberg for having raised and taught me well, to my friends **Steven** and **Hendrik** for tolerating my unbearable behaviour and to the one I love, **Griet**, for her unique implementation of loving me.

The Baekeland mandate is organized through 2 supervisors, an academic one and an industrial one. I would like to thank my academic supervisor, **Prof. Joris Dirckx**, for giving me the opportunity and for having the confidence to let me conduct a doctoral research project under the wings of his laboratory. His guidance has been much appreciated.

Prof. Paul Govaerts served as the industrial supervisor of my PhD research, but that title does not quite cover the impact of his presence. He has not only been my guide during this endeavour but also an inspiration for both my professional contentment and my personal perspectives. His passion for the Eargroup is extraordinary, his challenging mind-set irresistible. The hospitality with which he invited me to work and life is something that I can only admire. I believe it was the ground for bringing out the best in me. These 8 years of working together have been turbulent at times but always fascinating and definitely rewarding. Merci Paul.

B. Vaerenberg received a PhD grant for this work from the IWT (Agency for Innovation by Science and Technology, Baekeland-mandaat IWT090287).

SUMMARY

A cochlear implant (CI) replaces the function of the inner ear (cochlea) through direct electrical stimulation of the auditory nerve. Since its introduction 30 years ago [1] [2], the multichannel cochlear implant has become routine therapy for severe perceptive hearing loss and deafness in both adults and children. The positive outcome of cochlear implantation has been proven in a convincing manner, in terms of audiological results and, for children also in terms of speech and language development and scholar integration [3] [4] [5] [6] [7] [8]. The CI attempts to mimic as reliably as possible the physiological process of hearing as found in the healthy ear. For this, it is important that the configuration parameters (called the "map") of the CI speech processor are tuned to the individual recipient. This way an optimal electrical activation pattern is provided to the neural interface and a maximally functional hearing is restored. This process of tuning the speech processor is commonly called CI programming or "fitting" [9].

At present, the programming of CIs is being perceived as complex and as a significant constraint on the further development and introduction of CIs worldwide. The people responsible for fitting (usually audiologists or engineers) having extensive expertise are rare [10]. Even for those "expert fitters" it seems almost impossible to master all programming parameters and their interactions, and more importantly to predict in a reliable manner the impact on a recipient's auditory performance of changing them [11] [12] [13]. The current practice of CI fitting therefore is constrained to the manipulation of a limited set of parameters (mainly the current levels that cause liminal (THR) en comfortable (MCL) perception for each electrode) [14] [15] [16].

Today, fitting is a manual process in which parameter changes are justified through patient feedback. This is by definition a subjective approach, usually targeted at auditory comfort ("does this sound more pleasant or not?") and often not in line with a recipient's auditory performance as it would be expressed through psychoacoustical measurements [17] [18]. The use of objective measurements (ECAP [19] [20] [21], ESRT [22], etc.) to adjust the processor has gained popularity in recent years. However, the correlation of these measurements to the actual optimal settings was revealed to be limited. Moreover, these methods still remain focused only on finding appropriate EDRs and other fitting parameters are often neglected.

The Eargroup has chosen, since many years, to adopt an outcome-driven approach, in which CI fitting is motivated by and tested for auditory performance, as measured by behavioural psychoacoustic tests. Prior to this PhD project however, fitting at the Eargroup was done manually and empirically, based on years of their experience. In order to make this more evidence-based, it was necessary to investigate the relation between electrical stimulation and auditory performance in a more fundamental way.

Given the limited availability of documentation on how a CI processes sound and how it converts it to electrical stimulation levels, this matter was further investigated in collaboration with engineers

from the CI manufacturers. The results regarding the coding of sound intensity can be found in “Intensity coding in current generation CI systems”.

To construct a global inventory of the methodology used for CI fitting, 47 experts from international CI centres were interrogated on their methods. The resulting inventory represents a total of 47,600 CI users (> 15% of CI users worldwide), making it an unprecedented synthesis of the current state of the art. The results were processed and compiled in the manuscript “A global survey on the state of the art of CI fitting”. In summary, this inventory shows that many different approaches exist for finding an optimal program for the individual CI recipient. Although several of those approaches in the hands of different experts may lead to similarly good results, it is hard to compare/benchmark them in the absence of commonly agreed upon targets for outcome.

Basically, a CI replaces the function of the cochlea as a receptor of sound. The responsibility of the cochlea is to deliver signals to the brain in a way that they carry maximal information about the sound that is present. To achieve this goal the cochlea needs to make sure that different sounds are transduced into different electrical patterns such that the brain can tell them apart. This is referred to as the discriminative power (i.e., intensity sensitivity, frequency selectivity and temporal resolution) of the cochlea. When attempting to improve a CI user's speech perception or even quality of life by fitting a good map, it is reasonable to focus on maximizing the CI's discriminative power with respect to the available peripheral neural interface.

This leads to the need for a set of outcome measures that reflects real-life auditory performance, but at the same time allows for an analytical interpretation. In other words, a good set of outcome measures to be used to drive CI fitting should depend highly on the functioning of the peripheral auditory system (in our cases replaced by a cochlear implant) and as little as possible on anything else. In “Measurements & outcome”, such a set of outcome measures is developed. These measurements are implemented in the A§E Psychoacoustics Test Suite and they are used to assess the state of the auditory system at the levels of detection, discrimination and identification and evaluate the auditory system’s ability to process intensity, spectral and temporal features of sound. For the fitting paradigm that has been developed during this project, 4 of these psychoacoustical tests are of particular importance: Audiometry (sound field detection thresholds), Speech Audiometry (speech recognition scores), Phoneme Discrimination (distinguishing between trivial speech sounds) and Loudness Scaling (the growth of loudness sensation).

“Modelling the impact of fitting on outcome” elaborates on the efforts of constructing a model for optimizing maps based on deficits in these outcome measures. A deterministic rule-based model, built on clinical heuristics and called the Eargroup’s advice, has been developed and made available to clinicians through a software tool called Fox (Fitting to Outcome Expert). This process is described in the manuscript "Development of Fox". The details of the model are, for reasons of intellectual property and commercial exploitation, only made available to the private doctoral committee. In a later stage it was also explored how probabilistic modelling techniques, like Bayesian networks, could be put to use to handle some of the intrinsic uncertainty related to psychoacoustical measures and the complexity of electrical hearing. The papers “A probabilistic graphical model” and “The tuning model” report on these efforts. At present the resulting probabilistic network is being evaluated at the Eargroup.

The clinical application of the Eargroup’s advice through Fox is described in "Clinical implementation". It is applied through a fitting paradigm that could be described as "adapt first, tune later". It comprises 2 distinct stages:

(1) Automaps: allow the recipient to adapt to increasing levels of electrical stimulation and to (re)gain a reference frame for loudness during the first few weeks after switch-on.
(2) Tuning: the iterative optimization of the recipient's map by measuring outcome and adjusting the map in order to minimize the recipient's distance to predefined targets for each outcome measure.

The process of Automaps is described in detail in the manuscript "Experiences of the use of Fox in new users". This report outlines the fitting protocol that is typically followed at the Eargroup for postlingually deafened adult CI recipients using the Fox system from switch-on onwards. The timing of 4 sessions in the first six months was found to be adequate to optimize the subjects’ maps in the great majority of cases. Across these 4 sessions the total time spent is of the order of 2.5 hours, which includes all “audiological” issues, i.e. technical explanations, device programming and performance measures. This compares favourably with fitting times reported by traditional methods. This report demonstrates that good results can already be obtained with a relatively small clinical workload and that a systematic outcome-driven approach, with the assistance of an intelligent agent like Fox, is capable of selectively improving test results.

The manuscript “Evaluation of Fox with established cochlear implant users” reports on a study to evaluate whether Fox is able to complement standard clinical procedures in clinics other than the Eargroup. Ten adult postlingually deafened and unilateral long term CI users underwent speech perception assessment with their current clinical program. One iteration of Fox optimization was performed and the map adjusted accordingly. After a month of using both clinical and Fox programs a second iteration of Fox optimization was performed. Following this, the assessments were repeated without further acclimatization. Sound field aided thresholds were found to be significantly better for the Fox than for the clinical program. Group speech scores in noise were not significantly different between the two programs but three individual subjects had improved speech scores with the Fox map, two had worse speech scores and five were the same.

“Multicentre assessment of Fox in new cochlear implant users” reports on a controlled, randomised, clinical study conducted in CI centres in Germany, the United Kingdom and France. The aim was to compare the overall fitting time and the overall speech perception performance, between Fox and standard clinical fitting procedures (Control group). The results showed a significant improvement in word scores in quiet (35%, $p = 0.02$) and sentences in +5dB signal to noise (23%, $p=0.04$) for the Fox group compared to the Control group at six months. The fitting time for Fox was also significantly reduced at 14 days ($p<0.001$) and equivalent over the six month period. There was much less overall variance in the Fox results. From this it is concluded that the use of Fox produced results that were at least equivalent to conventional fitting methods for all the outcome measures tested. Despite including more testing of outcomes during fitting and the adjustment of a greater range of parameters, Fox does not add to the fitting time. Fox appears highly efficient and effective in providing an optimal map.

The paper “Setting and reaching targets with computer-assisted CI fitting” contains a retrospective data analysis on 255 adults and children in 14 participating centres. The paper aims at demonstrating the feasibility of defining a substantial set of psychoacoustic outcome measures with preset targets and to adopt a systematic methodology for reaching these targets. For each patient, 66 measurable psychoacoustical outcomes were recorded several times after cochlear implantation: free field audiometry (6 measures), speech audiometry (4), spectral discrimination (20) and loudness growth (36). These outcomes were pooled to 22 summary variables. The initial results were compared with the latest results. Results showed that the use of Fox significantly improved the proportion of the 22 variables on target. When recipients used the automated maps provided at switch-on, more than half (57%) of the 22 targets were already achieved before any further optimisation took place. Once the Fox system was applied there was a significant 24% ($p < 0.001$) increase in the number of targets achieved.

This study demonstrates that it is feasible to set targets and to report on the effectiveness of a fitting strategy in terms of these targets. Fox provides an effective tool for achieving a systematic approach to programming, allowing for better optimisation of recipients' maps. The setting of well defined outcome targets, allowed a range of different centres to successfully apply a systematic methodology to monitoring the quality of the programming provided.

SAMENVATTING

Een cochleair implantaat (CI) vervangt de functie van het slakkenhuis (cochlea) van het oor ten dele door directe elektrische stimulatie van de gehoorzenuw. Sinds de introductie 30 jaar geleden [1] [2], behoort het multikanaals intracochleaire implantaat nu bij zowel kinderen als volwassenen met ernstige perceptieve slechthorendheid c.q. doofheid tot de routine therapieën en zijn de positieve resultaten overtuigend aangetoond, zowel op audiologisch als voor kinderen ook op linguïstisch en schools niveau [3] [4] [5] [6] [7] [8]. Het CI beoogt het fysiologische proces dat normaalhorenden in staat stelt te horen zo getrouw mogelijk na te bootsen. Het is belangrijk dat de instelbare parameters (de “map” genoemd) van de processor van het CI hierbij per individu afgeregeld worden om het complexe elektrische activatiepatroon in de cochlea zo natuurlijk als mogelijk te laten verlopen en zo een optimaal gehoor te realiseren. Dit afregelen wordt gemeenzaam “fitting” genoemd [9].

De fitting van CI’s wordt ervaren als complex en als een belangrijke rem op de verdere introductie en ontwikkeling van CI’s wereldwijd. De personen, meestal audiologen of ingenieurs, met doorgedreven expertise, zijn zeldzaam [23]. Zelfs voor deze “expert fitters” blijkt het quasi onmogelijk om alle paramaters van het implantaat en hun interacties te beheersen en vooral om de impact van wijzigingen ervan op de auditieve performantie op een gecontroleerde manier in te schatten [24] [25] [26]. De fitting van CI’s zoals deze momenteel in de wereld gebeurt, bestaat er dan ook in slechts een heel beperkt aantal paramaters van de processor te manipuleren (doorgaans de stroomniveaus per elektrode die tot liminaire (THR) en comfortabele (MCL) perceptie leiden) [27] [15] [28].

CI fitting is vandaag nog een manueel proces waarbij de verandering van parameters wordt getoetst aan de feedback van de patiënt. Dit is per definitie subjectief, vaak gericht op comfort (“klinkt het aangenamer of niet?”) en dikwijls niet in overeenstemming met de auditieve performantie zoals die o.a. blijkt uit psychoakoestische metingen [29] [30]. Het gebruik van objectieve metingen (ECAP [31] [32] [33], ESRT [34], etc.) voor het afstellen van de processor heeft de laatste jaren aan populariteit gewonnen. Nochtans is de correlatie van deze meetresultaten tot de eigenlijke optimale instellingen beperkt gebleken. Bovendien blijven deze methoden nog steeds gericht op het vinden van de adequate THR en MCL levels en worden andere parameters dikwijls verwaarloosd.

De Oorgroep kiest al jaren voor een *outcome-driven* aanpak, waarbij de fitting gemotiveerd en getoetst wordt aan prestaties, gemeten aan de hand van gedragsmatige, psychoakoestische testen. Het fitten gebeurde, voorafgaand aan dit doctoraat, echter nog manueel en empirisch, gebaseerd op de jarenlange ervaring. Om dit meer *evidence-based* te maken, was het nodig om de relatie tussen elektrische input en performantionele output fundamenteler in kaart te brengen.

Gezien de documentatie over hoe een CI geluid verwerkt en omzet naar electrische stimulatieniveaus erg beperkt bleek, werd dit in samenwerking met ingenieurs van de CI fabrikanten verder onderzocht. De resultaten met betrekking tot intensiteitscodering worden beschreven in "Intensity coding in current generation CI systems".

Om in kaart te brengen welke methodes vandaag gebruikt worden voor CI fitting, werden 47 experts uit binnen- en buitenlandse CI centra ondervraagd over hun werkwijze en methodes. De opgestelde inventaris vertegenwoordigt in totaal 47.600 CI gebruikers (meer dan 15% van de CI gebruikers wereldwijd). Het is in die zin een unieke synthese van de huidige state of the art. De resultaten werden verwerkt en gebundeld in het manuscript "A global survey on the state of the art of CI fitting". Dit onderzoek toont aan dat de methodes ten velde, voor het vinden van een optimale map, heel uiteenlopend zijn. Een aantal van deze methodes leidt ongetwijfeld tot goede resultaten. Maar door elke afwezigheid van unaniem aanvaarde "targets" blijft een onderlinge vergelijking moeilijk.

In essentie vervangt het CI de functie van de cochlea als receptor van geluid. Dier taak bestaat erin een signaal aan te bieden aan de hersenen dat voldoende informatie over het geluid draagt. Het slakkenhuis moet er dus voor zorgen dat verschillende geluiden worden omgezet in verschillende elektrische patronen, die door de hersenen onderscheiden kunnen worden. Dit heet het discriminatief vermogen van de cochlea en bestaat uit 3 componenten: intensiteitsgevoeligheid, frequentieselectiviteit en temporele resolutie. Wanneer getracht wordt het spraakverstaan van een CI-drager te verbeteren door het aanpassen van zijn/haar map, is het dus redelijk om daarbij te streven naar een configuratie van het CI die een maximaal discriminatief vermogen realiseert, rekening houdend met de beschikbare perifere neurale interface.

Er is dus een duidelijke nood aan instrumenten die het gehoor kunnen opmeten, zoals het functioneert in het dagelijkse leven, maar die tegelijkertijd ook een analytische interpretatie toelaten op het niveau van de cochleaire functie. Een testbatterij om CI fitting te sturen dient dus metingen te bevatten die sterk afhankelijk zijn van de werking van het perifere gehoorsysteem (in dit geval vervangen door een cochleair implantaat) en zo weinig mogelijk van andere factoren. In "Measurements & outcome" wordt zulk een verzameling gehoortesten ontwikkeld. Ze werd vorm gegeven in de A§E Psychoacoustics Test Suite en als dusdanig gebruikt om het gehoor te evalueren op niveau van detectie, discriminatie en identificatie van intensiteit en spectrale en temporele aspecten van geluid. Voor het fitting paradigma dat tijdens dit project ontwikkeld werd, zijn 4 van deze psychoakoestische tests van bijzonder belang: Audiometrie (detectiedrempels in het vrije veld), Spraakaudiometrie (spraakverstaan gemeten aan de hand van korte woorden), Foneemdiscriminatie (onderscheid maken tussen triviale spraakklanken) en Luidheidsschaling (de aangroei van luidheidssensatie).

"Modelling the impact of fitting on outcome" behandelt het model dat maps optimaliseert op basis van deze gehoortesten. Dit deterministisch model is gebaseerd op klinische heuristiek en wordt het Oorgroep advies genoemd. Het werd beschikbaar gemaakt voor clinici aan de hand van de Fox software (Fitting to Outcome Expert). Een uitgebreide beschrijving vindt u in "Development of Fox".

De details van dit model zijn, om redenen van intellectuele eigendom en commerciële exploitatie, enkel beschikbaar voor de private doctoraatscommissie.

In een later stadium werd ook onderzocht hoe probabilistische modelleringstechnieken, zoals Bayesiaanse netwerken, kunnen ingezet worden om beter om te gaan met de onzekerheid die inherent is aan psychoakoestische testen en aan de complexiteit van elektrisch horen. De papers "A probabilistic graphical model" en "The tuning model" brengen verslag uit over deze inspanningen. Op dit moment wordt een eerste versie van een probabilistisch netwerk voor mapoptimalisatie uitgetest in de Oorgroep.

De klinische toepassing van het Oorgroep advies via Fox wordt beschreven in "Clinical implementation". Dit gebeurt aan de hand van een fitting paradigma dat kan worden omschreven als "eerst gewenning, dan afregelen". Het bestaat uit 2 fasen:

(1) Automaps: de CI gebruiker laten wennen aan toenemende niveaus van elektrische stimulatie, zodanig dat een referentiekader voor geluidssterkte (opnieuw) verworven wordt tijdens de eerste paar weken na switch-on.
(2) Tuning: de iteratieve optimalisatie van de map door het meten van gehoor en het aanpassen van de map opdat de CI gebruiker dichter bij het bereiken van vooraf bepaalde targets voor iedere gehoormeting zou komen.

De werkwijze van Automaps wordt beschreven in het manuscript "Experiences of the use of Fox in new users". Dit rapport schetst het fitting protocol dat voor post-linguaal dove volwassen CI gebruikers in de Oorgroep meestal gevolgd wordt vanaf switch-on. Het inplannen van 4 sessies tijdens de eerste zes maanden bleek in de grote meerderheid van de gevallen voldoende om de maps van de proefpersonen te optimaliseren. De totale tijd die gespendeerd werd tijdens deze 4 sessies was in de orde van 2,5 uur en omvat alle "audiologische" kwesties (technische uitleg, gehoormetingen en programmatie van het CI). Dit is typisch minder dan wanneer traditionele fitting methodes gebruikt worden. De studie toont ook aan dat men goede resultaten kan verkrijgen met een relatief beperkte klinische workload en dat een systematische resultaat-gedreven aanpak, met de hulp van een "intelligent agent" zoals Fox, ervoor zorgt dat testresultaten selectief verbeterd kunnen worden.

Het manuscript "Evaluation of Fox with established cochlear implant users" beschrijft een studie die beoordeelt of Fox in staat is om de procedures die in klinieken buiten de Oorgroep gebruikt worden, te verbeteren. Bij tien volwassen, post-linguaal dove, unilaterale CI gebruikers werd eerst het spraakverstaan opgemeten met hun huidige klinische map. Deze map werd vervolgens aangepast volgens de suggesties van een Fox-iteratie. Zowel hun klinische als hun Fox-programma werden een maand lang afwisselend gebruikt door de CI dragers, waarna een tweede Fox-iteratie werd uitgevoerd. Op dat moment werd spraakverstaan opnieuw opgemeten, zonder verdere acclimatisatie. De audiometrische drempels bleken beduidend beter te liggen met het Fox-programma. Spraakverstaan in ruis was niet significant verschillend tussen de twee programma's.

Drie proefpersonen vertoonden wel een verbeterd spraakverstaan met de Fox-map, twee personen hadden slechtere spraakscores en vijf bleven er onveranderd.

“Multicentre assessment of Fox in new cochlear implant users” rapporteert over een gecontroleerde, gerandomiseerde klinische studie, uitgevoerd in CI-centra in Duitsland, het Verenigd Koninkrijk en Frankrijk. Het doel was om de totale tijd gespendeerd aan fitting, en het algemene spraakverstaan te vergelijken tussen Fox en standaard klinische fittingprocedures (de controlegroep). De resultaten toonden significant betere woordscores in stilte (35%, p=0,02) aan na zes maanden en ook op de zinnentest in +5dB signaalruisverhouding werden significant betere scores (23%, p=0,04) vastgesteld in de Fox groep. De tijd gespendeerd aan fitten was significant korter in de Fox-groep tijdens de sessie 2 weken na switch-on (p < 0,001) en equivalent aan de controlegroep wanneer over de ganse periode van zes maanden bekeken. De totale variatie in de Fox-resultaten was beduidend lager. De conclusie is dat het gebruik van Fox resultaten behaalt die ten minste gelijkwaardig zijn aan die bij conventionele fittingpraktijken, voor alle tests die werden afgenomen. Ondanks het veelvuldiger testen tijdens het fitten en het manipuleren van een groter aantal map parameters, zorgt Fox niet voor een verlenging van de tijd die nodig is voor fitten. Fox blijkt zeer efficiënt en effectief in het verschaffen van een optimale map.

De paper “Setting and reaching targets with computer-assisted CI fitting” beschrijft een retrospectieve analyse van gegevens verkregen bij 255 volwassenen en kinderen in 14 deelnemende centra. De paper onderzoekt of het haalbaar is om een testbatterij met bijbehorende targets te definiëren en vervolgens een systematische methodologie voor het bereiken van deze targets toe te passen. Bij elke patiënt werden na cochleaire implantatie 66 psychoakoestische targets herhaaldelijk opgemeten: audiogram (6 targets), spraakaudiometrie (4), spectrale discriminatie (20) en luidheidsaangroei (36). Deze 66 targets werden teruggebracht tot 22 samenvattende variabelen. De initiele testresultaten werden vergeleken met de laatst verkregen resultaten. Er werd aangetoond dat het gebruik van Fox een aanzienlijke verbetering in het behalen van de 22 targetvariabelen teweegbrengt. Na switch-on met behulp van Automaps, werden reeds meer dan de helft (57 %) van de 22 doelstellingen bereikt, vóór enige verdere optimalisatie plaatsvond. Zodra Fox werd ingeschakeld, was er een significante toename 24% (p<0,001) in het aantal bereikte targets.

Deze studie toont dus aan dat het haalbaar is om targets te stellen voor auditieve performantie en te rapporteren over de effectiviteit van een fittingstrategie aan de hand van deze targets. Fox biedt op deze manier een effectief instrument ter optimalisatie van maps van CI dragers. Fox zorgt er ook voor dat een systematische aanpak van CI programmatie geinstalleerd wordt, waardoor de variatie in outcome na implantatie relatief klein is. Het vastleggen van een aantal goed gedefiniëerde targets, heeft verschillende CI centra in staat gesteld een systematische methode toe te passen die toezicht houdt op de kwaliteit van hun CI fitting.

CHAPTER 1: THE EAR & COCHLEAR HEARING LOSS

Scanning electron micrograph of inner hair cells and outer hair cells in the organ of Corti. [adapted from M. Lenoir]

1.1. THE HUMAN EAR

1.1.1. STRUCTURE AND FUNCTION

The human peripheral auditory system is composed out of 3 parts: the outer ear, the middle ear and the inner ear (Figure 1). The external auditory canal together with the pinna makes the outer ear, which is responsible for guiding sound waves to the tympanic membrane. The shape and material properties of the outer ear feature a relatively wide resonance, favouring the incoming sound waves in a region of about 1 kHz to 5 kHz, a band that contributes heavily to speech perception. This gain is maximal around 2.5 kHz and would typically be between 12 and 15 dB [35]. The pinna also introduces a filter of which the characteristics are highly dependent on the direction of the incoming sound [36], contributing to the ability to localise sound sources.

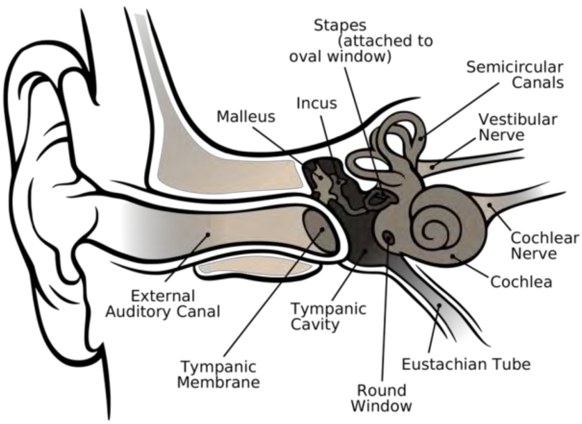


Figure 1: The human peripheral auditory system showing the outer ear, middle ear and inner ear. [adapted from Chittka & Brockmann]

The transition from outer to middle ear happens at the tympanic membrane (eardrum), where sound pressure is converted into mechanical motion. The middle ear is a chain of 3 ossicles (malleus, incus and stapes) attached to the eardrum at one side (malleus) and the oval window at the other side (stapes). This structure makes that a sound's energy in air is transformed efficiently, matching the significantly different impedance of the fluids in the inner ear. This transformation is most efficient for frequencies between 0.5 kHz and 5 kHz [37], again the region in which important information of speech is conveyed. The stapes is attached to the oval window, the boundary between the middle and the inner ear, and its vibration causes a travelling wave to propagate through the fluids in the spiral-shaped structure of the inner ear (cochlea) [38]. Tiny muscles attached to the ossicles contract at high sound levels, restraining the movement of the stapes such that low frequency sounds are attenuated significantly [39]. This phenomenon is known as the stapedius reflex and is driven by the central nervous system. The reflex may prevent damage to the inner ear when we are exposed to loud sounds but its reaction is too slow to protect against intense impulse sounds. A similar feature exists for the eardrum. The tensor tympani muscle is able to pull the malleus medially, tensing the tympanic membrane, damping vibration in the ear ossicles and thereby reducing the amplitude of sounds [40]. This muscle however, is contracted primarily to dampen the noise produced by chewing.

The inner ear is contained in a bony labyrinth cavity (Figure 2) in the temporal bone and comprises two main functional parts: the vestibular system, responsible for sensations of balance and motion and the cochlea, responsible for hearing.

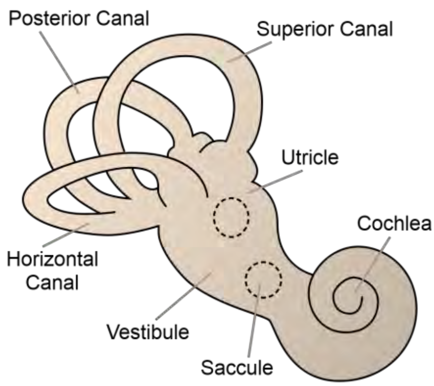


Figure 2: The bony labyrinth of the inner ear containing the cochlea and the vestibular system with the utricle, saccule and its three semicircular canals. [adapted from F. Rogers]

The vestibular system consists of three orthogonal semicircular canals and two vestibular sacs (utricle and saccule). These structures contain the same kinds of fluids and sensory cells (hair cells) as found in the cochlea. They provide sensory information about motion, equilibrium, and spatial orientation [41]. The utricle and saccule detect gravity (vertical orientation) and linear movement. The semicircular canals detect rotational movement. The anatomical vicinity and similarity of the vestibular system and the cochlea make that these systems are also related physiologically and that certain disorders (e.g. Menière's disease) affect both sensory functions [42].

1.1.2. THE COCHLEA

The bony spiral structure of the cochlea is filled with nearly incompressible fluids. Two membranes along the length of the cochlea divide it into 3 compartments: Reissner's membrane separates the scala vestibuli from the scala media and the basilar membrane separates the scala media from the scala tympani (Figure 3) [43].

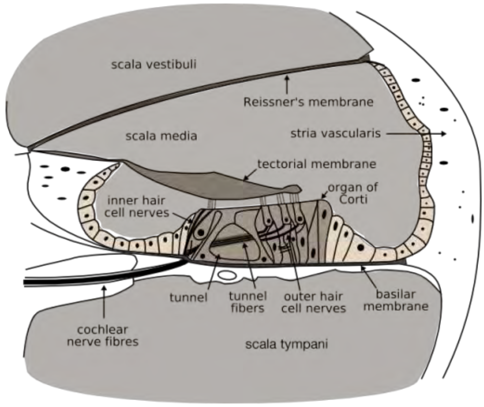


Figure 3: Cross section through a winding of the bony structure of the cochlea showing the basilar membrane, Reissner's membrane, the three scala and the organ of Corti. [adapted from O. Ropshkow]

The side of the cochlea close to the vestibule, where the oval window is found, is called the base. The tip of the cochlea is known as the apex, where there is an opening between the basilar membrane and the wall of the cochlea, called the helicotrema. The helicotrema allows fluid to flow from scala vestibuli into scala tympani and vice versa. An inward movement of the stapes into the oval window causes a corresponding outward movement of the round window, which is a membrane located at the basal part of the scala tympani [44]. The resulting pressure differences between the two sides of the basilar membrane cause it to move perpendicularly in response to the travelling pressure wave traversing the cochlea as shown in Figure 4. For very low frequency pressure waves (e.g., caused by chewing or changing altitude), the presence of the helicotrema prevents an overall pressure difference from building up between both scala vestibuli and tympani [38].

On top of the basilar membrane, in the scala media, the actual receptor organ for hearing is found, called the organ of Corti. This structure contains the receptor cells responsible for converting mechanical waves into electrical signals for transmission to the brain. These receptor cells are called hair cells because of their stereocilia, hair-like mechanosensing structures that respond to the mechanical motion of the basilar membrane [45]. Two distinct groups of hair cells exist: inner hair cells (IHC) and outer hair cells (OHC). The OHCs are located more closely to the outside of the spiral and are organized into 3 rows that run along the length of the cochlear duct (scala media). The IHCs are lined up in a single row located more closely to the modiolus, the central axis of the cochlea containing the spiral ganglion (Figure 5). A human ear has about 12000 OHCs and 3500 IHCs [46]. Through a gelatinous layer, hair cell stereocilia are more (OHC) or less (IHC) connected to a third membrane that runs along the length of the cochlea: the tectorial membrane. Such movement of the basilar membrane will cause the stereocilia to deflect, leading to a depolarization of the hair cell from its resting potential to its receptor potential. In the IHC this results in the release of the neurotransmitter L-glutamate, eliciting action potentials in the ganglion cell dendrites connected to these hair cells [47]. The signal is then transmitted through the auditory nerve fibres to the brainstem.

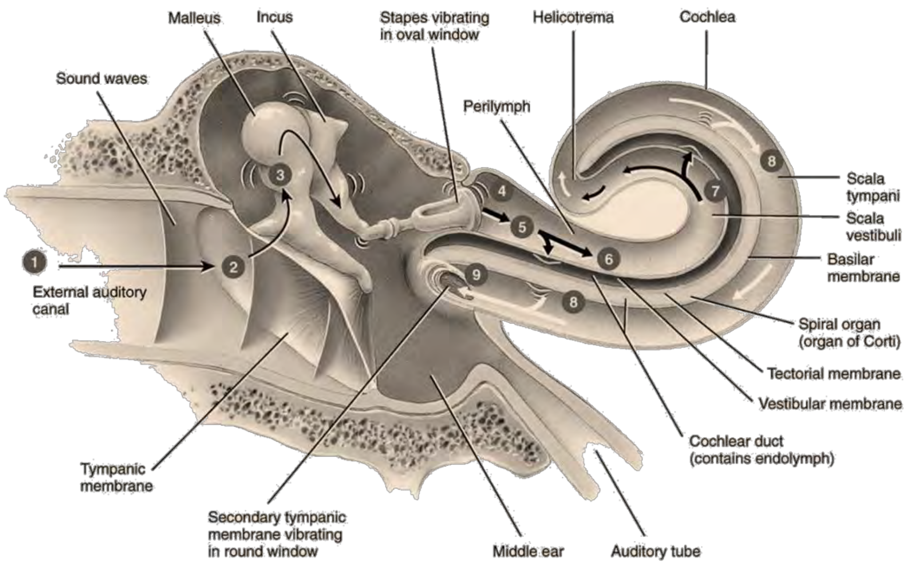


Figure 4: Schematic representation of the structures and workings of the human peripheral auditory system. Sound waves enter the external auditory canal (1) and arrive at the tympanic membrane which converts them into a mechanical motion (2). This motion is transmitted and impedance matched by the ossicles (3) resulting in the stapes vibrating in the oval window (4), which causes a travelling wave to propagate the scala vestibuli from base to apex (5, 6, 7, 8). The travelling wave displaces the basilar membrane maximally at sites that depend on the frequencies contained in the input signal. Either through the helicotrema or by basilar membrane displacement is pressure transferred to scala tympani causing a vibration response of the round window (9) opposite to that of the stapes. [adapted from Tortora, J. Wiley]

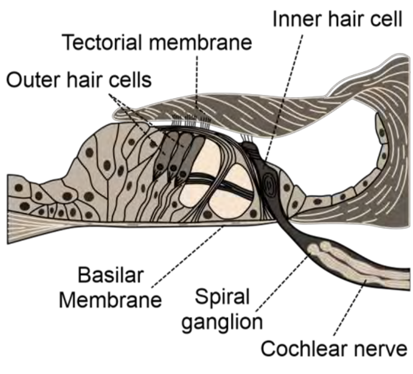


Figure 5: A close up cross section of the organ of Corti showing the tectorial and basilar membranes, inner and outer hair cells, spiral (cochlear) ganglion and the cochlear nerve. [adapted from K.H. Maen]

The place on the basilar membrane where there is maximal displacement depends on the frequencies contained in the travelling wave that originates from stapes vibrations and therefore depends on the sound that is captured by the ear. This is a direct result of the changing mechanical properties of the basilar membrane when spiralling upward from base to apex. At the base the basilar membrane is relatively narrow (100 µm) and rigid, tightly squeezed in supporting cells. Towards the apex, however, it becomes wider (up to 500µm), less stiff and less firmly enclosed by supporting cells. That change in mechanical properties makes that each frequency has a specific site on the basilar membrane that displaces maximally (Figure 6 A) [45]. The place dependent frequency sensitivity of the cochlea is called tonotopy: every site on the basilar membrane, and by extension the hair cells in that vicinity and their associated neurons have a particular frequency to which they are most sensitive (their characteristic frequency, CF). At the base the basilar membrane is most sensitive for high frequencies, at the apex for low frequencies. The relation is approximately logarithmic, which means that each octave is represented by about the same distance (roughly 3mm) on the basilar membrane (Figure 6 B). This mechanism for translating frequencies to sites on the basilar membrane is merely based on the material properties of the basilar membrane. It is therefore a "passive" process contributing to the tonotopic organization of the ear, which can also be observed in a dead cochlea [38].

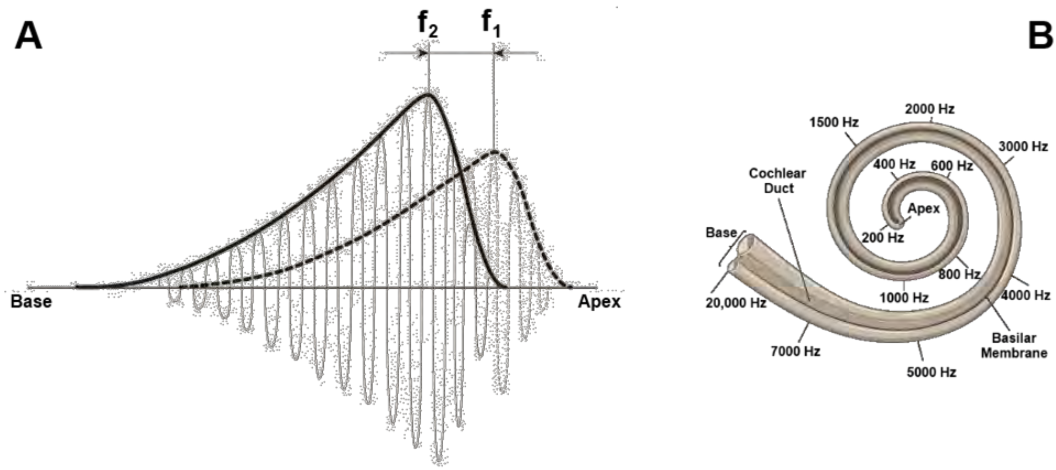


Figure 6: A: schematic of basilar membrane displacement for pure tones of 2 frequencies f_1 (dashed line) and f_2 (solid line) where $f_1 < f_2$. The travelling wave originating from stapes vibration causes maximal displacement of the basilar membrane nearer to the base of the cochlea for higher frequencies and nearer to the apex for lower frequencies. Receptor cells at the place of maximal displacement are stimulated the strongest. After reaching its maximum the magnitude of the displacement decreases relatively quickly. B: the tonotopy of the cochlear is logarithmically organized, spanning about 3mm per octave. [A adapted from Frijns & Schoonhoven, B from Encyclopaedia Britannica, Inc.]

Despite some similarities, IHCs and OHCs serve quite different functions. Afferent neurons, which transmit information from the cochlea to the brain, are connected to the IHCs (each IHC connects to about 20 neurons). Therefore it are the IHCs, converting mechanical waves into electrical current, that generate action potentials in the neurons of the cochlear nerve that eventually delivers the perception of sound [46]. The OHCs on the contrary, do not send signals to the brain. Their function is that of a local, frequency sensitive amplifier. OHCs have a motor function allowing them to change their stiffness, shape and length. In response to the smallest of vibrations in the cochlear duct, OHCs will start contracting and expanding, amplifying these vibrations. Through their connection with the tectorial membrane, the OHCs are able to actively influence the mechanics of the cochlea, more specifically amplify the displacement of the basilar membrane, in such a way that IHCs are able to detect and convert this movement to electrical signals [46]. This active mechanism is also observed when a sensitive microphone is positioned in the ear canal. Sound waves originating from the inner ear can be recorded this way. This phenomenon is known as otoacoustic emissions (OAE) and may be evoked by applying a stimulus, but is also observable when no signal is present at all (spontaneous OAE). These sounds originating from the inner ear are attributed to OHC activity resulting in oscillations on the basilar membrane and vibrations travelling back through the middle ear to the ear canal [48].

In addition to an amplifying role, the OHCs also support the tonotopical tuning of the cochlea. Each hair cell has its own characteristic frequency, at which it will provide maximum amplification [49] (in case of OHCs, maximum transduction efficiency in case of IHCs). This "active" mechanism of OHCs adjusting the mechanics of the cochlea contributes highly to the large range of sound levels that can be processed by the human ear and the high frequency resolution it exhibits. There are also about 1800 efferent nerve fibres (transmitting information from the brain to the cochlea) that are mostly connected to the OHCs. These signals arriving at the OHCs originate from the higher centres of the auditory system and may allow central processes to influence (attenuate or tune) the amplifying behavior of OHCs [50].

1.1.3. HAIR CELLS

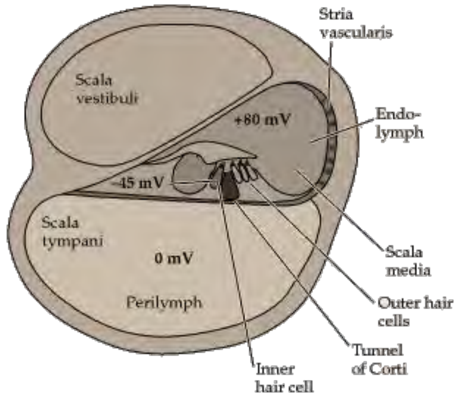


Figure 7: Cross section of a cochlear winding, showing scala vestibuli and tympani (perilymph) and scala media (endolymph). The stereocilia of the hair cells protrude into the endolymph, which is high in K^+ and has an electrical potential of +80 mV relative to the perilymph. [adapted from Purves, Augustine, Fitzpatrick , et al.]

The fluids in scala vestibuli and scale tympani are called perilymph, the scala media is filled with endolymph (Figure 7). Perilymph and endolymph have unique ionic compositions suited to their functions in regulating electrochemical impulses of hair cells. The electric potential of endolymph is about 80 mV more positive than perilymph due to a higher concentration of potassium (K) compared to sodium (Na) [51].

Hair cells can convert the displacement of the stereociliary bundle into an electrical potential in as little as 10 microseconds [52]. Such speed contributes to the faithful transduction of high frequency signals and enables the accurate localization of the source of sounds (using interaural time differences). The need for microsecond resolution places certain constraints on the transduction mechanism, ruling out the relatively slow second messenger pathways used in visual and olfactory transduction. A direct, mechanically gated transduction channel is needed to operate this quickly. The filamentous structures that connect the tips of adjacent stereocilia, known as tip links, directly open cation-selective transduction channels when stretched, allowing K^+ ions to flow into the cell. As the linked stereocilia pivot from side to side, the tension on the tip link varies, modulating the ionic flow and resulting in a graded receptor potential that follows the movements of the stereocilia (Figure 8 C).

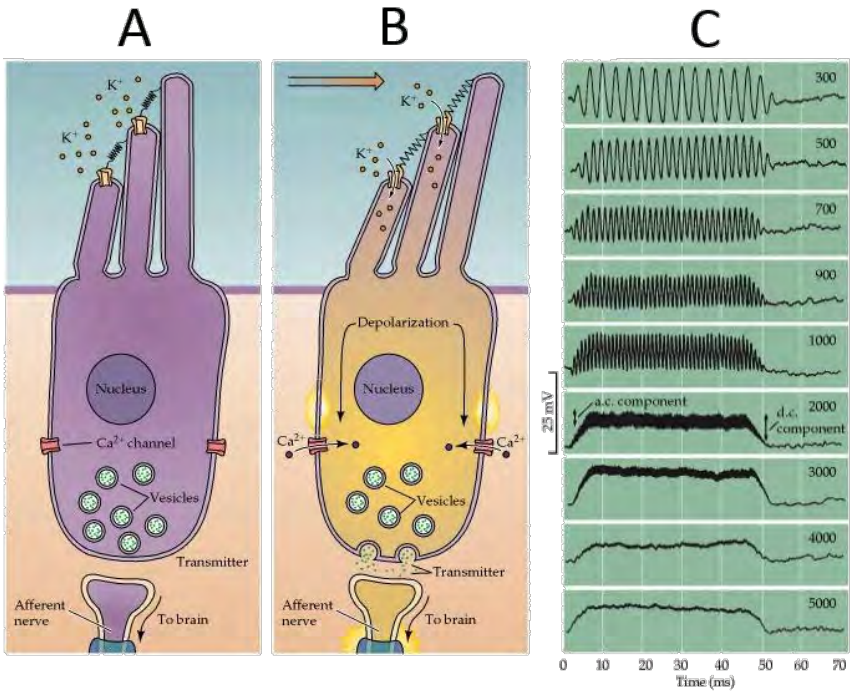


Figure 8: Mechanoelectrical transduction mediated by hair cells. A, B: when the hair bundle is deflected toward the tallest stereocilium, cation-selective channels open near the tips of the stereocilia, allowing K^+ ions to flow into the hair cell. The resulting depolarization of the hair cell opens voltage-gated Ca^{2+} channels in the cell soma, allowing calcium entry and release of neurotransmitter onto the nerve endings of the auditory nerve. C: receptor potentials generated by an individual hair cell in the cochlea in response to pure tones (indicated in Hz at the right of the tracings). Note that the hair cell potential faithfully follows the waveform of the stimulating sinusoids for low frequencies (< 3 kHz), and responds with a DC offset to higher frequencies. [A,B adapted from Lewis & Hudspeth; C from Palmer & Russell]

The hair cell has a resting potential between -45 and -60 mV relative to the fluid (perilymph) that bathes the basal end of the cell. At the resting potential, only a small fraction of the transduction channels are open. When the hair bundle is displaced in the direction of the tallest stereocilium, more transduction channels open, causing depolarization as K^+ enters the cell. Depolarization in turn opens voltage-gated calcium channels in the hair cell membrane, and the resultant Ca^{2+} influx causes transmitter release from the basal end of the cell onto the auditory nerve endings (Figure 8 A, B) [53]. Because some of the transduction channels are open at rest, the receptor potential is biphasic: movement toward the tallest stereocilia depolarizes the cell, while movement in the opposite direction leads to hyperpolarization. This situation allows the hair cell to generate a sinusoidal receptor potential in response to a sinusoidal stimulus, thus preserving the temporal information present in the original signal up to frequencies of around 3 kHz (Figure 8 C) [54]. Hair cell repolarization occurs by positive ions flowing through channels to the perilymph in scala tympani, where concentrations of positive ions are very low.

Unlike many other electrically active cells, the hair cell itself behaves like a typical receptor cell and does not fire action potentials (spike) itself. Instead, the released neurotransmitter diffuses across the narrow space between the hair cell and nearby nerve terminals, triggering action potentials in the nerve [55]. A single auditory nerve fiber generates an action potential when the cell's membrane is depolarized to a threshold value, after which a spike occurs (Figure 9). Moreover, as long as the neuron remains depolarized beyond its threshold level, action potentials, or spikes, will continue to occur. The firing rate (spikes per second) is dependent on the magnitude of the depolarizing current. The greater the current, the faster the spike rate. Many hair cells' resting potentials are of a magnitude that makes associated neurons spike continuously even when there is no external acoustic stimulus [45]. This is known as the spontaneous firing rate.

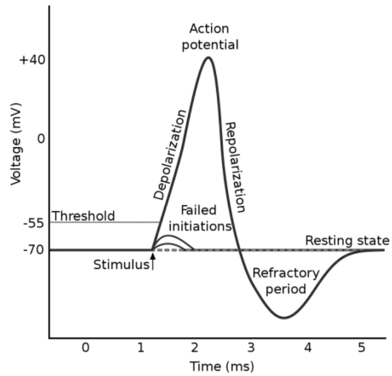


Figure 9: Schematic of an electrophysiological recording of an action potential showing the various phases that occur as a stimulus is applied. [adapted from D. Iberri]

There is obviously a limit to the rate at which action potentials can be generated. For most neurons, the maximum rate is about 1000 spikes per second, although inner hair cell neurons saturate even more quickly (at about 300 to 500 s/s) [56]. In other words, once a spike occurs, it is impossible to generate another one for minimally 1 millisecond. That time is called the absolute refractory period. The period in which the initiation of a next action potential is inhibited (but not impossible, e.g. by applying a larger stimulus) is known as the relative refractory period and may last several milliseconds. As a consequence a single nerve fibre would not be able to keep up with receptor potential frequencies that may be as high as 3 kHz. The fact that every hair cell has a dozen neurons at its disposal compensates for this limited firing rate.

1.1.4. AUDITORY PATHWAYS

The structure of the central auditory system is shown in Figure 10. Unlike most other sensory systems, the auditory system comprises a multitude of complex parallel pathways below the level of the thalamus. The cochlear nerve contains central processes of neurons in the spiral ganglion, which is located in the modiolus of the inner ear [57]. The axons in the cochlear nerve project to dorsal and ventral cochlear nuclei in the brain stem (Figure 12). They do this in a tonotopical manner (i.e., the axons originating in the basal turns of the cochlea, corresponding to high frequency sound, project to the deepest part of the nucleus, whereas the axons arising from the apical turns of the cochlea, mediating low frequency sound, project to the superficial part of the nucleus). Both cochlear nuclei (dorsal and ventral) receive input from the ipsilateral cochlear nerve. These tonotopically organized second order neurons are the source of all central auditory pathways. Like the cochlear nerve and the rest of the auditory pathways, second order neurons exhibit continuous background firing that is increased/decreased by sound driven excursions of the basilar membrane and spiral organ [58].

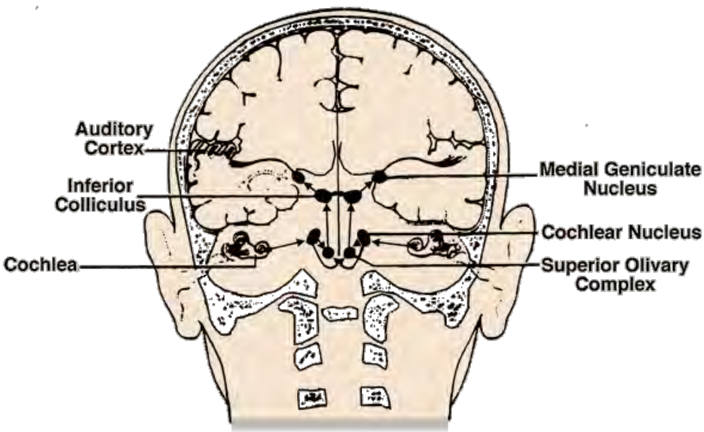


Figure 10: Structure of the central auditory system showing the main pathways as they would appear in a section through the centre of the head cut from ear to ear. [adapted from E. Covey]

Cells in the cochlear nuclei (and other brainstem auditory nuclei) may transform the incoming signal by changing the response pattern (distribution of action potentials over time), increasing the range of latencies, or changing excitatory input to inhibitory output (Figure 11). The changes in discharge pattern are due to differences in the types of ion channels present in the membranes of different cell types, as well as other factors. These transformations are necessary for the different analytical and computational tasks that are performed at higher levels [59].

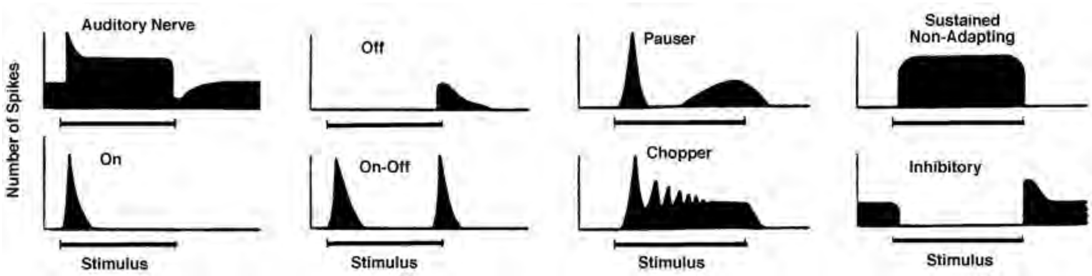


Figure 11: In response to a tone or noise burst, distribution of action potentials over time follows approximately the same pattern in every auditory nerve fibre (upper left). In the absence of sound, auditory nerve fibres discharge continually at a low rate (spontaneous activity). At the onset of a sound, the firing rate increases abruptly to reach a peak, then quickly decays to a lower steady state that is maintained for as long as the sound is present. At the offset of the sound, the firing rate drops below spontaneous rate, but soon recovers. Neurons in the cochlear nucleus (other graphs) exhibit a variety of different discharge patterns in response to the same sound. [adapted from E. Covey]

The superior olivary nuclei on either side of the brain are tonotopically organized too, and receive bilateral auditory inputs from the cochlear nuclei. This nuclear complex can localize sound in acoustic space by discriminating the differences in the time of arrival of sound or the differences in the intensity of sound at each ear. The dorsal nucleus of the trapezoid body sends output to cranial nerves V and VII for reflex contraction of tensor tympani and stapedius muscles to dampen loud sound. This nucleus is also the source of the efferent axons which selectively "tune" the spiral organ for frequency discrimination by affecting the working of OHCs. Although the axons of the neurons from the superior olivary complex and nucleus of the trapezoid body ascend bilaterally in the lateral lemniscus, a majority of these axons ascend in the contralateral lateral lemniscus and project to the nucleus of the lateral lemniscus at the level of the pons-midbrain junction. The nuclei of the lateral lemniscus are important for analyzing temporal patterns of sound in order to identify what type of sound it is and, ultimately, what it means. The neurons in the nucleus of lateral lemniscus, in turn, project to the inferior colliculus (located in the caudal midbrain) [58].

The dorsal portion of the inferior colliculus receives projections from neurons that are responding to low frequencies of sound, whereas the ventral portion receives projections from those neurons responding to high frequencies of sound. The auditory information is then processed (selecting specific features of sound) and relayed by the inferior colliculus to the medial geniculate nucleus of the thalamus, where some form of awareness of sound takes place. The medial geniculate nucleus, also tonotopically arranged, then relays precise information regarding the intensity, frequency, and binaural properties of sound. The axons of these neurons, in turn, project to the primary auditory cortex, responsible for recognizing temporal patterns of sound and direction of pitch change (i.e., elements of melody, speech, etc.). The cortex has separate tonotopic maps for detecting pitch and direction (pitch and direction information is relayed to the cortex by separate pathways). The tonotopic organization of the auditory relay nuclei is maintained in the auditory cortex as well. The auditory association cortex surrounds the primary auditory cortex from which it receives input. The association cortex is required to extract meanings of sound patterns and associate learned significance with a particular sound pattern [58].

It is clear that tonotopy is maintained along the larger part of the auditory pathways, right up to the cortex. In addition the auditory system includes special circuits for computing location of a sound source in space and circuits for analyzing and selecting temporal patterns of sound. The different stages along the path indicate that part of the processing is done sequentially. But the great diversity of cell types, response patterns and neural connections, especially when travelling higher up onto the pathways, suggest that there also are different kinds of processes that are executed in parallel.

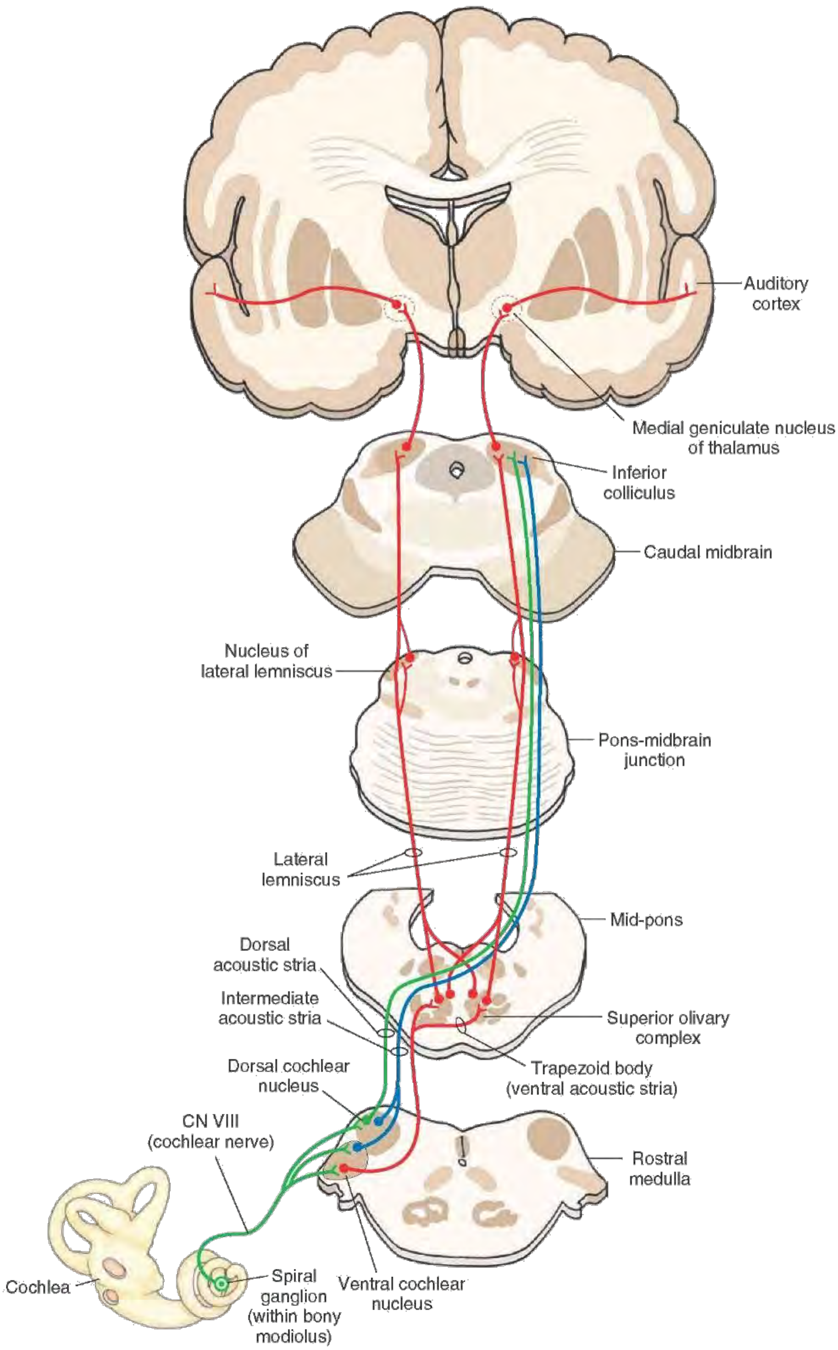


Figure 12: Central auditory pathways. The different stages along the path indicate that part of the processing is done sequentially. The great diversity of cell types, response patterns and neural connections, especially when travelling higher up onto the pathways, suggest that there also are different kinds of processes that are executed in parallel. [adapted from MF Bear et al.]

1.2. CODING OF SOUND FEATURES

1.2.1. SOUND WAVES

Sound is a pressure wave transmitted through a medium, like air. Such a wave has a number of properties (Figure 13). The amplitude is the maximal deviation from the resting position (atmospheric pressure) and relates to the strength of the signal. For periodic (i.e., repetitive) waveforms, the number of cycles per second is called the frequency of the signal, which relates to the pitch of a sound. The inverse of a wave's frequency is the period (the time it takes for one cycle to occur). The distance a wave travels during one period is called the wavelength.

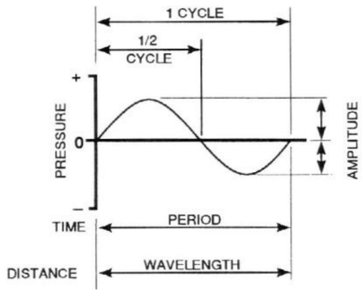


Figure 13: Properties of a sound wave, showing cycle, amplitude, period and wavelength.

Most sounds in real life, speech for instance, contain multiple frequencies. Their waveforms are complex and only locally more or less periodic. It is often more useful to display these sounds through their frequency spectrum, which is the decomposition (Fourier transform) of a complex signal into its pure tone addends (Figure 14).

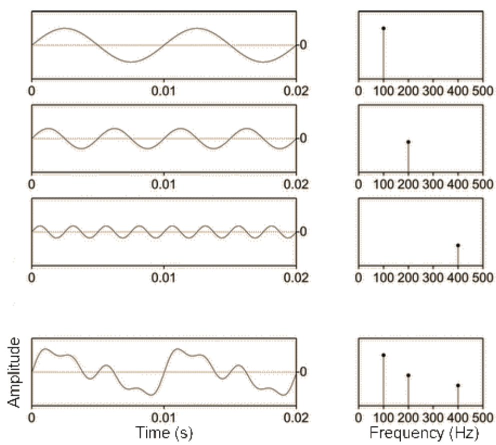


Figure 14: Sine waves of frequencies 100, 200 and 400 Hz respectively, and the sum of these signals (bottom). At the left the waveforms are depicted, at the right their spectra.

1.2.2. INTENSITY & LOUDNESS

In clinical settings it is common to use the term intensity as a measure for the magnitude of a stimulus. Sound intensity (sound power per unit area, expressed in W/m^2) however is a physical quantity that is not very useful for describing the magnitude of an acoustic stimulus, since it is not sound intensity but sound pressure (in Pa) to which the ear is directly sensitive. Sound intensity $\vec{I}$ is a vector that is the product of sound pressure p and particle velocity $\vec{v}$, as shown in Equation (1), which makes that it cannot even be measured with a simple microphone.

$$\vec{I} = p\vec{v} \tag{1}$$

where:

- $\vec{I}$ the sound intensity vector of which magnitude is expressed in W/m^2;
- p is the RMS sound pressure;
- $\vec{v}$ is the particle velocity vector of which magnitude is expressed in m/s.

The magnitude of sound intensity is proportional to sound pressure squared. So there is a direct and straightforward relation between the two quantities, and when the term stimulus/sound intensity is used in this dissertation one should think of the physical magnitude of sound, that is obtained by measuring sound pressure.

Loudness on the other hand is not a physical quantity that can be measured objectively. Instead it is a subjective measure for the psychological correlate of sound pressure/intensity. It is defined as "that attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from quiet to loud". Because the perceived loudness of a sound relates to sound pressure logarithmically, sound pressure is usually expressed in decibels and referred to as sound pressure level (dB SPL). Since the decibel unit expresses a ratio between two values of a physical quantity, SPL is referenced to a standard sound pressure of 20 µPa, which is considered the threshold of hearing for a 1 kHz pure tone. To measure the SPL of an acoustic stimulus, the root mean square value of the instantaneous sound pressures over the time of measurement is taken, as shown in Equation (2).

$$dB\ SPL = 20\log_{10}\left(\frac{p_{rms}}{20\ \mu Pa}\right) \tag{2}$$

where:

- dB SPL is the Sound Pressure Level in dB;
- p_{rms} is the root mean square value of the instantaneous sound pressures over the time of measurement.

The sensitivity of the human ear changes as a function of frequency, as shown in the equal-loudness graph (Figure 15). Each line on this graph shows the SPL required for frequencies to be perceived as equally loud. The ear is most sensitive to sounds around 2 to 4 kHz, with sensitivity declining to either side of this region.

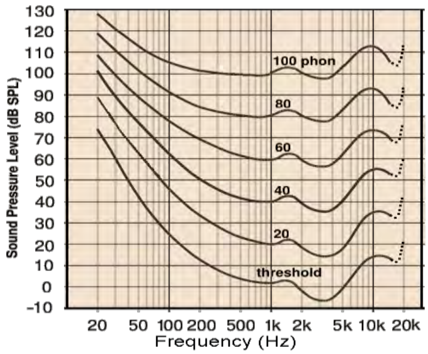


Figure 15: Equal loudness contours (ISO 226:2003), displaying the sound pressure levels of pure tones of different frequencies that cause an equal loudness percept (in phon, 1 phon is equal to 1 dB SPL at a frequency of 1 kHz). [adapted from Lindosland]

When measuring absolute hearing thresholds, clinicians use the dB HL (Hearing Level) scale to express the presentation level of narrow band stimuli. For a particular frequency, 0 dB HL is the minimal sound pressure level at which the population of normal hearing listeners detects the presented signal. In some experiments the dB SL (Sensation Level) scale is used, which indicates a stimulus level referenced to the listener's individual detection threshold for that particular signal.

The perception of loudness is also related to the duration of a sound. The human auditory system averages the effects of sound pressure level (SPL) over a 600 - 1000 ms interval. A sound of constant SPL will be perceived to increase in loudness as samples of duration 20, 50, 100, 200 ms are heard, up to a duration of about 1 second at which point the perception of loudness will stabilize. For sounds of duration greater than 1 second, the moment-by-moment perception of loudness will be related to the average loudness during the preceding 600 - 1000 ms [60].

The human ear has a large dynamic range. The sound pressure at hearing thresholds causes stereocilia to deflect no more than about 0.04 nm [61]. Still, that same ear is also capable of processing sounds with amplitudes 10^5 times greater (hence, yielding a dynamic range of 100 dB), without pain or immediate damage. This is possible because the cochlea amplifies soft sounds more than it does louder sounds, such that an input range of 100 dB is compressed into a response of the basilar membrane of about 50 dB (Figure 16). This compressive nonlinearity is mostly observed in mid-range sound levels (the range of sound levels from 30 to 80 dB SPL is compressed into a response of the basilar membrane corresponding to about 10 dB). This is the result of the active mechanism of OHCs. This mechanism is most effective for low and moderate sound levels, where its amplification may be 50 dB or more [62]. As sound levels increase the amplification reduces gradually, from 50 dB gain for sound levels smaller than 20-30 dB to 0 dB gain for sound levels beyond 90 dB. After that the response of the basilar membrane becomes linear again, as the active mechanism is no longer effective. In listeners where OHC function is lost, the basilar membrane response would be entirely linear (as depicted by the dashed line in Figure 16), resulting in elevated thresholds (+ 50 dB) and a rapid loudness growth for sound levels above threshold.

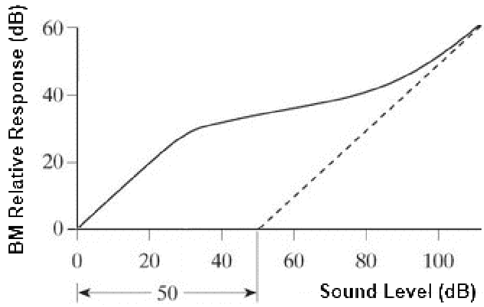


Figure 16: Schematic input-output function of the basilar membrane for a sine tone at characteristic frequency. An arbitrary decibel scale is used for the vertical axis, so that an input of 0 dB yields 0 dB output. The dashed line shows the slope that would be obtained if the response were linear. [adapted from Frijns & Schoonhoven]

Nerve fibres in the cochlea show an increased firing rate when inner hair cells are stimulated. These auditory neurons however do not all have the same spontaneous firing rate. Some of them generate fewer spikes per second in the absence of a stimulus than others. It has been shown that those neurons with low spontaneous discharge rates are also less sensitive [63]. They need a stronger stimulus before their firing rate increases and they do not saturate (reaching their maximum discharge rate) as quickly as their more sensitive, high spontaneous firing rate companions (Figure 17). Due to this variety of more and less sensitive neurons at the disposal of inner hair cells, the large range of sound levels can be mapped into a neural signal carried by neurons that by themselves can only count for level increases over ranges of about 20 (highly sensitive neurons) to 50 dB (low sensitivity neurons). An increase in sound pressure will lead to more neurons increasing their firing rate, gradually recruiting less sensitive neurons. This construction enables the transmitted signal to carry two cues for the brain to decode loudness: firstly, the amount of neurons driven by the IHCs and their compound firing rate, but secondly the notion of which neurons are firing at which rate (less sensitive neurons firing above spontaneous rate suggest an input signal of considerable level).

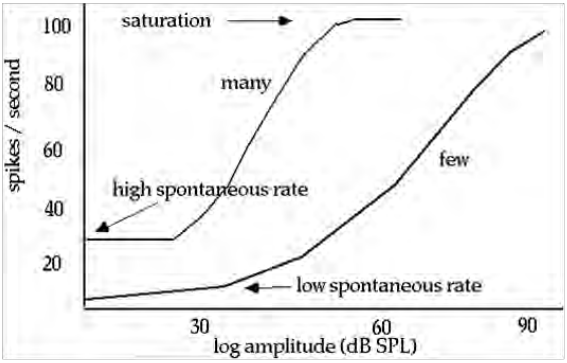


Figure 17: Responses of neurons with low and high spontaneous discharge rates to different stimulus levels. The more sensitive neurons, which are also more common, exhibit higher spontaneous firing rates, but are also the quickest to saturate. [adapted from C. Darwin]

In addition to the fact that, with increasing sound pressure, hair cells having characteristic frequencies that are contained in the signal, will cause more and more neurons to discharge at higher and higher rates, there is also the fact that the displacement of the basilar membrane itself will get broader, triggering more "off frequency" inner hair cells, which in turn leads to an increase in the discharge rate of their neurons. This is a third cue for the brain to estimate loudness.

1.2.3. FREQUENCY & PITCH

Like loudness being a subjective percept related to sound intensity, pitch is the perceptual property that is related to the frequencies contained in a signal. By allowing us to order sounds on the low-high dimension, pitch carries essential information about the tonality and melody in music and about the linguistic context of words and sentences in spoken language (e.g. clause typing) [64] [65]. In the relation between spectral content and pitch, the fundamental frequency of periodic signals is the most important factor. The fundamental frequencies of several competing voices in a noisy environment for example, allow us to distinguish between separate speakers [66]. Pitch perception results from 2 distinct but inseparable cochlear coding mechanisms. One of them is based on place of excitation: tonotopy. The other, called phase locking, is based on temporal coding.

1.2.4. TONOTOPY

One could picture the tonotopical organization of the auditory system as a filter bank. Each site on the basilar membrane would represent a band pass filter that only allows a particular range of frequencies to be passed to the brain. Ideally, any single frequency would only be allowed to pass through one single filter, and no 2 different frequencies would be passed by the same filter. This would require an infinite number of non-overlapping filters of infinitely narrow bandwidth. In the ear, this is obviously not the case. Many different sounds end up in the same filter, disturbing each other's signal, for example when one is unable to hear the phone ringing when taking a shower. This is called masking: the process by which the detection threshold for one sound (the signal) is elevated by the presence of another sound (the masker). Through masking experiments it is possible to uncover the limits of the auditory system's frequency selectivity. In case a signal of particular frequency is masked by a masker of another frequency, then the system has failed to resolve those two frequencies. By observing the conditions necessary to just have a signal be masked, it is possible to obtain an image of the bandwidth and shape of auditory filters. The results of one such an experiment are shown in Figure 18 A and are known as psychophysical tuning curves. The experiment involves presenting the signal (probe tone) at a low level (10 dB SL) and then finding the level for a number of neighbouring masking frequencies (maskers) at which the listener is no longer able to detect the signal. In this experiment probe and masker are presented simultaneously. Masker duration was 500 ms. The probe tone had a duration of 250 ms, and was temporally centred within the masking tone [67].

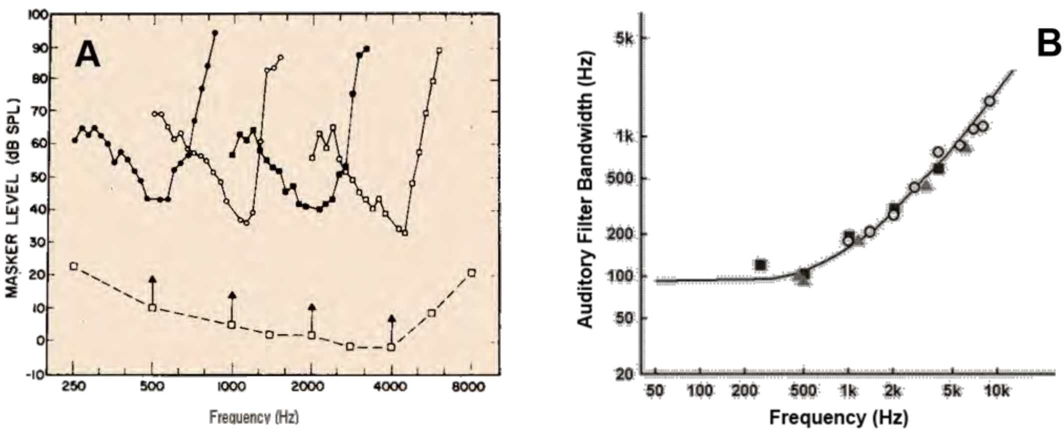


Figure 18: A: psychophysical tuning curves obtained in a normal hearing listener using low-level (10 dB SL) probe tones (signal). The lower function, indicated by open squares, is the pure-tone sensitivity curve in dB SPL. Arrows above the sensitivity curve indicate the level and the frequency of each probe tone. B: critical bandwidth in function of centre frequency as determined by various types of psychophysical experiments. [A adapted from Carney & Nelson, B from B. Sharf]

It is clear from this and other experiments that the bandwidth of auditory filters increases with centre frequency (Figure 18 B). This means that on a linear frequency scale the auditory system is better at resolving 2 low frequency tones than 2 high frequency tones. An auditory filter's bandwidth is also known as a critical band (CB) [68] and refers to the concept of a band of frequencies within which a second tone will interfere with the perception of a first tone by auditory masking. A commonly used approximation [69] of critical bandwidth is given by its Equivalent Rectangular Bandwidth (ERB, Figure 19 A). The approximation is applicable at moderate sound levels and for frequencies between 100 Hz and 10 kHz. Equation (3) shows how to calculate the ERB bandwidth for a particular frequency.

$$ERB(f) = 108f + 24.7 \qquad (3)$$

where:

- ERB(f) is the ERB bandwidth in Hz, for an auditory filter centered at frequency f;
- f is the center frequency in kHz, of the auditory filter.

The auditory filter shape is asymmetric in the sense that the slope towards higher frequencies is steeper than the slope at the low frequency side. The asymmetry causes lower frequency maskers to be more effective on higher frequency signals than vice versa. This phenomenon is known as upward spread of masking and its effect increases with increasing signal levels. Because of these properties auditory filters are often modeled by a gammatone filter bank (Figure 19 B).

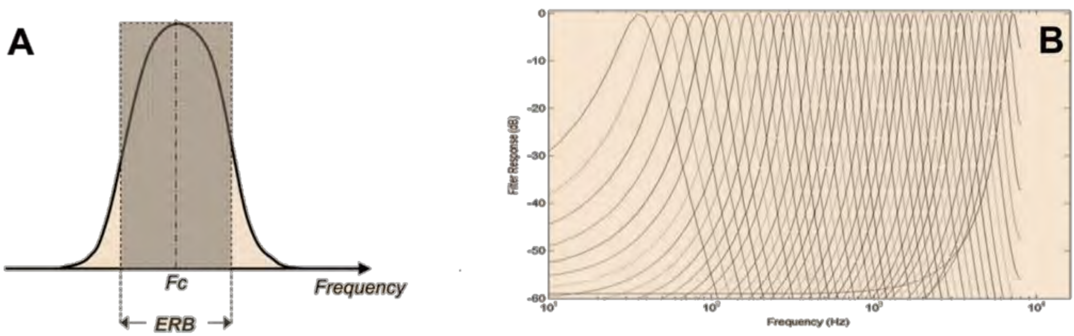


Figure 19: A: an auditory filter (yellow area) and its ERB filter (gray area). They are differently shaped but have equal height and total area (both pass the same amount of energy). B: filter responses of a model of the auditory filter bank, using 40 gammatone filters having 1 ERB bandwidth and approximating the asymmetric shape of the auditory filter. [adapted from J.P. Mason]

1.2.5. PHASE LOCKING

Spectral coding in the cochlea is linked to tonotopy, which codes frequencies by place. However, for sounds with low frequencies (under 3-5 kHz), frequency coding depends on a second, temporal mechanism: phase locking at the level of the IHCs [45]. When the basilar membrane vibrates in response to low frequency signals, the IHCs in the region of vibration exhibit an alternating excitation-inhibition at the frequency of vibration. This, in turn, generates action potentials in auditory nerve fibres attached to those hair cells. The action potentials in the nerve reflect the time pattern of excitation and inhibition in the hair cell. The result is a train of nerve impulses time locked to the individual cycles of the acoustic stimulus. For a sine wave, the impulses are generated around a particular point on the sine-wave cycle, a process that is referred to as phase locking. Because of its refractory period, an auditory nerve fibre may not be able to respond to every successive cycle of a stimulus. When it responds, however, it does so around a constant phase angle of the stimulus [70]. Consequently, the impulses occur around integral multiples of the period of the sine wave stimulus. A population of auditory nerve fibres, all phase locking to the same stimulus, represent in their combined discharge pattern the complete temporal representation of the stimulus (Figure 20).

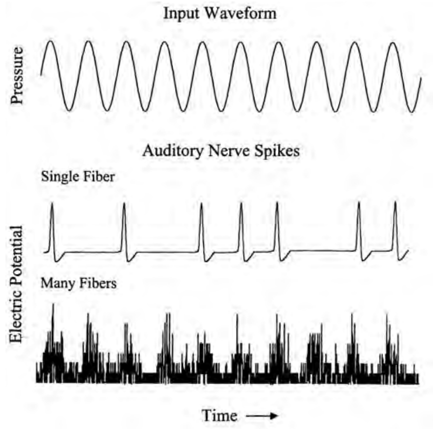


Figure 20: Auditory nerve synchronicity to the phase of an input signal. The biphasic receptor potential of the IHCs and the CF of auditory neurons make that nerve fibres discharge with maximum probability at a time of maximal displacement of the basilar membrane, transmitting a temporal code to the brain that is synchronized to the temporal fine structure of the input signal. [adapted from C.J. Plack]

A time domain representation of sound (the waveform) can be broken down (e.g., by the Hilbert transform) into a slowly varying temporal envelope (TE, the amplitude modulation), and a rapidly varying temporal fine structure (TFS, a frequency modulated carrier). Figure 21 shows an example of such factorization for a filtered speech signal.

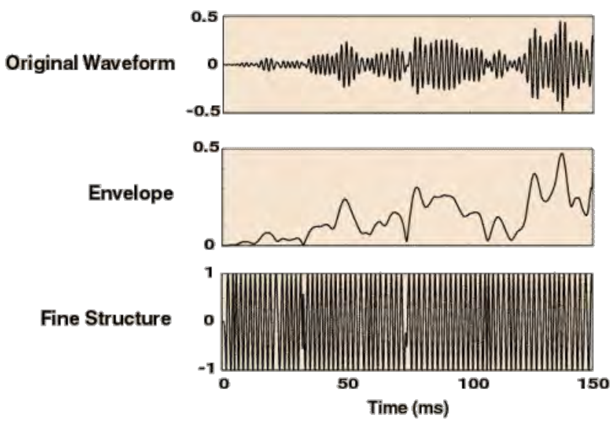


Figure 21: Sound can be presented as the product of a low frequency, amplitude modulating temporal envelope and a frequency modulated carrier (temporal fine structure). [adapted from Swaminathan & Heinz]

One could think of tonotopy (place coding) as the increase in discharge rate of a neural population at a particular site on the basilar membrane in response to and proportional to the amplitude of the temporal envelope at the output of an auditory filter with a centre frequency corresponding to the characteristic frequency of that site. Phase locking would then be viewed as the additional mechanism that makes that population of neurons discharge in synchrony with the temporal fine structure of the band limited auditory filter output.

When a complex sound, like speech is processed by the cochlea, the outputs of the auditory filters present a series of band limited signals to the neural population. As auditory filters broaden towards the base of the cochlea (high frequencies), multiple frequency components, such as the formants in speech sounds, may end up in the same auditory filter, as shown in Figure 22 for an example harmonic complex with fundamental frequency (F0) of 200 Hz.

When 2 or more frequencies in a signal are relatively close together, they cause an alternating constructive and destructive interference on each other, resulting in an amplitude modulated signal having a carrier frequency that is the average of those frequencies and a modulation depth that depends on the relative amplitudes of the individual frequencies. This is called temporal beating and makes that also high frequencies (where synchronicity to the TFS is lost, as explained earlier) may present a temporal code, based on the TE, to their neural population.

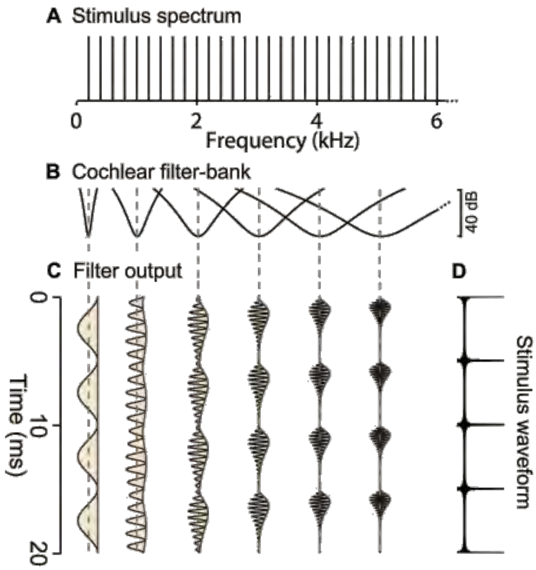


Figure 22: Schematic of Cochlear Filtering, Temporal Fine Structure, and Temporal Envelope Modulation. A: schematic spectrum of a harmonic complex (F0 = 200 Hz). B: cochlear filter bank, with filters centred at 0.2, 1, 2, 3, 4, and 5 kHz. C: waveforms at the output of the corresponding filters of (B) in response to the stimulus (A). For each waveform, the temporal fine structure is drawn centrally, with the temporal envelope (from the Hilbert transform of the signal) running along its peaks at the right. As filter centre frequency increases, the temporal fine structure oscillations become faster, and the temporal envelope becomes increasingly modulating. D: waveform of the 200 Hz F0 harmonic complex. [adapted from C.J. Plack]

1.3. COCHLEAR HEARING LOSS

1.3.1. TYPES AND DEGREE OF HEARING LOSS

Hearing impairment is the most frequent sensory deficit in human populations, affecting more than 350 million people in the world. Consequences of hearing impairment include inability to interpret speech sounds, often producing a reduced ability to communicate, delay in language acquisition, economic and educational disadvantage, social isolation and stigmatization. About 1 to 1.5 out of every 1000 children is born with a permanent hearing loss greater than 40 dB [71] [72] [73] [74]. Most congenital and childhood-onset hearing loss follows from various disease and injury causes. Examples include otitis media, meningitis, rubella, congenital anomalies and nonsyndromic inherited hearing loss. The leading causes of adult-onset hearing loss are presbyacusis (age related hearing loss) followed by noise-induced hearing loss.

Hearing loss is often classified by its degree, measured as the average elevation in pure tone detection thresholds. There is no real consensus, neither on the number of classes to use nor on their names and exact ranges, but generally the idea behind and purpose of these classifications is similar to the system proposed by Clark in 1981 [75], shown in Table 1.

Table 1: Classification of degree of hearing loss, based on average pure tone detection thresholds and proposed by Clark in 1981.

Degree of hearing loss	Hearing loss range (dB HL)
Normal	-10 to 15
Slight	16 to 25
Mild	26 to 40
Moderate	41 to 55
Moderately severe	56 to 70
Severe	71 to 90
Profound	91+

There are two basic types of hearing loss: conductive and sensorineural. Conductive hearing loss occurs when the efficiency of sound transmission through the outer and/or middle ear is reduced. Common causes include too much earwax in the ear canal, damage to the eardrum, the presence of fluids (e.g., from infection) in the middle ear, and damage to or stiffening of the ossicular chain. These result in an attenuation of the signal that arrives at the cochlea, which makes sounds softer and thereby harder to detect. The attenuation may vary with frequency which may result in a somewhat different tonal balance, but in general there are no other perceptual consequences to conductive hearing loss. This type of hearing loss can often be corrected with drugs, or surgically.

1.3.2. SENSORINEURAL HEARING LOSS

Sensorineural hearing loss (SNHL) is the most common type of permanent hearing loss and happens when there is damage to the inner ear (cochlear hearing loss) or to the auditory nerve pathways (retrocochlear hearing loss). Causes of SNHL may be genetic or the result of ototoxic drugs, aging, head trauma, noise-induced trauma, malformations of the inner ear or a tumour in the vicinity of auditory neural structures. The majority of SNHL is due to damage to the structures within the cochlea. Cochlear hearing loss not only reduces the ability to hear soft sounds, but even when speech is loud enough to hear, it may sound unclear, distorted or muffled. Another common symptom of cochlear hearing loss is that, while having a reduced audibility for soft sounds, sounds at high levels are often perceived as having the same loudness as they would for an unimpaired listener. This phenomenon is called "loudness recruitment" and implies that loudness, for people with cochlear hearing loss, grows more rapidly with increasing stimulus levels, which results from the fact that they have a reduced dynamic range for the transduction of the wide range of intensities encountered in our environment (Figure 16) [45].

In general, cochlear hearing loss is related to damage to OHCs and/or IHCs. Their stereocilia may be distorted or destroyed (Figure 23), or the cells themselves may have died. OHCs are the most vulnerable, and when damaged, the effect of the active mechanism in the cochlea may be reduced or lost, which affects the nonlinearity of the basilar membrane response. This results in a decrease in the amplification of soft sounds and the frequency selectivity for which the OHCs are responsible. In a cochlea without functioning OHCs, the basilar membrane would respond (mechanically) as if the cochlea was dead (Figure 24 A).

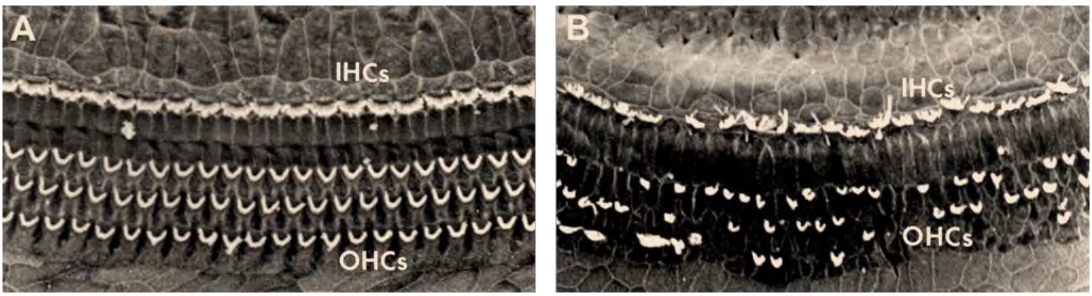


Figure 23: Scanning electron micrographs of the normal (A) and damaged (B) cochlear sensory epithelium. In the normal cochlea, the stereocilia of a single row of inner hair cells (IHCs) and three rows of outer hair cells (OHCs) are present in an orderly array. In the damaged cochlea, hair cells are missing, and stereocilia are abnormal, leading to hearing loss. [adapted from E.M. Keithley]

A signal input to the dead cochlea needs to be more powerful to cause a given displacement of the basilar membrane and the sharp tuning curves found in the in-vivo cochlea, can no longer be observed (they have broadened significantly) [76]. Figure 24 B illustrates the effects of total OHC loss on the neural response activity for a specific site along the basilar membrane. Response thresholds are elevated due to decreased sensitivity. Tuning has broadened due to loss in OHC frequency selectivity and the characteristic frequency of the site on the basilar membrane has shifted towards a lower frequency. The similarity between the effects measured at the mechanical level, and the effects measured at the neural interface, suggest that auditory nerve tuning reflects, at least partially, basilar membrane tuning.

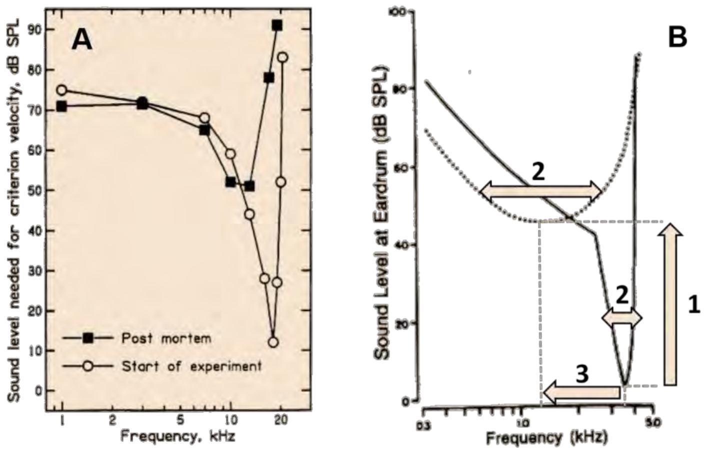


Figure 24: A: measurements of basilar membrane motion in a guinea pig, comparing in vivo and post-mortem basilar membrane tuning curves for the same cochlear region. B: schematic of the effects of total OHC loss with IHCs intact (dotted line), compared to a normal cochlea (solid line), on the sound pressure level needed to elicit neural responses from a specific site in the cochlea: (1) elevated response threshold due to decreased sensitivity; (2) broader tuning due to loss in OHC frequency selectivity and (3) a shift of the site's characteristic frequency towards a lower frequency. [A adapted from Sellick et al., B from B.C.J. Moore]

Inner hair cells represent the transduction mechanism of the cochlea. When IHCs are damaged, this mechanism's sensitivity to basilar membrane displacement is reduced, resulting in elevated detection thresholds (Figure 25 A). As long as OHCs are still functioning normally (which is rare), the sharp tuning remains, although shifted towards higher sound levels. When both IHCs and OHCs are severely damaged, both sensitivity and tuning are greatly reduced (Figure 25 B).

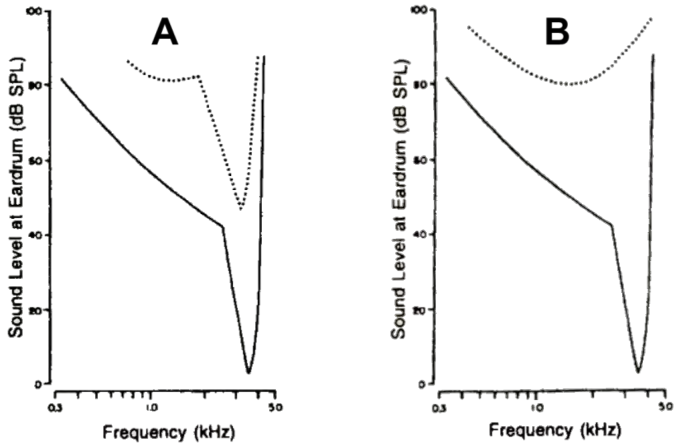


Figure 25: A: the effects of moderate damage to IHCs and OHCs minimally damaged (dotted line), compared to a healthy cochlea (solid line), on the sound pressure level needed to elicit neural responses from a specific site in the cochlea. Detection thresholds are shifted 40 dB upwards but the tuning (OHC function) remains. B: as A, but with severe damage to both IHCs and OHCs. Both sensitivity and tuning are heavily reduced, resulting in threshold shifts of 90 dB or more. [adapted from B.C.J. Moore]

These observations suggest that the primary cause of most types of acquired cochlear hearing loss is damage to OHCs and/or IHCs. For moderate losses, impairment of the active mechanism (OHCs) may be the main cause, resulting in thresholds elevated by up to 50 dB. The loss of absolute sensitivity in such cases is the result of decreased basilar membrane displacement in response to low level sounds. For more severe hearing losses, it is likely that both IHC and OHC functions are impaired. In such cases, on top of losing the active mechanism, the response of the basilar membrane needs to be larger to produce the minimal amount of neural activity for detecting a signal.

At places in the cochlea where IHC function is completely non-existent, there is no transduction of basilar membrane motion at all. Those regions are called dead regions.

There are indications that phase locking may also be affected by damage to hair cells, but the exact reasons for this are unclear. The precision of synchronicity is reduced in such cases and this may result in important perceptual deficits [45].

1.3.3. THERAPIES

Conductive hearing losses can often be treated with drugs or surgery. In some cases the use of a hearing aid or a bone anchored hearing aid (BAHA, Figure 26 A) may be beneficial. A BAHA uses bone conduction to transmit sound directly to the cochlea, bypassing the outer and middle ear. With permanent sensorineural hearing loss, medical and surgical options are very limited. People with moderate to severe SNHL, may benefit from the acoustical amplification of conventional hearing aids (Figure 26 B). These devices use a small loudspeaker positioned near the eardrum and provide acoustical amplification to compensate for elevated detection thresholds.

Figure 26: A: a bone anchored hearing aid, a semi-implantable percutaneous bone conduction hearing device coupled to the skull by an osseointegrated titanium fixture. The system transfers sound to the inner ear through the bone, thereby bypassing problems in the outer or middle ear. B: a conventional acoustical hearing aid, which amplifies sound before it is delivered to the eardrum. It has limited benefit for severe and profound sensorineural hearing losses, since it does not improve the frequency selectivity of the impaired inner ear. [A from Entific Medical Systems, B from Senlan Electronic Factory]

Nowadays also implantable systems exist to provide vibratory stimulation to the middle ear. These middle ear implants (e.g., MED-EL Vibrant Soundbridge, Figure 27) use a floating mass transducer (FMT) surgically fixated to the ossicular chain or directly onto the round window and may alleviate mild to severe sensorineural hearing loss or conductive and mixed hearing losses.

Figure 27: MED-EL's Vibrant Soundbridge middle ear implant system. On the left the externally worn audioprocessor and the implant with floating mass transducer (FMT) are shown. The illustration on the right shows the FMT attached to the incus. [adapted from MED-EL G.m.b.H.]

The benefit of these mechanical/acoustical amplifiers for people with severe to profound SNHL however, is limited because the signal still needs to travel through the heavily impaired inner ear. So while they may improve audibility in these cases, the reduced frequency selectivity of the cochlea remains a huge restriction on speech intelligibility. For severe to profound hearing losses, direct electrical stimulation of the auditory nerve is often the only option. Since almost 30 years now, this can be accomplished by a device known as a cochlear implant.

CHAPTER 2: COCHLEAR IMPLANTS & FITTING

1980: George Watson, Cochlear's second CI recipient, gets his 'portable' speech processor fitted. [from Cochlear Ltd.]

2.1. COCHLEAR IMPLANT DESIGN

2.1.1. HISTORY

In 1800 the Italian physicist Alessandro Volta, also known as the inventor of the battery, inserted a pair of metal rods into his ears. He then applied 50 Volts to them, resulting in an uncomfortable, but auditory, sensation which he described as "uno shok nella testa" (a shock in the head) followed by "un rumore simile a una zuppa densa che ribolle" (the sound of boiling soup). Since then, electrical stimulation of the auditory system has come a long way, thanks to a number of bold and visionary individuals believing that the concept may eventually lead to the restoration of hearing to the deaf. In 1972 for instance, following the work of Lundberg, Djourno and Eyries, Doyle et al., Simmons, Michelson and others, an otologist named Dr. William House and an engineer named Jack Urban teamed up to develop the House 3M single-electrode implant (Figure 28). The device had a portable speech processor and could be taken home. It provided the valuable awareness of sound and cues that aided with lip reading. In most cases however, it was unable to provide open set word recognition. It was approved by the FDA in 1984 and implanted in over a thousand recipients.

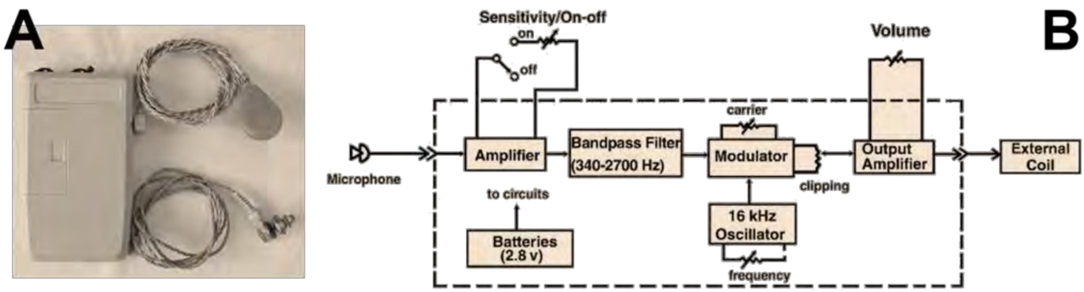


Figure 28: The House 3M single-channel cochlear implant. A: the body worn speech processor unit with transmitter coil and microphone cable. B: signal processing diagram of the House 3M implant system. [A adapted from HearingAidMuseum.com, B from Fretz RJ, Fravel RP]

Also during the late seventies and early eighties, Graeme Clark (professor in Otolaryngology in Melbourne, Australia) and his co-workers, made notable progress in developing a multi-channel cochlear implant. Their device was put to market by the company Cochlear Ltd. in 1984 under the name "Nucleus 22". The multi-channel implant improved spectral perception and speech recognition capabilities significantly. In the mean time, similar efforts were conducted in Austria (the MED-EL implant), California (the AB Clarion implant), France (Neurelec) and also in Antwerp (the LAURA implant).

Today, 30 years later, implantation of the multichannel CI has become a routine intervention. For people with severe to profound hearing losses, in whom the acoustical stimulation of a hearing aid is unable to provide sufficient information for adequate speech perception, a cochlear implant (CI) is often the only solution. The contemporary CI has a series of electrodes surgically implanted in the inner ear and provides direct electrical stimulation of the auditory nerve. This way, the entire outer, middle and large part of the inner ear, such as damaged hair cells, are bypassed. Still, a CI needs a population of auditory nerve fibres to stimulate, but in the larger part of candidates the neural interface is sufficiently present for restoring a useful hearing sensation.

2.1.2. DEVICE STRUCTURE AND FUNCTIONS

The modern cochlear implant consists of an internal, surgically implanted part (Figure 29 B) and an external part (Figure 29 A) that is usually worn behind the ear. The external part is called the speech processor (SP), which has a microphone to pick up sounds from the environment. The processor converts these analog sounds into a digitally coded signal, which it sends via an RF (radio frequency) transmitting coil to the receiver-stimulator implanted under the skin. The receiver-stimulator decodes the signals transmitted by the sound processor into an electrical stimulation pattern to be delivered to its electrode array. This array is usually positioned in the scala tympani by means of a cochleostomy or an insertion through the round window. The array consists of about 20 electrode contacts running along the length (up to 2 whorls) of the cochlea (Figure 29 C) and provides direct stimulation of spiral ganglion cells in the modiolus.

Figure 29: The main parts of a cochlear implant. Its external parts (A) are a speech processor (1) that can be worn behind the ear and a transmitter coil (2). Internally (B) there is a receiver-stimulator (3) and the electrode array (4) positioned in the inner ear (C). [A,B adapted from Cochlear Ltd., C from sacic.com.au]

Four major CI manufacturers have commercially available systems on the market (Figure 30) anno 2013. Market leader is Cochlear Ltd. based in Australia. MED-EL is an Austrian CI manufacturer. Advanced Bionics' headquarters has long been in California, but the company (AB) has been acquired by the Swiss holding Sonova in 2009. Neurelec, the smallest player, is located in France.

Figure 30: Current generation speech processors of the 4 CI manufacturers: from left to right: Cochlear's Nucleus System, MED-EL's OPUS 2 processor, AB's Harmony processor and Neurelec's Saphyr processor.

Over the years, many types of electrode arrays have been produced by the different manufacturers. The number of intracochlear electrode contacts ranges from 12 in the MED-EL implant to 22 in Cochlear's system (Figure 31 A). In addition to the intracochlear electrodes (Figure 31 B), 1 or 2 electrodes are positioned outside of the cochlea (often as part of the casing of the internal receiver-stimulator unit). These electrodes serve as reference or 'ground' electrodes. Basically 2 modes of stimulation can be accomplished with these implants: bipolar, in which a current circuit is established between the stimulating electrode and one of its neighbours, and monopolar, in which one (or both) of the extracochlear electrodes serves as ground electrode (Figure 31 C).

Figure 31: A: a cochlear implant electrode array (Cochlear's Nucleus Contour electrode). Its shape is precurved to minimize damage to cochlear structures and tissues during insertion. In addition, this type of perimodiolar (or "modiolus hugging") electrode minimizes the distance between electrode contacts and spiral ganglion cells that are target for stimulation. B: a phase-contrast X-ray image of Cochlear's Nucleus Contour electrode, a perimodiolar banded precurved electrode array, inserted into a cochlea. C: schematic of bipolar and monopolar electrode coupling. Solid lines indicate current paths of the actual stimulation cycle (time slice). Dotted lines represent a next time slice (assuming a sequential stimulation strategy). In bipolar stimulation mode a current circuit is established between the stimulating electrode and one of its neighbours, resulting in a locally constrained current flow. In monopolar modes, current flows from the actual stimulation contact to the extracochlear electrode(s). [A adapted from cochlear Ltd., B from Xu J et al., C from Clark GM]

Bipolar stimulation modes were more frequently used in the past. They enable a more locally constrained flow of current, aiming at stimulating a specific site of cochlea and thus improving the tonotopical resolution. A drawback is that, because of its relatively narrow current spread, bipolar stimulation requires higher current levels to reach/stimulate spiral ganglion cells and consequently to bring forth a given sound percept when compared to monopolar mode. This has its implications on battery usage and in some cases, when electrode impedances (resulting from resistive properties of the electrode contact, cochlear fluids and adjacent tissue) are high, the system may not be able to supply the voltage required to obtain the desired current level. Because of these reasons, and supported by a tendency towards the use of perimodiolar ("modiolus hugging") electrodes, all manufacturers now provide monopolar electrode coupling as their default stimulation mode. The proximity of those electrodes to the modiolus and thus of their electrode contacts to spiral ganglion cells, makes that with monopolar stimulation, despite its wider current spread, the target population of auditory nerve fibres is specific enough to provide the speech processor with an adequately tonotopical interface.

2.1.3. STIMULATION STRATEGIES

To mimic the tonotopical organization of the normal cochlea, a CI aims at providing different stimulation sites for different frequency bands. So an essential step in every multichannel cochlear implant is the distribution of the input signal spectrum over its different channels. This is accomplished by implementing a series of band pass filters such that they each pass a portion of the spectral input to their corresponding channel (Figure 32). Such a filter bank may be implemented by a Fast Fourier Transform (FFT), as in Cochlear's and Neurelec's default strategies, or a series of Finite Impulse Response (FIR) filters, which AB and MED-EL use for their default strategies. Most FIR based strategies use the envelopes of the outputs of the filter bank to modulate the amplitudes of pulses delivered to their corresponding electrodes. Envelope detection is typically implemented by low pass filtering and half- or full-wave rectification of the filter output, or by using the Hilbert Transform. FFT based strategies have no need for dedicated envelope detection because the output of the FFT is no longer a time domain representation of the signal. Instead the amplitudes of the FFT output bins are combined into channels according to a frequency allocation table (FAT), which results in a fluctuating band limited energy for each channel, which is conceptually equivalent to the band limited signal envelopes obtained in FIR filter based strategies. In any case, those band limited signal envelopes are then compressed into ranges that are suitable for electrical stimulation by their corresponding electrodes (i.e., the Electrical Dynamic Range, EDR). A pulse train's amplitude is modulated with the compressed envelope for each channel and delivered through the RF transmitter coil to the internal receiver, which in turn applies the appropriate voltage to the electrode contacts to obtain the desired current levels.

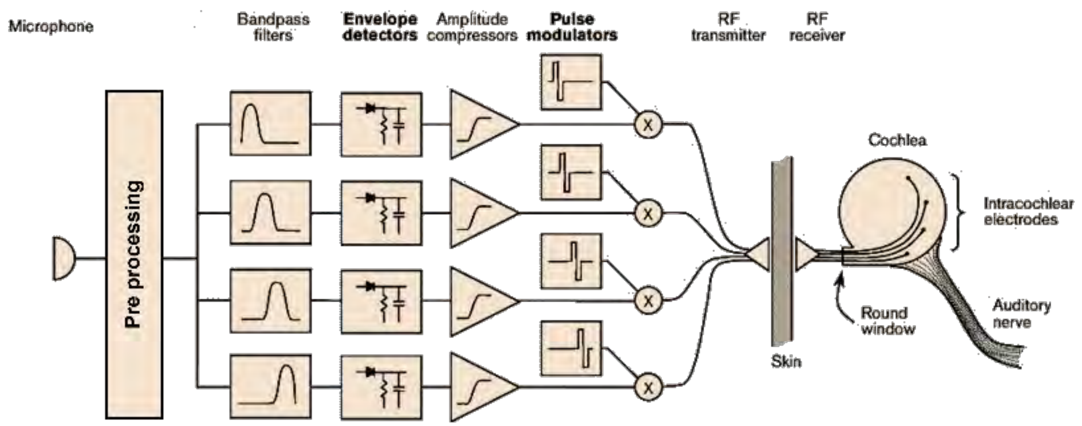


Figure 32: simplified schematic of the signal processing in current generation CI systems, showing 4 channels. An acoustic signal is captured by the microphone and pre-processed. It is then divided over the different channels by a bandpass filter bank. The envelope of the signal at the output of each filter is then compressed into a range suited for electrical stimulation. A pulse train's amplitude is modulated with this compressed envelope and delivered through the RF transmitter coil to the internal receiver, which in turn applies the appropriate voltage to the electrode contacts to obtain the desired current levels. [adapted from Rubinstein]

Although both legacy and experimental stimulation strategies exist that use other pulse types, the most common pulse shape used in current CI systems is a biphasic rectangular pulse (Figure 33 A). Biphasic pulses have the advantage of causing the net current through the tissue to be zero, avoiding unwanted long-term electrochemical effects. These pulses are delivered to the electrode contacts according to a specific stimulation strategy (also known as speech coding strategy).

Speech coding strategies may either use a sequential stimulation or a (partially) simultaneous stimulation. Sequential stimulation strategies like Continuous Interleaved Sampling (CIS) stimulate electrodes one after the other, such that no 2 electrodes are active at any instant in time (Figure 33 B). Stimulating multiple electrodes at the same time may yield an unpredictable loudness percept and reduced tonotopical resolution (less focused stimulation) because of channel interactions (addition of voltage fields). Most current coding strategies therefore use sequential stimulation, or almost sequential stimulation, like the HiRes-P strategy from AB, in which electrodes are stimulated in pairs that are at a considerable (half-array) physical distance from each other (like electrode 1 and 9, 2 and 10, etc.) in their 16 electrode array.

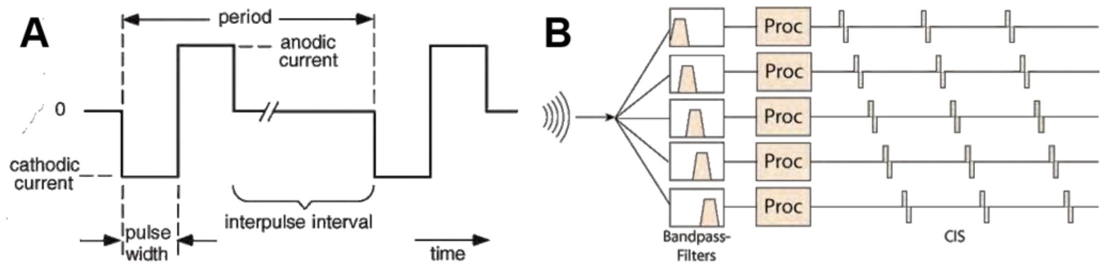


Figure 33: A: current generation CIs use balanced biphasic pulses to make sure the net current through the tissue is zero. The pulse width is defined as the time of a single pulse phase and the inter-pulse interval is the time between 2 pulses. B: a simplified schematic of sequential stimulation, showing 4 channels. The commonly used Continuous Interleaved Sampling (CIS) strategy and its offspring use sequential stimulation to minimize channel interaction. [A adapted from May F, B from Haslwanter T]

Seeing that a CI only has about 20 electrode contacts, the number of distinct sites along the cochlea to stimulate is very limited. This in contrast to a healthy cochlea with its 3500 IHCs enabling signals to be transduced with high spectral resolution. In an attempt to increase this resolution for CI stimulation, some manufacturers have provided strategies that use "virtual channels". Those strategies (e.g. AB's HiRes Fidelity 120 strategy) stimulate 2 adjacent electrodes at a time, in such a way that the proportion of current between the 2 represents the position of spectral peaks in relation to those channels' centre frequencies. This way, virtual channels are created that allow the number of distinct pitch percepts for a CI recipient to extend beyond the number of physical electrodes. Studies [77] have shown that even when electrodes are stimulated sequentially (as in MED-EL's FSP strategy, which implements the concept of virtual channels by using overlapping bell-shaped bandpass filters), the perceived pitch of a frequency in-between 2 bandpass filters' centre frequencies is intermediate to the single-electrode pitches.

Early CI systems featured relatively slow stimulation rates (the number of pulses delivered to an electrode per second), like 250 pulses per second (pps) or even less. Today's stimulation rates are typically around 1000 pps and may be as high as 5000 pps. Stimulation at these rates is not perceived as a burst of pulses, but rather as a continuous signal. Changes in the rate of stimulation can affect a CI recipient's perception of loudness and/or pitch. Increasing the rate often results in a louder percept, because of temporal summation of the stimulus. Changes in rates below 500 pps may also affect the pitch percept, due to changes in the temporal code (similar to phase locking) that is presented to the auditory nerve. Above 500 pps, further increases are unlikely to result in the perception of a higher pitch. Because of these effects, most strategies keep the stimulation rate fixed, once it has been set. An exception to this is MED-EL's FSP strategy family (Figure 34). They intentionally adapt the rate of stimulation on a subset of channels (apical channels) to changes in the temporal fine structure of the input signal (pulses for these channels are triggered by zero-crossings in the bandpass filter's output), aiming at transmitting temporal fine structure cues, such as fluctuations of the fundamental frequency of a signal.

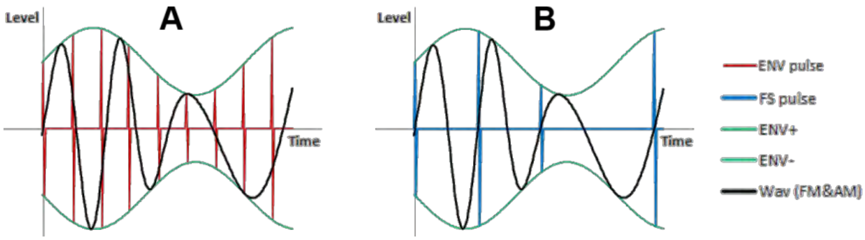


Figure 34: Illustration of MED-EL's FSP strategy (B) compared to a purely envelope based strategy (A). The temporal envelope (green lines) is used to modulate the amplitude of pulses (red and blue lines). In the FSP strategy, in addition to using the envelope, the temporal fine structure is used to determine the stimulation rate in a subset of apical channels. Pulses for these channels are triggered by zero-crossings in the bandpass filter's output (black lines). [Adapted from MED-EL G.m.b.H]

Some strategies, known as "n-of-m" strategies (e.g., Cochlear's Advanced Combination Encoder (ACE) strategy), stimulate only a subset of the electrode array at any given stimulation cycle (frame). M is typically the number of enabled channels, and n (<= m) is a value that can be configured in the speech processor. The choice of which n channels to stimulate during a given stimulation cycle is usually determined by the position of spectral peaks in the input signal. Only the n channels showing the highest signal levels in the signal analysis frame at that time are selected for stimulation, a process known as "Maxima Selection". Because the total stimulation rate of all channels combined is limited by the implant hardware, n of m strategies allow for a higher stimulation rate to be set per channel. Also, the amount of channel interaction is reduced and in some listening conditions, it may be the case that maxima selection improves the signal to noise ratio, as frequency bands containing relatively low energy (presumably the noise) do not lead to stimulation of their electrodes.

Temporal envelopes of speech signals convey important cues for speech understanding. CI recipients may theoretically benefit from a relatively high stimulation rate because it allows for a higher temporal accuracy when representing the input signal envelopes as pulse trains. Figure 35 shows the effect of stimulation rate in processing the syllable /ti/. The bottom waveform is the original speech envelope of channel 5 for this syllable. As seen in the 200 pps stimulation rate condition, pulses are spaced relatively far apart, so this sort of processing may not be able to extract all of the important temporal information contained in the original waveform. When a higher pulse rate is used, the pulses are placed more closely to one another so they can carry the temporal fine structure more precisely [78].

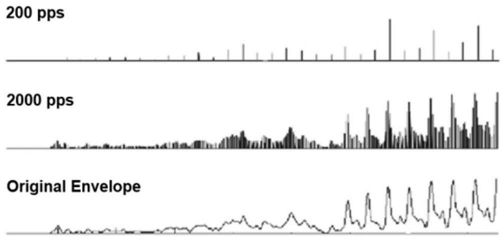


Figure 35: The pulse trains for channel 5 (centre frequency of 3316 Hz) of the syllable /ti/ with stimulation rates of 200 pps and 2000 pps. The syllable /ti/ was band pass filtered into six channels and the output was rectified and sampled at the rates indicated in this figure. The bottom panel shows the speech envelope of channel 5 for syllable /ti/. [adapted from Loizou et al.]

Stimulating the auditory nerve with electrical pulses at rates less than 2000 pps, causes the fibres in the vicinity of an electrode contact to discharge in a highly synchronous fashion. This is in contrast with a normal cochlea, where the receptor potentials of IHCs of varying sensitivity (different spontaneous firing rates) trigger action potentials in nerve fibres rather stochastically. The synchronous discharge when stimulating electrically results in a simultaneous refractory period for most of the targeted fibres, during which the auditory system is unable to respond to subsequent stimulation. This poses a constraint on temporal accuracy and consequently on the perception of rapid fluctuations in a stimulus. With rates above 2000 pps auditory nerve fibres are more likely to assume a stochastic discharge pattern, which is expected to improve the conveyance of temporal fine structure cues.

So, from signal processing and physiological points of view it seems reasonable to expect a performance improvement with higher stimulation rates, but in practice speech intelligibility of CI users is often not improved with increased rates. Instead, various research has shown that the optimal stimulation rate varies greatly within individual CI recipients and may range from a few hundred pps to 5000 pps [79] [80] [81] [82].

2.1.4. SIGNAL PROCESSING

To obtain a tonotopical organization for electrical stimulation, CI electrodes are positioned along the length of the cochlea. This means that they may all face very different conditions when targeting auditory nerve fibres for stimulation. The surviving neural population in the vicinity of specific electrode contacts, or the physical distance to spiral ganglion cells in the modiolus may vary considerably throughout the array. This requires that every electrode operates within its own range of stimulation levels (EDR). That range is typically configured between the smallest level that is detectable (EDR Minimum) and the level that causes a loud but still comfortable percept (EDR Maximum). A typical EDR size for a CI electrode is about 6 to 12 dB (of possible increase in charge per pulse phase). This means that the auditory nerve has a very limited dynamic range for electrical stimulation. For that reason a CI needs to compress the >100 dB acoustical range available to normal hearing into an electrical range that is a factor 10^4 smaller. This is usually accomplished by 2 types/stages of compression: (1) a long term compression of the broadband input signal and (2) an instantaneous compression of the band limited signal within each channel.

Long term compression is typically performed by automatic gain control (AGC) systems at the CI's front end processing stage (Figure 32). They act as volume controls that adapt the overall sensitivity of the system such that the range of sound levels to be processed at any given time is optimal for each environmental signal level (Figure 36 A). The range of input levels that is processed at any given time typically has a size somewhere between 40 and 60 dB, and is a configurable parameter in the speech processors of most CI systems. This means that at any instant in time, only about half of the >100 dB of acoustical range is considered for processing and as a consequence, candidate for contribution to the perception of loudness growth. Different manufacturers refer to this limited range of instantaneous input with different names (Cochlear uses Instantaneous Input Dynamic Range (IIDR), MED-EL uses Adaptive Sound Window (ASW) and AB just refers to this range as Input Dynamic Range (IDR), while IDR in the terminology of other manufacturers is the entire range of inputs that is covered when the instantaneous input range is shifted by AGC systems in response to the overall environmental sound level). To avoid confusion, in this dissertation the term Instantaneous Mapping Range (IMR) will be used, because it is manufacturer-neutral and reflects the concept of a limited range of levels that is mapped into a channel's EDR at any given instant in time quite well.

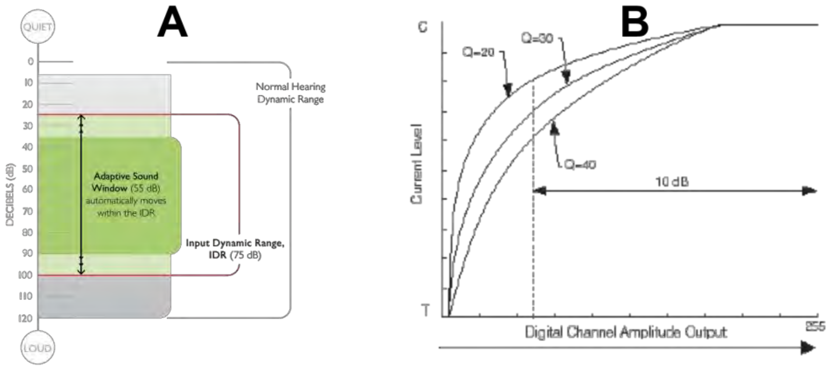


Figure 36: The two stages of signal level compression in contemporary CI systems. A: schematic of MED-EL's Automatic Sound Management. It accomplishes long term compression of the broadband input signal by limiting the range of levels that is processed at any given time to 55dB (in this dissertation referred to as the Instantaneous Mapping Range (IMR)). MED-EL calls this range the Adaptive Sound Window and it is automatically positioned within MED-EL's 75dB input range (IDR), based on the overall environmental sound level as analyzed by the AGC system over the past few hundred milliseconds. In quiet environments this means that the system becomes more sensitive, while in loud/noisy environments the sensitivity is reduced. B: Cochlear's instantaneous compression of band limited signal amplitude depicted as the Current Level delivered to an electrode in function of the envelope amplitude at that channel's bandpass filter's output. T (for threshold) level is Cochlear's proprietary name for EDR Minimum, and C (for comfort) for EDR Maximum. The compression can be configured in the processor by means of the Loudness Growth (Q) parameter, which controls the steepness of the amplitude growth function and determines the percentage of a recipient's EDR that is allocated to the top 10 dB of the sound processor's input dynamic range. A low Q-value makes the loudness growth function steeper and has the effect of making soft sounds perceptually louder. [A adapted from MED-EL G.m.b.H., B from Cochlear Ltd.]

Instantaneous compression is performed at the level of the band limited signal within every channel. When considering the amplitudes at the outputs of band pass filters in a linear scale (proportional to sound pressure in Pa), all CI systems use highly compressive logarithmic mapping functions (e.g., Figure 36 B shows Cochlear's compression) to make sure an IMR in which the largest amplitudes are more than 300 times (50 dB) larger than the smallest amplitudes it needs to cover, is properly mapped into an EDR in which the largest current value that can be delivered without discomfort is only about 5 times (10 dB) bigger than the smallest current that is perceivable. However, when the filter output is expressed in dB (proportional to sound pressure level, dB SPL) the mapping function in all CI systems is approximately linear (the loudness growth expressed in units of charge per dB remains roughly constant across the EDR) when the default processor configuration is used. Equation (4) shows the basic mapping function that is common to all CI systems:

$$Q = EDR_{min} + (A_{ch} + G - IMR_{\min}) * \frac{DR}{IMR} \qquad (4)$$

where:

- Q is the charge per pulse phase in nanoCoulomb (nC) that is output by the channel;
- EDR is the size of the Electrical Dynamic range in nC of the channel;
- EDR_{min} is the channel's EDR minimum value (threshold of electrical stimulation level) in nC;
- A_{Ch} is the envelope amplitude at the channel's bandpass filter output in dB;
- G is the channel gain in dB that is configured for the channel in the speech processor;
- IMR is the size of the Instantaneous Mapping range in dB;
- IMR_{min} is the IMR minimum value (the smallest amplitude that is mapped into the EDR) in dB

It is clear from Equation (4) that, within a channel, instantaneous compression is increased when either the EDR is reduced or the IMR is enlarged, or both. Cochlear, MED-EL and Neurelec do provide an additional parameter to adjust this relation in the speech processor (assigning more/less loudness growth to higher/lower intensity ranges), but in most cases it is left untouched [83]. The EDR is configurable per channel (a channel parameter), while the IMR is common to all channels (a global parameter). Most CI systems also provide a global parameter to adjust the microphone's sensitivity and a channel parameter to add a fixed gain to the input (or output, e.g., in AB's system) of a channel's mapping function. Other processes commonly found in CI systems that impact the magnitude of electrical stimulation for a particular channel given an input signal level include: (1) a pre-emphasis filter, which resembles an A-weighting filter for the purpose of attenuating low and very high frequency sounds (similar to the equal-loudness curves of Figure 15); (2) the gain proposed by the AGC system as appropriate for the given overall environmental sound level and obviously (3) the channel's band pass filter characteristics in relation to the input signal. In summary, these effects can be described by Equation (5).

$$A_{Ch} = S + A + PRE(s) + AGC(s) + BPF(s) \tag{5}$$

where:

- A_{Ch} is the envelope amplitude in dB at the channel's bandpass filter output;
- S is the microphone sensitivity in dB;
- A is the amplitude of the input signal in dB;
- PRE(s) is the attenuation resulting from the pre-emphasis filter (given the spectrum of the input signal s) in dB;
- AGC(s) is the gain in dB proposed by the AGC system and determined from the overall level of the input signal s;
- BPF(s) is the attenuation in dB resulting from the channel's band pass filter in response to the input signal s.

2.1.5. TECHNOLOGICAL EVOLUTIONS

Over the years CI processors have evolved towards more and more sophisticated signal processing devices, aiming at providing a stimulation pattern that results in an optimal auditory percept. During the past few years, this evolution has been further stimulated by a number of mergers between CI manufacturers and hearing aid companies. The advanced technology found in the front end signal processing of acoustical hearing aids is being ported to CI processors at a high pace. The use of multiple microphones for beamforming, algorithms for noise reduction, de-reverberation and wind noise suppression and auto-classifiers for selecting optimal processing parameter values based on environmental or listening conditions are examples of stages in a CI's signal processing path that also affect the relation between an input signal's intensity and the electrical output that is delivered to the electrodes. Given that these processes are highly dependent on the temporal and spectral features of the input signal, it is difficult to predict their effect on the output of a CI.

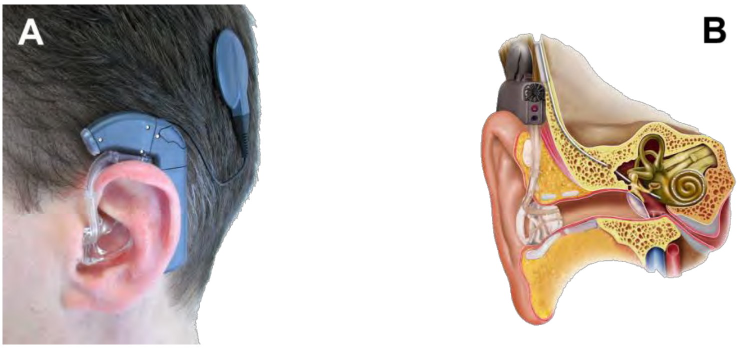


Figure 37: the MED-EL DUET 2 processor providing electroacoustic stimulation (EAS). The device is very similar to a regular CI system, except for the acoustical amplification it provides to stimulate residual low frequency hair cells. B: drawing of an implanted EAS system, showing the ear mould positioned in the external ear canal for acoustical stimulation and the electrode array inserted in the cochlea for electrical stimulation. The speech processor is worn behind the ear and provides acoustical amplification of low frequency sounds while transmitting high frequency sound information via the RF coil to the implant. [adapted from Zipfer]

Another recent evolution in cochlear implantation is the use of electroacoustic stimulation (EAS, Figure 37). In subjects that have some low-frequency hearing left, CI surgery is now often conducted with the preservation of that residual hearing in mind, aiming at the possibility of acoustically stimulating the remaining hair cells with low frequency sound while electrically stimulating the nerve fibres that transmit mid and high frequency sounds.

2.2. COCHLEAR IMPLANT FITTING

The central auditory system undergoes a huge development during the first years after birth. Stimulation of the auditory system during that period makes that the brain organizes itself to process these kinds of information. This is called brain plasticity and this ability is gradually reduced after the first years of infancy. In the absence of auditory stimuli, parts of the brain that are normally assigned to hearing will largely be taken over by other functions (Figure 38 A). In general, the longer the auditory system has been deprived from stimulation, the harder it is for the brain to extract meaningful information from auditory stimulation once it has been restored (e.g., by cochlear implantation). It is one of the reasons for explaining the huge variation in outcome (speech and language skills, social integration, etc.) in CI recipients. Today more and more implantations are done at a young age, to minimize sensory deprivation and to maximally exploit brain plasticity. Whereas in the year 2000 congenitally deaf children rarely received an implant in their first year of life, this has become rather common practice no more than ten years later, with even a majority of CIs being implanted before the age of 2 (Figure 38 B).

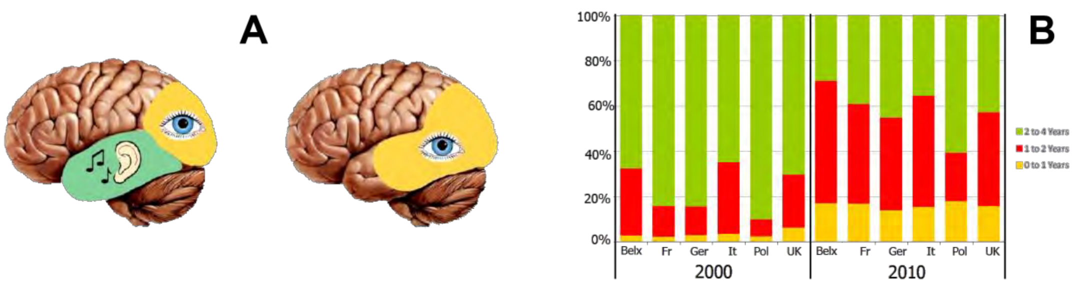
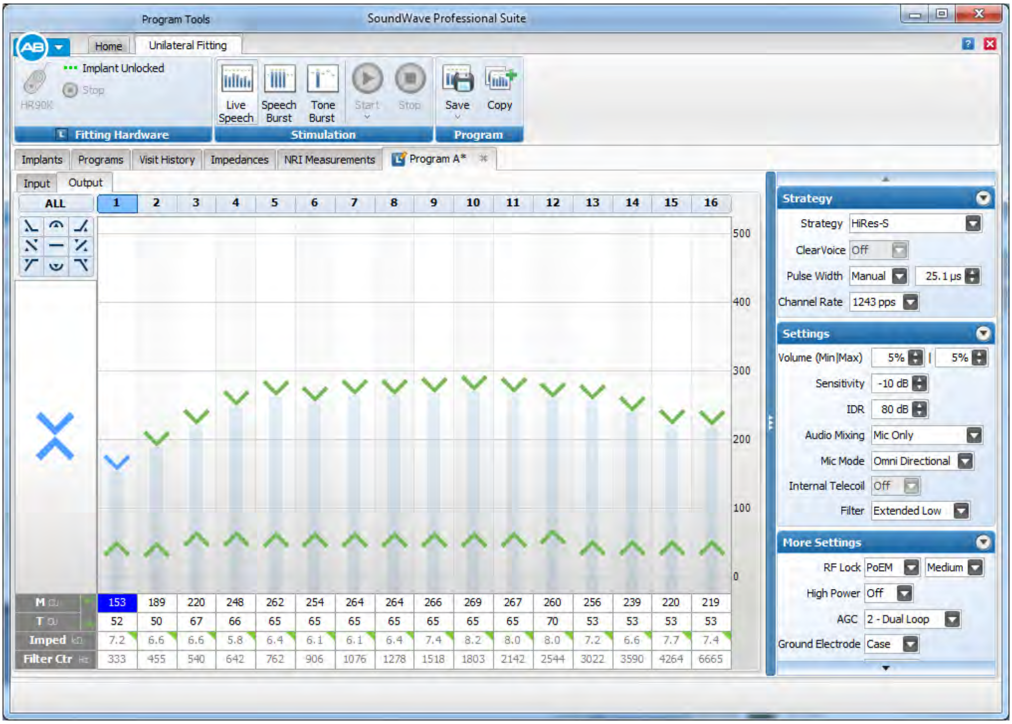


Figure 38: A: an illustration of brain plasticity and how the auditory cortex is assigned to visual functions in the absence of auditory stimulation. B: the age at implantation for 0 to 4 year old children in different European countries, comparing the age at implantation in the year 2000 with ages at implantation in the year 2010. Source: Cochlear Ltd. database.

Following surgical implantation, the CI sound processor must be appropriately programmed and customized for the individual, which is commonly called programming, mapping or fitting. The aim of fitting is to set a number of parameters to ensure that the electrical pattern generated by the internal device in response to sound, yields an optimal auditory percept [84] [85] for that individual recipient. Several tuning parameters are available and all their values together are commonly called the "map". Fitting is performed by professionals (e.g., audiologists) using proprietary fitting software provided by the CI manufacturer (Figure 39).



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
M	153	189	220	248	262	254	264	264	266	269	267	260	256	239	220	219
T	52	50	67	66	65	65	65	65	65	65	65	70	53	53	53	53
Imped	7.2	6.6	6.6	5.8	6.4	6.1	6.1	6.4	7.4	8.2	8.0	8.0	7.2	6.6	7.7	7.4
Filter Ctr	333	455	540	642	762	906	1076	1278	1518	1803	2142	2544	3022	3590	4264	6665

Figure 39: Screenshot of AB's fitting software SoundWave, showing 16 channels exposing T (EDR minimum) and M (EDR maximum) parameters. Global (channel independent) parameters are shown at the right side of the screen.

The first fitting session, called the "switch-on", is typically scheduled a few weeks after the surgical intervention. During that session the speech processor is activated for the first time. This first electrical stimulation is experienced quite differently by different recipients and depends on age, duration of deafness, speech and language skills, personal expectations and other factors. One can hardly imagine that the auditory percept caused by electrical stimulation of only 20 separate cochlear sites could ever compare to what is perceived through normal ears. Yet, for certain people who have heard well in the past (e.g., with help of hearing aids), and have developed speech and language skills, and have not been deprived from auditory stimulation for too long prior to implantation, a cochlear implant may already provide identification of sounds and even open set word recognition on the very first day of CI activation. On the other side, there are also persons who are hardly able to attribute anything meaningful to the sounds they now perceive for the first time in their lives. In any case, the road ahead is one of counseling by professionals, learning to handle the device technology, speech therapy and rehabilitation, repeated fitting and evaluation sessions, and learning to process that new world of sounds.

The map used for initial stimulation (during the switch-on session) is often based on the default parameter values (stimulation strategy, filter bank, pulse width, etc.) recommended by the manufacturer and stimulation levels that are clearly detectable but not uncomfortable (somewhat below EDR Maximum). The electrodes' EDRs are usually first measured behaviorally (i.e., by psychophysical loudness assessment) using electrical tone bursts on separate electrodes and some sort of Hughson-Westlake method [86] to find loud but comfortable stimulation levels (EDR Maxima) and to a lesser extent also the detection threshold levels (EDR Minima). The levels are also often (re)adjusted globally (i.e., in group) while in "Live" mode (i.e., using microphone input of live speech and environmental sounds) or through loudness balancing of individual electrodes. The resulting EDR Minimum and Maximum levels may be very different for each individual recipient, due to variations in the state of the peripheral (e.g., the position of electrodes and the state of surviving neural structures) and the central auditory system (e.g., duration of deprivation) as explained earlier. Due to adaptation to the electrical stimulation, CI recipients need to have their processor reprogrammed repeatedly over time. Typically there are several fittings sessions scheduled in the first few months after switch-on, during which maps are further tuned. After that, CI recipients tend to stabilize and in general only visit the CI clinic once a year to undergo an extensive evaluation, possibly resulting in (minor) adjustments to the CI processor.

The methods that are being used by clinical professionals in the field to search for an optimal map show a large variation. That variation results from variations between recipients (e.g., infant/child/adult, prelingually/postlingually deafened, etc.), technological differences between different CI systems, but also for a large part from the specific methodology that has been adopted in individual CI clinics/centres [83]. Most clinicians focus hardest on finding an appropriate EDR for each electrode. In the past this was usually done by measuring a maximum (and in many cases also a minimum) stimulation level for each and every electrode in the array. Since these measurements are performed behaviorally, this is a time consuming process that requires a prolonged attention of the recipient.

Figure 40: Screenshots of Cochlear's Custom Sound 4.0 fitting software. A: the default programming approach involves setting levels for only 5 electrodes out of 22 and interpolating those levels to the electrodes in between. B: programming of maps into the 4 available slots, each for a specific listening situation.

Today, many CI centres follow hundreds of CI recipients and the number of new implantations keeps on increasing every year. Also, in many countries the available resources for CI fitting are limited, which makes it financially difficult to spend large amounts of time programming each individual recipient. For those reasons CI manufacturers provide the option to measure only a few electrodes, and derive levels for other electrodes by interpolation. It has been shown that such an approach, often called Streamlined Programming (Figure 40 A), does not decrease performance significantly [14]. Other time-saving approaches involve deriving Minimum levels from Maximum levels (or vice versa), setting Minimum levels to zero, or using a predefined profile. Predefined profiles may be the result of statistical approaches or may be based on objective measures such as electrically evoked stapedius reflex thresholds (eSRT, used to derive Maximum levels) or electrically evoked compound action potentials (eCAP, used to derive Minimum Levels) in the individual recipient. Those methods are also often applied for recipients who are unable to cooperate in behavioral measurements, like young infants.

Other fitting parameters (stimulation strategy, number of maxima, rate of stimulation, noise reduction, etc.) are often left untouched. Some centres have a policy of letting the recipient explore different options in a take-home experience (a CI processor has multiple slots that can hold different maps for the recipient to choose from) and keep the maps that are reported to be most suitable for everyday listening and listening in specific conditions (e.g., noisy environments, listening to music or in class rooms) as illustrated in Figure 40 B.

2.3. INVESTIGATIONS IN CI SIGNAL PROCESSING AND FITTING

Given the limited availability of documentation on how a CI processes sound and how it converts it to electrical stimulation levels, this matter was further investigated in collaboration with engineers from the CI manufacturers. It has been a 4 year process of iteratively building a knowledge base on the behaviour of CI systems. The results regarding the coding of sound intensity can be found in “Intensity coding in current generation CI systems”. This manuscript is unique in the sense that, for the first time, the behaviour of the various CI systems is described in a coherent and unambiguous manner (Figure 41) and published with co-authorship of each of the four CI manufacturers. The knowledge obtained has been crucial for constructing and tweaking our model for CI fitting.

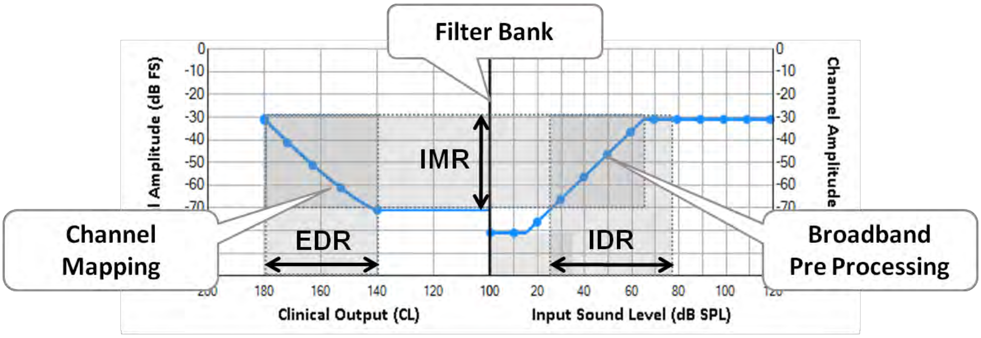


Figure 41: The Intensity Coding Function (ICF) plots: a uniform way of visualising the behaviour of CI speech processors of the 4 CI manufacturers.

In the absence of exhaustive literature on the current practice of CI fitting, we also set out to construct a global inventory of the methodology used in the world today. Through an extensive questionnaire, 47 experts from international CI centres were interrogated on their methods. Subsequently we organized a two-day symposium, specifically on the subject of CI fitting. Over one hundred experts from CI centres around the world came to Antwerp and explained their methods (Figure 42). On top of that, CI fitting experts from 29 international centres were subjected to a 90 minute telephone interview to verify and expand the data set. The resulting inventory represents a total of 47,600 CI users (> 15% of CI users worldwide), making it an unprecedented synthesis of the current state of the art. The results were processed and compiled in the manuscript “A global survey on the state of the art of CI fitting”.

Figure 42: A picture taken during the International Debate on Cochlear Implant Fitting, Antwerp, 2012.

In summary, this inventory shows that many different approaches exist for finding an optimal program for the individual CI recipient. Although several of those approaches in the hands of different experts may lead to similarly good results, it is hard to compare/benchmark them in the absence of commonly agreed upon targets for outcome. It is remarkable that most of the fitting procedures involve tuning map parameters based on subjective feedback from the CI recipient. This subjective feedback is often the expression of some level of comfort of listening as experienced by the recipient. One could argue that this (providing the map that is most comfortable) cannot be the ultimate goal of fitting. Rather the auditory performance that can be achieved with a particular map should be, eventually, the criterion to evaluate the efficacy of fitting, with the condition that a map must not cause discomfort or pain. However, there is no real agreement on which targets to set for such auditory performance, and how to measure it.

With a clearly defined set of outcome parameters and targets, one could start optimizing the process of CI fitting in function of those parameters. In addition it can be anticipated that this would lead to more systematic policies and reduce the time spent at fitting as well as the variabilty across centres in the amount of time and resources that is invested in CI programming (some clinics now spend more than tenfold the time on programming their recipients when compared to other clinics).

2.4. INTENSITY CODING IN CURRENT GENERATION CI SYSTEMS

A uniform graphical representation of intensity coding in current generation cochlear implant systems

Ear & Hearing, 2014

Bart Vaerenberg[1,2], Paul Govaerts[1], Thomas Stainsby[3], Peter Nopp[4], Alexandre Gault[5], Dan Gnansia[6]

[1] The Eargroup, Antwerp-Deurne, Belgium
[2] Laboratory of Biomedical Physics, University of Antwerp, Belgium
[3] Cochlear Ltd, Sydney, Australia
[4] MED-EL, Innsbruck, Austria
[5] Advanced Bionics, Stäfa, Switzerland
[6] Neurelec, Vallauris, France

Abstract

Understanding and predicting the impact of MAP changes on the electrical current delivered at the level of cochlear implant electrodes is challenging. However it is an important prerequisite for programming these devices in the clinical practice. This paper describes a graphical representation to illustrate the intensity coding behavior of four cochlear implant systems (Cochlear, MED-EL, AB and Neurelec). For this we have broken down the intensity coding into 2 separate transformations: (1) from broadband acoustical input to band limited channel amplitude and (2) the mapping function within a single channel. These functions have been synthesized and presented in a uniform plot across brands. The plot describes the output of a CI channel in response to different input signals. This has been incorporated in an interactive software application which illustrates the different stages of intensity coding and the impact of the relevant fitting parameters for each CI brand. It provides the clinician with an assistive tool to better understand and predict the behavior of cochlear implants which may lead to more knowledgeable interpretation and CI programming.

2.4.1. INTRODUCTION

Cochlear implants are now widely accepted as an effective treatment for profound deafness [87]. Following surgical implantation, the sound processor must be appropriately programmed and customized for the individual, which is commonly called fitting. The aim of this is to set a number of parameters to ensure that the electrical pattern generated by the internal device in response to sound, yields an optimal auditory percept [84] [85]. Several tuning parameters are available and all their values together are commonly called the MAP.

The main focus in current generation CI systems lies on compressing the wide range of intensities present in acoustical input signals into the limited range that is available for electrical stimulation. Hence most of the MAP parameters relate to the coding of intensity while only few relate to other sound coding features, like the spectral mapping. A recent global survey on CI fitting practices has shown that in most cases fitting is restricted to setting the threshold of audibility for electrical stimulation, and a level of upper tolerance limit, for each electrode separately [83]. Those levels define the electrical dynamic range (EDR) of each electrode and will be referred to as EDR Minimum and EDR Maximum respectively. Other MAP parameters may also affect the system's mapping of intensities and adjusting them has been shown to produce better outcomes for individual recipients [83]. Yet, these additional parameters are left at default values in most cases [83]. The authors believe that the reasons for this are multiple. A major reason may lie in the intrinsic complexity of the CI system and its sound processing and the differences, often subtle, in the underlying technologies used by the different CI devices. This makes it difficult to predict the impact of a specific MAP parameter on the behavior of a given CI system. In addition, features/parameters with similar names across brands may be implemented differently. For instance, the channel gains in Cochlear’s system are applied at the input of a channel, while in AB devices they are added to the channel’s output. Although this has a similar effect on loudness in both systems, it may induce different effects on threshold, maximum stimulation level, etc. as will become apparent below.

In this paper we have undertaken to present a uniform graphical representation which illustrates the effects of parameter changes on the CI's output for all currently available CI systems. This representation has been incorporated in an interactive software application which allows a dynamic visualization in function of chosen MAP settings. We believe that such comprehensive summary of the behavior of CI systems, represented in a uniform way across brands, may assist the audiologist in gaining more insight into the clinical behavior of these systems and in further optimizing the fitting process.

It is beyond the scope of this paper to explain the meaning of all possible MAP parameters found in the various CI systems. For this information, the reader is referred to the manufacturer’s user manuals and clinical guidelines, and to existing comprehensive overviews [88] [85].

2.4.2. MATERIAL & METHODS

2.4.2.1. VISUALIZATION OF THE INTENSITY CODING FUNCTION

To visualize the input-output relation of a CI system, we have established a three-axial graphical representation to reflect the three major stages which can be indentified in the signal processing path of all current generation systems (Figure 43): (1) an acoustical stage where the broadband signal is captured by a microphone and pre-processed; (2) a digital stage where the signal has been digitized and its energy is distributed over a number of channels and finally (3) an electrical stage where the energy in each channel is mapped to an electrode activation level.

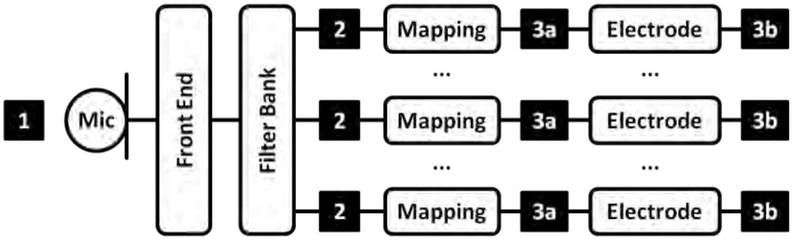


Figure 43: CI processing path block diagram showing the different stages at which signal levels are considered. (1) the intensity of the acoustical signal; (2) the amplitude of the band limited digital signal at the output of the filter bank and (3) the magnitude of the electrical stimulation, either as displayed in the fitting software (3a) or as the equivalent amount of charge delivered by the electrode (3b).

Figure 44 depicts the acoustical stage at the right horizontal axis (marked as 1), the digital stage at the vertical axis (marked as 2) and the electrical stage at the left horizontal axis (marked as 3).

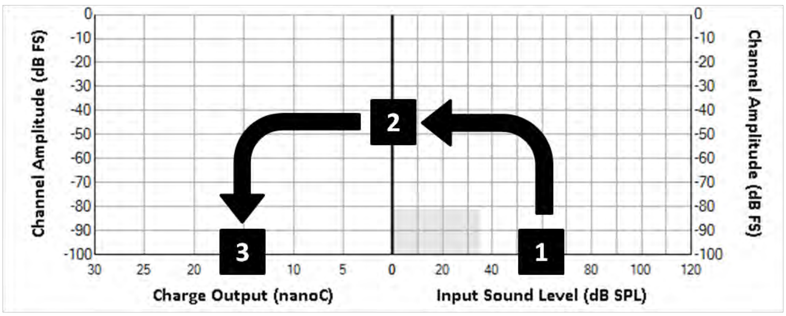


Figure 44: The empty ICF plot showing the 3 dimensions relating to the 3 signal processing stages: (1) the acoustical dimension on the right horizontal axis, representing the broadband input sound level (dB SPL); (2) the digital dimension on the central vertical axis, representing the narrowband channel amplitude (dB FS); (3) the electrical dimension on the left horizontal axis, representing the charge output (nC).

At any of these stages the signal level is expressed in an appropriate unit. The acoustical input signal level is presented in broadband dB SPL (stage 1). The signal is then digitized and pre-processed (e.g. beamforming, noise reduction, wind noise reduction, de-reverberation, etc.) which often also includes the application of a pre-emphasis filter and a gain factor, before it is split into separate frequency bands. These processes make that at the output of the filter bank it is no longer feasible to express the signal level in terms of acoustical sound pressure. Instead, this signal level is expressed relative to a digital maximal value known as Full Scale (0 dB FS, i.e., the largest signal amplitude that can be expressed by the internals of the CI system, stage 2). It is essential to be aware that at this point the energy of the input signal is divided over multiple channels which makes the dB FS scale represent a narrow band energy in a single channel. Finally, using the values of EDR

Minimum and EDR Maximum, this energy is mapped to an electrode activation level. That level may be expressed in a 'clinical' unit that is displayed to the user of the fitting software (stage 3a) or it may be expressed in a unit of the equivalent charge per pulse phase at the channel's electrode contact (i.e., nanoCoulomb (nC), stage 3b).

Hence, the acoustical axis of Figure 44 displays the Input Sound Level ranging from 0 dB SPL to 120 dB SPL, the digital axis displays the Channel Amplitude ranging from -100 dB FS to 0 dB FS and the electrical axis displays the Charge Output ranging from 0 to 30 nC. As such the 3 axes in Figure 44 allow the visualization of 2 distinct transformations: on the right one can see the pre-processing that operates on the broad band signal, and on the left one can see the mapping function that is used to transform the narrow band energy within a single channel to an electrode activation level. The combined graph is called the Intensity Coding Function (ICF) plot and the interactive software application which allows the dynamic visualization of these plots in function of chosen MAP settings is called Intensity Coding in Cochlear Implants (ICCI). Both are explained in more detail further in the text.

2.4.2.2. SCOPE, CONSTRAINTS AND DISCLAIMER

The different CI systems currently available all have their particular signal processing strategies and features, which makes it hard to produce a general model that covers them all. For this reason some constraints have been applied that allow the graphs to have a well-defined scope in which they should be interpreted. Firstly, the plots are only reflecting the current generation (anno 2013) CI systems using their default speech coding strategies. For Cochlear this is the CP810 processor with ACE strategy and CI512 implant; for MED-EL this is the OPUS 2 processor with FS4 strategy and CONCERTO or SONATA implant; for Advanced Bionics (AB) this is the Harmony processor with HiRes strategy and HiRes90k implant; for Neurelec this is the Saphyr processor with Digisonic SP implant. Secondly, a number of features have been excluded from the analysis, because they are either of little relevance to the essentials of intensity coding or too dynamic (dependent on temporal and spatial aspects of the input signal) to be visualized on a static graph. These features include: input mixing and beamforming (the use of multiple microphones, optionally in combination with telecoil, aux, etc.), volume control options that can be manipulated by the recipient, noise reduction algorithms, temporal aspects of speech coding strategies, etc. These features have not been included as variable parameters in ICCI and their effect on the device behavior is ignored. The features that have been included are: EDR Minimum, EDR Maximum, Instantaneous Mapping Range (IMR), Input Gain, Output Gain, Input Compression and Output Compression. Those features and the parameters they relate to are summarized in Table 2 for each of the manufacturers. EDR Minimum and Maximum define the range of stimulation levels for each electrode. IMR (expressed in dB) is the range of input sound levels that is being mapped into the EDR at any given instant in time. Input Gain is the application of a gain factor at the broad band input signal. Output Gain is the application of a gain factor per electrode. Input Compression is the long-term compression of the broad band input signal due to AGC systems and alike. Output compression is the instantaneous compression applied per channel (in the mapping function).

Even for those MAP parameters that are included, not all technical details are implemented by the ICCI application. As such it should be considered and used as an assistive tool which provides an indication of the behavior of these systems rather than an exact simulation of their signal processing algorithms. This is important since it may give rise to inconsistencies between the technical documentation provided by the CI companies and the output of the ICCI application. The reader should understand that for technical accuracy, the documentation from the CI companies always prevails. Nonetheless, the authors are convinced that the possible inconsistencies in the ICCI application are compatible with the principle aim of this paper and the ICCI application, namely to grant a principal understanding of the intensity coding in cochlear implants. Some of the authors have responsible positions at CI companies. Their contribution has been essential to the development of the graphical representations and the content of this manuscript. But neither they nor the CI companies can be held legally responsible for any such inconsistencies which have been unavoidable in the interest of the comprehensibility of this paper and the associated application. The use of ICCI and interpretation of the ICF plots is intended as an assistive tool for the competent clinical CI-programmer who remains fully and solely responsible for their use in the clinic.

Table 2: Mapping features that are taken into consideration in the ICF with their corresponding parameter names across brands

	Cochlear	MED-EL	AB	Neurelec
EDR Minimum	T (CL)	THR (QU)	T (CU)	Min (µs)
EDR Maximum	C (CL)	MCL (QU)	M (CU)	Max (µs)
Instantaneous Mapping Range	C-SPL - T-SPL (15 - 75 dB)	(55 dB)	IDR (20 – 80 dB)	(85 dB)
Input Gain	Sensitivity (dB)	AGC Sensitivity (%), AGC Compression Ratio	Sensitivity (dB)	Analog Gain (dB)
Output Gain	Gain (dB) or ADRO		Gain (dB)	Gain (dB)
Input (long term) Compression	ASC	AGC Compression Ratio	AGC	
Output (instant) Compression	Loudness Growth (Q)	Maplaw Compression		Volume

2.4.2.3. SOURCES OF INFORMATION

The ICFs were synthesized from a number of existing documentation sources and verified through interviews with the manufacturers' engineers. Since all manufacturers provide general descriptions of their CI systems through their fitting software, basic concepts have been used from Custom Sound 3.2 Help contents and the Cochlear™ Clinical Guidance Document [89], the Maestro 4.0 Help contents and a FocusOnFineHearing™ Technology document [90] the SoundWave 2.1 Help contents [91] and the Digimap 3.4 Help contents [92]. These sources are designed to provide assistance in performing specific programming tasks, but often lack detail and cohesion with regard to the composite signal processing chain as a whole. To be able to construct the ICFs additional information was required. That information has been partly obtained from published articles [93] [94], presentations at conferences and fragmented documentation that has been collected by the authors over the years. Individual interviews with company engineers were conducted to complete and validate the required sources.

2.4.2.4. THE IFC PLOT

INPUT SIGNALS

The ICF is plotted in response to any of 3 types of signals: a pure tone (adjustable in frequency between 100 and 8000 Hz), a speech signal or a white noise signal. It is thereby assumed that: (i) the frequency of the pure tone signal equals the centre frequency of the observed channel's filter band, and as such completely falls into that channel (there is 0 dB attenuation with regard to the broadband energy); (ii) the speech signal does not fall into a single channel, it is distributed over multiple channels in such a way that the observed channel receives energy from the speech signal that is equal to the broadband energy attenuated by 12 dB; (iii) the white noise signal is distributed over all channels such that the observed channel receives energy from the white noise signal that is equal to the broadband energy attenuated by 24 dB. The attenuation of 24 dB relates to the fact that the average CI channel has a bandwidth of 493 Hz which is 1/16 of the total bandwidth (8 kHz) of current generation systems. For the speech signal, having at any given moment a band width between those of a pure tone and white noise, the attenuation inflicted by the filter bank was chosen to be the mean of the attenuations of the pure tone (0 dB) and the white noise (24 dB) signals, hence 12dB. It must be noted that the choice of these attenuations is a simplification and that in reality the attenuation is highly dependent on how the channel band pass filter is organized in relation to the input signal. It is also assumed that all input signals feature a stable long term intensity by which AGC systems reach convergence (i.e. broadband intensity is maintained stable for a time longer than the system's attack time). This response to long term intensities incorporates, amongst others, the static gain function of AGC systems (a new gain is determined from this function when the long term intensity of the input signal changes). Nonetheless the ICF plots in ICCI, as illustrated in Figure 45, allow showing the response to rapid fluctuations (open symbols) around this long term intensity (filled symbols), of which it is assumed that they do not trigger the slow detectors of AGC systems. “Appendix B: ICF plots with fast responses” displays these kinds of plots for the 4 CI brands at 3 different intensities (35, 65 and 95 dB SPL).

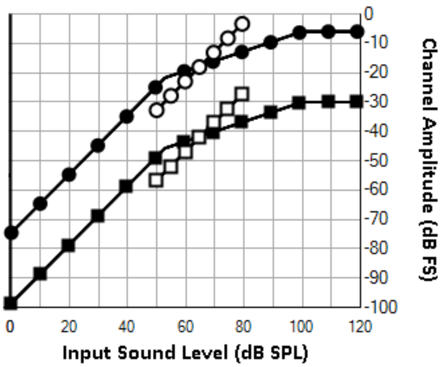


Figure 45: Example plot displaying MED-EL's transformation of acoustical input to channel amplitude for both pure tone (circles) and white noise (squares) signals. The response to long term intensity is plotted in filled symbols. The response to fast deviations (+/- 15dB) from a long term average intensity of 65 dB SPL is plotted in outline symbols.

MICROPHONE, PRE-EMPHASIS, SYSTEM NOISE FLOOR

In the ICF plots the front end pre-processing and filter bank steps are combined and depicted on the right chart (Figure 44). The microphone's frequency response is not considered (i.e. it is assumed to be flat). The system noise (primarily determined by the microphone noise floor) is assumed to be white within the range being processed (0.1 to 8 kHz) and equivalent to 35 dB SPL within an acoustical band of that same range. As with the white noise input signal, the energy per channel is assumed to be equal to the broadband energy attenuated by 24 dB, and would therefore be approximately 11 dB SPL per channel on average, as illustrated in Figure 46. For all devices, the pre-emphasis filter is assumed to be an A-weighting filter. The ICCI application allows adjusting the observed channel's centre frequency, such that the effect of pre-emphasis becomes apparent.

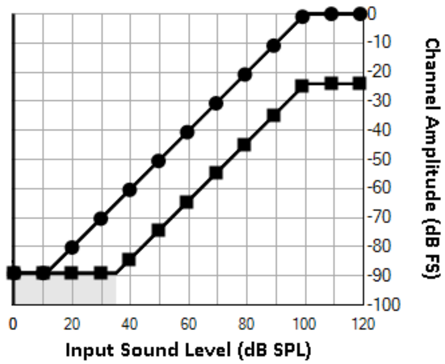


Figure 46: Example plot showing Neurelec's system noise floor (equivalent to 35dB SPL) as the light gray area at the bottom left. It causes a 11 dB SPL equivalent floor effect on the channel amplitude of pure tones (circles) and a 35dB SPL equivalent floor effect on white (0.1 - 8 kHz) noise signals (squares).

MAPPING

The mapping step is depicted on the left chart of the ICF plots (Figure 44). It maps the channel's amplitude into the "Electrode Activation Level" expressed in the unit that is presented to the clinician in the manufacturer's fitting software. Strategy-specific features (e.g., current steering [94]) are not considered in the ICF plots. There is an option to convert the manufacturer-specific level to its general charge equivalent, expressed in nanoCoulomb per pulse phase, as illustrated in Figure 47.

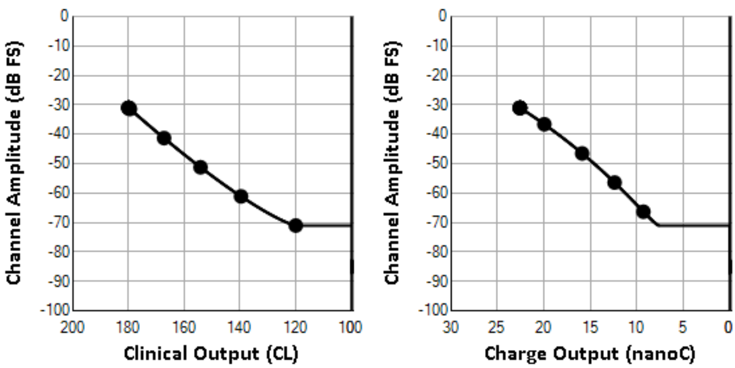


Figure 47: Example plots showing Cochlear's mapping function expressed in both their clinical unit (left chart) and its equivalent charge per pulse phase (using a Pulse Width of 50µs) (right chart).

2.4.3. RESULTS

We have established a three-stage graphical representation of the intensity coding for the four brands of current generation CI systems. This representation is available as static graphs as appended in “Appendix A: ICF plots for map parameters”, but it is more useful in the interactive dynamic software application ICCI (see further). It illustrates the impact of MAP parameters on the coding of intensity and allows understanding and predicting the transition from free field acoustical signals through digital band limited energies into electrical stimuli at the output of the CI channels/electrodes. This approach reveals a number of MAP-features which are common to all four implant brands and a number of device specific features and how they impact the coding of sound.

2.4.3.1. COMMON FEATURES

INPUT SENSITIVITY / GAIN

All brands expose a parameter that can be used to adjust the microphone sensitivity. In Cochlear and AB devices this is called Sensitivity. In Neurelec devices this is called Analog Gain. These three operate as a fixed gain applied to the input signal. In MED-EL, microphone sensitivity is defined by a combination of the AGC Sensitivity and the AGC Compression Ratio. Increasing the AGC Sensitivity shifts the AGC knee-point towards low level sounds, resulting in softer level sounds to be represented within the EDR. Increasing the Compression Ratio reduces the part of the EDR that is assigned to signals above the AGC knee-point, also resulting in softer level sounds to be mapped into the EDR.

In addition to static microphone sensitivity, some CI systems have a mechanism that automatically adjusts the microphone sensitivity based on environmental sound levels. In Cochlear devices this is called Autosensitivity (ASC) which uses the environmental noise level and signal-to-noise ratio to determine an appropriate input gain. In MED-EL and AB devices this is called Automatic Gain Control (AGC) which are dual-loop AGC systems. Although they are implemented quite differently, they both use the input signal level to determine an appropriate gain. They contain a slow detector with a relatively long attack time which means that they work as a volume control. Neurelec processors do not have an AGC feature.

NOISE FLOOR, IDR, IMR AND SATURATION

All manufacturers provide default values for their fitting parameters that cause the system noise (as described earlier) to be excluded from the EDR and therefore to not be perceived by the recipient. This is accomplished by effectively limiting the range of signal levels (Instantaneous Mapping Range, IMR) that is mapped into the EDR at any given time. In Cochlear's device the position and size of the IMR can be set using the T-SPL and C-SPL parameters (Figure 48). AB provides an IDR parameter that affects the size of the IMR only. Neurelec and MED-EL have a fixed IMR. The settings for input sensitivity/gain as described in 3.1.1 however, also impact the actual (i.e., with regard to sound pressure level) position of the IMR in all brands. Moreover, the use of AGC systems introduces the need to distinguish 2 different approaches to the input range: (1) the range that is instantaneously considered (IMR) and (2) the entire range (IDR) that is covered when shifting this instantaneous range in response to the automated adjustment of an input gain/sensitivity factor.

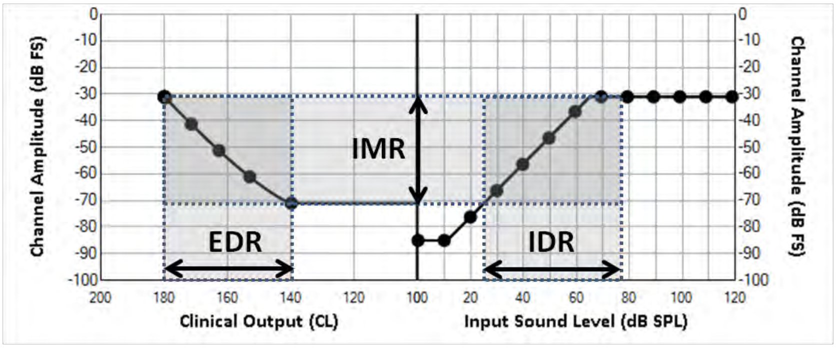


Figure 48: Example plots showing Cochlear's mapping of IDR into an EDR between 140 and 180 CL when using default parameter values (T-SPL=25, C-SPL=65, Sensitivity=12). The IMR of 40 dB (between T-SPL and C-SPL) is potentially extended to an IDR of 52 dB when ASC is enabled and Sensitivity is at 12.

All 4 brands have an upper limit for input signal level around 100 dB SPL. Beyond that level the systems are saturated. Given the 11 dB SPL noise floor within a single channel, this means that CI systems could provide IDRs of up to 90 dB. This 90 dB however, is to be mapped into an EDR that is an order of magnitude lower (typically below 10 dB), meaning that there is a trade-off between range and resolution. Cochlear maximizes the electrical resolution by limiting their IMR to a default of 40 dB. In MED-EL devices a fixed IMR of 55 dB is used. AB uses 60dB by default. Neurelec maximizes range and maps 85 dB into their EDR.

2.4.3.2. COCHLEAR

Figure 49 shows ICF plot for Cochlear with the impact of its MAP parameters. Cochlear uses the term IIDR to refer to instantaneous mapping range (IMR). Threshold (T) and maximum comfort (C) levels determine the range and position of the EDR. The IMR/IIDR is set by the T-SPL and C-SPL parameters, but changes in microphone sensitivity also impact its position (they alter the softest level sound that is mapped into the EDR) and they also alter the automatic gain control (AGC) knee-point (i.e. the input level corresponding to C-level stimulation, not to be confused with AB's or MED-EL's AGC systems, which serve another purpose). As a consequence C-SPL and T-SPL parameter values only reflect actual input levels if the Sensitivity parameter value is taken into account (i.e., a C-SPL value of 65 only maps 65 dB SPL to C level if Sensitivity is set to 12 (default); if Sensitivity is set at 0 then a C-SPL value of 65 would map 77 dB SPL to C level). In addition, the ASC might adjust the sensitivity based on environmental sound levels. In quiet, a program with ASC acts similarly to a program with no additional input processing, with a fixed sensitivity setting. But if the level of background noise is above the ASC Breakpoint (57 dB by default), sensitivity is reduced according to the level of the noise so that the peaks of speech exceed the long-term average noise spectrum by at least 15 dB

[89]. ASC can only reduce, not increase, the Sensitivity value set in the MAP (if a Sensitivity of 0 is set, ASC has no effect).

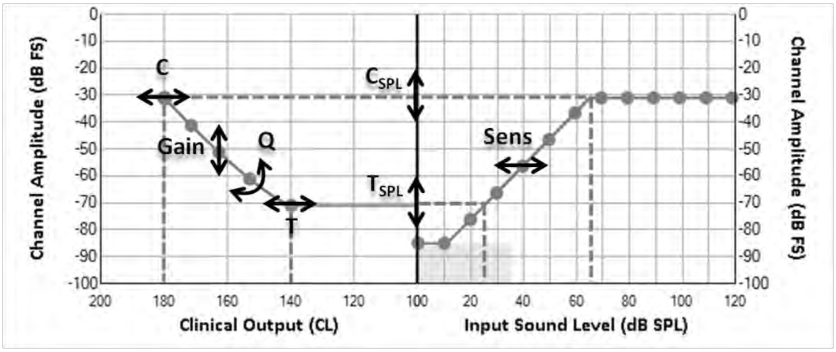


Figure 49: ICF plots for Cochlear devices in response to a pure tone, showing pre-processing impacted by Sensitivity (Sens) at the right and mapping impacted by T-SPL, C-SPL, T, C, Gain and Loudness Growth (Q) at the left. Straight arrows depict translations; curved arrows indicate changes in the shape of the curve.

An AGC system with infinite compression keeps the broadband input signal within range. Signals above the C-SPL are not clipped but their peak output is effectively limited to C-SPL (65dB) by the AGC (Attack Time = 5ms; Release Time = 75 ms). The AGC acts to prevent clipping by its fast attack and slower decay.

Either clinical gains as defined per channel by the MAP, or ADRO (Adaptive Dynamic Range Optimization) gains are applied to the channel input (i.e., filter bank output). Clinical Gains range from -12dB to +10dB and have the effect of amplifying the channel's input with a fixed value (set in the MAP), but they are ignored when ADRO is enabled.

Stimulation above C-level can never occur, even if clinical gains of +10 dB are set. When applying positive gain to a channel, saturation occurs for that channel at C-SPL minus Gain, e.g. 55 dB SPL when C-SPL is 65 and channel gain is +10. When applying negative gains, saturation remains in hands of the AGC (and occurs at C-SPL). Amplitudes below T-SPL do not produce any stimulation. Indeed signals below the T-SPL are effectively ignored. Setting T-SPL below 25dB may result in excessive electrical (system) noise being perceived constantly.

Increasing C-SPL (even more so when combined with an increase in C levels) may cause a different psychophysical loudness perception. To compensate for this, the Custom Sound software automatically adjusts the Loudness Growth (Q) parameter. The Q parameter sets the % of dynamic range (EDR) that is allocated to the top 10 dB mapping input (IMR). A higher Q value is more compressive and curves the function more. According to the manufacturer, Q is not meant to be seen as compression and Cochlear would not advise making major changes to Q.

The pulse width is not considered in the CL unit used in the fitting software. As a consequence, doubling the pulse width while keeping the same T and C levels will cause the channel to deliver twice the amount of charge (per pulse phase) to the electrode. The accuracy of the pulse amplitude is limited to the CL values between T and C. As a consequence, in a channel with T = 140 and C = 170, every channel amplitude is mapped to one of the 31 distinct pulse amplitudes available between 140 and 170.

2.4.3.3. MED-EL

Figure 50 shows the ICF plot for MED-EL with the impact of its MAP parameters. MED-EL's AGC system is a dual loop AGC [95]. The fast and slow detectors work in parallel on the same input signal. The resulting outputs of the two detectors are weighted to determine the corresponding gain. With default compression of 3:1 and sensitivity of 75%, the static gain of the dual loop AGC has its knee-point at 52.7 dB SPL. The fast detector has 4 ms attack and 16 ms release time; slow detector has 100 ms attack and 400 ms release time. The knee-point is shifted (+14 dB to -5 dB with respect to the default value) when changing the AGC Sensitivity parameter. The AGC system keeps the broadband input signal within range. The AGC fast detector acts to prevent clipping by its fast attack and slower decay. Input of 106 dB SPL causes stimulation at MCL Level. The upper limit for signal level is 100 dB SPL with 6 dB head room available to allow for rapid fluctuations (peaks), see "

Appendix B: ICF plots with fast responses" for ICFs showing rapid fluctuations.

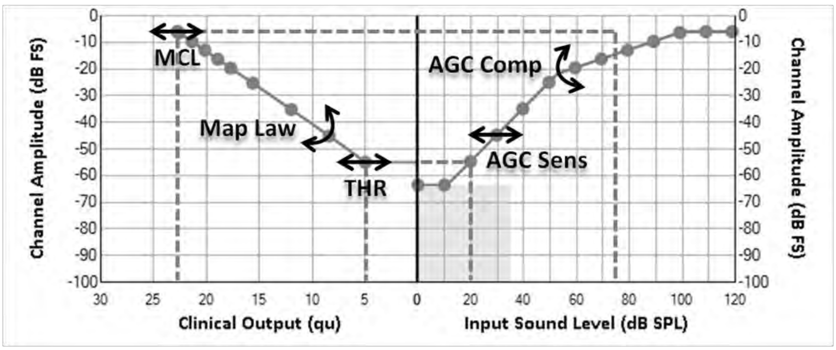


Figure 50: ICF plots for MED-EL devices in response to a pure tone, showing pre-processing impacted by the AGC's Sensitivity (AGC Sens) and Compression Ratio (AGC Comp) at the right and mapping impacted by THR, MCL, Gain and Map Law compression at the left. Straight arrows depict translations; curved arrows indicate changes in the shape of the curve.

MED-EL uses the term Adaptive Sound Window for their 55 dB IMR. The position of this window is managed by the AGC system. THR and MCL levels determine the range and position of the EDR. Stimulation above MCL level can never occur. In the commonly used IBK volume mode, no stimulation below THR can occur for any volume setting (0 to 100%) and THR is the minimum stimulation level for any enabled channel, meaning that the system will deliver charge to all enabled channels having THR greater than 0, even if there is no input signal present.

The Maplaw Compression parameter defines the loudness growth function to map channel (envelope) amplitudes into the EDR. A higher compression value assigns a larger portion of the EDR to softer sounds. A lower compression assigns a larger portion of the EDR to louder sounds.

The amplitude of a stimulation pulse is given in current units (CU). One current unit is approximately 1 µA. The charge of one phase is defined as the product of stimulation current in CU and pulse phase duration, which is displayed in µs in the fitting software. One charge unit (QU) is approximately 1 nC. The MED-EL fitting software works charge-based so that Pulse Amplitude (CU) cannot be manually adjusted. Rather, Pulse Charge (QU) is set in the fitting software. Doubling the phase duration (e.g. by increasing Min. Duration) while keeping the same THR and MCL (QU) levels, will not cause the channel to deliver more charge (per pulse phase) to the electrode, instead the pulse amplitude (CU) is decreased automatically.

2.4.3.4. ADVANCED BIONICS

Figure 51 shows the ICF plot for AB with the impact of its MAP parameters. AB's AGC system also features two detectors (fast and slow). The resulting outputs of the two detectors are compared to determine the corresponding gain. The slow detector has a long attack time, meaning that the system allows keeping response linearity. It features a compression factor of 12:1, a 240 ms attack and 1500 ms release time and a threshold of 61 dB. The fast detector has a threshold of 72 dB and is used to prevent clipping. AB's upper limit signal level is 97 dB SPL; beyond that level peaks are clipped using a "soft clipping" mechanism that is comparable to an extremely fast AGC system (attack time < 1ms).

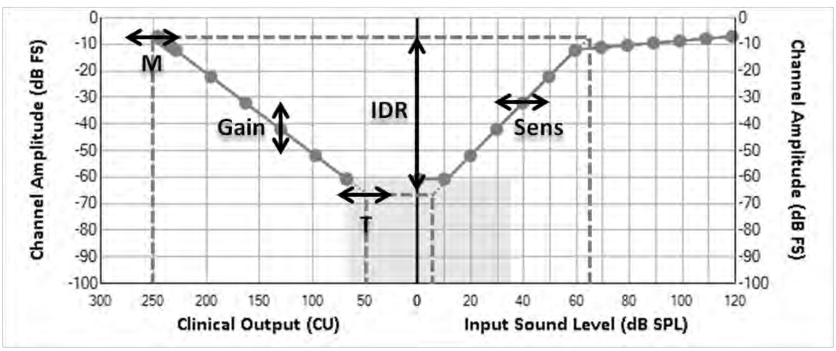


Figure 51: ICF plots for AB devices in response to a pure tone, showing pre-processing impacted by the Sensitivity (Sens) parameter at the right and mapping impacted by T, M, Gain and IDR parameters at the left.

AB uses an IMR window of which the position is managed by their AGC system. The IDR parameter determines the size of the IMR window. T and M levels determine the range and position of the EDR. Stimulation above M level is possible. When AGC is disabled, a 5000 Hz pure tone input signal of just over 60dB SPL already causes stimulation at M level. Increasing the input level further causes stimulation to go over M. Stimulation continues linearly below T Level. Setting a large IDR value or high Sensitivity in combination with correct (measured) T levels, may result in electrical (system) noise being perceived constantly. This relates to the microphone noise as described earlier, which is then mapped into the EDR.

AB does not provide a parameter to change the shape of the mapping function. A channel's (envelope) amplitude expressed in dB is always mapped linearly into the EDR (i.e., an increase of 1dB causes an increase in x CU). The Clinical Unit (CU) used in the fitting software incorporates both Pulse Amplitude and Duration and is therefore related to charge (Coulomb). Pulse Amplitude (µA) cannot be manually adjusted using the fitting software. Doubling the pulse width while keeping the same T and M levels, will not cause the channel to deliver more charge (per pulse phase) to the electrode, instead the pulse amplitude (µA) is decreased automatically.

2.4.3.5. NEURELEC

Figure 52 shows the ICF plot for Neurelec with the impact of its MAP parameters. The Neurelec system has no automatic gain or sensitivity control. The implant recipient has the option to adjust the sensitivity manually. With an Analog Gain of 0 dB, peaks above 100 dB SPL are clipped by the A/D converter. Neurelec uses a fixed IMR of 85 dB. Min and Max levels determine the range and position of the EDR. Signals of 100 dB SPL are mapped to Max Level, 15 dB SPL corresponds to Min Level. Stimulation above Max level can never occur, even if the Volume parameter is at maximum. Channel amplitudes below 15 dB SPL do not produce any stimulation. Indeed signals below this level are effectively ignored.

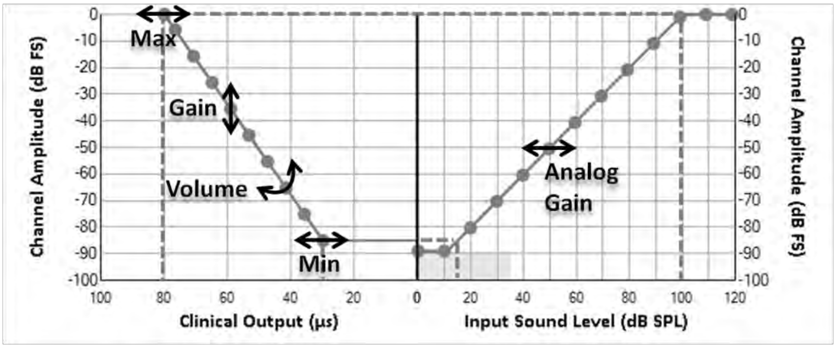


Figure 52: ICF plots for Neurelec devices in response to a pure tone, showing preprocessing impacted by the Analog Gain parameter at the right and mapping impacted by Min, Max, Gain and Volume parameters at the left. Straight arrows depict translations, curved arrows indicate changes in the shape of the curve.

Neurelec allows setting gains (ranging from 0 to -10 dB) on each of the 63 bands of the FFT output. These gains are applied before the combination into channels and may therefore affect the channels that are selected for stimulation (Maxima Selection). The Volume parameter defines the loudness growth of the mapping function. A higher Volume value assigns a larger portion of the EDR to softer sounds. A lower Volume assigns a larger portion of the EDR to louder sounds. Neurelec keeps the pulse amplitude fixed and adjusts the pulse width to code for loudness. The Clinical unit for Min and Max levels, displayed in the fitting software, therefore relates to a dimension of time/duration. The pulse amplitude is not considered in the clinical unit used in the fitting software. As a consequence, doubling the pulse amplitude (Amplitude parameter) while keeping the same Min and Max levels, will cause the channel to deliver twice the amount of charge (per pulse phase) to the electrode.

2.4.3.6. INTERACTIVE PLOTS IN THE ICCI APPLICATION

An interactive application called ICCI that allows plotting the ICF's for the 4 brands is available on request from the first author. A web based version is also available at http://www.otoconsult.com/fitting/icci. As shown in Figure 53, the application allows to adjust the fitting parameters and to view the resulting changes on the plots. As explained previously, the top plots depict both transformations from acoustical level into digital level (preprocessing plus filter bank) and from digital level into electrical level (mapping). In addition, the result of merging these processes into a single transformation from acoustical input level into electrical output level is depicted by the application in its bottom graph. Static plots illustrating the effect of each parameter are included in "Appendix A: ICF plots for map parameters".

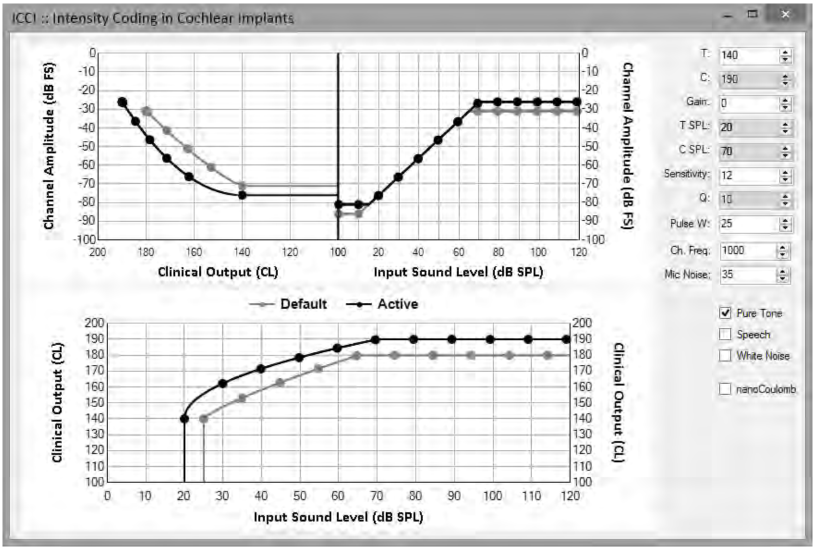


Figure 53: A screenshot of the ICCI application showing the ICF for Cochlear's device. A default map where T equals 140 CL and C equals 180 CL is plotted in gray. A map where C is changed to 190 CL, T-SPL to 20, C-SPL to 70 and Q to 10 is plotted in black. The top charts show the separate preprocessing and mapping functions, The bottom chart is the result of merging the top charts into a single transformation of acoustical input level into electrical output level.

2.4.4. DISCUSSION

The fitting of cochlear implants is a technical procedure which requires a thorough insight in these systems' sound processing. With the ever increasing complexity of the underlying technology and given the fact that many CI centers offer and program several CI brands to date, the professional fitter is facing a greater than ever challenge to fully predict the impact of parameter changes on the implant's behavior. One way to cope with this is to simplify the act of fitting by limiting the number of parameters to adjust and by adopting approximations and rules of thumb to make global profile optimizations. There are indications however that addressing more of the many parameters available may lead to better outcome in specific, if not most cases [96] [82] [97]. For this to be done in a knowledgeable manner, the level of understanding and predicting these devices' behavior is certainly less than what engineers need when designing them, but is likely to be more than what is commonly available.

The effect of parameter changes is explained in clinical guidance documents, but very often this information is fragmented, limited to a single parameter at a time and not integrated in the complete input-output behavior of the system. In addition, every manufacturer has its own way of presenting their system's behavior, and uses proprietary names for parameters which basically do the same thing (e.g. Input Sensitivity/Gain, IDR/IIDR/Adaptive Sound Window). This increases the load on clinicians to understand the behavior of the CI systems they are programming on a daily basis.

By synthesizing the behavior of these different CI systems into a uniform graphical representation, the specific features of a particular CI system are accessible to clinicians in a more transparent way, which in turn may assist them in the programming of these systems. Using the ICCI application, the authors themselves have come to a number of remarkable observations. For example, it is clear that manufacturers handle the compression of 90 dB of input range into a couple of dB's of electrical range quite differently. If we consider the default settings for each of the devices, then we see that IMRs of 40, 55, 60 and 85 dB are chosen by Cochlear, MED-EL, AB and Neurelec respectively. It may be that Neurelec, in the absence of an AGC function, opted for such a large IDR to cover the diverse listening situations in daily life. The other brands have automated gain controls, allowing them to maximize the intensity resolution at any given environmental noise level. The drawback of such an approach may be that the loudness growth in the upper intensity range is limited to some extent, which introduces a phenomenon that is not known to be present in a normal auditory system.

Another interesting observation is that the default mapping functions in all brands are more or less linear when considering the channel input amplitude expressed in dB and the channel output level expressed nC. While the other manufacturers use a clinical unit that linearly relates to charge output, Cochlear uses a clinical unit (CL) that relates to charge (nC) exponentially (an increase of 1 unit in CL has the effect of increasing the current amplitude by approximately 2 %). However, the plots show that Cochlear uses a mapping function that compensates for this. They do not map channel amplitude dB's to CL entirely linearly. In a typical map, where Q is 20, the ICF is slightly curved. But when converted to charge (nC) the function becomes approximately linear again. All together their mapping is not very different from the other manufacturers', when default parameters are used. So, the differences in output compression techniques between manufacturers are more attributable to limiting IMR and using automatic gain controllers than they are to compressive mapping functions.

Although this information is not readily available, we assume that the microphone noises, and therefore the noise floors of all systems are rather similar. It should not surprise to learn that all brands use similar microphones on their CI processors. The fact that all systems are in theory configurable to saturate around 100dB SPL also has to do with all brands facing the same technological limitations in analog to digital conversion and sampling range.

AB is the only brand in which stimulation continues below T level by default. Cochlear and Neurelec implants cease stimulation when the input to a channel falls below the IMR. MED-EL keeps stimulating at THR. This implies that in AB, from a technical point of view, the IDR and T parameters are mutually redundant (e.g. one could achieve the very same effect of adjusting IDR, by adjusting T).

For setting T levels, the 4 brands use 2 distinct approaches: MED-EL and AB would not advise against setting THR/T at 0 or at a fixed fraction of MCL/M; Cochlear and Neurelec recommend measuring T levels for each CI recipient. Indeed the plots show that in Cochlear, setting T at 0 would increase the audibility threshold dramatically, as illustrated in Figure 54. In MED-EL and AB, this effect is considerably smaller, due to the nature of their mapping function.

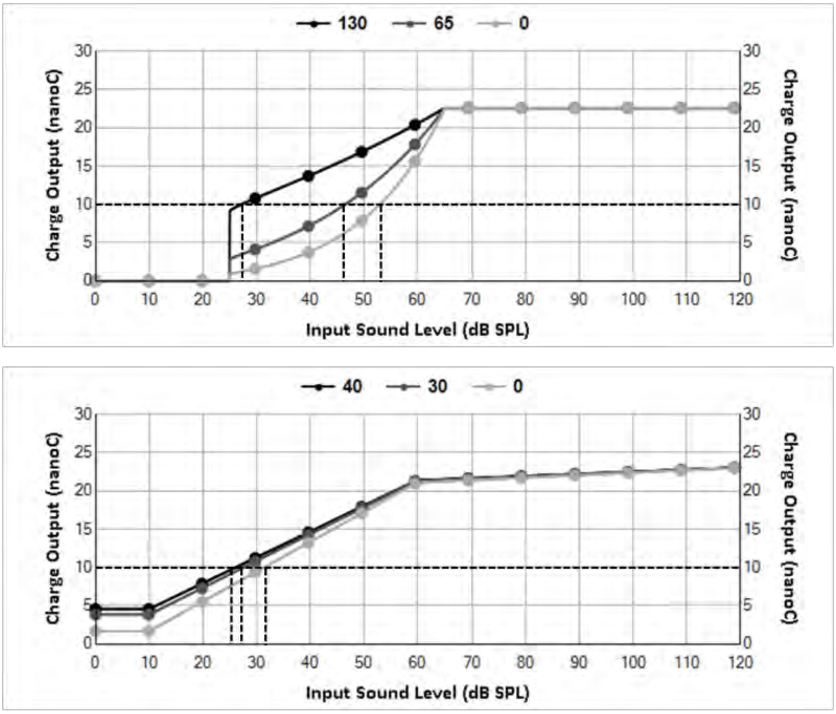


Figure 54: The effect of setting T, assuming 10 nC is the recipient's detection threshold. In Cochlear (upper graph) a measured value of 130 CL using a 50µs pulse width, results in an audiometric threshold of 27 dB. Setting T to half of that or to 0 CL increases the audiometric threshold to 46 and 53 dB respectively. In AB (lower graph) the audiometric threshold is far more stable. A measured T of 40 CU results in a audiometric threshold of 25 dB. Setting T to 10% (30 CU) of M or 0 CU increases the audiometric threshold only by 2 and 7 dB respectively.

The plots can be used to give an indication of how to resolve issues related to fitting. For example, if audiometric thresholds are higher (worse) than target, the plots assist in identifying the MAP parameters and the direction and magnitude by which they can be adjusted to improve this outcome. A typical intervention might be to increase the EDR Minimum. Taking Cochlear's device for an example, the plots show that this is indeed effective. It is however not the only way to reach this goal. As illustrated in Figure 55, the audibility threshold may also be improved by either increasing Sensitivity or Gain or decreasing Q, or a combination of those adjustments. Decreasing T-SPL on the other hand does not improve the audibility very much.

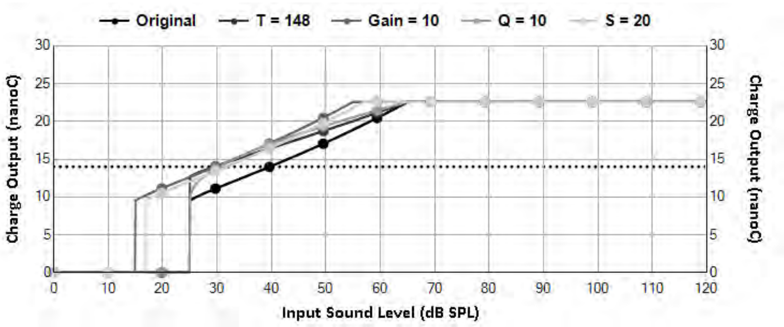


Figure 55: Improving an audiometric threshold from 40 dB to 30 dB, can be done in several ways. In Cochlear's system, changing T from 132 to 148 CL, Gain from 0 to 10 dB, Q from 20 to 10 or Sensitivity from 12 to 20 all have approximately the same effect on the detection threshold. All of these manipulations however do have different effects on other properties of the mapping function.

The same manipulations may also be used to increase speech perception at low intensities (< 50 dB SPL). To improve perception of loud speech, when for example a roll-over effect (i.e., a decrease in speech intelligibility at higher presentation levels) is observed, one may adjust parameters related to AGC compression, or decrease EDR Maximum levels. If in the context of loudness scaling or other psychoacoustic measures one would aim at improving the difference limen of intensity (DLI) decreasing the IMR to maximize the mapping resolution, may be a first approach. In any case, all of these adjustments are very likely to not only have the intended effect, but also induce side effects on other aspects of the intensity coding. The plots are instrumental in uncovering these side effects, so that one may choose that adjustment that is most effective in resolving an issue without compromising other important requirements for an optimal fitting.

2.4.5. ACKNOWLEDGEMENTS

B. Vaerenberg received a PhD grant for this work from the IWT (Agency for Innovation by Science and Technology, Baekeland-mandaat IWT090287).

2.5. A GLOBAL SURVEY ON THE STATE OF THE ART OF CI FITTING

Cochlear implant programming:
a global survey on the state of the art

The Scientific World Journal, vol. 2014, Article ID 501738, 12 pages, 2014

B Vaerenberg[1,2], C Smits[3], G De Ceulaer[1], E Zir[4], S Harman[5], N Jaspers[6], Y Tam[7], M Dillon[8], T Wesarg[9], D Martin-Bonniot[10], L Gärtner[11], S Cozma[12], J Kosaner[13], S Prentiss[14], P Sasidharan[15], J Briaire[16], J Bradley[17], J Debruyne[18], R Hollow[19], R Patadia[20], L Mens[21], K Veekmans[22], R Greisiger[23], E Harboun-Cohen[24], S Borel[25], D Tavora-Vieira[26], P Mancini[27], H Cullington[28], A Ng[29], A Walkowiak[30], WH Shapiro[31], PJ Govaerts[1]

[1] The Eargroup, Antwerp-Deurne, Belgium
[2] Laboratory of Biomedical Physics, University of Antwerp, Belgium
[3] Amsterdam (Netherlands): VU University Medical Center
[4] Beirut (Lebanon): Hôpital Sacré Cœur
[5] Bradford (UK): Yorkshire CI Service
[6] Brussels (Belgium): University Hospital St-Luc
[7] Cambridge (UK): Emmeline Centre
[8] Chapel Hill (NC, USA): University of North Carolina at Chapel Hill
[9] Freiburg (Germany): University Medical Center Freiburg
[10] Grenoble (France): CHU de Grenoble
[11] Hannover (Germany) : Medizinische Hochschule
[12] Iasi (Romania): University of Medicine and Pharamcy Grigore T. Popa
[13] Istanbul (Turkey) : Meders Hearing and Speech Center
[14] Kansas City (MO, USA): University of Kansas Medical Center
[15] Calicut, Kerala (India): Dr Manoj's ENT Superspeciality Institute & Research Centre
[16] Leiden (Netherlands): LUMC
[17] London (UK): Royal National Throat Nose and Ear Hospital
[18] Maastricht (Netherlands): University Hospital
[19] Melbourne (Australia): Royal Victorian Eye and Ear Hospital
[20] Mumbai (India): PD Hinduja Hospital National & MRC
[21] Nijmegen (Netherlands): UMC St. Radboud
[22] Nottingham (UK): Nottingham Auditory Implant Programme
[23] Oslo (Norway): University Hospital
[24] Paris (France): Hôpital Rothschild
[25] Paris (France): Hôpital Beaujon
[26] Perth (Australia): Medical Audiology Services
[27] Rome (Italy): University Sapienza
[28] Southampton (UK): South of England CI Centre

[29] *Toronto (Canada) : Sunnybrook*
[30] *Warsaw (Poland): World Hearing Center*
[31] *New York (NY, USA): NYU Langone Medical Center*

Abstract

The programming of cochlear implants (CI) is essential for good performance. However, to date still no commonly accepted Good Clinical Practice guidelines exist. This paper reports on the results of an exhaustive inventory of the current practice worldwide. A questionnaire on current practices was distributed to CI centers worldwide. The results were discussed during an International Debate and checked and verified by means of individual interviews during the months after the Debate. In addition all centers were invited to participate in a cross-sectional study logging the details of 5 consecutive CI fitting sessions in 5 different CI recipients. Descriptive statistics are used to present the results in terms of 5 parameters (median, quartiles, and extremes), histograms and box and whisker plots. Forty seven CI centers filled out the questionnaire. All together they follow 47000 CI recipients in 17 countries and 5 continents. Sixty-two percent of the results were double-checked by individual interviews and 72% of the centers returned the cross-sectional data for verification. Data indicate that general practice starts with a single switch-on session, followed by three monthly sessions, three quarterly sessions and then annual sessions, all containing one hour of programming and testing. The main focus lies on setting maximum and to a lesser extent minimum current levels per electrode. These levels are often determined on a few electrodes and extrapolated for the others. They are mainly based on subjective loudness perception by the CI user and to a lesser extent on free field audiometry and speech audiometry. Objective measures play a small role as indication of the global MAP profile. Other MAP parameters are rarely modified. Measurable targets are only defined for free field audiometry. Huge variation exists between centers on all aspects of the fitting practice.

2.5.1. INTRODUCTION

Cochlear implants (CI) processors must be appropriately programmed and customized for the recipient [84] [85]. The aim of this is to set a number of parameters to ensure that the electrical pattern generated by the device in response to sound, yields optimal speech intelligibility. Several electrical parameters are available and all their values together are commonly called the MAP. Finding and programming the optimal values for a recipient is commonly called the act of fitting. It is achieved using proprietary software and a hardware interface connected to the processor, and depends on behavioral responses from the CI recipient.

After the initial switch-on or activation of the processor, several fitting sessions are normally required [98]. Most of the MAP adjustments take place over these first few months, until levels remain relatively stable [99] [100] [98]. Following stabilization of electrical dynamic range, fitting sessions are usually limited to periodical checks, typically annually, as long as progress remains satisfactory.

Training in fitting is usually provided primarily by the CI manufacturers, and although there are guidelines and recommendations, no standardized methodology exists. There are no agreed standards or targets for what should be adjusted, or the outcomes expected; as a consequence the MAP a recipient receives could be very different depending on the center visited and the individual heuristics of the audiologist responsible. Most implant teams have an expert opinion of what the expected level of performance for an individual recipient should be and more detailed adjustments are made to the MAP if this target is not reached.

This paper attempts to describe the current state of the art by providing a comprehensive inventory of the fitting strategies in a substantial number of CI centers worldwide. It is beyond the scope of this paper to explain the meaning of all possible MAP parameters or settings. For this information, the reader is referred to the companies' user manuals and to existing comprehensive overviews [88] [85].

2.5.2. MATERIAL AND METHODS

In preparation to an international debate which was organized in Antwerp, Belgium in October 2012, a questionnaire was distributed to 47 CI centers worldwide. All questionnaires were returned. All responses were analyzed and the data were discussed during the 2 day debate. After this debate all centers were invited to a remote interview (telephone or Skype) to clarify and correct the answers where needed. In addition, the participating centers were invited to log one single fitting session in 5 consecutive recipients of one same CI brand in the months of September – October 2012. This yielded a prospective cross-sectional snapshot of the actual fitting procedure which served as verification for the questionnaire statistics.

The questionnaire was available online. Briefly the questions focused on the following topics:

- Number of implant recipients being followed and the annual increase
- Brands of implants being implanted and fitted
- MAP parameters being modified from default at switch-on and during the follow-up
- Assessments undertaken (subjective, objective, psychoacoustic) and used to steer the MAP modifications
- Well defined targets used

The cross-sectional log files contained for each subject the actual values of the different MAP parameters and whether they had been modified during the session under study. In addition, they also contained the information on whether or not objective or psychoacoustic measures were executed during the session.

We summarized all answers either numerically (counts, percentages) or categorically (for instance: the categories never, exceptionally, sometimes, regularly, always). In addition, all supplementary information, nuances, specifications were recorded when relevant.

Descriptive statistics were used and the results are presented graphically by means of histograms or box and whisker plots. Distributions are described by medians, quartile ranges (QR: between 25th and 75th percentile), extremes (minimum and maximum) and outliers.

The term Cochlear is used for the Nucleus device (Cochlear Corporation, Sydney Australia), Med-El for the Med-El device (Med-El, Innsbruck Austria), AB for the Advanced Bionics device (Advanced Bionics Corporation, Valencia California) and Neurelec for the Digisonic device (Neurelec, Vallauris France). Throughout the text, the term minimum level is used for the T, THR, T or MIN parameters of Cochlear, Med-El, AB and Neurelec respectively. The term maximum level is used for the C, MCL, M or MAX parameters. In this paper the term eCAP (electrically evoked compound action potential) is interchangeable with eCAP threshold measurements and refers to (t)NRT, (t)ART, (t)NRI for Cochlear, Med-El and AB respectively.

2.5.3. RESULTS

2.5.3.1. PARTICIPATING CENTERS

Forty-seven centers from 17 different countries (Australia, Belgium, Canada, France, Germany, India, Italy, Lebanon, Morocco, Norway, Poland, Romania, Spain, The Netherlands, Turkey, United Kingdom and USA) and 5 different continents (Europe with 60% of centers, North-America 11%, Asia 4%, Australia 4%, Africa 2%) filled out the paper survey (see full list at the end of the paper). All together they were following 47600 CI users with an annual increase of 4800. Twenty-nine centers had a representative being interviewed. They were following 37000 CI recipients with an annual increase of 3700. This means that the responses of 62% of the participating centers were double-checked covering 78% of the CI-recipients being followed. The cross-sectional snapshot yielded data from 255 fitting sessions of 34 centers.

The participating centers have an average experience of 21 years (median start-up in 1991; QR: 1987-2000) and a median number of 625 implants (QR: 338-1300) with 62 new implants last year (QR 50-123).

On average each center provides 3 CI brands (Cochlear, Med-El and AB); 10.5% provide only 1 brand; 10.5% 2 brands; 55% 3 brands and 24% 4 brands. The predominant device is Cochlear in 43% of the centers, Med-El in 29%, AB in 25% and Neurelec in 4%. For all three major brands we received responses from at least 26 centers of which at least 15 were interviewed afterwards. Only Neurelec was underrepresented with 4 centers responding on paper of which 3 were interviewed. For the cross-sectional verification, at least 14 centers returned the log files of 5 consecutive CI users for each of the major brands. For Neurelec 7 centers returned the log files.

79.5% of centers in the study provide implants to both children and adults, 17% to adults only and 3.5% to children only.

2.5.3.2. SWITCH-ON PROCEDURES

On average, the CI processor is switched on after 28 days (QR: 21-30) with some centers starting after 2 weeks (Perth, Melbourne, Chapel Hill) while one center only hooks up the processor after 6 weeks (Cambridge).

All centers (100%) start with impedance measurements and if short or open, most (60%) deactivate the corresponding electrodes immediately. Two centers (Brussels, Freiburg since 2013) indicate to systematically execute a pure tone audiometry prior to switch-on to assess possible residual hearing, while another centre (Hannover) does this during the switch-on week (see further).

Most, if not all centers' focus goes to the setting of the minimum and the maximum current level of the electrodes. Med-El has a default THR level of 0 and 70% of centers don't change this. AB recommends setting the T level at 10% of the M level and 22% of centers do so. A majority of centers (55%) only determine either the minimum (31%) or the maximum (24%) level and make the other level depend on the first one. Forty-five percent of centers determine both the minimum and the maximum level behaviorally.

2.5.3.3. DETERMINE MINIMUM LEVEL ALONE

If only the minimum level is determined, this is either done behaviorally (56%) or by means of intraoperative or postoperative eCAP thresholds (44%). The eCAP measures are mostly followed by behavioral verification and adjustment if need be. Most centers (78%) only determine the minimum levels on a few electrodes and interpolate the values obtained to the other electrodes. Maximum levels are then positioned at one or more intervals above the minimum levels and most centers (67%) perform some form of loudness balancing before switching on the microphone. One centre (Leiden) uses a preset profile of maximum levels which is positioned above the determined minimum levels.

2.5.3.4. DETERMINE MAXIMUM LEVEL ALONE

Determining only the maximum level is restricted to Med-El and AB implants where the minimum level is then set at 0 or 10% of the maximum level. The maximum level is either determined behaviorally (71%) or by means of objective measures (eCAP in 29%, which is combined with or replaced by ESRT (electrically evoked stapedius reflex thresholds) in 14%). If objective measures are used, behavioral verification is done by half of the centers. Interpolation is used in only a minority of centers (29%) and so is loudness balancing (43%).

2.5.3.5. DETERMINE BOTH MINIMUM AND MAXIMUM LEVEL

Many centers determine both the minimum and the maximum levels and they all do this behaviorally. Only 15% of these combine this with eCAP measures. One center (Antwerp) has a particular way of using preset MAPs with minimum and maximum levels based on statistical analysis of MAPs which have provided good results in other recipients [101] [102]. These preset MAPs are given without any prior behavioral or other evaluation. Most centers (69%) measure the levels on a number of electrodes and interpolate the levels on the other electrodes. In some cases this can be as few as 3 electrodes (Southampton, Iasi), the results of which are then used to shift a preset profile towards the measured levels. Most centers perform some kind of loudness balancing (62%).

In general, if the maximum levels are measured rather than derived from the minimum levels, half of the centers (50%) reduce these levels before switching on the microphone. Just after switching to live mode, almost all centers (93%) increase or decrease the maximum levels based on the recipient's perception and some (45%) also shift the minimum levels. A small number of centers perform some kind of psychoacoustic test immediately after switch-on, e.g. filtered Ling sounds loudness scaling (Nijmegen), Ling sounds detection (Perth), or closed or open set word understanding (Paris, Chapel Hill).

Most centers (76%) send the CI user home with incremental MAPS after the switch-on session. These MAPS contain progressively higher maximum levels allowing the CI user to accommodate to each MAP before switching to the next one. Some centers (17%) set a large volume range and instruct the CI recipient to increase the volume progressively over time. One center (Hannover) replied that they don't systematically increase the maximum level over time.

2.5.3.6. OTHER MAP PARAMETERS

Figure 56 A shows that other MAP parameters are rarely modified from default during the switch-on session.

COCHLEAR

Thirteen percent of centers prefer more than the default 8 Maxima (9, 12 or 14) and 6% combine this with a higher than default Channel Rate (1200 pps). The Autosensitivity function is switched off by 13% of centers at switch-on. The Eargroup in Antwerp sets different Gains (statistically defined profile), Analysis T-SPL (20), Analysis C-SPL (70) and switches off the ADRO function. The latter is also done by Nottingham where T-SPL is set to 25 dB and C-SPL set to 75 dB at switch-on, in combination with a Q-factor of 16; both ADRO and ASC are deactivated. Paris also sets the Loudness Growth Function (Q-factor) at 16. The Volume Adjustment is set to 0 by 10% of centers, all located in the UK.

MED-EL

With the Med-El device, 23% of centers start with a different strategy than the default FS4 strategy. Chapel Hill provides the patients with two strategies, HDCIS or FSP, which are the two strategies approved for use in the USA. Perth lets the patients chose between FS4 and FS4p and has experienced that 90% of recipients prefer FS4p. Paris-Avicenne gives FS4p as startup strategy and York, Paris-Beaujon and Kansas City give FSP as startup strategy. The lowest filter frequency is set to 70 Hz by 23% of the centers. Paris-Avicenne overrules the default settings for Highest Frequency (set to 8000 Hz), AGC Sensitivity (set to 85%) and MapLaw (set to 1000), Nottingham overrules the default Minimum Pulse Width Duration (set to 20µs) and Nijmegen uses a high MapLaw setting (1000).

ADVANCED BIONICS

With the AB device a majority of centers overrule the default strategy (HiRes-P) and start with the HiRes-S strategy (72%) and of those two thirds select the Fidelity 120 strategy compared to one third who sticks to the default setting with Fidelity 120 switched off. This is in contrast to the centers who keep the HiRes-P strategy, of which 78% also keep the default setting with Fidelity 120 off. Some centers (20%) switch on Clearvoice systematically and some centers (30%) change the default Pulse Width setting of 10.8 µsec to either a higher value or to the automatic Pulse Width algorithm II (APW2). The default input dynamic range (IDR=60 dB) is changed by 24% of centers. Some lower it to 50 dB (Las Palmas, Paris-Avicenne) or 54 dB (Naples) while others increase it to 70 dB (London St-Thomas, Beirut, Kerala) or to 80 dB (Antwerp). Antwerp also sets the sensitivity to -10 dB and the Gains to a preset profile which differs from the default values (0 dB).

NEURELEC

The statistics of Neurelec's Digisonic device are not solid since they are derived from merely four centers, one of which (Southampton) only uses the binaural version. Half of them change the default number of maxima from 12 to 11 (Antwerp) or 6 (Southampton) and one center (Antwerp) switches the stimulation rate systematically from 600 pps to 500 pps and the pre-emphasis (égalisation de sonie) to -1.

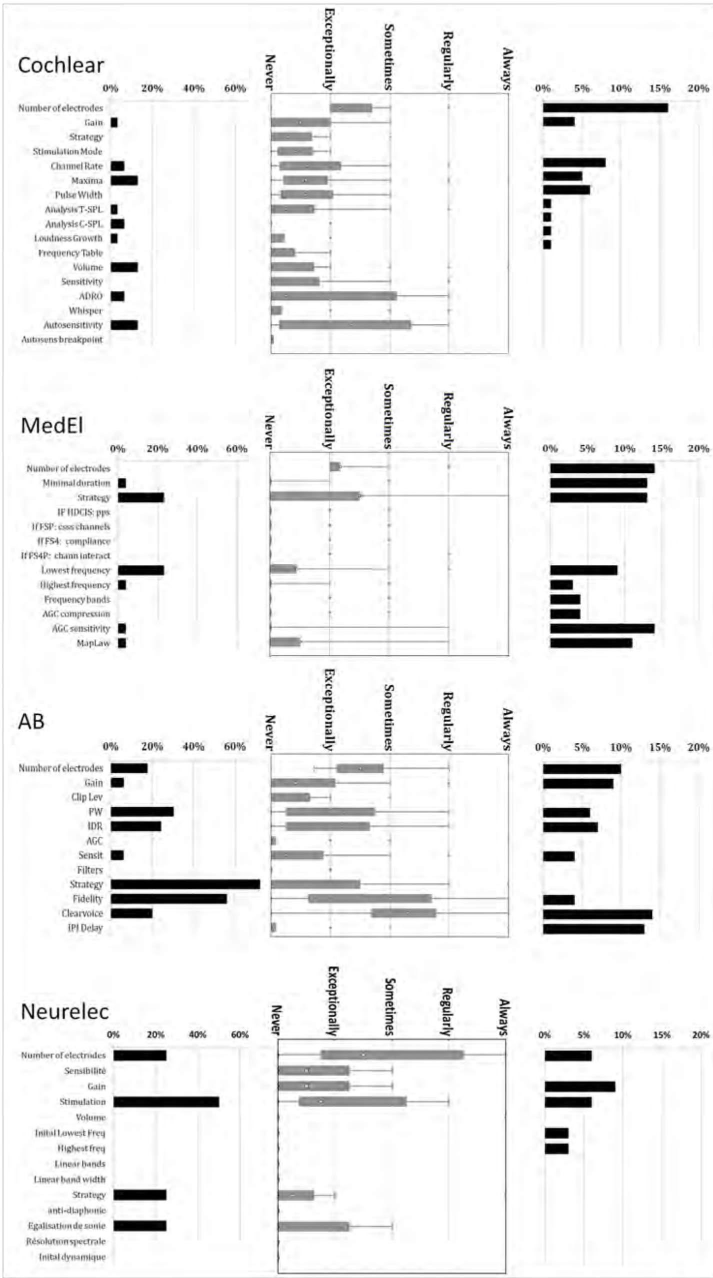


Figure 56: Occurrence of MAP changes for the 4 brands (Cochlear, Med-El, AB, Neurelec). (A) The left pane shows the frequency of changing the default settings at switch-on, as retrieved from the questionnaire and the interview; (B) the mid pane shows the distribution of the frequencies of changing the MAP parameters during the follow-up sessions, as retrieved from the questionnaire and the interview (Box and Whisker plots with the central dot depicting the median value, the box the quartile range and the whiskers the range); (C) the right pane shows the occurrence of MAP changes as observed in the cross sectional snapshot.

2.5.3.7. TIME

Figure 57 shows that 71% of the centers consider the switch-on as 1 single session. Eleven percent spread the switch-on over 4 sessions or more, often on consecutive days. One center organizes the switch-on over 7 sessions (Oslo). The median cumulative time spent at is 1 hour. This does not take into consideration the time spent for counseling or instructing the patient. As said earlier, testing is rarely undertaken during the switch-on session. There are outliers who spend more than 3 hours on fitting (Coimbra, London St-Thomas, Freiburg, Southampton, Kiel, Oslo) or more than 1 hour at testing (Oslo, Montpellier, Kiel). One center reports spending no more than 5 minutes in total, which is the result of a fully automated switch-on (Antwerp). There are no statistically significant differences between centers who perform cochlear implantations mainly in children compared to mainly in adults (Mann-Whitney U-test p>0.05).

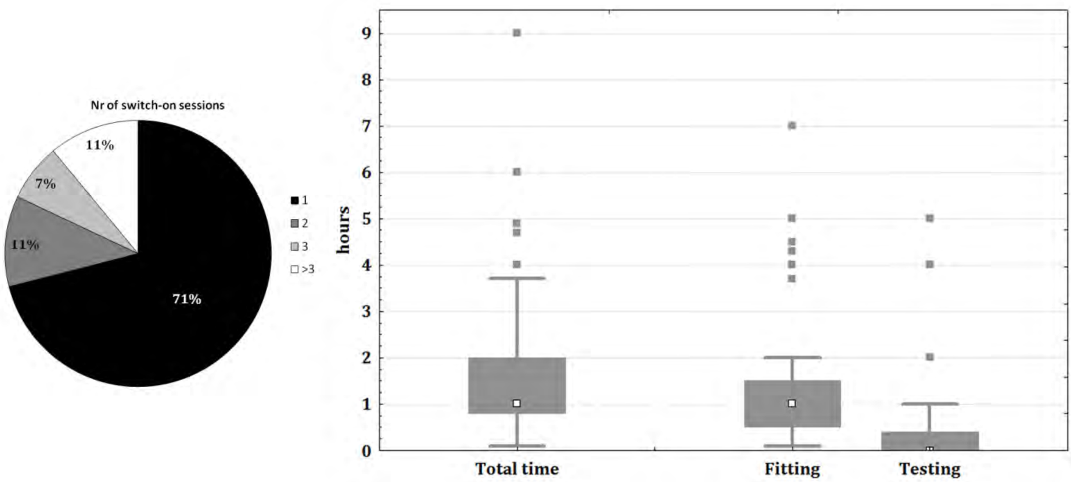


Figure 57: Time analysis of the switch-on session, showing the number of sessions at daily intervals which are considered to constitute the switch-on procedure (pie chart at the left) and the time spent at the switch-on session (box and whisker plot at the right), both the total time and its break-down into time spent at fitting and at testing. Time for counseling has not been enquired in this study. See caption of Figure 56 to interpret the box and whisker plots.

2.5.3.8. FOLLOW-UP PROCEDURES

After the switch-on session, all centers schedule a number of consecutive sessions to reach stable MAP settings. The average center schedules 3 sessions in the first quarter, 3 sessions in the following 3 quarters and 1 annual session afterwards (see further). Attention goes mostly to the verification and adjustments of minimum and/or maximum levels to optimize loudness and almost half of the centers (46%) explicitly say that the follow-up sessions are roughly the same as the switch-on session.

ADJUSTMENT OF MINIMUM AND MAXIMUM LEVELS

All centers adjust maximum levels and many (61%) also adjust minimum levels. Global shifting of the maximum profile is very common (96%) while tilting is done by less than half of the centers (39%). One centre lets the CI-user set and balance his/her own maximum level to most comfortable (Grenoble). All centers perform some kind of loudness balancing across individual electrodes and some centers perform pitch ranking (17%).

Psycho-acoustical tests (tonal audiometry, speech audiometry) or objective measures (eCAP, ESRT) are commonly performed (see below). Fourteen percent of the centers report to use these early stage sessions to try out different strategies or different settings of other MAP parameters than minimum and maximum levels.

ADJUSTING OTHER MAP PARAMETERS

Figure 56 B shows that MAP parameters other than minimum and maximum levels are rarely modified. This is further illustrated by Figure 56 C showing the cross-sectional observations. Deactivation of electrodes is one of the more common actions, but centers still report to do this only every now and then (median response value is between exceptionally and sometimes, corresponding to approximately 10-15% in the cross-sectional data). Figure 58 shows the reasons reported to inactivate electrodes. The most commonly reported reason is abnormal impedances, which is reported to occur ‘sometimes’. Electrodes are also deactivated for other reasons such as when there is an indication of extracochlear location, or if they cause non-auditory stimulation, uncomfortable or no perception, if the maximum levels are exceptionally high or if tonotopical tests such as pitch ranking, channel separation or spectral discrimination show unexpected results. These situations are reported to occur almost never. Electrodes are hardly ever deactivated based on loudness assessment or objective measures. Other exceptional reasons of electrode deactivation are negative results on integrity tests or the desire to increase the stimulation rate. One centre used to systematically start with one or more inactivated electrodes (Leiden) [103], a habit which has only recently been abandoned.

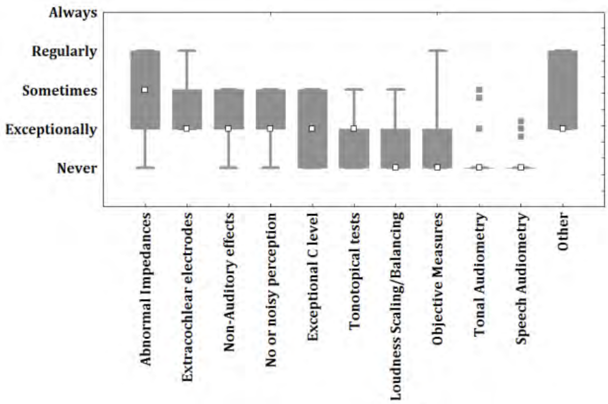


Figure 58: Alleged reasons for deactivating electrodes and the frequency they are reported to be really responsible for electrode deactivation in daily live.

COCHLEAR

With the Cochlear device, the additional MAP parameter which is modified most, though still only exceptionally is the Autosensitivity feature, which is then deactivated. In the cross-sectional data, also channel rate, number of maxima and pulse width were modified in 5-8% of cases.

MED-EL

With the Med-El device, the Strategy is reported to be changed in ‘some’ cases. Some centers change the default strategy (FS4 except in the USA) to FSP or FS4p in exceptional or some cases. One centre routinely sets the strategy to HDCIS in the primary program (Chapel Hill) and lets the patient choose between this strategy and FSP. This was confirmed in the cross-sectional data, which also showed that AGC Sensitivity, Minimum Duration and MapLaw were changed in 11-14% of the cases and by many centers (36-79%).

ADVANCED BIONICS

Advanced Bionics has more MAP parameters modified by a substantial number of centers in the course of the early follow-up period. The Clearvoice feature is activated sometimes to regularly (14% of cases in the cross-sectional study and 36% of the centers), and also the Fidelity 120 feature is sometimes changed. Pulse width and IDR are next in line, but they are only changed in exceptional cases. This is confirmed by the cross-sectional data where these MAP parameters were only changed in 4-7% of the cases and by less than 25% of the centers. In the cross-sectional data the pulse rate was more often changed (IPI delay, 13% of cases, and 29% of centers).

NEURELEC

Neurelec again has too few data to allow any reliable statements. The results are nevertheless included in the graphs for completeness and illustration.

TIME

Figure 59 shows that most centers schedule between 5 and 8 additional sessions during the first year (median=6; QR: 5-8, range: 3-15). The median cumulative time spent at the acts of fitting and testing during the first year after switch-on is 6 hours (median for fitting = 3.3 hours and for testing = 2.0 hours). There are no significant differences between centers who perform cochlear implantations mainly in children compared to mainly in adults (Mann-Whitney U-test p>0.05).

After the first year, the median number of sessions per year is 1 (QR: 1-1, range 0.3-1 with one outlier with 3 annual sessions). The median time spent is 1.3 hours (QR: 0.9-2.0 hours, range 0.5-4 hours with one outlier of 8 hours per year) of which 0.5 hours for fitting and 0.8 hours for testing.

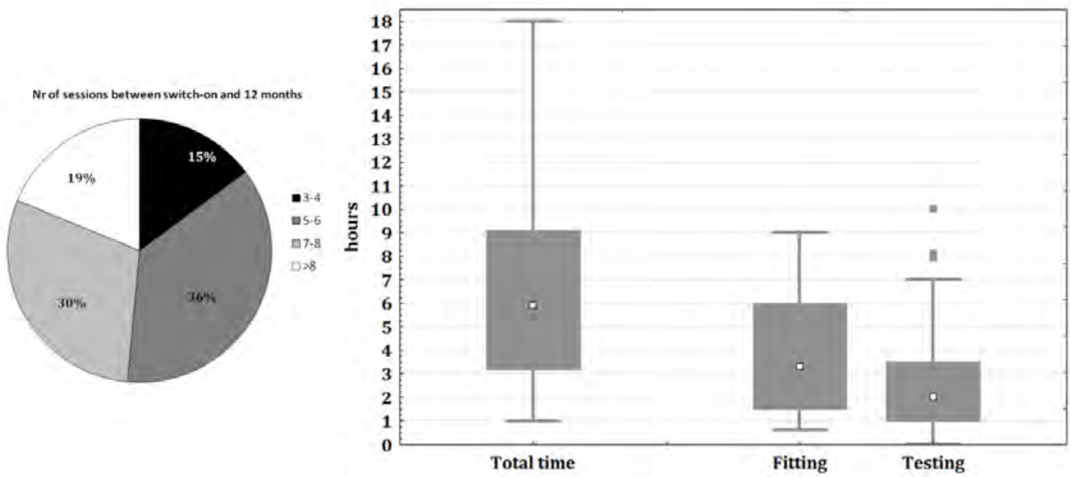


Figure 59: Time analysis of the follow-on sessions during the first year after switch-on, showing the number of sessions (pie chart at the left) and the cumulative time spent at them (box and whisker plots at the right), both the total time and its break down into time spent at fitting and at testing. Time for counseling has not been enquired in this study. See caption of Figure 56 to interpret the box and whisker plots.

OUTCOME MEASUREMENTS

Figure 60 shows that most centers report to assess subjective features and use them for fitting. Overall comfort (93%), auditory comfort (83%) and the presence of non auditory sensations (83%) are used by most centers. None of the centers report well defined and measurable targets for any of these features. Non auditory satisfaction, such as contentment, quality of life, implant use are commonly assessed (87%) but only used by 41% of the centers to change the MAP settings.

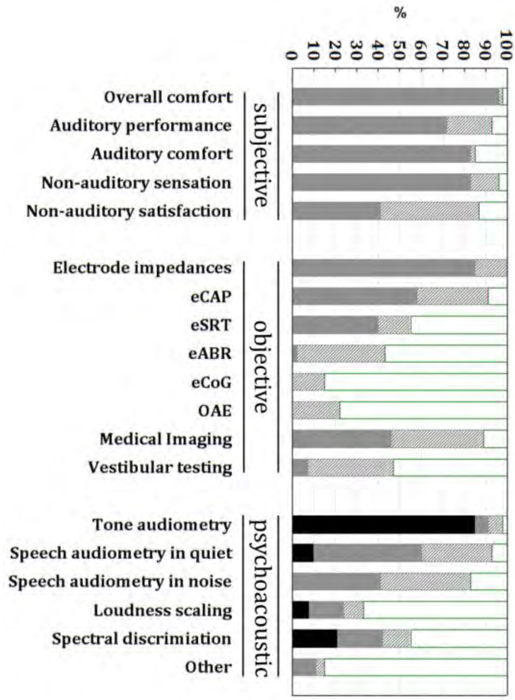


Figure 60: the different outcome assessments which were enquired in the questionnaire together with the frequencies of the responses. The outcomes are grouped into 3 groups (subjective, objective and psychoacoustic outcomes). The possible answers were (1) yes we assess this and use it to optimize the fitting (solid black and grey bars), (2) yes we assess this but for other reasons than steering the fitting, like for documentation or longitudinal follow-up (shaded bars) or (3) no we don't use to assess this (white bars). For the solid bars (assess and use it) a distinction was made into whether they have well defined targets to reach (black) or not (grey).

Of the objective measures [104], electrode impedances are measured by 100% and used by 85% of the centers. They are used to deactivate electrodes in case of short or open circuit. Thresholds based on eCAP [105] or eSRT [106] measurements are used by 59% and 39% of the centers respectively. They are mainly used to set the MAP profiles. Medical imaging is used by 46% of the centers to change the fitting, mainly to deactivate electrodes which are believed to be extracochlear. Other objective measures may be performed, but they are not used to drive the fitting. None of the centers report to use objective measures to reach well defined targets during the fitting, except for Nottingham, where ESRT measures are used to loudness balance the MAPs. The cross-sectional data confirmed that besides impedance measurements, no objective measure was performed in more than 5% of the cases.

Psychoacoustic measures are the only outcome measures for which a number of centers have well defined targets. This holds mainly for free field audiometry (85%) with targets set between 20 and 40 dB HL (median 30 dB HL, QR: 25-35 dB HL, see Figure 61). Spectral discrimination tests are used to drive the fitting by 41% of the centers of which 20% use well defined targets (either 100% if the A§E phoneme discrimination test [107] is used or 83-100% if Ling sounds are used). Speech audiometry in quiet or in noise are reported to be used to change the MAP parameters by 61% and 41% respectively, but only 11% of the centers have set well defined targets and this is only for speech audiometry in quiet. No two centers have set the same target for this measure however. Acoustical loudness scaling is used to change the fitting by 24%, but only 8% have well defined targets, which are the same across centers, namely results falling in the normal zone (of hearing listeners).

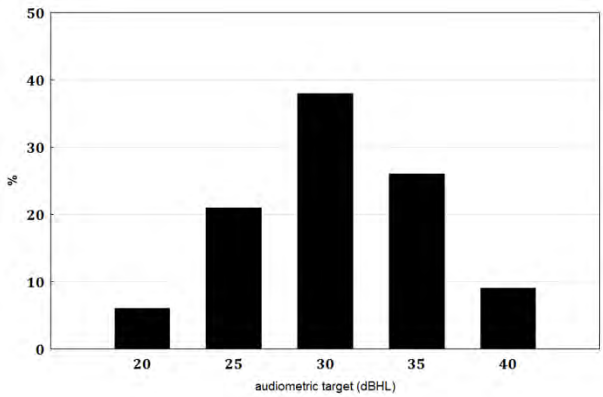


Figure 61: Histogram showing the frequency of the reported audiometric targets (dB HL) at different centers.

The cross-sectional data confirm that free field audiometry was performed in 60% of the cases, speech audiometry in quiet in 45%, speech audiometry in noise in 19%, loudness scaling in 11% and spectral discrimination tests in 15% of the cases. Other tests used were speech tracking and Ling sounds detection, discrimination and loudness scaling tests, but these were very rare.

2.5.4. DISCUSSION

Multichannel intracochlear implants have been clinically available for more than 25 years now. The fitting of the processors to the individual recipient is considered to be crucial in obtaining good results. To date there is no well described and commonly adopted Good Clinical Practice (GCP) for this act nor is there evidence based material to distinguish efficient procedures from less efficient ones. Over these 25 years, fitting a CI has been carried out by competent clinicians who have established their own heuristics, good practices and empirical knowledge. It seems reasonable to believe that a critical analysis of the cumulative knowledge acquired over the years may serve as a first step towards a definition of GCP. This report attempts to give an inventory of the current state of the art as it is based on a vast number of CI centers worldwide. All together they represent over

47000 CI-recipients and 93% of the participating centers have more than 10 years of experience. 65% of the centers are European, which may cause a bias towards an overrepresentation of European habits. Altogether this is an unprecedented inventory and we believe it gives a representative view on the current practices in CI fitting, which may be considered as the benchmark of CI fitting anno 2013.

It is important to consider that the conclusions are based on a compilation of a written questionnaire, an oral interview and a cross-sectional fitting data snapshot. An intrinsic weakness of such an approach is that it lacks precision. It is based on the anamnestic summary of centers' practices as provided by only one representative per center, whereas different practices may exist within one center. Yet in the absence of hard evidence or more accurate overall data, an exploratory inventory like this one is a very legitimate and necessary first step towards a better understanding of the field. The cross sectional sample serves as verification and substantially improves the validity of the data. We advise the reader not to interpret the presented numerical data as ultimately accurate but rather as indicative while always keeping a confidence interval in mind.

A first observation is that most centers now offer 3 CI brands and perform cochlear implantation in both children and adults. This is different from years ago when many CI centers only offered one brand and only performed CI in adults.

A second observation is that despite the huge variability across centers (see further), some common practices can be extracted and they would seem to be as follows.

The typical switch-on procedure takes one session comprising counseling and 1 hour of fitting. Testing is not performed at this stage. The fitting procedure is as follows.

1. Connect the processor 4 weeks after surgery;
2. Measure impedances and deactivate electrodes in case of short or open circuits;
3. Measure the maximum level behaviorally on a number of electrodes along the electrode array, and interpolate the others;
4. Set the minimum level at 0 for Med-El, 10% of M for AB; for the other implants measure the minimum level behaviorally on a few electrodes and interpolate the others;
5. Perform loudness balancing by presenting a signal on all electrodes successively;
6. Reduce the maximum level and switch on the microphone;
7. Let the CI recipient accommodate for a few minutes and ask whether sounds, including loud sounds, are tolerated; decrease or increase the entire profile of maxima in order to make loudness tolerable or comfortable;
8. Put a number of progressive MAPs in the processor;
9. Instruct the patient to accommodate to each program for a couple of days and switch to the next one afterwards.

The typical first year follow-up would comprise three monthly sessions followed by three quarterly sessions of one hour each. The sessions would typically look like this:

1. Perform free field audiometry and speech audiometry (in quiet);
2. Measure impedances and deactivate electrodes in case of short or open circuits;
3. Verify the levels on individual electrodes by loudness balancing;
4. Shift the profiles of the maximum and if necessary also of the minimum levels globally;
5. If deemed necessary, tilt the maximum levels globally;
6. Define own criteria to identify selected and exceptional cases in whom other MAP parameters are modified.

We believe that the description of these typical procedures is the common denominator of current practices and could serve as guidelines for newcomers in the field, as backbone of instructional courses, etc.

But, as said, it is remarkable to observe the substantial variability across centers and this holds for virtually all aspects of CI fitting and follow-up. Each CI center has its own policy in terms of timing, content and methodology.

The switch-on of the processor is scheduled between 2 and 6 weeks after surgery. It is obvious that concern about wound healing is a reason not to activate the processor too early. On the other hand one does not like to deprive the CI recipient from audition for too long a period and it has been shown that the natural and progressive increase in electrode impedances after surgery is discontinued by electrical stimulation [108] [100]. These may be factors in favor of early device switch-on. The observation that some centers commonly activate the processor as soon as 2 weeks after surgery, seems to suggest that 2 weeks may still be well within the safe time window.

On average, CI recipients undergo one switch-on session followed by six fitting sessions in the first year, each taking approximately one hour of technical interaction (fitting and testing) plus a considerable amount of counseling time which has not been enquired in this survey. Behind this average there are huge between-center differences. Even within one center there may be important differences between different clinicians and between different patients (for instance children compared to adults). Most centers have one switch-on session followed by a take-home experience for accommodation. Some centers however schedule 5 to 7 consecutive sessions at daily intervals (Hannover, Freiburg, Oslo). This may be based on the experience that such intensive schemes lead to stable MAPS fast or it may be for practical reasons, for instance for patients who are living at a distance from the CI center. In some cases this is compensated by less follow-up sessions in the first year, like in Hannover and Oslo, where no more than 4 follow-up sessions are scheduled in the year after switch-on, but not in Freiburg, where 10 more sessions are planned in the first year, a scheme which fits in a well established rehabilitation concept. In the year following switch-on, some centers spend no more than approximately 1.5 to 2.5 hours (Paris-Avicenne, Casablanca, Ghent, Pune, Mumbai, Hannover, Berlin, Valencia, Lyon) and one center schedules only 3 sessions (Warsaw), while other centers spend at least 12 hours (Las Palmas, Leiden, London St-Thomas, Amsterdam) or as much as 15 sessions (Brussels). After the first year, there is more consistency in terms of follow-up. Almost all centers have one session per year which takes between 1 and 2 hours of technical interaction (fitting and testing) with the CI user. Three centers have less than 1 annual session (Hannover and London-RNTNE every second year, Nijmegen every third year and Mumbai on patient's request). It seems that these annual sessions are merely planned for verification and to reassure that the performance has not deteriorated, rather than for modifying the MAPs which remain rather stable after the initial months [109]. From that perspective it seems justified to increase the interval of one year. However, informal feedback from centers has revealed that these annual visits are also felt to be important for technical check-up of the processor and the microphone function and for ongoing counseling. When asked about this during the above mentioned International Debate, 62% of the participants voted that annual visits were essential during the first 5 years and this figure dropped to 28% after 5 years.

When it comes to the content of fitting and follow-up, most attention goes to the setting of minimum and maximum levels per electrode. Every center appears to have its own policy on how to determine these levels. Behavioral assessment is commonly used, but whereas this was performed for each individual electrode in the past, it now seems common to assess the levels on a few electrodes only and to deduce them by interpolation for the remaining electrodes. This probably coincides with the change from bipolar to monopolar stimulation and is based on growing evidence that such approach yields equally good results [110]. Evoked potentials (mainly eCAP thresholds) are used by an important minority of the centers, but they appear to be used as global indication of minimum or maximum levels rather than as strict anchor points. The levels set this way are preliminary anyhow, since they are shifted and to a lesser extend tilted in live mode, mainly based on subjective appreciation of loudness. Other MAP parameters are rarely modified. Some default settings are systematically overruled by a large number of centers, which probably reflects their conviction that the default settings do not necessarily give the best results. It seems obvious for CI companies to take this into consideration and to change some of their default settings. Deactivating electrodes is the most frequent next MAP modification, although this remains rare. This may be subject for reflection since indications exists that selectively deactivating electrodes may substantially improve auditory performance. When asked whether the selective dropping of one electrode may cause a significant improvement on speech understanding, 95% of the participants in the debate voted affirmatively. However, it remains difficult to identify such electrodes. The current survey demonstrates that centers have their own and often different methods to do so (Figure 58). Finding a valid method to identify electrodes which, when deactivated, cause a significant improvement in auditory performance, may be a very legitimate subject for future research.

Most striking is the observation that the centers rely mostly on the recipient's subjective feedback to drive the MAP changes. This is remarkable since many CI users have no clear reference point to estimate the subjective quality of sound, either because they have never had normal hearing before or because they have been deprived of normal hearing for many years and have got used to hearing aid sound over the years. In addition many experienced clinicians indicated that patient's subjective judgment may not coincide with optimal performance. Objective measures are only used to get a prior estimate of the shape of the minimum or maximum levels, but they almost never serve the fine tuning of the device. One would expect that after more than 25 years of cochlear implantation, the field had developed psychoacoustic targets to steer the device fitting, but this survey shows that targets only exist in terms of audiometric thresholds. One of the reasons for this may be the fact that it is not always obvious how to modify the MAP parameters if targets are not met. Current CI systems are complex and predicting their behavior after changing the MAP is not obvious. On the other hand, such reasoning is also vicious and defining targets may be an incentive to explore and develop procedures to achieve target in the most efficient ways. Most centers agree on a target of 30 dB HL (± 5dB) for most audiometric frequencies, and this is achievable with current microphones and front-end processing. Auditory performance, however, hardly depends on thresholds, but rather on supraliminal sound processing. The core function of the cochlea is discriminating the different features of sound, such as loudness, spectral content and temporal content and it is striking to see that less than 50% of the centers report to base their fitting on measures to assess this and that less than 25% report to have targets in this domain [107] [111]. Speech audiometry in quiet or in noise relates to the daily auditory performance but depends on more than just the cochlear processing of sound. Therefore speech audiometry is only partly indicative of the quality of cochlear functioning. Speech audiometry is used by approximately half of the centers but most use it to monitor performance, i.e. to detect any undesired deterioration over time. Only 11% report to have well defined speech audiometrical targets when it comes to CI fitting. This is in line with instructional literature which extensively explains the available methodology and how to use it to determine the minimum and maximum levels but which avoids mentioning measurable targets [112] [113] [114] [88] [85]. Shapiro coined the term 'common lethargy' when referring to the CI audiologists' willingness to consider changes in device programming and he correctly stated that device programming is not a goal per se, but that the absolute goal is to provide the patient with a comfortable program which ensures maximum performance [85].

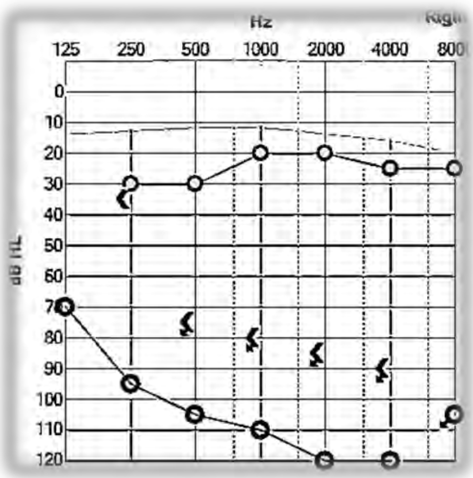
In conclusion, it seems fair to summarize the current state of CI fitting as setting global profiles of maximum current levels and to a lesser extent of minimum current levels, mainly based on subjective feedback from the CI user. Many different approaches exist and in the absence of targets or well defined outcome measures, it seems impossible to compare all these differences and to judge whether some yield better results or are more efficient than others. It is likely that several approaches in the hands of different experts may lead to similarly good results. It is equally likely that defining common measurable targets may be a next step to be taken towards the optimization of the art of fitting.

2.5.5. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the following CI centers who returned the questionnaire on fitting practices (the asterisk indicates CI centers who also accepted to be interviewed by the first author): *Amsterdam (Netherlands): VU University Medical Center, *Antwerp (Belgium): the Eargroup, *Beirut (Lebanon): Hôpital Sacré Cœur, Berlin (Germany): Unfallkrankenhaus Berlin, Birmingham (UK): Midlands Hearing Implant Programme, *Bradford (UK): Yorkshire CI Service, *Brussels (Belgium): University Hospital St-Luc, *Calicut, Kerala (India): Dr Manoj's ENT Superspeciality Institute & Research, *Cambridge (UK): Emmeline Centre, Casablanca (Morocco): RIFMED, *Chapel Hill (NC, USA): University of North Carolina at Chapel Hill, Frankfurt am Main (Germany): University Hospital, *Freiburg (Germany): University Medical Center Freiburg, Ghent (Belgium): University Hospital, *Grenoble (France): CHU de Grenoble, *Hannover (Germany) : Medizinische Hochschule, *Iasi (Romania): University of Medicine and Pharmacy Grigore T. Popa, *Istanbul (Turkey) : Meders Hearing and Speech Center, *Kansas City (MO, USA): University of Kansas Medical Center, Kiel (Germany) : CIC Schleswig-Kiel, Kilmarnock (Scotland, UK): Scottish CI Programme, Las Palmas (Spain): University Hospital Gran Canaria, *Leiden (Netherlands): LUMC, London (Canada) : London Health Sciences Centre, *London (UK): Royal National Throat Nose and Ear Hospital, London (UK): St Thomas' Hospital, Lyon (France): Hôpital Edouard Herriot, *Maastricht (Netherlands): University Hospital, *Melbourne (Australia): Royal Victorian Eye and Ear Hospital, Montpellier (France): Institut St-Pierre, *Mumbai (India): PD Hinduja Hospital National & MRC, Naples (Italy): University Hospital Federico II, New York (NY, USA): NYU CI Center, *Nijmegen (Netherlands): UMC St. Radboud, *Nottingham (UK): Nottingham Auditory Implant Programme, Oldenburg (Germany): EVK Oldenburg, *Oslo (Norway): University Hospital, *Paris (France): Hôpital Avicenne, *Paris (France): Hôpital Rothschild, *Perth (Australia): Medical Audiology Services, Pune (India): KEM Hospital, *Rome (Italy): University Sapienza, *Southampton (UK): South of England CI Centre, *Toronto (Canada) : Sunnybrook, Valencia (Spain): Hospital La Fe, *Warsaw (Poland): World Hearing Center.

BV received a PhD grant for this work from the IWT (Agency for Innovation by Science and Technology, Baekeland-mandaat IWT090287). Part of the work is supported by a 7th Framework for SME grant ("Opti-Fox") from the European Commission.

CHAPTER 3: MEASUREMENTS & OUTCOME



The audiogram: the most common audiological measurement. This one shows air and bone conduction thresholds prior to cochlear implantation and sound field thresholds after cochlear implantation.

3.1. INTRODUCTION

One could argue that the ultimate outcome measure for cochlear implantation is one that can be expressed in terms of quality of life (QOL). It is after all the requirement most desirable to fulfil. Surely auditory capabilities influence a CI recipient's everyday experiences such as social interaction, school adjustment, employment and other constituents of QOL. Given that fitting affects auditory performance, a good map contributes to QOL. But, although measuring QOL is doable, its holistic nature does not allow for much analytical interpretation that is useful for tuning a map.

Basically, a CI replaces the function of the cochlea as a receptor of sound. The responsibility of the cochlea is to deliver signals to the brain in a way that they carry maximal information about the sound that is present. To achieve this goal the cochlea needs to make sure that different sounds are transduced into different electrical patterns such that the brain can tell them apart. This is referred to as the discriminative power (i.e., intensity sensitivity, frequency selectivity and temporal resolution) of the cochlea. When attempting to improve a CI user's speech perception or even quality of life by fitting a good map, it is reasonable to focus on maximizing the CI's discriminative power with respect to the available peripheral neural interface, rather than hoping to resolve issues that lie outside the inner ear, like central auditory processing, high level cognitive processes, speech perception, social skills and quality of life. There are other instances and practices, like auditory rehabilitation programs and psychological counselling, that are more suited to handle these issues.

The assessment of speech perception, known to be highly dependent on cognition, is nonetheless routinely performed by many CI clinics. Many factors contribute to speech intelligibility and the map being used constitutes only a small part of them, which makes it difficult to take effective actions, through fitting, against the observed deficits. Because of this indirect relation between speech intelligibility and cochlear performance, the results on these tests are seldom used to tune the map. Speech perception tests are rather conducted as some kind of formal evaluation of the outcome of cochlear implantation in an individual patient.

Sound field thresholds in response to pure tones (i.e., the audiogram) are also often measured. They do allow for a more analytical approach since they represent the detection ability of the auditory system, and do not depend on cognition and speech/language skills, as speech audiometry does. One could make meaningful adjustments to a map based on the observation of elevated thresholds at particular frequencies, and achieve an instant improvement in that outcome. The audiogram however, relates little to everyday auditory performance since most real-life stimuli are complex signals that present themselves supraliminally (i.e. well above threshold).

This leads to the need for a set of outcome measures that reflects real-life auditory performance, but at the same time allows for an analytical interpretation. In other words, a good set of outcome measures to be used to drive CI fitting should depend highly on the functioning of the peripheral auditory system (in our cases replaced by a cochlear implant) and as little as possible on anything else.

In addition, most outcome measurements are time and resource intensive. They require qualified personnel and calibrated equipment in order to obtain reliable results and often constitute of lengthy and repetitive processes of presenting stimuli and recording responses. To be feasible in a clinical setting, a set of outcome measures therefore also needs to be assessable in reasonable amount of time.

In today's clinical practice, two main types of measurements are used in the context of hearing assessment, and more specifically, CI fitting: objective and subjective (i.e. psychophysical). What follows is an overview of commonly used measurements in the field of cochlear implants.

3.1.1. OBJECTIVE MEASUREMENTS

Objective measurements, as opposed to psychophysical experiments, do not require participation of the subject and can therefore typically be conducted in subjects of all ages and in most instances also intraoperatively, when under general anaesthesia. These measures often assess physiological phenomena, but not necessarily (e.g. measuring CI electrode impedances). Debate on the usefulness of objective measurements for the purpose of CI fitting is ongoing, as they are put in competition with behavioral measurements (e.g., psychophysical loudness assessment) to find optimal stimulation levels. Nonetheless, clinicians have often relied on objective measures as a guide to determining appropriate stimulation levels for young children and recipients with multiple disabilities, in whom it is difficult to obtain reliable behavioral responses. Even adults who have been deaf for a long period of time could be considered unreliable reporters, since they may not be able to realistically scale the magnitude of a stimulus due to their inexperienced or eroded auditory perception.

3.1.1.1. ELECTRODE IMPEDANCE TELEMETRY

Electrode impedance measurements are performed intraoperatively, as well as routinely during fitting sessions, to verify whether all electrodes are functioning properly. This implant test is performed very quickly using the fitting software of the CI system (Figure 62). Impedances within the normal range suggest that current flow occurs in the tissue and fluids of the cochlea (common ground impedances) and between intra- and extracochlear electrodes (monopolar impedances). Low impedances (when measured using common ground coupling), indicate particular electrodes being shorted together, whereas high impedances might be due to a broken electrode wire or an electrode only in contact with air. An increase in impedances over time may also indicate inflammation or scar formation of tissues within or around the inner ear.

Electrode impedances should also be considered by the CI system when supplying the necessary voltage to electrodes in order to reach the desired current level. The current sources of the internal stimulator are limited in the amount of voltage they can generate (a technological limitation, also dependent on battery capacity). That maximal voltage is called the compliance voltage. If an electrode has a high impedance such that the voltage required to obtain the desired current level cannot be delivered, it is said to be no longer "within compliance". In such a case, the only way of increasing the amount of charge per pulse phase is to increase the pulse width used by the map, which typically affects all channels (and may adversely affect the rate of stimulation). It should be noted that impedances usually fluctuate over time, mainly between the time of surgery, when they are typically low, the moment of switch-on, when they are typically increased, and the period thereafter, during which they usually decrease again as a result of frequent electrical stimulation.

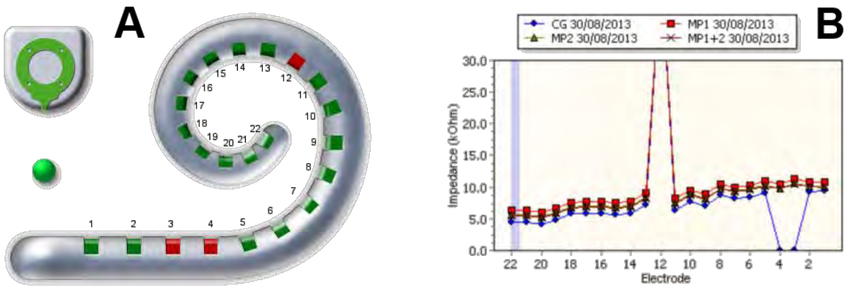


Figure 62: A: Electrode impedance telemetry performed in Cochlear's Custom Sound fitting software. By default the systems measures impedances in 4 electrode coupling modes: Monopolar 1 (MP1), Monopolar 2 (MP2), Monopolar 1+2 (MP1+2) and Common Ground (CG). B: In this case, the impedances for electrode 12 were found to be exceedingly high using any coupling, indicating an open circuit (the electrode wire may be broken). In Common Ground coupling the impedance measured between electrodes 3 and 4 was found to be near zero, indicating on short circuit (electrodes 3 and 4 may be shorted together somewhere in the array).

Other non-physiological objective measures, that are less commonly used, include the measurement of Electrical Field Potentials (EFP) produced by the electrodes and Averaged Electrode Voltages (AEV).

3.1.1.2. ELECTRICALLY EVOKED AUDITORY BRAINSTEM RESPONSE (EABR)

Auditory brainstem responses (ABR) can be evoked by stimulating the auditory nerve with a recipient's CI system [115]. The same can be achieved by using a transtympanic needle electrode to stimulate the promontory of the cochlea. EABR responses are captured using surface-potential electrodes just like for acoustically evoked ABR (Figure 63 A), while the stimulus is delivered by the CI programming software.

The potentials recorded represent ongoing electrical activity in the auditory pathway during the first 10 milliseconds after the start of stimulation. The amount of activity is wave-like, as the potentials travel through the different structures of the auditory system. The peaks in the waveform, as shown in Figure 63 B, are suggested to represent the following neural structures [116]:

- Wave I – peripheral portion of cranial nerve VIII (leaving the cochlea);
- Wave II – central portion of cranial nerve VIII (entering the brain stem);
- Wave III – cochlear nucleus (caudal portion of auditory pons);
- Wave IV – superior olivary complex/lateral lemniscus;
- Wave V – lateral lemniscus/inferior colliculus;
- Waves VI and VII – thalamic (medial geniculate body) origin is suggested, but the actual site of generation is uncertain.

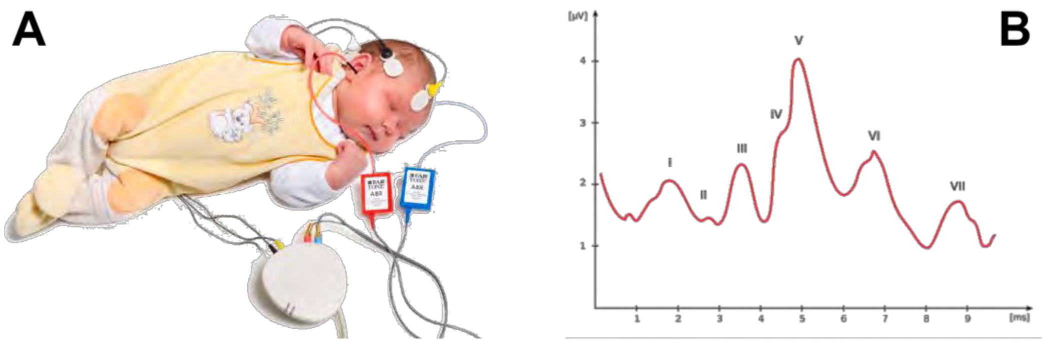


Figure 63: A: ABR measurement conducted in a newborn using acoustic stimuli and surface electrodes. B: schematic of evoked auditory brainstem responses in a normal hearing person. The different peaks in the waveform are assumed to represent different structures along the auditory pathways. [A adapted from Isham M, B from Sutashu]

The results (threshold for responses and peak latencies) of an (E)ABR measurement may be useful for predicting spiral ganglion cell survival, pre-operative ear selection, prediction of post-operative benefit [117], device integrity (e.g., electrode output) testing and for guiding the CI fitting in difficult cases [118]. However, the ability of EABR to predict post CI speech perception scores was found to be very limited. Also the correlation between EABR thresholds and behavioral EDR Minima was found to be poor [119]. EABR measurements are disrupted by movement, which makes them difficult to obtain in young children, and the procedure to measure responses for each electrode takes a considerable amount of time. These disadvantages make EABR of limited use to the routine fitting, but rather to be considered for special cases only.

3.1.1.3. ELECTRICALLY EVOKED STAPEDIAL REFLEX (ESR)

ESRs are measured visually during surgery or by monitoring the impedance of the middle ear, similar to acoustically derived stapedial reflexes. When neural reflex pathways are intact, the stapedius muscle contracts in response to high stimulus levels, also when administered electrically (through the CI). When this contraction is observed, the operator would typically decrease the stimulus level, gradually until the reflex can no longer be recognized. That level is referred to as the ESR threshold (ESRT) and is being used by some clinicians to set EDR Maxima during the fitting. A limitation to the use of ESRTs for fitting is that they are not always present or observable, even in the normal hearing population, or they may have been lost during CI surgery. In the latter case, it may still be possible to record the ESRT from the non-implanted ear, since the reflex presents itself bilaterally. Measuring ESRTs is typically a straightforward procedure, taking about 1 minute per electrode.

ESRs have been shown to correlate strongly with EDR Maxima measured behaviorally in experienced adult implant users [120], of which the majority had ESRT levels below the levels of uncomfortable loudness as determined behaviorally. Studies have found that speech recognition and listening comfort were not decreased when using ESRTs to set EDR Maxima [121] [120]

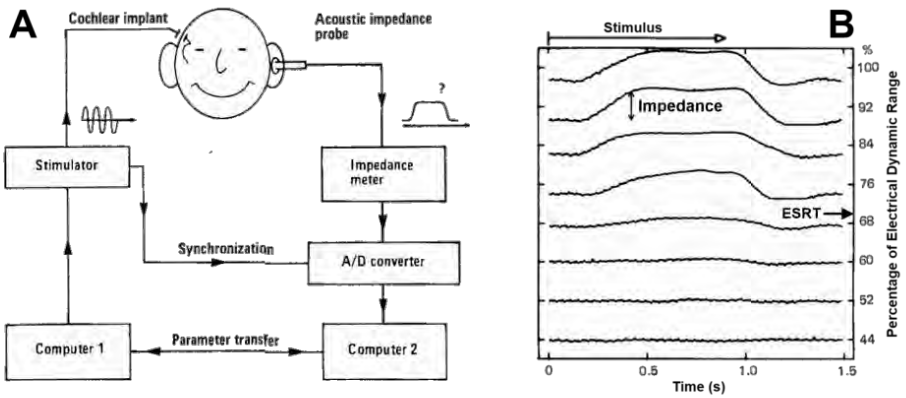


Figure 64: A: setup for automated ESRT measurement. The stimulus is delivered through the fitting software to the implant. Meanwhile the impedance at the contralateral ear is measured using a middle ear analyzer. Sampled values are analyzed for responses and stimulation parameters are adjusted accordingly in the fitting software. B: typical pattern of contralateral stapedius reflexes elicited by CI stimulation at different levels (each line represents to response to one stimulation level (expressed as a percentage of the EDR set in the fitting software). A reduced compliance (an increase in middle ear impedance) in response to the electrical stimulus indicates contraction of the stapedius muscle. The stimulation level where the reflex is no longer observable is the ESR threshold (ESRT).

3.1.1.4. ELECTRICALLY EVOKED COMPOUND ACTION POTENTIAL (ECAP)

The compound action potential generated by spiral ganglion fibers in response to stimulation of a specific electrode can be recorded through a neighboring electrode [122]. Cochlear company was the first (in 1998) to release this technology as Neural Response Telemetry (NRT) with their Nucleus 24 implant, but other manufacturers followed: Advanced Bionics with their Neural Response Imaging (NRI) and MED-EL with Auditory Nerve Response Telemetry (ART). ECAP recordings provide significant advantage over EABR in terms of measurement time. There is much less need for averaging responses because the ECAP is not disturbed by movement and electrical artifacts are controlled better. No additional surface-electrodes need to be positioned and the measurement is run automatically from within the fitting software (Figure 65 A).

Unfortunately ECAPs cannot be obtained in every recipient (the prevalence is similar to that of ESRT, around 70%, [123]). Also, the relation between ECAP and behavioral thresholds is dependent on the individual [124] and ECAP thresholds have been found to change over time [125] which limits their use for setting stimulation levels. When ECAP measurements are combined with behavioral measurements on a few selected electrodes however, it has been shown that they can be useful in estimating appropriate stimulation levels [124] [126] [127] [128] [129].

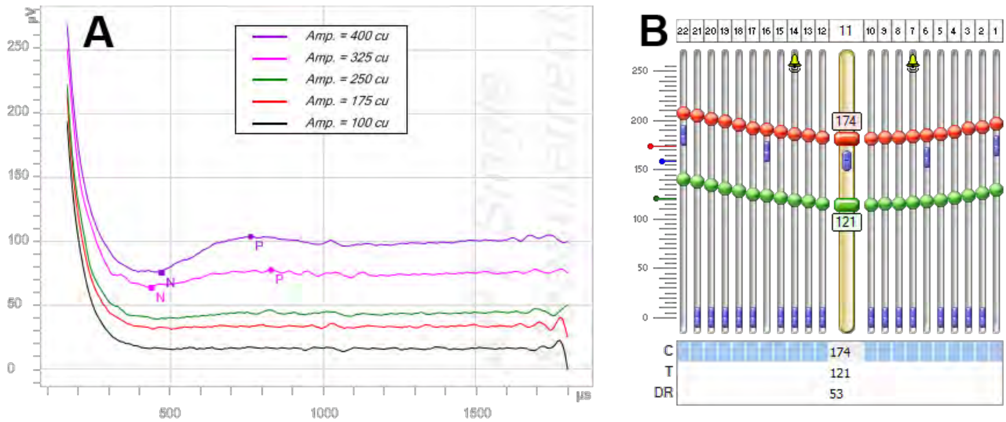


Figure 65: A: screenshot of ART (MED-EL Maestro System Software v4.1.1) ECAP measurements at different stimulation levels. N and P symbols indicate the detection of a compound action potential. B: screenshot of Cochlear's NRT/Objective Offset programming method (Custom Sound v4.0 software). This method involves measuring NRT ECAP thresholds at 5 (number is configurable) electrodes (indicated by blue symbols). Profiles for both EDR Minima (green symbols) and Maxima (red symbols) are created by interpolating the ECAP thresholds. By behaviorally measuring a threshold and a maximal comfort level at a single electrode (nr 11 in this case, illustrated by the enlarged channel component width) those profiles are shifted globally.

The ECAP based fitting approaches, which are nowadays readily available in the programming software provided by the manufacturers (Figure 65 B), may be a means to save time when compared to measuring all levels behaviorally and may provide a usable level estimation in cases where psychophysical loudness assessment is unreliable. But for reliable responders, behavioural measurements most often yield more optimal stimulation levels [19].

3.1.2. SUBJECTIVE MEASUREMENTS

Subjective measurements require active participation of the subject. A task is explained to the subject, who then needs to interpret and execute it. The goal is to assess the sensations and perceptions of the subject. Those sensations may be expressed and recorded qualitatively, for example when asked to describe sound clarity, to fill out a questionnaire on the level of satisfaction experienced by cochlear implant usage, or to report on the comfort of auditory sensation. In this dissertation, the responses to such qualitative enquiries will be referred to as subjective feedback. The global survey [83] shows that subjective feedback is often used for programming a CI and that a lot of attention is given to whether auditory sensation is comfortable in the opinion of the recipient.

As opposed to eliciting subjective feedback, psychophysical experiments investigate the relationship between physical stimuli and the sensations and perceptions they affect in a quantitative manner. In case sound is used for stimulation, those experiments are also referred to as psychoacoustical tests. A listener's response to stimuli that are presented at different levels is typically probabilistic. It can be described by a psychometric function (e.g., cumulative Gaussian or logistic functions) showing that the probability of positive responses increases from 0% (or chance level, depending on the test task) to 100% with increasing stimulus intensity (Figure 66). The perceptive threshold as defined by the presentation level that yields a positive response in 50% of the presentations is also referred to as equilibrium point or Just Noticeable Difference (JND).

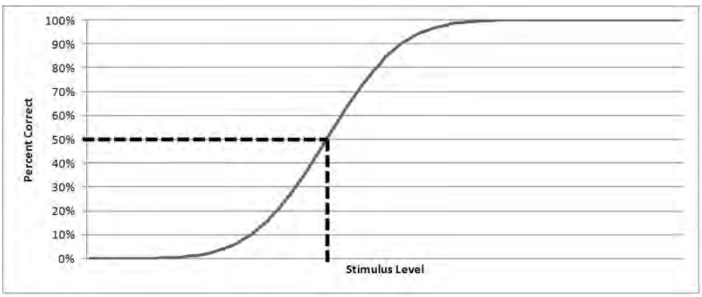


Figure 66: A typical psychometric function showing the probability of a correct response in function of the presentation level. The equilibrium point is defined as the point along the curve where 50% of the subject's answers are correct. The stimulus level at this point is the subject's threshold or JND.

Typically the test protocol consists of changing the properties of a stimulus (such as intensity, spectral content or temporal features) in a systematic way and studying the subject's behaviour in response to those changes (which explains the term "behavioural", which is also often used for this type of testing). There is a gray zone between subjective feedback and psychoacoustical testing as one may consider the interpretation of responses more or less quantitative. For example when asking a subject to describe the perceived loudness of a particular stimulus, the response may be interpreted qualitatively or recorded quantitatively (e.g., using a visual-analog scale).

Figure 67: a typical setup for conducting psychoacoustical measurements, consisting of a calibrated audiometer to generate stimuli and/or control their presentation level, calibrated transducers (e.g. headphones or loudspeakers), a sound treated room to exclude interfering ambient noises and a computer operated by the examiner to execute the test protocol. [adapted from Banús CM]

To maximize the reliability of results obtained from psychoacoustical experiments their test protocols should be well defined and the delivery of the stimulus well controlled (i.e. using calibrated equipment and a sound treated room, Figure 67). Also important is to be concerned with biased interpretation of responses by the experimenter. Test protocol, stimulus delivery and experimenter bias are three factors that need to be controlled to minimize the test-retest variability and ensure reproducibility. It is not uncommon to disregard these aspects to some extent when eliciting ad-hoc subjective feedback from a subject during a fitting session. Clinicians often present live speech or selected phonemes (e.g. the Ling 6 sound test, [130]) in an office room to evaluate their fitting. It is clear that the conditions in which this happens vary significantly, over time (e.g., variations in the audiologist's voice), by place (e.g., differences in ambient noise level and spectrum) and by audiologist (e.g., accent and pronunciation differences).

3.1.3. MEASURING AUDITORY COMFORT

The behavioural measurements most frequently used in the context of CI fitting use narrowband stimuli generated by the fitting software. Common practice amongst clinicians to find the optimal stimulation range (EDR) for each electrode in the array consists of presenting tone bursts (pulse trains) on a single electrode or on a small group of electrodes. While this may produce insight into the range of levels between detection and discomfort for isolated electrodes, it is questionable whether these measurements assess anything that is relevant for everyday cochlear implant use.

The world of sounds around us does not contain signals that cause stimulation of a single electrode, instead it is filled with complex sounds and broad spectra stimulating many (if not all) electrodes at the same time. Considering the summation effects when multiple electrodes are stimulated together (i.e., during the same stimulation cycle, hence also applicable to sequential stimulation strategies), the perceived loudness is likely to be greater than what would be reported when stimulating a single electrode at a comparable level. This may lead to the situation where broadband sounds cause discomfort while the EDR Maxima for each separate electrode are well within the comfortable range. The same holds for EDR Minima, which may be set too high by measuring responses to single electrode activation. Due to summation, broadband sounds could possibly be detected with lower stimulation levels. Setting EDR Minima higher than necessary means sacrificing part of the already pocket-sized EDR available for electrical stimulation. Summation may well be one of the reasons why many clinicians do not keep the levels as measured per electrode when fitting initial maps. Instead, they switch to "live" mode (microphone input) and readjust Maxima (and to lesser extent Minima) globally, by shifting these profiles up or down based on feedback on comfort (and/or audibility) given by the recipient. The stimulation levels found behaviourally per electrode (which is a time consuming procedure) are then repositioned, assuming that the shape of the profile still reflects the variation in comfort/audibility along the electrode array. However, loudness does not necessarily grow equally fast at each electrode, meaning that the precisely measured profile shape loses its intended benefit to some degree.

Another, even more critical, concern with regard to measuring stimulation levels based on comfort is that most CI recipients have not retained, or did never acquire, a frame of reference for estimating comfortable loudness that can be put to use efficiently at initial fitting sessions. The sensations resulting from electrical stimulation of an auditory system that has never been stimulated before, or has been deprived from adequate stimulation for a long time, cannot be expected to be reported by the recipient in a realistic (i.e., similar to sensations in normal hearing people) manner when questioned on comfort. The auditory system needs to adapt to this new kind of stimulation and the recipient needs to (re)build a frame of reference for loudness and comfort. Considerable amounts of time and effort are spent at finding optimal stimulation levels for each electrode during the early stages after switch-on. Based on the reasoning above, it can be questioned whether this is a very efficient approach. That question has been, in the context of the research reported here, one of the main driving forces behind the exploration of fitting paradigms other than those based on behavioural comfort measures per electrode.

3.2. FITTING FOR PERFORMANCE

The approach to CI fitting that has been developed during the course of the present project could be labelled as "adapt first, tune later". It proposes a procedure for CI fitting comprising 2 stages:

(1) the adaptation to loudness using a series of statistically inferred start-up maps of increasing stimulation levels during the first few weeks after switch-on and
(2) the iterative tuning of the recipient's map towards maximal auditory performance as measured by psychoacoustical experiments.

The first stage allows the recipient to (re)gain some form of reference frame for loudness and other auditory sensations resulting from CI stimulation. Only after such is acquired, can maps be efficiently tuned for performance. This process of tuning a map involves:

(1) defining targets of performance,
(2) measuring a recipient's distance from those targets and
(3) using that information to make meaningful adjustments that are expected to bring the map (hence the recipient) closer to target.

As explained earlier, to serve as a usable target in clinical practice, a set of outcome measures needs to produce reliable results in a reasonable amount of time. These results should reflect a recipient's auditory performance in daily life and at the same time be specific enough to unveil auditory deficits on which actions may be taken (i.e., by fitting) to reduce or resolve them. Following that reasoning, auditory comfort cannot serve as a primary target for fitting. Instead, in the fitting paradigm developed here, comfort is taken for granted. It is not considered a performance metric, but rather as something that is acquired, unless there are clear signs of discomfort.

3.2.1. MEASURING AUDITORY PERFORMANCE

In psychophysics one can distinguish several levels of perception: detection, discrimination and identification. When applied to hearing, these levels translate to a hierarchy of auditory skills. In psychoacoustical tests, these levels represent different types of tasks the subject may be instructed to perform.

Figure 68: Psychophysical levels of perception: detection (perceiving the presence of a signal), discrimination (perceiving a difference between 2 distinct signals) and identification (being able to attribute a meaningful label to a signal).

Detection is the most basic level of sound perception [131]. It is the awareness of the presence or absence of sound, which in psychoacoustical experiments is usually assessed by asking the subject to respond when they hear a sound. Most detection tasks seek for the smallest stimulus level at which the sound is still detected some proportion p of the time (50% is usually used for p) by the listener (i.e. the subject's detection threshold). Detection of isolated sounds could in theory be accomplished by a single channel cochlear implant limited to binary output only. Obviously detection is an essential process, prerequisite to higher auditory skills.

Discrimination is the ability to tell if two sounds are the same or different. To discriminate between two sounds, the listener first needs to be able to detect the two sounds. Therefore, discrimination is a higher level task than is detection. In some discrimination tasks, the goal is to find the smallest difference between two sounds, which the listener can perceive. Such a task results in a difference threshold (also called difference limen or just noticeable difference (JND)). There is some room for interpretation when classifying psychophysical tasks. In detection tasks it is assumed that the signal presented is not contrasted with another signal, but obviously there is always some sort of background noise. In that sense a detection task could be considered a discrimination task in which the goal is to discriminate the signal from the noise. Also the other way around, discrimination tasks in which either of the presented signals could be considered as a standard or reference signal (e.g. in a same-different discrimination task for intonation where one of both stimuli presented does not feature intonation at all) may be classified as detection tasks.

Identification is the ability to attribute a sound with a meaningful label. This level of perception requires that the listener first be able to detect and discriminate the stimuli before they can be uniquely identified (i.e., selected from a library of competitive sounds available in the (auditory) memory). Identification therefore is a higher level task dependent on detection and discrimination. But it also depends on cognitive skills, while cognition is not directly affected by the cochlear function. So psychoacoustical identification tasks only offer an indirect view on the functioning of the peripheral auditory system. Good results on an identification task are an indication for proper cochlear functioning, but inadequate results may be the result of deficits in the cochlear function and/or central processing. Despite their less direct relation to the cochlear function, identification tasks are often used in clinical practice because they are in general more appealing, less abstract of a task and therefore more easily conducted in everyday patients.

One could argue that enabling the discrimination of sounds is the core task of the cochlea. A CI system therefore needs to present the neural interface with a stimulation pattern that maximizes a perceiver's ability to discriminate between the variety of sounds that are present in the world around us. The properties of the input stimulus (and thus of the resulting stimulation pattern) that allow a listener to extract the necessary information to distinguish between different sounds are called an auditory cue. In that sense it can be said that the task of a CI is to preserve as much of the cue when converting an acoustical input signal to an electrical stimulation pattern. Through psychoacoustical tests one can investigate this preservation of cues. Stimuli used in psychophysical experiments are often designed in such a way that the cue needed to perform the task results from a specific property of the stimulus (e.g., intensity, frequency spectrum, temporal envelope and fine structure). This way one can assess deficits in specific coding mechanisms (e.g. loudness coding, tonotopy and phase locking) related to the processing of those properties. Figure 69 shows a schematic representation of how these cues and coding mechanisms can be tested at the different levels of perception.

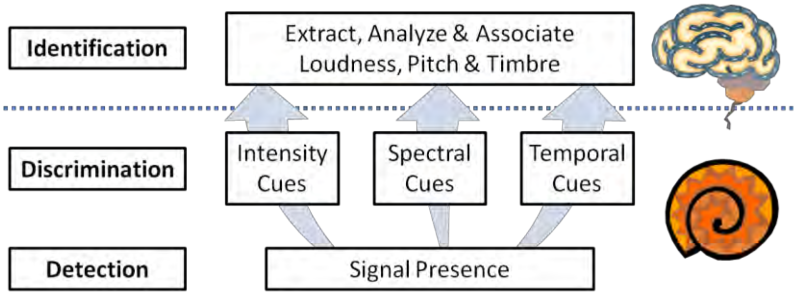


Figure 69: Psychoacoustical tests can be designed to assess auditory abilities at the different levels of perception. Detection tasks seek the threshold for awareness of signal presence. Discrimination tasks can be used to assess the processing/preservation of cues related to specific signal properties (intensity, spectral and temporal). Detection and discrimination are core responsibilities of the peripheral auditory system (cochlea). Identification tasks require higher level processing (central auditory/nervous system) to extract, analyze and associate patterns of loudness, pitch and timbre to meaningful concepts.

Following this reasoning, the Eargroup has designed, throughout the years, a battery of psychoacoustical tests aimed at maximizing the analytical interpretation of auditory deficits related to specific cochlear coding mechanisms. To address the concerns related to reliability of psychophysical experiments (as mentioned earlier) and to maximize time/resource efficiency, those tests have been implemented in a software package: the A§E Psychoacoustic Test Suite [107]. Figure 70 gives an overview of the test modules that are now routinely used in the Eargroup's clinical practice. Pure tone Audiometry covers the detection layer. Phoneme Detection is sometimes performed to verify audibility of phonemes prior to presenting them in a discrimination task. Phoneme Discrimination assesses spectral processing at the discrimination level. The Disharmonic Intonation test assesses temporal fine structure coding (through phase locking) while the Harmonic Intonation test also contains cues resulting from place coding (tonotopy). These tests are described in more detail in "Clinical assessment of pitch perception". The Loudness Scaling Test assesses loudness coding at the identification level. The Phoneme Identification test is a closed set picture pointing task using onomatopoeia or mouth images to assess spectral and temporal fine structure coding in young children. Speech Audiometry is a high level identification task that requires adequate processing of a variety of spectral and temporal (envelope and fine structure) cues.

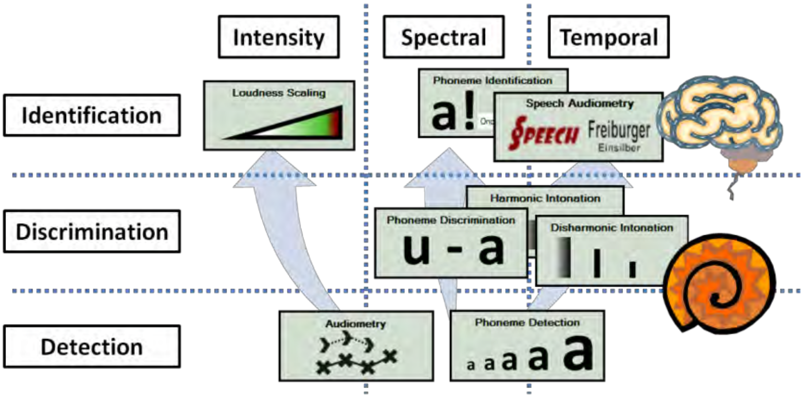


Figure 70: The modules of the A§E Test Suite that are routinely performed in the Eargroup's clinical practice. The detection layer is covered by pure tone Audiometry, and Phoneme Detection is sometimes performed to verify audibility of phonemes prior to presenting them in a discrimination task. At the discrimination level spectral processing is assessed through the Phoneme Discrimination test. The Disharmonic Intonation test assesses temporal fine structure coding (through phase locking) while the Harmonic Intonation test also contains cues resulting from place coding (tonotopy). Loudness coding is assessed at the identification level. The Phoneme Identification test is a closed set picture pointing task using onomatopoeia or mouth images to assess spectral and temporal fine structure coding in young children. Speech Audiometry is a high level identification task that requires adequate processing of a variety of spectral and temporal (envelope and fine structure) cues.

For the fitting model that has been developed during this project, 4 of these psychoacoustical tests are of particular importance: Audiometry (sound field detection thresholds), Speech Audiometry (speech recognition scores), Phoneme Discrimination (distinguishing between trivial speech sounds) and Loudness Scaling (the growth of loudness sensation). A brief description of those tests follows in the next sections.

3.2.2. AUDIOMETRY

Pure tone audiometry is used in clinical settings to identify the hearing thresholds of a listener at different frequencies. It is a detection task in which pure tones are presented through headphones, insert phones, bone conduction or loudspeakers. A popular method for conducting audiometry is the Hughson-Westlake procedure or one of its modifications [132]. It uses a descending familiarization trial that starts at a level presumed to be well above threshold and decreases intensity in steps of 10 dB. Afterwards, a threshold is sought using ascending trials, increasing stimulus level by 5 dB steps. Usually the threshold is defined as the lowest intensity at which positive responses were obtained in 50% of the trials. The result of the test is depicted on an audiogram (Figure 71). When audiometry is performed on an aided ear (e.g., using a hearing aid or CI) the stimulus (warble tones are used to avoid standing waves) is most often delivered through loudspeakers, as headphones are not suited for presenting signals to the microphone of a sound processor worn behind the ear. The resulting detection thresholds in such cases are often referred to as aided sound field thresholds.

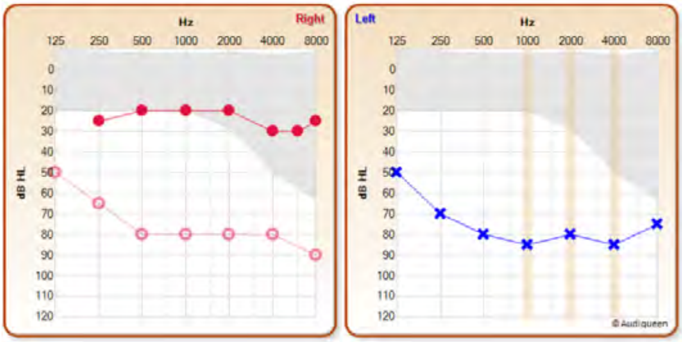


Figure 71: Audiograms for right and left ear, showing an individual's pure tone detection thresholds (in dB HL, vertical axis) at different frequencies (in Hz, horizontal axis) prior to cochlear implantation (circles and crosses) and after cochlear implantation at the right ear (dots).

The median test time (unpublished data from 17884 time tracked sound field threshold measurements performed with A§E at the Eargroup) is 41 seconds per threshold. The median number of thresholds measured in those audiograms is 6 (typically at 250, 500, 1000, 2000, 4000 and 8000 Hz) which means about 250 seconds in total to obtain the audiogram for an ear.

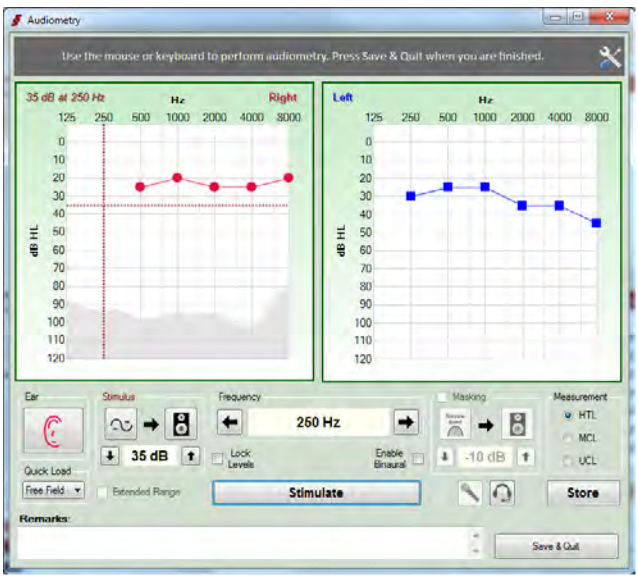


Figure 72: Screenshot of the A§E Audiometry Module. The software controls a digital audiometer to perform typical pure tone audiometry.

3.2.3. SPEECH AUDIOMETRY

In clinical settings formal assessment of speech perception skills is performed through speech audiometry. The test involves the acoustical presentation of speech stimuli at certain presentation levels. The subject is instructed to repeat what he/she has heard and an audiologist judges whether the response was correct. The result of such a test is depicted on a speech audiogram (Figure 73). Stimulus materials used in clinics typically consist of predefined lists of isolated words or complete sentences. Sentences have the advantage that they relate more to everyday speech, which makes them suited for the evaluation of an individual's auditory performance in daily life. But sentences have the disadvantage that they contain more redundant information. Because of this higher redundancy, the result is more dependent on language and cognitive skills. Auditory deficits that cause parts of a sentence to be misunderstood can be compensated by understanding of the contextual meaning of the sentence. The redundancy in short words is lower, which makes them more suited for diagnosing auditory deficits. Whenever available, monosyllables with a consonant-vowel-consonant (CVC) structure are often preferred. Some speech audiometric tests use nonsensical stimuli to reduce the dependency on language skills.

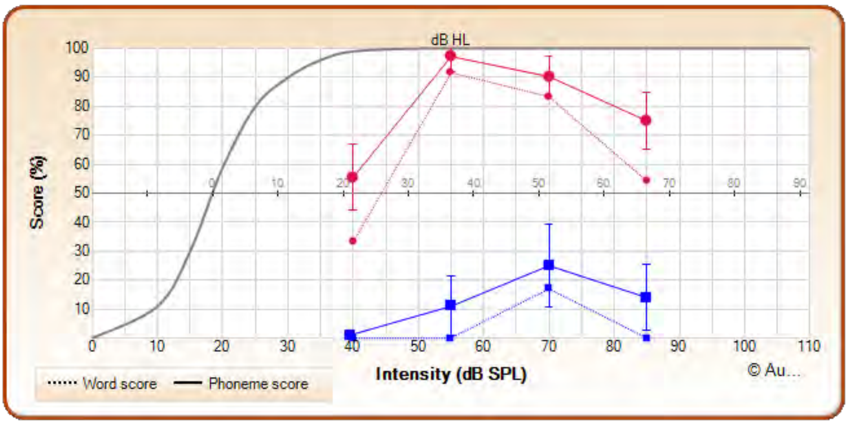


Figure 73: Speech audiograms for an individual's right (dots) and left ear (squares). The right ear was aided with a cochlear implant, the left ear with a conventional hearing aid. Lists of monosyllabic words were presented at different intensities (40, 55, 70 and 85 dB SPL, horizontal axis) and both word (small symbols) and phoneme (large symbols) scores (percentage of correctly repeated words/phonemes) were recorded by an audiologist. The gray line depicts the normal hearing population's phoneme scores. Error bars show a confidence interval for the recorded score.

In current clinical practice 2 types of speech material are being used: recorded (typically a CD) and live voice (the examiner's voice picked up by a microphone). Recorded material is in general considered to be superior to live voice since it is less subject to variation in terms of presentation level and pronunciation details, and it excludes lip-reading by definition.

The scoring of responses may either be performed at the sentence, word or phoneme level. The more fine-grained the scoring, the less subject it is to variability (test-retest), because of the larger amount of scoreable items. Test protocols may either present a fixed amount of stimuli at each presentation level, or adaptively adjust the presentation level in function of the percentage of correctly repeated speech tokens. The result of a speech audiometric test is often summarized through its Speech Reception Threshold (SRT), which is the presentation level at which the listener correctly repeats 50 % of the presented items.

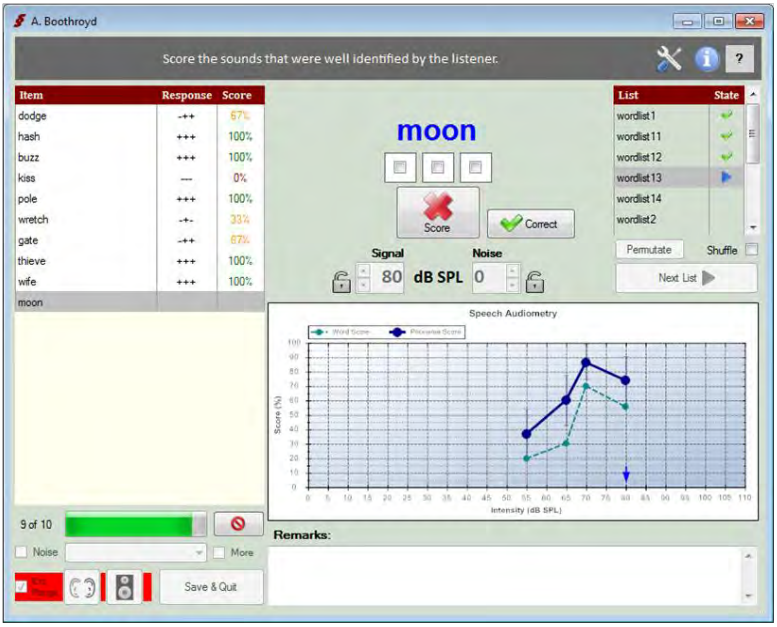


Figure 74: Screenshot of the A§E Speech Audiometry module. Lists of the Arthur Boothroyd [133] monosyllabic CVC words are presented at different intensities and responses are scored on the phoneme level by an audiologist.

The median test time (unpublished data from 291742 time tracked speech audiometry trials performed at the Eargroup) is 1.3 seconds per speech token. The median number of speech tokens presented in a speech audiometry test run is 48, typically as lists of 12 words at 40, 55, 70 and 85 dB SPL. A typical speech audiogram therefore is obtained in about 75 seconds.

3.2.4. PHONEME DISCRIMINATION

The Phoneme Discrimination test as implemented in the A§E Test Suite [107] is a discrimination task in which phoneme contrasts (typically 20, as illustrated in Figure 76) are presented in an oddity paradigm. The presentation level is set well supraliminally (by default at a level that in normal hearing subjects is equally loud as a 70 dB SPL 1 kHz narrow band noise of a third of an octave in bandwidth). The phonemes are constructed in such a way that the only difference between different phonemes is the spectral content (Figure 75). Both duration and loudness are equal. During the test procedure the background phoneme is repeated with fixed intervals of typically 850 milliseconds. Every now and then the background phoneme is replaced with the stimulus phoneme, which is also called the odd phoneme. The listener should react to the odd phoneme. He or she should not recognize the phonemes but just notice that they are different from one another.

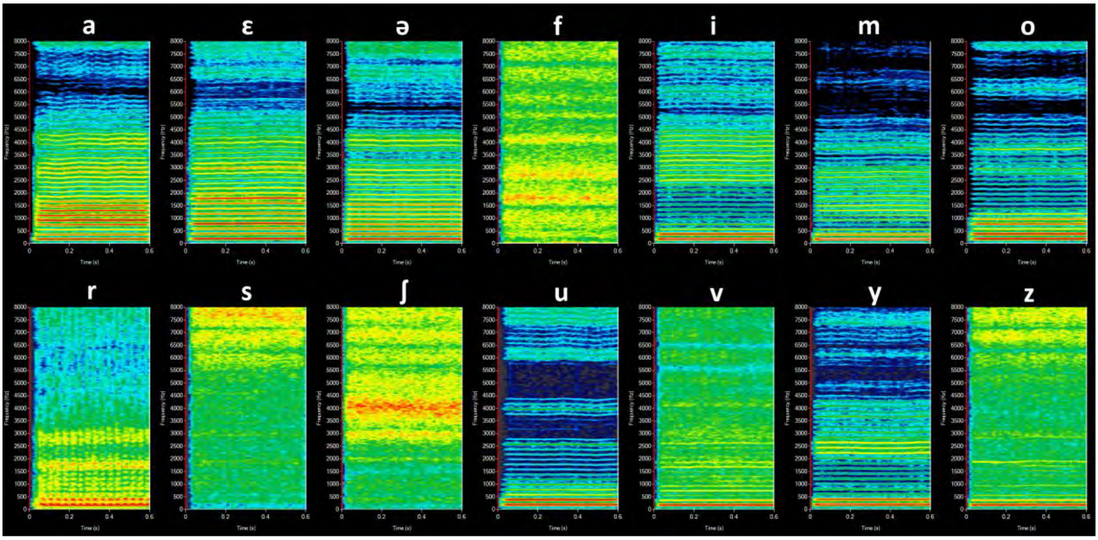


Figure 75: Spectrograms of the 14 phonemes contained in the A§E Phoneme Discrimination test.

This test assesses what is presumably one of the most essential features of the cochlea, namely its frequency resolving power. Therefore, it is particularly useful in the evaluation of the deficient cochlea and of cochlear therapies. Its usefulness has been proven in the selection of cochlear implant candidacy, the evaluation of cochlear implants, and the outcome-driven fitting of cochlear implants. Since it is essentially a discrimination task, this test requires little or no cognitive cooperation. It can be done in infants, toddlers, children and adults. In infants and toddlers, special techniques may be required to elicit the responses. These are the same techniques that are used in paediatric audiometry, like conditioned orientation reflexes, visual reinforcement, play audiometry, etc.

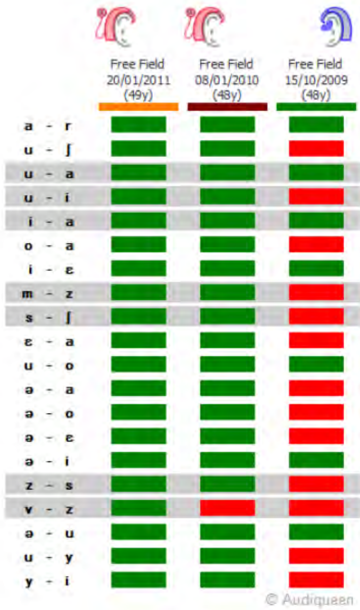


Figure 76: Results of Phoneme Discrimination obtained in an individual on three different dates. The left column shows the phoneme contrast presented. Three columns of coloured rectangles show whether the phonemes in the contrast could be discriminated (green) from each other or not (red). The oldest result (right column) was measured from the subject's left ear aided with a conventional hearing aid. The other 2 results were obtained in the right ear, aided with a cochlear implant. Greyed contrasts represent the shortlist of 7 contrasts, for use in subjects with a limited attention span.

Hearing listeners should have no difficulty in discriminating all the contrasts. In case of sensorineural hearing loss, the aided cochlea may still experience great difficulty discriminating sounds that are played well above threshold. If too many contrasts can no longer be discriminated, this may add to the indication for cochlear implantation. After cochlear implantation the discrimination should be restored to a great extent. This is usually the case immediately after switch-on.

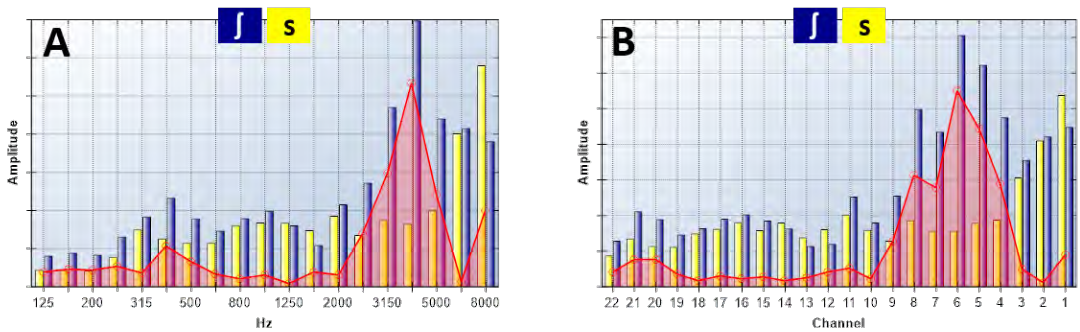


Figure 77: Spectral analysis of the A§E Phoneme discrimination, showing frequency spectra of phonemes / ʃ / in blue and / s / in yellow. Spectra are plotted either on a frequency scale (e.g., in ISO266 third octave bands, A) or in function of a CI filter bank (e.g. Cochlear's Nucleus CI filter cut-offs, B). The difference between the spectra is plotted in red, pointing to the frequency bands or channels in which both phonemes differ most (with regard to energy).

If the cochlear implant wearer fails to discriminate two phonemes, A§E provides the spectral analysis of both phonemes and illustrates in which frequency bands their energies differ the most (Figure 77). This difference can be shown as a function of the audiometric frequencies but also as a function of the electrode numbers of the CI array. This analysis pinpoints the electrodes that code for the spectral cue that is not perceived by the CI-wearer. This information may help the CI programmer to change the settings of the device in order to provide a better spectral discrimination.

Figure 78: Screenshot of the A§E Phoneme Discrimination module. Phoneme contrasts are presented in an oddity paradigm and detection of the odd phoneme is scored by an audiologists.

The median test time (unpublished data from 19340 time tracked phoneme contrast discrimination measurements performed with A§E at the Eargroup) is 35 seconds per contrast. The median number of contrasts measured in the phoneme discrimination test runs of that data set is 11. The time typically needed to obtain results for the full set of 20 contrasts is 700 seconds.

3.2.5. LOUDNESS SCALING

The A§E Loudness Scaling test consists of a typical loudness growth assessment. It is an identification task to assess the intensity coding of sound. Narrow band noises of 250, 1000, or 4000 Hz are presented at different intensities, going from 10 to 100 dB HL (depending on the stimulus frequency). The listener should score the perceived loudness on a scale ranging from 0 to 6, corresponding to inaudible, very soft, soft, normal, loud, very loud and too loud. The individual listener does not receive the entire range from 10 to 100 dB HL, but only the intensities between the lower and the upper fences, which are set during the training session. The result of this test is referenced to the median values in hearing listeners, as well as the 95 percent confidence interval between percentile 2.5 and percentile 97.5. This represents the normal zone.

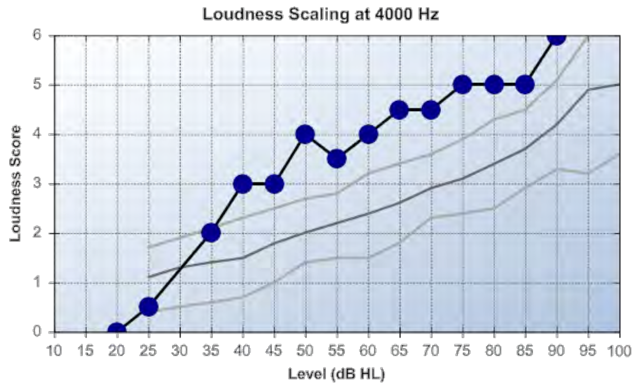


Figure 79: An individual's result on the A§E Loudness Scaling test using a narrow band noise centred at 4000 Hz. The graph shows the median responses for the presented intensities plotted on top of the normal zone (defined as the median values in hearing listeners (dark gray line), as well as the 95 percent confidence interval (demarcated by the light gray lines) between percentile 2.5 and percentile 97.5)

The Loudness Scaling test is particularly useful in patients with hearing aids or cochlear implants. Since it is essentially a low level identification task, this test requires some adaptation to the device. Therefore, this test should not be used at switch-on. It is rather recommended to do this test after a couple of weeks or months of adaptation to the device. If test results fall beyond the limits of the normal zone, this indicates that sounds of these intensities are not well coded by the device. For instance, if scores at a moderate intensities fall above the normal zone, sounds within that range of intensities are perceived too loud. One may want to change the program of the device to cope with this. In cochlear implants, this may typically influence the settings of the EDR Maxima. Likewise, scores at low intensities that fall below the normal zone are perceived too soft. In cochlear implants this may typically influence the settings related to the EDR Minima or microphone sensitivity.

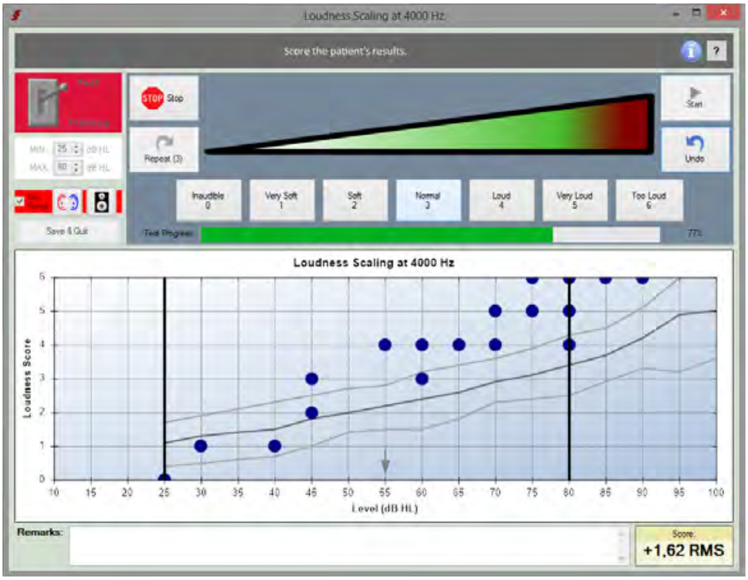


Figure 80: Screenshot of the A§E Loudness Scaling module. A 4000 Hz narrow band noise is presented at different intensities and their subjective loudness perception is scored on a scale ranging from Inaudible (0) to Too Loud (6).

The median test time (unpublished data from 52538 time tracked loudness assessment trials performed at the Eargroup) is 4.1 seconds per trial. The median number of trials per test run in that data set is 24. So a loudness scaling result for one frequency is typically obtained in about 100s.

3.3. IMPROVEMENTS IN MEASURING HEARING PERFORMANCE

Running the full set of outcome measurements (i.e., Audiometry at 6 frequencies, Phoneme Discrimination of 20 contrasts, Loudness Scaling at 3 frequencies and Speech Audiometry at 4 intensities) typically takes 22 minutes. Given that additional time is needed to setup the test environment, install and instruct the subject and get him/her acquainted with the task at hand, the total amount of time required easily extends to twice the pure test time. During this time, both qualified personnel (audiologists) and infrastructure (sound treated rooms, audiometers) are kept occupied. In a performance based approach to CI fitting it is essential that the state of the auditory system is measured repeatedly. This iterative testing quickly becomes a significant burden on the resources of a clinical CI team. To address these concerns, considerable amount of attention has been given to the optimization of the amount of resources required to conduct these outcome measures, both in terms of time and of infrastructure. The A§E Test Suite has therefore been entirely refurbished during the course and as a part of this PhD research, while keeping the reliability of psychophysical experiments as the primary requirement. The A§E software has also been extended with additional tests to explore the effects of deficits in temporal and fine spectral coding.

A first optimization consists of extensive integration with existing systems to ensure maximal ease of clinical workflow execution. To that aim the A§E software has, for instance, been extended with operational interfaces to control digital audiometers such as the GN Otometrics Aurical and the Interacoustics Equinox/Affinity. The automated selection of input sources, output transducers and presentation levels by the A§E software increases the efficiency and accuracy of the test procedures as the operator should not be concerned with configuring the audiometer manually (which is subject to human mistakes/errors). Another example is the integration in Hospital Information Systems (HIS) and Electronic Medical Records (EMR), such that the audiologist should not be concerned with the input of patient related data into the A§E software (i.e., maintaining referential integrity between data in the HIS/EMR and in A§E), nor with the storage of test results (e.g., printing them or saving them to a file repository). Test results are automatically stored on the A§E storage servers in the cloud in an anonymous way.

Another significant contribution to the efficiency of outcome measurement is the automation of test protocols. For this reason, we have developed an improved adaptive algorithm (TEMA) for finding detection or discrimination thresholds. This development is described in detail in “Managed estimation of psychophysical thresholds” and allows the reliable estimation of thresholds, also in non-robust responders. This makes that test protocols using this algorithm may be suited for use in clinical settings, where it is not uncommon to encounter malingering or aggravating subjects. At present, TEMA is used for instance in the Harmonic Intonation (HI) and Disharmonic Intonation (DI) tests, which are described in more detail in “Clinical assessment of pitch perception”.

The HI and DI tests are fairly new tests and are conceived as a first attempt to systematically assess pitch perception and, as part of this project, to explore the optimization of temporal coding in CIs.

Today they are not yet part of the model, but their validity and usefulness has been investigated in a number of research projects. For instance, they have been used amongst other outcome measures to evaluate electroacoustic (EAS) hearing with the Neurelec Zebra processor. This study is reported on in “Combined electric and acoustic hearing performance”. Also MED-EL’s FS4 speech coding strategy has been tested with HI and DI. Details on this study are found in “Pitch perception & speech in noise”.

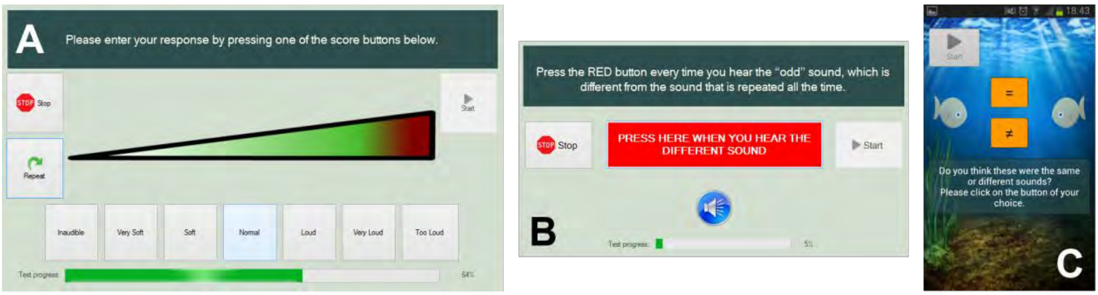


Figure 81: Screenshots of the Loudness Scaling (A) and Phoneme Discrimination (B) self tests on a Windows computer with touch screen. The Disharmonic Intonation test is implemented for Android devices as the Fish Ears app (C), which is available for free in the Google Play store.

A next step in outcome measurement automation is the evolution towards self-testing (Figure 81). If subjects are able to perform certain test procedures autonomously, an audiologist is no longer required to be present during the entire test execution. Eventually this may even lead to remote and home testing, which has the potential to free up large amounts of clinical time. At present self testing is performed on a daily basis as a part of the Eargroup’s clinical routines. One of the more challenging tests to present in a self test paradigm is speech audiometry, as the scoring of responses requires judgment on the phonemic correctness of the subject’s utterances. In “Automated language-universal speech audiometry” a self-test for speech perception is developed, in which the responses of the subject are recorded with a microphone and scored by a machine.

Typical clinical sound treated rooms are relatively expensive in terms of construction, required equipment (an audiometer with transducers), maintenance (regular calibration) and size (they occupy space in the clinic that could be put to other use). Yet they are essential in an outcome based fitting approach to ensure reliable measuring conditions. In many clinics sound treated rooms are over-occupied and the repeated testing for fitting CI recipients would put strain on the availability of these rooms even more. To address these issues, the idea of a portable desktop test box to facilitate psychoacoustical measures in CI recipients has been conceived. The section “An audiometric test box for hearing assessment in CI recipients” reports on the development and specifications of such a device.

3.4. MANAGED ESTIMATION OF PSYCHOPHYSICAL THRESHOLDS

Managed estimation of psychophysical thresholds

The Journal of Hearing Science, 2013, vol. 3, nr. 3, pages 19-31

B Vaerenberg[1,2], WFL Heeren[3], PJ Govaerts[1]

[1]The Eargroup, Antwerp, Belgium;
[2]Laboratory of Biomedical Physics, University of Antwerp, Antwerp, Belgium;
[3]Leiden University Centre for Linguistics, Leiden, The Netherlands;

Abstract

The estimation of perceptive thresholds is a basic element of psychoacoustics. One of the drawbacks of commonly used adaptive algorithms is the lack of reliability when the behavioral response is not robust. To address this issue an adaptive algorithm TEMA (Threshold Estimation by Managed Algorithm) has been developed. TEMA seeks the 50% point on the psychometric curve based on an up-down staircase procedure. Internal controls and stochastic processes aim at enhancing the reliability. The development of TEMA is described, together with its validations with reference to common adaptive procedures. Both Monte Carlo simulations and real subject testing were performed to assess the psychoacoustic threshold in intonation perception tests and the number of stimulus presentations needed. Twenty-nine adult listeners participated in the within-subjects comparison. Nineteen listeners had normal-hearing, the other ten were hearing impaired (5 aided, 5 unaided). The results show that TEMA outperforms the commonly used algorithms in non-robust responders, with a minimal cost in terms of test duration. TEMA's adaptive algorithm was shown to be significantly more resistant to gambling or cheating behavior and threshold drift than traditional, reversal-based algorithms. TEMA increases the accuracy of threshold estimation and the test reliability in non-robust responders. This makes TEMA applicable for automated threshold measurements in clinical settings.

3.4.1. INTRODUCTION

Different procedures have been developed over time to seek the perceptive threshold of a variety of sound features. The most common application in clinical practice is found in pure tone audiometry, a widely used evaluation of a listener's hearing capacity. A popular method for conducting audiometry is the Hughson-Westlake procedure or one of its modifications [132]. It uses a descending familiarization trial that starts at a level presumed to be well above threshold and decreases intensity in steps of 10 dB. Afterwards, a threshold is sought using ascending trials, increasing stimulus level by 5 dB steps. Usually the threshold is defined as the lowest intensity at which positive responses were obtained in 50% of the trials. The definition of threshold as used in audiometry has lead to the fact that clinicians, in general, when speaking of auditory thresholds, naturally think of the stimulus level at which a subject responds correctly in 50% of the trials.

The execution of the Hughson-Westlake and similar procedures in clinical environments typically requires manual manipulation of the stimulus level by a competent tester, which is often time-consuming and subject to intra- and inter-tester variability. Many attempts have been made to automate this for all sorts of scientific and clinical settings, for instance with Bekesy audiometry [134] or AMTAS [135]. These attempts are generally based on the principle of a stimulus level that is automatically adapted to the listener's responses. Hence, the latter methods are called adaptive methods.

A listener's response to stimuli that are presented at different levels is typically probabilistic. It can be described by a psychometric function (e.g., cumulative Gaussian or logistic functions) showing that the probability of positive responses increases from 0% (or chance level, depending on the test task) to 100% with increasing stimulus intensity (Figure 82). The perceptive threshold as defined by the presentation level that yields a positive response in 50% of the presentations is also referred to as equilibrium point or Just Noticeable Difference (JND).

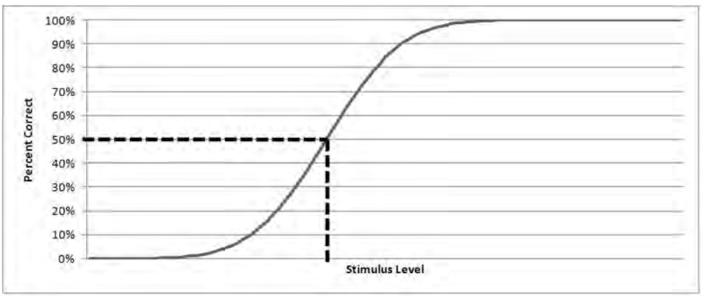


Figure 82: A typical psychometric function showing the probability of a correct response in function of the presentation level. The equilibrium point is defined as the point along the curve where 50% of the subject's answers are correct. The stimulus level at this point is the subject's threshold or JND.

Currently, three types of adaptive methods are being used in psychophysical research as well as in clinical practice: (i) parameter estimation by sequential testing (PEST), [136], (ii) staircase procedures [137], and (iii) maximum likelihood estimation procedures (MLE) [138] [139]. For more extensive overviews and comparisons of adaptive methods in psychophysical research the reader is referred to [137], [140] and [141]. All adaptive procedures require choices to be made by the developer, with respect to stimulus selection, the initial stimulus level, step size, stop criterion, threshold estimation etc. These choices and therefore all existing procedures have advantages and disadvantages, and discussions exist in literature as to under which circumstances one is superior to the other (e.g., [140] [142] [141] [143]. By choosing the right method for a particular experiment or setup, efficiency can be increased considerably.

Of the three types of methods mentioned earlier, up-down staircase procedures are commonly used and can be called the standard of current art [144] [145] [146]. In up-down procedures the presentation level of any given stimulus depends on the participant's response to one or more preceding stimuli. In the 1-up 1-down procedure the stimulus level is decreased after one positive response, and increased after one negative response. The step size typically decreases as the test proceeds, and this can be either on a discrete or a continuous scale. A run is defined by one or more stimuli yielding the same response (either positive or negative), and a reversal occurs when the direction changes from decreasing to increasing (i.e. the response changes from positive to negative), or vice versa. By gradually decreasing the step size the stimulus level in a 1-up 1-down procedure converges to the 50% correct point, the participant's threshold. Typically the test is terminated when a preset number of reversals is reached, and threshold estimation is done by averaging either the minima and maxima of all runs or the mid-run estimates of every second run [147]. Usually, the first reversals are discarded in this computation. In transformed n-up m-down procedures the stimulus level is changed only after a certain sequence of responses. For instance, the 1-up 2-down method increases the stimulus level after each incorrect response, but it only decreases the stimulus level after 2 consecutive correct answers. The transformed procedures converge at other points along the psychometric function, such as 70.7% for the 1-up 2-down version (see [137], Table 1).

This paper introduces the TEMA (Threshold Estimation by Managed Algorithm) algorithm. It was developed for a new module of the Auditory Speech Sounds Evaluation (A§E) test [107] which was originally designed to assess speech sound detection, discrimination and identification in hearing impaired listeners. The A§E is now being extended with a cross-linguistically usable module that includes prosodic stimuli (pseudo-sentences and pseudo-words) and synthetic stimuli (harmonic complexes) to assess the coding of low frequency (< 500 Hz) sound by the aided or unaided ear. These modules will be described in a separate publication.

Since it aims at being used in clinical practice, TEMA should be relatively short in duration, easy to understand for testers and participants, and place minimal requirements on the participants' memory load. In addition it should either produce a result that is reliable or produce no result at all. In contrast to most scientific research methods, where results of high numbers of experiments are statistically analyzed to draw conclusions, the outcome of a single experiment on an individual subject is clinically relevant, making reliability an important requirement. To address this requirement TEMA aims at improving over commonly used up-down procedures in the following aspects. 1) the arbitrariness of using a predefined number of reversals as a stop criterion and for threshold estimation, 2) the use of non-intuitive thresholds (e.g. 70%) and 3) the lack of detecting non-robust (misinformed, incapable or malingering) responders. As such, the procedure should allow for complete automated appliance in clinical practice, without the need for a clinical professional to be present during the procedure to detect whether a subject has misunderstood instructions, shows unstable response behavior or is determined to fake a poor result.

The following sections first describe the development and implementation of the new algorithm and then its validation, followed by a discussion.

3.4.2. ALGORITHM DEVELOPMENT

3.4.2.1. DESCRIPTION OF THE ALGORITHM

An adaptive staircase algorithm was developed to seek the perceptive threshold or JND of stimuli presented in a variety of tasks (Yes/No, Same/Different, N Alternative Forced/Unforced Choice).

INITIALIZATION

The algorithm was designed for use with discrete stimulus levels, but it can be applied to continuous stimulus domains if a desired precision is supplied. The term ‘level’ refers in this context not necessarily to intensity level but to all possible level differences in the acoustic features of the signal (like spectral level). The stimulus domain ranges from reference level (i.e. no stimulus present) to a maximum level that is chosen to reflect the largest stimulus considered to be of interest. Discrete stimulus levels within this range are derived from the desired accuracy. They are ranked with rank 0 corresponding to the reference stimulus level and rank *M* to the maximum level. The initial stimulus level is set to the median of all available ranks. If the median is not a valid level (i.e. when the number of available levels is even), the first valid level greater than the median is selected. The step size *s* is expressed in terms of ranks. The initial step size is set to the highest integer less than one fourth of the total number of stimulus levels.

STIMULUS SELECTION

After a correct response the stimulus level is decreased by the step size, and after an incorrect response the stimulus level is increased by the step size. The selected stimulus level is never smaller than the minimum (i.e. reference) level and never larger than the maximum level. A reversal occurs when the subject’s response differs from the previous response; note that responses to internal control stimuli are ignored (see further).

STEP SIZE

After each reversal the step size is halved and rounded to the nearest integer, see equation (6).

$$\boldsymbol{s} = \boldsymbol{s_i} \times \left(\frac{1}{2}\right)^{\boldsymbol{R}} \text{ if } s \geq 1\text{, else } s = 1 \qquad (6)$$

where s is the step size (integer), s_i is the initial step size, and R is the number of reversals. Once the step size drops below 1 it is rounded up to 1.

The step size is recalculated after each trial. If the step size equals one, it is 'dithered' with one or two levels with a chance of 1 out of 3. This means that in one third of the cases where the calculated step size yields one, it is increased by either one or two units. The appliance of dither reduces the chance of a subject finding a pattern in the procedure.

INTERNAL CONTROLS

Presentations at zero stimulus level (also called reference level) are included as internal controls. They serve to check that the listener is not misunderstanding the task, e.g., using one response option only, and whether he or she is answering consistently, i.e. not just guessing. To not confuse the listener at the beginning of the task, internal controls are presented only if three or more responses have been recorded. After that, the chance of an internal control (p_{ref}) is 0.5. In this way the chance of total absence of controls in an experiment halves with each trial. As soon as the first internal control has been presented p_{ref} is determined according to equation (7), i.e. successive controls are presented with a chance relative to the ratio between false positive responses and the number of controls presented so far:

$$p_{ref} = \frac{\left(\mathbf{1} + \frac{\boldsymbol{F} \times \boldsymbol{N}}{\boldsymbol{C} \times \mathbf{6}}\right)}{\mathbf{1} + \frac{\boldsymbol{N}}{\mathbf{2}}} \qquad (7)$$

where F is the number of false positive responses, N is the total number of presentations, and C is the number of internal controls presented.

Figure 83 illustrates the regulation of internal control presentations for three false positive control ratios. When a listener passes all internal controls the chance of another internal control being presented converges to zero. The chance of a control stimulus being presented increases when the number of false positives increases relative to the number of control presentations. The chance of a control stimulus being presented decreases when the number of false positives decreases relative to the number of controls presented. For instance, when all responses to internal controls are false positives this chance converges to 1/3. When half of the presented controls are passed, chance converges to 1/6.

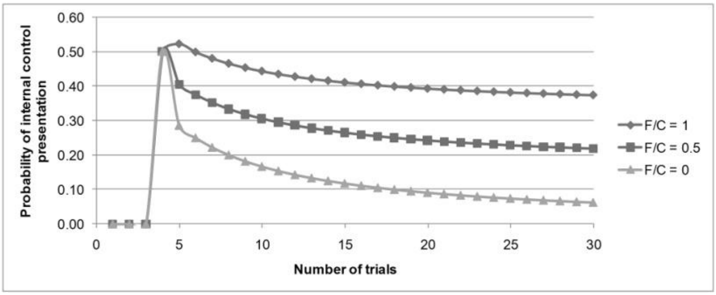


Figure 83: The probability p_{ref} of an internal control (presentation at reference level) being presented as a function of the test progress (number of trials). The probability depends on the ratio of false positive answers to the number of controls already presented (F/C). The first three presentations are never internal controls. The fourth presentation has a chance of 0.5 of being an internal control, etc.

CORRECTION OF THE ANSWER RATIO

After each response the ratio of correct to incorrect responses is calculated for each stimulus level in search for the threshold level. During this calculation, a correction is made based on the number of false positive responses. The basic assumption behind this correction is that if a subject responds positively when no stimulus is present, the percentage of correct (i.e. positive) responses at stimulus level will also be affected by this behavior.

For a stimulus level to be a candidate threshold we assume that 50% (range 35-65%) of the stimuli at that level are detected. However, based on the ratio of false positive responses to the number of internal controls (r_{fp} = F / C) we know if and how often the listener signals to detect a stimulus even when no stimulus is present at all. Therefore, the number of successes at stimulus level is decreased with the number of successes that are presumed to be created by this behavior. This number is based on the false positive ratio applied to half of the total count of the answers at this stimulus level (at threshold absolute guessing will occur in only half of presentations, i.e. the ones where the listener does not detect the stimulus).

Depending on the number of alternatives that are available to the listener, the chance of answering correctly when guessing at stimulus level might be smaller than the chance of answering positively when the stimulus is zero. For that reason the successes to discard are divided by the inverse chance of success minus one. This is the ratio between the probability of generating a false positive response at reference level and the probability of answering correctly at the higher stimulus level, all in a total guess scenario.

The corrected ratio is calculated according to equation (8).

$$r = \frac{S - \dfrac{r_{fp} \times \dfrac{N}{2}}{\dfrac{1}{p} - 1}}{N} \tag{8}$$

where r is the corrected ratio, S is the number of correct responses at the stimulus level, N is the total number of responses at the stimulus level, r_{fp} is the ratio of false positive responses to the number of internal controls, and p is the probability of success in the task.

STOP CRITERIA AND THRESHOLD ESTIMATION

After each trial the algorithm checks whether its stop criteria are met. The basic criterion is that a stimulus level must exist where the percentage of correct responses is between 35% and 65%. Equation (8) is used for this calculation. This level is adopted as the threshold level, and can either be a single stimulus level or be derived from two adjacent stimulus levels. Additional ‘adjacency’ criteria apply to the adjacent stimulus levels in both cases.

*Single Stimulus Level (*Figure 84 *A):* To be a threshold candidate, at least four responses have to be recorded at this level and at least three at both the upper and lower adjacent levels. Therefore the threshold level cannot be the minimum (i.e. reference) or maximum level. If more than one stimulus level has responses that meet these criteria, the threshold is estimated at the stimulus level where the percentage of correct responses is closest to 50%.

*Adjacent Stimulus Levels (*Figure 84 *B):* Two adjacent stimulus levels must exist, each containing at least four responses and where the upper one has more than 65% of answers correct and the lower one has less than 35% correct answers. The threshold level then is the mean of the two levels. Above the upper level at least three responses must have been recorded, except when the upper level is the maximum level. Below the lower level at least 3 responses must have been recorded, except when the lower level is the minimum level.

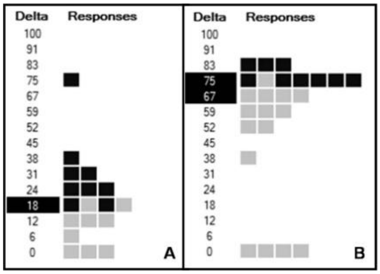


Figure 84: Histograms showing number of responses at different presentation levels (here displayed as the difference Delta between a reference signal and the stimulus). Each black square represents a positive response whereas each gray square represents a negative response. The lower line (Delta=0) depicts the responses to internal controls. TEMA estimates the JND as either a single stimulus level (A) or the average of two adjacent stimulus levels (B).

In both cases the whole set of answers given must meet additional criteria:

(1) Above threshold level, the total number of incorrect answers cannot be greater than the total number of correct answers.

(2) The probability of attaining at least the number of correct answers above threshold level through guessing is less than or equal to 10%. For this the cumulative binomial probability of the number of successes in the total number of responses above threshold is calculated.

(3) The false positive ratio should be less than 35%.

There are three types of alternative stopping criteria:

(1) *Maximum number of trials:* When a preset number of trials is exceeded the threshold is presumed to be non-existent. The default maximum is set arbitrarily to 100 trials.

(2) *Threshold above maximum level:* When at least three answers are recorded at maximum stimulus level and the percentage of correct responses is less than 35%, the threshold is estimated to be somewhere above the maximum level and therefore unknown.

(3) *Too many false positive responses:* When at least ten answers are recorded at reference level and five or more of them are positive, the procedure is aborted and the threshold presumed to be non-existent.

3.4.2.2. IMPLEMENTATION OF THE ALGORITHM

TEMA can be used in different test paradigms e.g. a two alternative discrimination task or a multiple choice identification task, etc. At present, it is being used to find JNDs with same-different discrimination and four-category identification tasks in the A§E 2009 prosodic test battery (developed by the Dual-Pro European consortium with a EC 7th Framework grant, for more information see http://otoconsult.com. Details will be published in a separate paper). Through these tasks thresholds for perception of low frequency information in linguistically relevant contexts are measured. The same-different task is used for detecting intonation in sentences, which is relevant for discriminating between statements and questions. The identification task is used for assessing a subject's perception of stress positions in words.

Since these tests use the TEMA algorithm in a 2-choice and a 4-choice test situation, which are typical clinical situations with specific consequences, we will briefly describe them and demonstrate the effect of the algorithm. Both tasks use reference stimuli consisting of pseudo-linguistic tokens spoken by a female voice. The fundamental frequency of the reference stimuli was adjusted to 200 Hz using Pitch Synchronous Overlap Add (PSOLA) resynthesis as built into the program Praat (v 5.1, [148]). The initial accuracy was set relatively high which resulted in a large amount of available stimulus levels. After test-retest validation on 87 human subjects the accuracy was decreased based on the test-retest variability to shorten the test duration. This resulted in 22 stimulus levels ranging from the 200 Hz reference to a 408 Hz maximum. Figure 85 shows how accuracy was kept constant at 0.5 semitones for levels up to 283 Hz, which is 6 semitones above reference level. From there on accuracy was decreased linearly and with respect to the stimulus level. Based on the available stimulus levels, the TEMA algorithm set the initial level to 275 Hz and applied an initial step size of 5 levels.

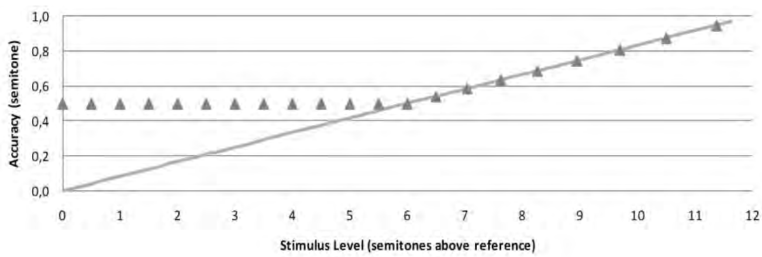


Figure 85: The stimulus domain (triangles) of both the Word Stress Pattern and the Sentence Intonation tests, showing the accuracy (spectral difference between stimulus and reference signals in semitones) as a function of the stimulus level. The spectral difference between adjacent stimulus levels is 1 semitone at stimulus level = 12. It decreases linearly with decreasing stimulus levels until it reaches a constant value of 0.5 semitones for stimulus levels lower than 6.

Both the Sentence Intonation test (SI) and the Word Stress Pattern test (WSP) feature a training mode, where the operator (audiologist) is able to present specific stimuli to the listener to get him or her acquainted with the task. It is important for the listener to be clearly instructed to only pay attention to intonation and to not use roving loudness cues to make decisions. The duration of the training is restricted to a maximum of 10 minutes.

During the test mode, the TEMA algorithm selects stimuli according to its dithered 1-up 1-down procedure and presents them in a timely fashion with roving intensity. According to its internal control mechanism, control stimuli are presented in a probabilistic manner. After a false positive response, a buzz sounds to discourage the listener's guessing behavior.

WORD STRESS PATTERN TEST

The Word Stress Pattern test is an identification task that uses three-syllable pseudo-words (see Figure 86 (TOP)). The listener is offered four response options; three of them are for indicating the presence of an intonation movement on one of the syllables, the fourth gives the listener the opportunity to indicate that he or she does not perceive any intonation or is unsure of its position.

SENTENCE INTONATION TEST

The Sentence Intonation test is a discrimination task presented in a same-different paradigm (see Figure 86 (BOTTOM)). Four- to six-syllable pseudo-sentences are presented in two intervals separated by 500 ms. One of two stimuli is the 200 Hz reference stimulus. The other features a rising intonation on the final syllable with stimulus level (Δf) as size. Each pseudo-sentence has a fixed pitch accent on the second syllable so as to mimic the presence of a sentence accent. The listener is offered two response options, one of them for signalling that the stimuli are different and one for indicating that he or she does not perceive a difference.

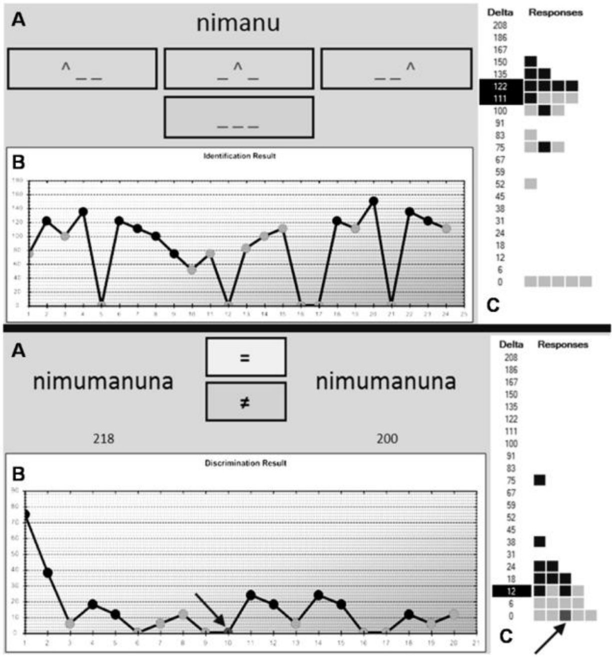


Figure 86: Screenshots of the A§E three choice identification task (top) and Same/Different discrimination task (bottom) using the TEMA algorithm. Each figure shows the presented signal(s) and scoring buttons (A), the sequence (B) and the histogram (C) of responses. Each black square/dot represents a positive response whereas each gray square/dot represents a negative response. The lower line (Delta=0) depicts the responses to internal controls. The dark gray square/dot (arrows) are false positive responses.

3.4.3. ALGORITHM VALIDATION

The algorithm was validated through Monte Carlo simulations as well as through listening tests with actual listeners. A traditional algorithm based on reversals was used as control.

3.4.3.1. METHODS

MONTE CARLO SIMULATIONS

The TEMA algorithm was tested for performance and accuracy using the Monte Carlo method. The response behavior of subjects with known thresholds was simulated by a computer algorithm based on pseudo-random sampling. In addition, for comparative reasons a more traditional adaptive procedure was simulated as this is widely accepted as a valid method for threshold estimation. The reference algorithm was chosen to be a standard 1up-1down procedure [137] using an identical initial value and step size calculation. Similar to the TEMA algorithm, this procedure results in convergence at the 50% correct point on the psychometric curve.

The reference algorithm was set to terminate when 10 reversals had occurred and to estimate the threshold as the arithmetic mean of the last 4 reversal points. These settings were chosen because they appeared optimal for the simulated tasks in terms of accuracy and duration. This was established in a pilot analysis based on the real responses of 178 human listeners in 1036 experiments, which were fed to all possible methods for threshold estimation based on E out of T reversals, where T is the total number of reversals at which the stop criterion is met, and E is the number of reversals that is used to estimate threshold. T was set to range from 4 to the number of reversals encountered in the experiment, and E was chosen to be an even number ranging from 4 to T. The values T=10 and E=4 yielded the optimal trade-off between test duration and threshold estimation stability. An additional stop criterion was included to abort the procedure when 4 consecutive negative responses were recorded at maximum stimulus level, or 4 consecutive positive responses at reference level.

To compare the TEMA algorithm with the reference algorithm, five categories of subjects were defined, and the response behavior of subjects in each category was modelled to investigate its impact on threshold estimation:

A Pure gamblers. These subjects do not react consistently to different stimulus levels. Two settings were used: subjects either respond at random after each presentation, or think a stimulus is always present. In both cases the chance of a correct response is constant and equal to the inverse of the number of alternatives in the task. The psychometric function of these subjects has zero slope.

B Cheaters. This category contains listeners who gain knowledge on the procedure being used and attempt to use this knowledge to manipulate threshold estimation. A number of configurations for each category was defined by adjusting the number of consecutive correct or incorrect answers.

C Perfect listeners. These listeners consistently answer correctly when stimulus level is above threshold and incorrectly when it drops below threshold. The slope of their psychometric curve is infinite.

D Normal listeners. To model normal response behavior a cumulative normal distribution function was sampled with a mean set to the assumed threshold and a standard deviation set to reflect the slope of each subject's psychometric function.

E Listeners with threshold drift. These subjects show a drift of threshold during the procedure. This reflects phenomena such as in-procedure training and temporary lapses. Several configurations were designed by varying initial threshold, the speed at which the drift from initial to target threshold took place, and the delay with which the threshold started drifting towards the target.

For each configuration the simulation was run 1,000 times on both the 2-alternative discrimination task and the 4-alternative identification task.

Categories A and B focus on threshold rejection ("no threshold found") when subjects respond inconsistently or manipulatively. In an optimal situation, the algorithm should reject all cases. For each configuration, the rejection rate was compared between the two algorithms by means of Chi-square tests with Yates' correction. The cut-off level of significance was set at 0.01.

The other categories assess the algorithm's accuracy and duration. In an optimal situation, the algorithm should find the exact threshold in as short a trajectory as possible. The number of trials until the stop criterion was reached as well as the threshold error (= estimated threshold – assumed threshold) were recorded for each simulated experiment. For each configuration, both variables were compared between the two algorithms by means of a t-test for independent samples. The cut-off level of significance was set at 0.01.

REAL TEST SUBJECTS

The performance of the TEMA algorithm was also compared to that of the traditional method using actual listeners, of whom informed consent was obtained.

The same-different discrimination task was used (Sentence Intonation test) with low pass filtered stimuli. Twenty-nine adult listeners participated in the within-subjects comparison. Nineteen listeners had normal-hearing, the other ten were hearing impaired (5 aided, 5 unaided). Each participant completed the task twice: once the TEMA algorithm steered JND estimation, the other time the traditional 4-out-of-10 reversals algorithm was used. As in the Monte Carlo simulations both JND estimation and test duration, i.e. the number of trials before the stop criterion was reached, were taken into account. A within-subject comparison was performed between the two algorithms for both variables by means of a Wilcoxon test for dependent samples. The cut-off level of significance was set at 0.01.

All statistics were performed using Statistica 7.0 software (StatSoft Inc, USA).

3.4.4. RESULTS

3.4.4.1. MONTE CARLO SIMULATIONS

The results are given in "Appendix D: TEMA Monte Carlo simulations" and summarized in Table 3. Gambling behavior was simulated in two configurations. The first setting generated a response from all available alternatives at random. The TEMA algorithm led to a threshold in 0.5% of the runs when two response alternatives were available, and in 12.0% of the runs when four alternatives were used. This is opposed to 38.6% and 93.3%, respectively, in simulations using the reference algorithm.

In addition, simulations of the behavior of a subject who thinks a stimulus is always present and therefore never chooses the 'I don't know' response yielded a threshold in 67.0% of the runs (4-alternative task) for the reference algorithm and 0.0% for the TEMA algorithm.

Table 3: Summary of simulated and real subject results comparing the TEMA and the REF algorithm.

Category	N	TEMA better	ratio			REF better
Gamblers	4	3	1.6-16			0
Cheaters	8	8	∞			0
Category	N	Better algorithm	Threshold		Nr trials	
			N	Diff	N	Diff
Perfects	6	TEMA	0		0	
		REF	0		6	2-98%
Normals	8	TEMA	2	3 to 6%	0	
		REF	1	5%	8	27-108%
Drifters	12	TEMA	6	29 to 190%	0	
		REF	4	9 to 37%	12	14-88%
Category	N	Algorithm	Threshold		Nr trials	
NH	19	TEMA	12		20	
		REF	13.5		18	
HI	10	TEMA	49		31	
		REF	34		21	

Simulations of cheating behavior were modelled in the most obvious way of tampering with an up-down procedure: by alternately answering correct and incorrect in a 1-up, 1-down procedure a number of reversal points will eventually be reached, and the procedure will converge to threshold. In a task with multiple choices the only way for a subject to respond in such a fashion would be when the stimulus level is well above the subject's actual threshold. This would correspond to malingering. In Yes/No tasks, like clinical pure tone audiometry, however the subject could easily respond Yes for a while, causing a simple 1-up 1-down procedure to select levels below the subject's actual threshold, at which point the subject could start alternating responses without the need to detect a stimulus and still reach a predefined number of reversals, causing the procedure to yield a threshold below the subject's actual detection threshold. The number of consecutive correct or incorrect answers (reversal rate) and the number of presentations after which the subject gets wind of the underlying procedure (delay) was adjusted to create different configurations. Whereas the reference algorithm converged to threshold in 100% of the simulations under all configurations, the TEMA algorithm only did in 7 to 29%.

The psychometric function with infinite slope, as in perfect listeners, produced a threshold in all runs, using either algorithm. Assumed thresholds of 2, 30 and 175 Hz were simulated. Both algorithms estimated the exact same thresholds. The reference algorithm needed 12 trials on average to meet its stop criteria, whereas an average of 20 trials was observed when using the TEMA algorithm.

Simulations of normal subjects were conducted in multiple configurations. Thresholds of 15, 50 and 150 Hz were used, and different slopes were applied by adjusting the standard deviation (σ) of the underlying normal distribution. Both algorithms showed similar and acceptable accuracy (error was less than 0.7 semitones). When simulating very mild slopes (σ = 150 Hz) the TEMA algorithm rejected up to 50% of the thresholds, and a small increase in accuracy was observed in comparison with the reference algorithm. The TEMA algorithm on average required 29 trials to reach threshold estimation or rejection, whereas the reference algorithm converged to threshold after an average of 19 trials.

For the simulation of unstable psychometric functions, i.e. threshold drift, a number of parameters was used. The initial threshold determined the mean of the underlying distribution at the start of the simulation. A configurable step size (on a Hz scale) was used to vary the speed at which the drift from initial to target threshold took place. A delay was set to determine the number of trials after which the threshold started drifting towards the target. As with the simulations of the normal subjects, the standard deviation was varied to mimic different psychometric slopes. Significant gains in accuracy where observed when using the TEMA algorithm to simulate drift from 100 Hz to 20 Hz starting after 10 trials with a decrease of threshold of 8 Hz per trial. This configuration led to estimation of a 24 Hz JND by the TEMA algorithm, whereas the reference algorithm yielded 62 Hz. When drifting from 50 Hz to 10 Hz at 2 Hz per trial the TEMA algorithm estimated a JND of 14 Hz, whereas the reference algorithm ceased at 25 Hz. Other configurations did not show a significant difference in accuracy between the algorithms.

3.4.4.2. REAL TEST SUBJECTS

Figure 87A compares JNDs obtained using the TEMA algorithm and the ones obtained using the traditional one for both the normal hearing and the hearing impaired listeners. The median difference was 0 Hz for normal hearing listeners ($p>0.05$) and 3 Hz for hearing impaired listeners ($p>0.05$).

Figure 87B shows the difference between the average number of trials needed to compute a threshold value with the TEMA algorithm in comparison with the traditional algorithm. The median difference was 5.7 trials for the normal hearing listeners ($p=0.02$) and 12.8 trials for the hearing impaired listeners ($p<0.01$).

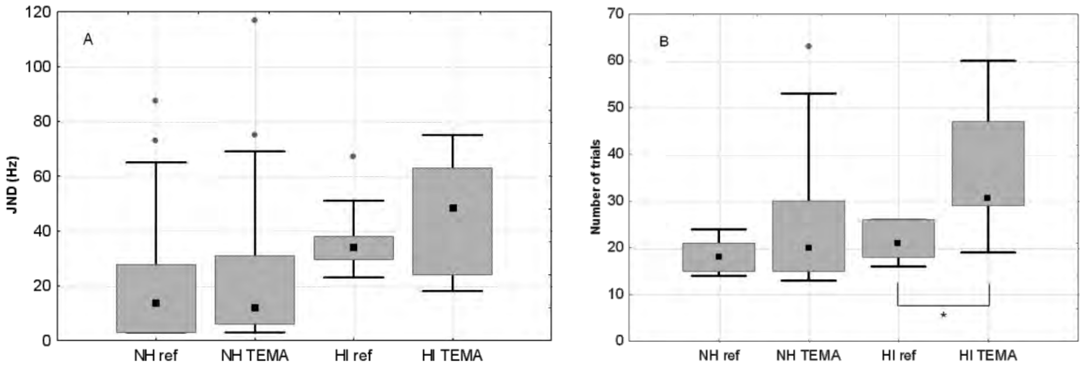


Figure 87: Summary statistics of the results with both algorithms in normal hearing and hearing impaired real test subjects, showing the JND (A) and the number of trials (B) needed to conclude the test. NH ref and NH TEMA: reference algorithm and TEMA in hearing subjects. HI ref and HI TEMA: reference algorithm and TEMA in hearing impaired subjects. The Box and Whisker plots represent the median (square), quartile range (box), range (whiskers) and outliers (dots) for each group.

Actual test durations varied between one and six minutes. Taking into account that the average test time in clinical practice is 6.5 seconds per trial (unpublished data based on 300 clinical test procedures), TEMA would increase the test duration by approximately 0.5 to 1.5 minutes.

3.4.5. DISCUSSION

Psychophysical threshold estimation is an important procedure in clinical and scientific practice. A perceptive threshold often distinguishes between normal or abnormal functioning; it is used to make therapeutic decisions, measure the effect of therapies or interventions and to follow up the course of disease or the evolution of a patient.

Although threshold measurement is common practice in the daily routine, the accuracy and reliability of the procedures used are not often questioned. In some cases strict and systematic instructions have been introduced to reduce the inter-tester variability. This for instance is the case for tone audiometry. But even then it is likely that these procedures are not entirely followed in everyday circumstances. Threshold estimations are time-consuming and accuracy and reliability are related to the amount of time spent at the measurement. This is specifically the case when the test subject’s responses are not fully consistent and subject to the interpretation and judgment by the competent tester.

Automation may be a way to systematize threshold measurements, improve the test quality and save time. First attempts to automate threshold measurements coincide with the introduction of desktop computers more than 30 years ago. This has yielded useful algorithms with acceptable accuracy and reliability and with a good cost/efficacy ratio.

For several reasons explained in the introduction, we have believed it to be worthwhile revisiting the existing algorithms and constructing a new one in an attempt to overcome some of their weaknesses and to optimize some of their features. This has led to the TEMA procedure, which was primarily developed to estimate the 50% point, i.e. the traditional threshold, for low frequency (< 500 Hz) perception in hearing impaired populations. It specifically addresses three challenges for up-down procedures that form the current standard of the art: the inherent arbitrariness of using a predefined number of reversals for threshold estimation, the use of non-intuitive thresholds, e.g., at 70% of the psychometric curve, and guessing behavior of subjects in simple procedures.

For stimulus selection the TEMA algorithm uses an up-down staircase procedure, which is in principle the simplest of the three methods discussed in the introduction. The sole assumption underlying a staircase procedure is the monotonicity of the psychometric function. A possible weakness is that the test subject may rely on expectation of the next stimulus instead of on perception. In the TEMA algorithm measures have been taken to actively discourage listeners from guessing as well as to diversify the selection of stimulus levels near the threshold.

The more popular methods of threshold estimation in psychophysical procedures do not converge at the 50% point along the psychometric curve, but generally at points over 70%. This has the advantage that they are more robust, i.e. have lower variance of the threshold estimate [149], but the disadvantages that those locations may be considered less intuitive, and estimate a point along the psychometric curve where upward and downward changes are more likely to be asymmetric. As TEMA was designed to estimate the 50% point (i.e. the traditional clinical threshold), a 1-up 1-down method was chosen. Alternative methods like the popular 1-up n-down however have the additional advantage that they require n consecutive correct responses for the stimulus level to be decreased, so picking responses randomly will make the staircase go up, more than it goes down. In case of unreliable subjects, the staircase will most likely hit its upper limit and the procedure will be aborted. In case of the 1-up 1-down method, the staircase does not feature such a preferred direction and will most likely stay within limits, even when subjects are responding randomly. To overcome this vulnerability to guessing a commonly used solution is to increase the number of observation intervals within trials. However this approach not only increases the dependency on the subject’s memory, but also prolongs the test duration. As the A§E tests target both the hearing impaired population and very young children, they require an algorithm to support reliable threshold estimation in even the simplest of tasks like Yes/No tasks and Same/Different tasks. For this reason the TEMA algorithm uses internal controls for detecting unreliable response behavior that can be used also in these kinds of tasks.

In adaptive methods, the procedure is traditionally stopped after a predetermined number of reversals have been reached, after which the threshold is calculated from another predefined number of these reversals. The problem with the use of a predetermined number of reversals for threshold estimation is its inherent arbitrariness. No matter how optimally the stop criterion in terms of reversals for a given test setup has been set, it will be suboptimal for the individual subject.

It is to be expected that in more experienced listeners, less reversals are needed for accurate threshold estimation than in the more naïve listeners. As today's computers allow for quick computation, the TEMA algorithm uses more advanced methods for dynamically setting the stop criterion as well as for threshold estimation.

Performance of the TEMA algorithm was analyzed through Monte Carlo simulations in which it was compared to the current standard. Moreover, listening tests with actual listeners were also run to compare the two algorithms.

The Monte Carlo simulations showed that both algorithms give perfect results in "perfect" test subjects. In these cases, the TEMA algorithm needs more trials to estimate the JND. On average, however, this takes only a few extra seconds (see further).

In "normal" subjects, both simulated and real, the results between the two algorithms are highly comparable and very accurate (close to the assumed threshold in the simulated cases). The TEMA algorithm again needs more trials to estimate a JND, corresponding to a few extra seconds test time. To the extent that the test subject's behavior approaches gambler's behavior, the TEMA algorithm clearly excels in accuracy, at the cost of substantially more time to come to conclusions.

But above all, the simulations showed that TEMA is significantly more resistant to gambling or cheating behavior and threshold drift than the traditional algorithm with reversals. As opposed to the traditional algorithm, acceptance scores in the case of gambling or cheating subjects were much lower for TEMA than for the traditional algorithm, which we take as evidence for the higher reliability of the new algorithm.

In the real subjects, TEMA also took somewhat longer. In hearing subjects, this was minimal. In the limited number of hearing impaired subjects however, the difference with the reference procedure was more pronounced, with a median of 12 additional trials. Three subjects out of 10 needed more than 25 extra trials. These also happened to be the ones showing the largest difference in threshold, the threshold found by TEMA being 21, 25 and 42 Hz higher than by the reference algorithm. It seems reasonable to speculate that the reference algorithm may have underestimated the threshold in these cases and that TEMA took more time to find more accurate a threshold. This would be in line with the Monte Carlo simulations. HI subjects show greater variability in JND when measured with TEMA than when measured with the reference algorithm. Although the HI sample size is small (10 subjects) and no significant difference was found between the thresholds obtained with both algorithms, this may illustrate a diversity within this subject group which is not fully expressed using the reference procedure, rather than an intrinsic variability caused by the TEMA itself. This is supported by the test-retest validation on 87 subjects which showed that the differences between test and retest TEMA thresholds are considerably smaller than the within subject differences between TEMA and reference algorithm thresholds.

We believe this to be of great clinical relevance. Gambling or cheating behavior exists in daily clinical practice. This is not only so for some rare malingering subjects, but also for subjects who with the best of intentions consider a test situation as a personal exam and who have the desire to succeed and to please the tester. In manual test procedures, the competent tester has the experience and capacity to judge the subject's behavior and to correct it by giving feedback or additional training and explanations. In automated procedures however, this judgment is lacking. Therefore the algorithm itself should contain internal controls and other processes to reduce the risk of overlooking gambling and cheating. Moreover, even if this type of behavior only influences the test result in a minority of the real subjects, it is a fact that treatment or intervention in this single subject does not depend on the group statistics, but merely on his or her own test result. As the outcome of a single experiment on this individual subject may have important clinical consequences, reliability of the result is of utmost importance. The downside of longer test durations appears to be limited to seconds. Both in the Monte Carlo simulation and the real patients, the additional number of trials required by TEMA was in the order of 10. Taking into account that the average test time in clinical practice is 6.5 seconds per trial (unpublished data based on 300 clinical test procedures), TEMA would thus increase the test duration by approximately 0.5 to 1.5 minutes, which may be considered acceptable.

In conclusion, we believe that the TEMA is an adaptive algorithm allowing automatic threshold measurement with a number of advantages over other procedures. The trade-off is that it slightly lengthens the test time but it is argued that this is of limited clinical burden and that this is outweighed largely by the gain in accuracy and test reliability.

3.4.6. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support and contributions by the members of the DUAL-PRO consortium: Otoconsult (Belgium), Neurelec (France), Del Bo Tecnologia Per L'Ascolto (Italy), Universiteit Antwerpen (Belgium), Leiden University (Netherlands), CNRS-LPP (France), Paris XII (France), Ca Foscari Venezia (Italy), Univ Bucharest (Romania). This work was supported by a grant from the European Commission (FP7-SME-222291 DUAL-PRO). BV received a PhD grant for this work from the IWT (agency for Innovation by Science and Technology), Baekeland-grant IWT090287

3.5. CLINICAL ASSESSMENT OF PITCH PERCEPTION

Clinical assessment of pitch perception

Otology & Neurotology, 32:736-741, © 2011, Otology & Neurotology, Inc.

Bart Vaerenberg[1,2], Alexandru Pascu[3], Luca Del Bo[4], Karen Schauwers[1], Geert De Ceulaer[1], Kristin Daemers[1], Martine Coene[5], and Paul J. Govaerts[1,5]

[1]The Eargroup, Antwerp, Belgium; [2]Laboratory of Biomedical Physics, University of Antwerp, Antwerp, Belgium; [3]University of Bucharest, Bucharest, Romania; [4]Fondazione Ascolta e Vivi, Milano, Italy; [5]Centre for Computational Linguistics and Psycholinguistics, University of Antwerp, Antwerp, Belgium

Abstract

The perception of pitch has recently gained attention. At present, clinical audiologic tests to assess this are hardly available. This article reports on the development of a clinical test using harmonic intonation (HI) and disharmonic intonation (DI). The study is designed as a prospective collection of normative data and pilot study in hearing-impaired subjects. Normative data were collected from 90 normal-hearing subjects recruited from 3 different language backgrounds. The pilot study was conducted on 18 hearing-impaired individuals who were selected into 3 pathologic groups: high-frequency hearing loss (HF), low-frequency hearing loss (LF), and cochlear implant users (CI). Exploratory diagnostics were conducted by means of the newly constructed HI/DI tests using intonation patterns to find the just noticeable difference (JND) for pitch discrimination in low-frequency harmonic complex sounds presented in a same-different task. JNDs for pitch discrimination using HI/DI tests were recorded in the hearing population and pathologic groups. Normative data are presented in 5 parameter statistics and box-and-whisker plots showing median JNDs of 2 (HI) and 3 Hz (DI). The results on both tests are statistically abnormal in LF and CI subjects, whereas they are not significantly abnormal in the HF group. The HI and DI tests allow the clinical assessment of low-frequency pitch perception. The data obtained in this study define the normal zone for both tests. Preliminary results indicate possible abnormal TFS perception in some hearing impaired subjects.

3.5.1. INTRODUCTION

Pitch is an attribute of sound that has been shown to be important for both music perception and the quality of speech perception [150] [151]. By allowing us to order sounds on the low-high dimension, pitch carries essential information about the tonality and melody in music and about the linguistic context of words and sentences in spoken language (e.g., clause typing) [64] [152]. Like loudness relates to sound intensity, pitch relates to the frequency content of sounds. In daily life, the relevant cues for voicing, melody, intonation, and other musically and linguistically important percepts are conveyed by relatively low frequency pitch, relating mainly to the fundamental frequency or F0. The fundamental frequencies of several competing voices in a noisy environment, for example, allow us to distinguish between separate speakers [153]. The way the cochlea codes spectral content of sound can be explained by 2 underlying mechanisms, place coding and phase locking. Both are complementary and overlapping. It is believed that for low-frequency signals, such as the fundamental frequencies of human voices, phase locking of the temporal pattern of nerve responses to the temporal fine structure of the signal is the more dominant cue for conveying pitch. With increasing frequencies, this neural synchronicity becomes more difficult to be maintained. Place coding then comes gradually into play and replaces the phase locking as mechanism for spectral discrimination [154]. In the clinic, hearing assessment often is restricted to measures of detection (e.g., tone audiometry) or identification (e.g., speech audiometry). Clinical tests allowing more fine-grained analysis of the coding of the different components of sound, like spectral discrimination, are rare, and to the best of our knowledge, no tests exist that focus on the capacity of the auditory system to discriminate pitch. The absence of such tests may not have been a problem so far. However, with the emergence of new therapeutic options for sensorineural hearing loss, like cochlear implants, electroacoustic stimulation, or even molecular or genetic therapies, the need for such tests may increase. For instance, cochlear implants (CIs) attempt to restore the tonotopic organization of the inner ear by inserting an array of electrode contacts into the cochlea. This way, the place coding mechanism of the auditory system is partially restored. We think that the A§E spectral discrimination task is helpful in assessing the spectral discrimination, and we use it daily in the selection of CI candidates and the programming of CI processors [107]. However, this test uses unfiltered phonemes as test items, and it therefore does not allow focusing on low-frequency discrimination. Poor low-frequency pitch perception may play a role in a number of frequently encountered complaints by current CI users, like poor music appreciation or poor spatial separation of multiple speakers [153] [155]. A clinical test focusing on pitch discrimination could potentially document and measure this. This is merely one illustration of the need of clinical tests to assess the coding of low-frequency pitch. This article presents the development of such clinical tests to assess the coding of low frequency pitch. They are believed to be relevant in gaining more detailed insight in the coding of sound by the unaided or aided auditory system, and they are expected to be indicative for the capability of the inner ear to use its phase-locking mechanism. Two distinct tests were designed: harmonic intonation (HI) and disharmonic intonation (DI). They both use low-frequency harmonic complexes presented in a same-different paradigm, to find the just noticeable difference (JND) for pitch discrimination in individual subjects. Intonation patterns are applied to the stimuli to maximize focus on temporal processing. For both tests, the construction of stimulus material, test-retest validation, normative data, and preliminary results in a number of hearing-impaired subjects and cochlear implant users are presented.

3.5.2. MATERIAL AND METHODS

In each trial of both the HI and DI tests, 2 stimuli are presented consecutively, one of which has an intonation, whereas the other one does not. The test task is a same-different discrimination task. The non-intonating stimulus that is one of both stimuli in all trials is a harmonic complex signal having a fundamental frequency (F0) of 200 Hz and 3 higher harmonics (with frequencies of 2F0, 3F0, and 4F0). The intensities of the harmonics decrease in comparison with F0 (-6 dB at 400 Hz, -12 dB at 600 Hz, and -18 dB at 800 Hz). A white noise was added to the stimuli (signal-to-noise ratio, +10.9 dB) to make them sound more natural and easy to listen to. Both in the HI and the DI test, the non-intonating sound is presented in contrast to an intonating sound. The intonating sounds used in the HI test feature a frequency sweep of all harmonics (including F0) from NF0 to N(F0 +ΔF), with N ranging from 1 to 4. In the DI test, however, the intonating sounds feature a sweep of the fundamental frequency only (F0 to F0 + ΔF), whereas the higher harmonics are kept fixed at their initial frequency, as shown in Figure 88. As a consequence, the harmonic separation of partial tones is distorted by the sweep, hence a disharmonic (or dissonant) intonation. For both stimulus types, the sweep is linear and introduced at 330 ms after the start of the signal. The sweep duration is 120 ms, and the total signal duration is 600 ms. The timings of the intonation were chosen to resemble the intonation pattern that is used in clause typing to form a question. Each trial thus consists of 2 consecutive stimuli separated by a 500 ms silence. One of 2 stimuli is the non-intonating sound, whereas the other sound is the intonating signal featuring a pitch change ΔF (imposed by either a harmonic intonation in the HI test or a disharmonic intonation in the DI test). The order of stimuli within a trial is randomized. Stimuli are presented to the listener in a same-different task. The listener indicates whether he perceives a difference between the presented sounds. A JND (also called difference limen or threshold) is sought using an adaptive staircase procedure. The details of this procedure are described elsewhere [156]. Briefly, after a training session to make the listener familiar with the task and the test sounds used, the test starts with a large ΔF of 41 Hz. In case the test person discriminates the 2 sounds, ΔF is reduced, and vice versa, according to a dithered one-up one-down procedure converging to the 50% point on the psychometric curve. Internal controls and stochastic processes are implemented to enhance the reliability and to sanction and correct for false-positive responses. Intensity roving (+/-2 dB) is applied to discourage listeners to use any possible loudness cues to discriminate between sounds. Reaction times are recorded on every trial, and total test duration is measured. Initially, the stimulus domain was constructed to contain 41 stimulus levels ranging from the 200 Hz reference to a 350 Hz maximum.

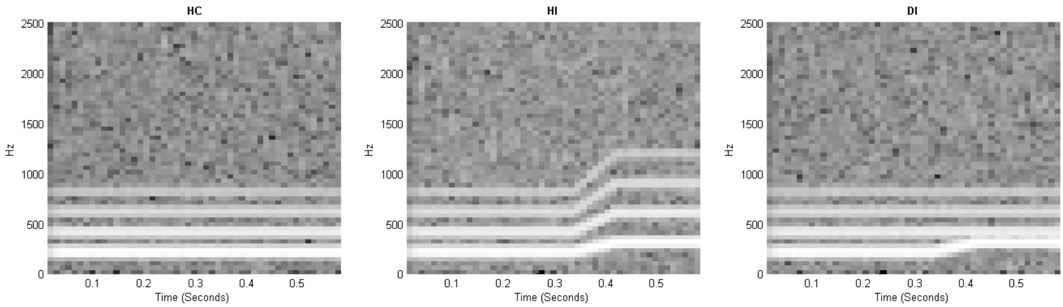


*Figure 88: Spectrograms of the stimuli: on the left, the non-intonating 200 Hz harmonic complex; in the middle, the harmonic intonating sound with all harmonics sweeping from N*200 to N*294 Hz; and on the right, the disharmonic intonation, with only the fundamental frequency sweeping from 200 to 294 Hz, keeping higher harmonics fixed at their initial frequency.*

Inter-level intervals were decreased stepwise from a 1/12 semitone interval in the 200 to 208 Hz range over a 1/6 semitone interval in the 208 to 229 Hz range to a 1/3 semitone interval in the 229 to 350 Hz range. To qualify the setup with the chosen stimulus domain as a robust and accurate measure of pitch perception, a test-retest validation was performed in 29 human subjects. The observed mean absolute difference (+/- SD) between test and retest was 0.043 (+/- 0.032) semitones for HI and 0.043 (+/- 0.63) semitones for DI. Based on this variability, the minimum inter-level interval for both tests was set to the 0.17 semitones, which represents the 97.5th percentile of the observed differences for the most variation-sensitive test (DI). The chosen minimum inter-level interval is expected to cause a test-retest variability of maximum 1 interval in 95% of test runs. The new stimulus domain is depicted in Figure 89 and contains 36 stimulus levels ranging from the 200 Hz reference to a new 414 Hz maximum. The inter-level interval is kept constant at 1/6 semitones for levels up to 224 Hz, which is 2 semitones above the reference level. From there on, the interval increases linearly and with respect to the stimulus level.

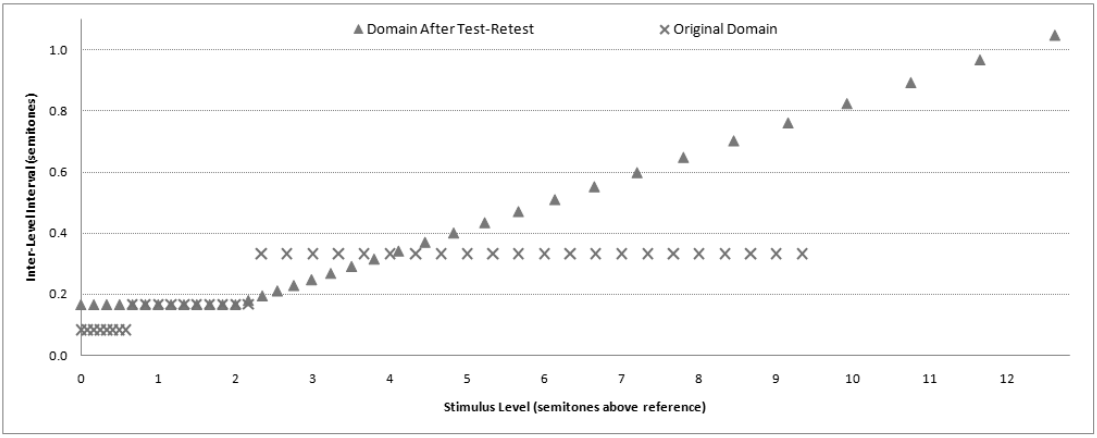


Figure 89: The possible stimulus levels in both versions of the tests. The initial stimulus domain (crosses) featured a stepwise increase in inter-level interval going up to a maximum of 350 Hz (i.e., 9.7 semitones above the reference stimulus level of 200 Hz). The domain after test-retest validation (triangles) features a constant inter-level interval of 1/6 semitones for levels up to 224 Hz (i.e., 2.0 semitones above reference). From there on, the inter-level interval increases linearly at a ratio of 1/12 semitone per 1 semitone increase in stimulus level up to a 414 Hz maximum (i.e., 12.6 semitones above reference).

By decreasing accuracy at higher JNDs, it is expected that the average duration of the test is decreased, in particular when performed in hearing-impaired subjects. Based on the available stimulus levels, the algorithm sets the initial level to 241 Hz and applies an initial step size of 9 levels. Whenever the procedure was unable to converge to a threshold (e.g., the subject's JND is not in the range of the stimulus domain), a JND of 220 Hz was coded for the current analysis. Both HI and DI tests are implemented in the A§E psychoacoustic test battery [107] [157]. The new test setup was then used to estimate JNDs for pitch perception in the normal-hearing population. Ninety subjects aged between 18 and 53 years were recruited from 3 different language groups (Dutch, Italian, and Romanian). All subjects had normal audiometric thresholds (< 20 dB HL at octave frequencies between 125 and 8,000 Hz) at both ears and reported no otologic history. Written informed consent was obtained for all participants. Both HI and DI test results were compared across language groups (Mann-Whitney U test). Results from different languages were then pooled to calculate the 95% confidence interval and to extract normative data for both tests. To explore the practical use of the intonation tests, a pilot study was set up with 18 hearing-impaired individuals who were selected into 3 pathologic groups of 6 subjects each: 1) high frequency hearing loss (HF) featuring audiometric thresholds (better ear) better than 25 dB HL at 250 and 500 Hz and worse than 40 dB HL at 2, 4, and 8 kHz (testing was done in the unaided condition); 2) low frequency hearing loss (LF) featuring audiometric thresholds (better ear) worse than 35 dB HL at 500 Hz and better than the threshold at 500 Hz at 2 and 4 kHz, and 3) cochlear implant users (CI) featuring a normal cochlear anatomy and unaided audiometric thresholds of more than 80 dB HL at the better ear, having been implanted (with full electrode insertion) with their first and only CI more than 6 months before the start of the experiments. Three of them were using the AB HiRes90k implant with Harmony processor (Advanced Bionics LLC, Valencia, CA, USA), the other 3 were using the Cochlear Nucleus 24 with Freedom processor (Cochlear Ltd., Sydney, Australia). The presentation level was 20 dB SL with a minimum of 70 dB SPL. Nonparametric statistics (box and whisker plots, Kruskall-Wallis, and Mann-Whitney U tests) were used to display the results and to compare the results between the 3 pathologic groups and between each group and the hearing subjects and paired nonparametric statistics (Wilcoxon test) to compare the differences between HI and DI results within subjects.

3.5.3. RESULTS

Figure 90 shows the results of normal-hearing subjects for HI (left hand side) and DI (right hand side). No statistically significant differences were found between different language groups, except for the HI results between Dutch (median JND, 1.5 Hz) and Italian (median JND, 2.5 Hz) speakers ($p < 0.01$). This difference of 0.09 semitones is clinically and linguistically irrelevant. Therefore, data from different language groups were pooled to obtain normative data, which are depicted in black (Figure 90). The HI test results (median JND, 2.0 Hz) seemed to be significantly different from the DI test results (median JND, 3.0 Hz) ($p < 0.001$).

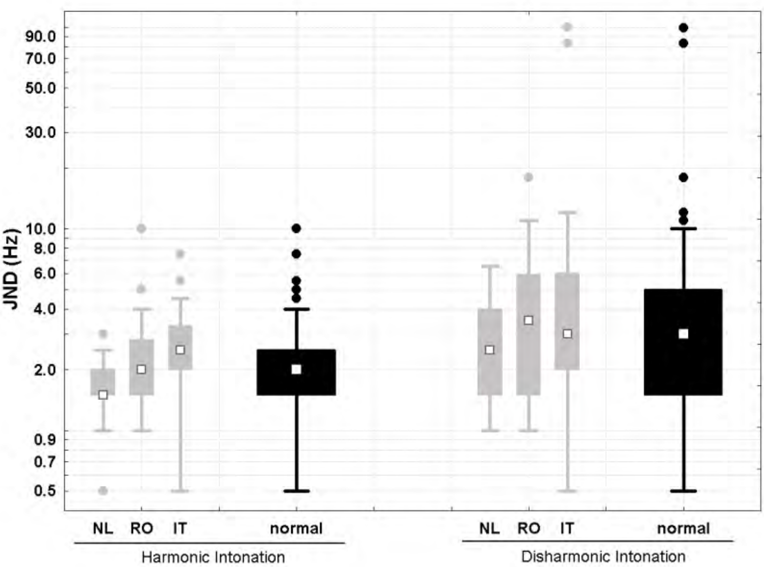


Figure 90: The results in hearing subjects on the HI (left graphs) and the DI (right graphs) test as box and whiskers plots, where the central dot represents the median, the box the inter-quartile region, the whiskers the range, and the separate dots the outliers). The results for each language group are depicted in gray (NL: Dutch, RO: Romanian, IT: Italian) and the pooled results in black.

Figure 91 shows the results of the pathologic groups. The HF group showed median JNDs of 2.0 Hz for HI and 5.0 Hz for DI with the majority of subjects having scores within the reference range. However, the LF and CI groups showed significant differences ($p < 0.01$) in both HI and DI tests when compared with the normative data. The LF group obtained median scores of 54.0 and 94.0 Hz and the CI group, 7.5 and 158.5 Hz, on HI and DI, respectively. Across all runs of both HI and DI, the average (+/- SD) test duration was 144 seconds (+/-105 s).

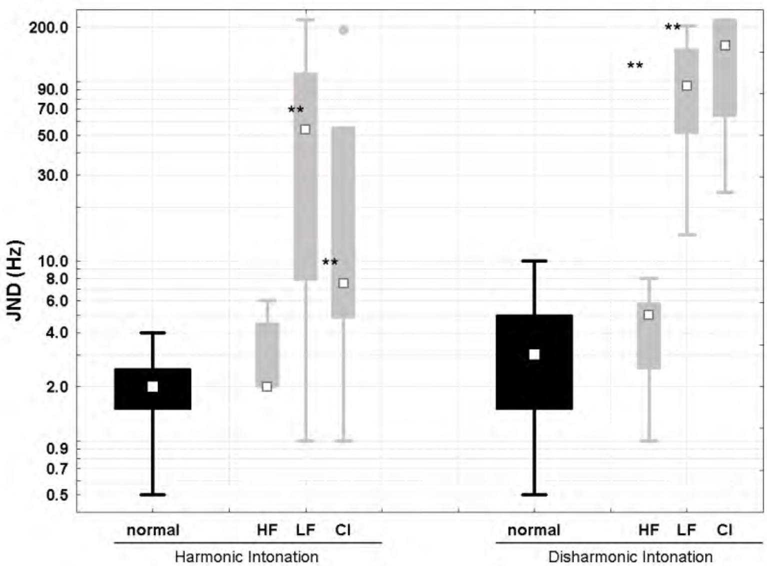


Figure 91: The HI (left graphs) and DI (right graphs) results of the pilot study in subjects with hearing loss (see legend to Figure 90 to understand the box and whisker plots). HF, high-frequency hearing loss; LF, low-frequency hearing loss; CI, cochlear implant users; and normal, the normal data (Figure 90).

3.5.4. DISCUSSION

The purpose of the harmonic and disharmonic intonation tests is to provide a clinical instrument to evaluate the spectral discrimination of the auditory system in the low-frequency range. They assess the perception of pitch changes in low-frequency complex tones. Two particular but inseparable peripheral auditory mechanisms are believed to lie at the origin of the spectral discriminative power of the cochlea. One of them is based on place of excitation (tonotopy) and conveys intonation through a spatial alteration of the population of active nerve fibres. The other is a time-based mechanism (phase locking) that locks onto the TFS of the signal to keep the nerve firings in sync with the fluctuations of sound pressure in time and conveys intonation by changing the auditory nerve fibres' firing rate, keeping it in pace with the instantaneous frequency of the signal. Although many experiments have indicated that the contribution of each of these mechanisms to the total of useful information that is centrally processed may vary according to the nature of the signal, no single experiment exists to isolate one of 2 mechanisms completely. Nonetheless, it is believed that in low frequencies, phase locking is the more important mechanism for conveying pitch. The HI and DI tests were designed to investigate pitch perception in a clinical situation. When comparing the stimuli, it is seen that the cue in HI is more salient: all harmonics are swept together with the fundamental. It is reasonable to assume that both place and time-based codes contribute to the accurate detection of this kind of intonation. However, in the DI stimuli, it is only the fundamental frequency that shifts. Looking at the critical bandwidth of auditory filters, it seems impossible to transfer an intonation as subtle as a few hertz in the 200-Hz region by a place-based code [158]. In consequence, time-based codes are likely to dominate the accurate detection of this kind of intonation. In theory, keeping the higher harmonics fixed while the fundamental sweeps causes beating, and this may introduce a new cue that could bias the results. Beating occurs when 2 sound waves of different frequency are presented simultaneously. This causes a modulation that is the result of the alternating constructive and destructive interference between the waves. However, this possible bias only comes into play for JNDs much higher than a couple of hertz. The beat frequency is equal to the absolute value of the difference in frequency of the 2 waves. So for instance, with F1 = 200 Hz and F2 = 320 Hz, the beat frequency will be $|200 - 320| = 120$ beats per second (bps). In our DI test, beating occurs when F0 interferes with the stationary 400-Hz harmonic. Hence, with F0 = 202.5 Hz, the beat frequency is $|400 - 202.5| = 197.5$ bps. However, temporal modulation transfer functions are known to be low pass with a cut-off frequency near 70 bps for normal hearing listeners [159]. This indicates that a beat frequency of 192.5 bps could not be a cue to distinguish 2 signals. Temporal beatings in the DI stimulus can only serve to distinguish a tone with a stationary F0 of 200 Hz from one with a gliding F0 from 200 to 330 Hz or higher. However, then it is no longer relevant for the clinical interpretation of the test results. As shown, JNDs above 4 or 10 Hz are outside the clinical normal zone. Another possible bias in the DI test results could come from the dissonance or loss of harmony in the signal that causes the percept of a split tone and a severely changing waveform that could lead to a lower JND. However, the results show that listeners are less sensitive to this loss of harmony than they are to the harmonic intonation. In normal-hearing subjects, differences between HI and DI, although statistically highly significant ($p < 0.001$), are so small (1 Hz) that they are unlikely to be clinically relevant. In conclusion, it seems fair to say that both tests are easily performed in normal-hearing subjects, that the results are in line with earlier findings of JNDs for pitch changes, which are approximately between 1 and 4 Hz in the 200-Hz range and that the 2 tests do not assess fully identical psychoacoustic phenomena. In addition, no relevant differences were found between language groups. We wanted to make sure that the normative data were not biased by the linguistic background of the listener. As said, pitch is used to convey linguistic information, but the importance of it can be different in different languages. For instance, the perception of syllable prominence in Dutch is predominantly cued by pitch and, to a lesser degree, by syllable duration, whereas in Italian, it is the other way around [160] [161]. It would be conceivable that Dutch listeners therefore have better acuity for pitch than Italian speakers. Because no differences were found between the Germanic and Romance language used, the tests seem largely language independent and applicable in different language groups. The adjusted stimulus domain (after validation) is expected to cause a test-retest variability less than 1 inter-level interval, which adds to the robustness of the tests. Test durations measured indicate that, on average, HI and DI together can be performed in a single subject in less than 5 minutes. Together with the included training mode, this makes that the tests are well feasible in clinical practice. Once a test is feasible in clinical practice and normative data have been obtained, the next step is to evaluate whether it is relevant in diagnostic, that is, pathologic situations. Although this is beyond the scope of the present article, preliminary results have been obtained in different groups with abnormal hearing. Although the numbers are too low to draw any robust conclusions, remarkable differences between results on HI and DI seem to exist in these groups. The results of the CI group show that the majority of these subjects are performing reasonably well on the HI task, presumably because they are still able to use the place cue caused by all harmonics sweeping to detect pitch changes. On the DI task, the only spectral cue consists of the 200 Hz component shifting. As current CI devices are mainly tonotopically organized and have a limited number of electrodes, it is not likely that a subtle change in a single-frequency component causes a different electrode to be stimulated [162]. The frequency bandwidth of the most apical channel was 250 to 416 Hz for the AB device and 188 to 313 Hz for the Nucleus device. Because no or only limited TFS is conveyed within one spectral band, the fundamental frequency needs to be analyzed into a different spectral band (causing a change in the physical stimulation site) for its sweep to be detected. As said before, high stimulus levels ($\Delta F > 150$ Hz) also cause temporal beatings in the DI signal that may serve as cue for CI users to discriminate between sounds. This hypothesis is in line with the high JNDs on the DI test observed in the CI group (median JND = 158.5 Hz). The different JNDs for HI and DI in the LF group could be attributed to the fact that the loss of audibility in the low-frequency region also impacts the spectral discrimination within this region. In general, this is attributed to the broadening of auditory filters as a result of malfunctioning hair cells. As the concept of filter bandwidth is not exclusively built on either place- or time-based coding, broadening may result from deficiencies in either or both of them. However, as discussed in the introduction, it is reasonable to assume that temporal coding is dominant in the DI task. Although speculative at this stage, it seems appealing to consider that this test might distinguish patients with perceptive hearing loss who have good low-frequency TFS coding (phase locking) from others who have not. An additional illustration comes from 1 particular subject having a low-frequency hearing loss, not included in the LF group.

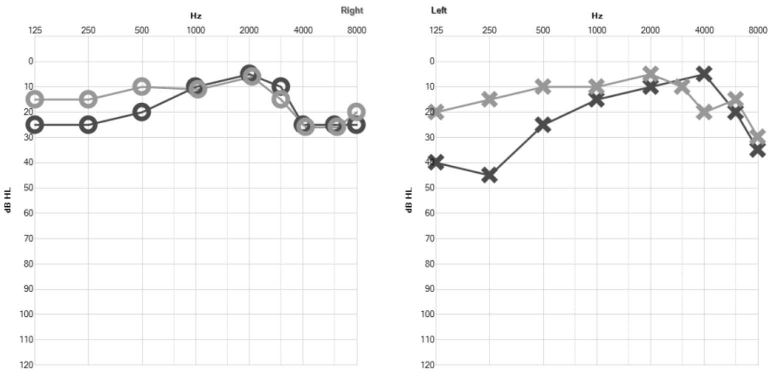


Figure 92: The musician's audiometric thresholds for right (left graph) and left (right graph) ears at different points in time. The dark gray curves show thresholds at initial measurement (T0). The light gray curves show thresholds measured 4 months later (T0+4).

This subject was a professional musician who experienced episodes of dizziness, loss of equilibrium, and left-sided tinnitus since more than 6 months. He presented with recently developed hearing loss, distorted sound perception, and fullness at the left ear. Pure tone audiogram showed a low-frequency perceptive hearing loss, mainly at the left side (Figure 92). He showed abnormal test results on both the HI and the DI test (Figure 93). He was given medical treatment for Ménière's disease (betahistine and antidepressants) for 4 months, and when he returned, the audiometric thresholds had normalized (Figure 92). Still, he felt unable to take up work again because, as a professional musician, he reported not to be able to follow the tone of his fellow musicians. When asked to specify, he said that "the harmonics sounded too loud, while the ground tone seemed to be missing". The test showed that the HI result had normalized, whereas the DI result had remained abnormal (Figure 93). Five months later, the man returned with the message that he had taken up work again and that, subjectively, the symptoms had disappeared. Test results confirmed normalization of the DI result as can be seen in Figure 93.

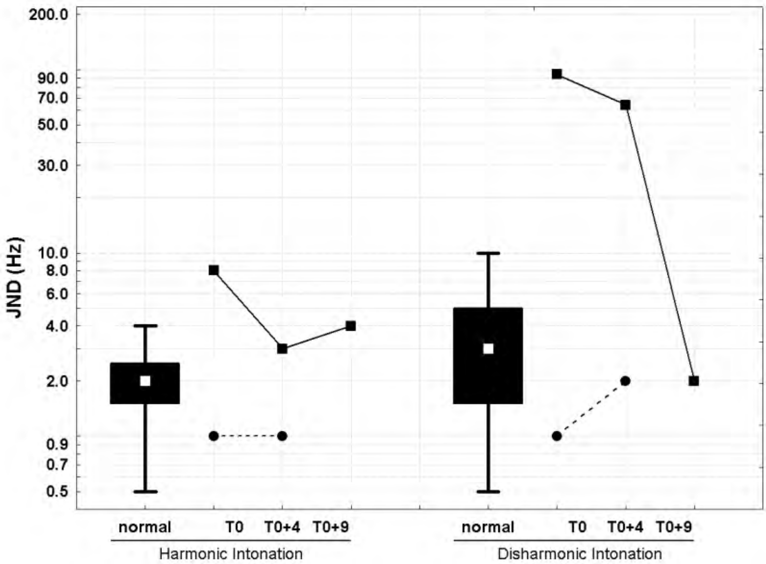


Figure 93: The HI (left graphs) and DI (right graphs) results for both the left ear (squares) and the right ear (dots) of the musician at different points in time and in comparison with normal data. The results for the left ear show normalization of HI after 4 months (T0+4) and of DI after 9 months (T0+9).

In conclusion, the HI and DI tests address the need for a more fine-grained and targeted clinical evaluation of the cochlear function. They provide clinicians with an instrument to assess the perception of low-frequency pitch perception, which is particularly important for understanding speech in multi-talker situations and also music appreciation. The tests have been shown to be clinically feasible with limited test duration and robust results. They also have been shown to be relevant because they are able to distinguish between different subpopulations and among individuals within subpopulations. This indicates that useful information could be extracted from application of the tests, and it is anticipated that they will enable clinicians to explore different pathologic conditions and that they may become instrumental in both diagnostic and therapeutic applications. They have been implemented in the A§E 2009 psychoacoustic test suite (http://www.otoconsult.com) and are available for further exploration and clinical use [163] [164] [165].

3.6. AUTOMATED LANGUAGE-UNIVERSAL SPEECH AUDIOMETRY

Language-universal speech audiometry with automated scoring

Proceedings of Interspeech 2013

Bart Vaerenberg[1,2], Louis ten Bosch[3], Wojtek Kowalczyk[4], Martine Coene[1,5], Herwig De Smet[6], Paul J Govaerts[1]

[1] The Eargroup, Antwerp-Deurne, Belgium
[2] Laboratory of Biomedical Physics, University of Antwerp, Belgium
[3] CLS/CLST, Radboud University Nijmegen, The Netherlands
[4] LIACS, Leiden University, The Netherlands
[5] Vrije Universiteit Amsterdam, The Netherlands
[6] Karel de Grote Hogeschool, Antwerp, Belgium

Abstract

The clinical assessment of speech discrimination by professional audiologists is resource intensive. Yet discrepancies in language or dialect between the test subject and the audiologist may cause a significant bias in the test result. To address these issues, a speech audiometric test (SAT) has been designed to be language/dialect independent and to allow automated scoring by means of an MFCC-based Dynamic Time Warping alignment measure. A Pearson correlation of 0.83 was found between the automatic scores and human phoneme scoring. Normative data were obtained and compared to conventional SATs which revealed differences in speech reception thresholds within 2 dB.

3.6.1. INTRODUCTION

Since the introduction of cochlear implants (CI) and similar therapies, assessment of speech perception has become more and more important in the clinical practice [87]. Speech audiometric results are interesting because they relate closely to the patient's hearing performance in daily life. As such speech audiometry is routinely performed for the selection of CI candidates, for the evaluation of outcome in CI recipients [166] and even to steer the programming (i.e., patient-dependent optimization) of CI speech processors [102]. From a patient's initial intake onwards and throughout the long term follow-up, speech audiometry is performed repeatedly, resulting in a substantial load on the clinic in terms of time and resources [167].

In conventional speech audiometry tests (SAT), words or sentences are presented acoustically to the subject at predefined intensities. The subject is instructed to repeat what he/she has heard and a trained professional (audiologist) scores the subject's responses. The scoring of responses depends on the stimulus material used and can either be a phoneme score (e.g., in CVC monosyllable tests) in which errors on the phoneme level are recorded, a word score in which each utterance is judged as entirely correct or incorrect, or a sentence score in which the subject's correct repetitions of a number of keywords in a sentence is counted. To minimize test-retest variability it is important that enough stimuli are presented in a single test run. Typically, 20 to 50 items are used per presentation level [168].

In addition, to obtain reliable and consistent results, it is essential that these speech perception tests are well designed in terms of stimulus quality, difficulty across lists, and output level calibration [169]. However, in many languages such standardized speech stimuli do not exist or there is no normative data available for them. Another drawback is that stimuli need to be representative for the subject's language in terms of vocabulary and phonemic content and that these linguistic variables show great variation across languages and dialects. As such, the scoring of responses may result in a significant error if the audiologist and the patient do not have the same native language or dialect [170].

In this study, we address these issues by developing a new type of test for speech understanding in quiet and in noise. The fundamentals of this test consist of the construction of a personal, yet language representative speech test for each individual subject based on his/her own lexicon and an automated, language and dialect independent scoring of responses, allowing subjects to perform the test on their own. During an initial session, words from a subject's daily readings are presented visually for him/her to pronounce. The subject' utterances are recorded and subsets of these words are presented acoustically, for the subject to repeat, during later test sessions. At that time, a scoring mechanism compares the original, visually prompted utterance to the repeated, acoustically prompted one. The former is assumed to be correctly pronounced while the latter may contain errors caused by deficits in the perception of the acoustic stimulus. The difference between both pronunciations may therefore be a measure for impaired speech perception.

In this paper we will particularly focus on the algorithms and the model used to allow for automated scoring. The remainder of this paper is organized as follows. In the next section, the assessment of pronunciation differences will be discussed. In the third section a model for converting these differences to humanly interpretable scores is developed. The fourth section handles the normative data and in section 5 our results are presented, which are discussed further in section 6.

3.6.2. ASSESSMENT OF PRONUNCIATION DIFFERENCES

There are different methods to assess the difference between two pronunciations. One way is to use methods based on Automatic Speech Recognition (ASR) [171]. Using ASR the difference between utterances can be measured in terms of different results from e.g., a free phone loop decoding, or by specific techniques used to measure differences between pronunciations from a learner and a teacher in Computer Assisted Language Learning [172]. The necessity of trained acoustic models, however, is a drawback for the application of the ASR method in a clinical setting, in particular in cases of under-resourced languages. It was therefore opted to use another way to assess differences between utterances: by using a language ignorant Dynamic Time Warping (DTW) approach [173]. Unlike ASR, DTW does not require substantial speech data for training; instead a few parameters must be chosen. The disadvantage of DTW is that it accumulates all acoustic deviations between the two utterances along the found best alignment path, irrespective whether these differences would make sense phonemically according to a listener. However, the use of DTW before ASR approaches came into fashion shows that the DTW alignment score can be a useful measure for distinguishing two pronunciations. This idea is further exploited in the descriptions below.

This particular application required a system for comparing 2 different pronunciations of a short word using a language-ignorant approach that does not involve individual training. Mel-Frequency Cepstral Coefficients (MFCC [174]) have been used for extracting perceptually relevant features from the short-term speech spectrum. Combined with DTW, the resulting MFCC DTW distance measure meets the defined requirements and was chosen to be used in this study.

As a first step, MFCC frames are calculated using a frame analysis window width of 25 ms and frame shift of 10 ms. The MFCC frames are augmented by their first and second order temporal derivatives (delta, delta-delta), and an utterance-based Cepstral Mean and Variance Normalization (CMVN) is applied to minimize the between-speaker differences and thereby to optimize the generalization of the speaker-independent DTW settings. The DTW then operates upon pairs of sequences consisting of these augmented MFCC vectors and the Euclidean distance is used to compute local scores. No additional costs are attributed to frame insertion and deletions. The total score of the best alignment path is normalized by dividing this score by the number of traversal steps making up that path and serves as the eventual DTW alignment score (henceforth 'DTW distance').

The applicability of the DTW distance for this particular application was validated by considering 4 different sub databases. These databases contain pairs of words pronounced by the same speaker: 1) 300 pairs of a same word (SAME, e.g., cat-cat); 2) 25 pairs of minimally different words (MINIMAL, e.g. lon-lom); 3) all pairs of typically different words (TYPICAL, e.g., cat-goes), chosen from a set of 300 short words and 4) 25 pairs of maximally different words (MAXIMAL, e.g., 'put'-'sil'). Each of these 4 data sets was collected from 47 speakers.

3.6.3. MAPPING DTW TO HUMAN SCORES

For the application to be usable in a clinical setting MFCC DTW distances had to be transformed to a psychometric scale ranging from 0 to 100, resembling the shape and range of the human scoring in a conventional SAT. This mapping is non-linear and it has been derived from experimental data.

To record the tuning data used to appropriately map DTW scores, a software application was developed to record, present and score speech utterances. Subjects were recruited from both normal hearing and hearing impaired populations of 4 different languages/regions (Dutch, Flemish, German and Portuguese, as listed in Table 4). For each subject a speech intelligibility rating (SIR)

[175] was assessed and only subjects with SIR1 (completely intelligible in conversation) were included in the experiment. During an initial session (SES1), subjects were asked to select a text source from their daily readings, for example a book or an online newspaper. From that text a set of 300 short words (3 to 5 characters) was extracted by the system. These words were presented visually on a computer monitor and pronounced by the subject. In addition, each subject pronounced 300 words extracted from a conventional SAT test for his/her native language (NVA [176] [177] for Dutch/Flemish, Freiburger Einsilber [178] for German and Crianças Dissilabica [179] for Portuguese, as listed in Table 4). All utterances were recorded at a 16 kHz sampling rate in a quiet office room environment.

Language	*Subjects*	*Conventional SAT*
Dutch (Belgium)	206	NVA Flemish
Dutch (Netherlands)	43	NVA Dutch
German	101	Freiburger Einsilber
Portuguese	21	Crianças Dissilabica
Total	371	

Table 4: Number of subjects and conventional SATs used for each language/region.

After a minimum interval of 2 weeks subjects returned for a second session (SES2), during which subsets of the recordings of both the conventional SAT words and the subject's own 300 short words were presented acoustically, in 2 different listening conditions. A speech shaped noise source was adjusted in intensity to create listening conditions that were expected to result in both easy (> 50% correct) and difficult (< 50% correct) speech understanding for each particular subject. At each of the 2 listening conditions the subject was asked to repeat the words that were presented and an audiologist performed a phoneme score on the responses, resulting in 4 human scores (own words in difficult conditions, own words in easy conditions, conventional SAT words in difficult conditions and conventional SAT words in easy conditions). Each of those scores was the result of 24 presentations (20 presentations for German). At the same time, the system also calculated the DTW distance for each pair of recordings, resulting in the respective 4 machine scores, defined as the average DTW distance across that set of utterance pairs. Correlation coefficients were calculated between the human and the DTW session scores for all languages separately and for all languages together.

3.6.4. NORMATIVE DATA

Conventional SATs often come with normative data (normal curves). These data consist of the average speech discrimination scores of normal hearing listeners at different intensities. To obtain comparable normative data for the DTW based SAT, a new data set (distinct from the tuning data set) has been obtained as follows.

Ten normal hearing Dutch speaking listeners were asked to pronounce and record 2 sets of words. The first set (OWN) contained 300 short words extracted from their daily readings, the second set (SAT) contained 300 words extracted from a conventional SAT's word lists (Brugse Monosyllable CVC Speech Lists [180]). The recordings were obtained after visual presentation, following the same procedure for the initial session (SES1) as described above. After a minimum of 3 days, speech perception was assessed in these subjects, using both sets of stimuli. The OWN words were scored by DTW; the SAT words were scored both by DTW and by a professional audiologist (HUMAN). The stimuli used and the scoring performed are summarized in Table 5.

Test	*Words*	*Scoring*
OWN	Daily Readings	DTW
SAT	Brugse CVC	DTW & HUMAN

Table 5: Speech perception tests performed in 10 normal hearing subjects.

Stimuli were presented monaurally under headphones (TDH39) in a clinical sound treated room. The desired output level was obtained for each individual stimulus by endpointing and RMS-equalizing the signal. The initial presentation level of 40 dB SPL was increased by 5 dB until a score of 90% or higher was obtained and then decreased from 40 dB SPL in steps of 5 dB until a 0% score was obtained.

3.6.5. RESULTS

In our DTW calculation, the input feature vectors consist of the 12 MFCC coefficients, which are augmented by their first and second order temporal derivatives of the MFCC vectors. In combination with the log(E) coefficient, this amounts to 3 * (12 + 1) = 39 parameters. In Figure 94, the cumulative distributions of DTW scores are presented for the four test sets (SAME, MINIMAL, TYPICAL and MAXIMAL)

The performance of the DTW distance was measured in terms of equal error rate (EER) of each class when compared to the SAME class, as shown in Table 6.

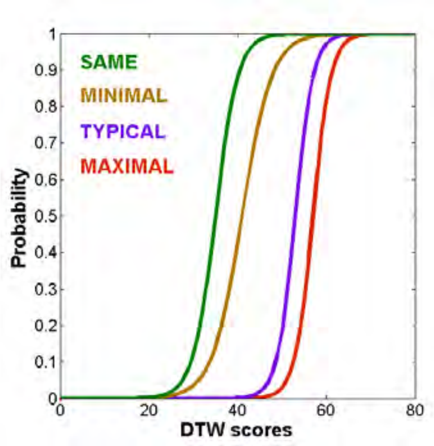


Figure 94: Cumulative distributions of DTW alignment scores for each of the classes (from left to right: SAME, MINIMAL, TYPICAL and MAXIMAL).

	SAME - TYPICAL	*SAME - MINIMAL*	*SAME - MAXIMAL*
EER	5.7	40.2	2.5

Table 6: The ability of the DTW distance to separate classes in terms of Equal Error Rate (in %) between the SAME class and the other 3 classes.

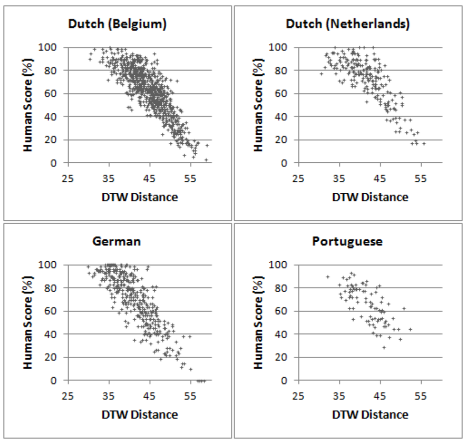


Figure 95: MFCC DTW distance vs. human score for each language/region (ref. Table 4). Each point represents the average of 24 (20 for German) stimulus-response pairs.

The average DTW distance, of a set of 24 (20 for German) presented words and their average phoneme score as judged by a professional audiologist are depicted in Figure 95. The scoring for Flemish (the Dutch dialect spoken in Belgium) was performed by 2 different audiologists, each of which produced half of the data set. Other languages/regions were scored by a single audiologist each. The correlations found between human scoring and DTW distance are listed in Table 7. When languages were pooled, an overall correlation of -0.83 was found.

Language	*Subjects*	*Correlation*
Dutch (Belgium)	206	-0.84
Dutch (Netherlands)	43	-0.79
German	101	-0.85
Portuguese	21	-0.70
All	371	-0.83

Table 7. Correlations between human phoneme scoring and MFCC DTW distance.

The scatter plots presented in Figure 95 show that the MFCC DTW distance is negatively correlated with the human score. Moreover, both measures take values in different ranges and the relation between them is non-linear. Therefore our transformation of the DTW distance into the human score involves two steps. First, we linearly transformed the MFCC DTW distance *d* into the DTW score *x* which is positively correlated with the human score and takes values in the range [0, 100]. Next, we assumed that a functional relation between the DTW score *x* and the human score *y* can be modeled with help of a generalized logistic function:

$$y(x) = \mathrm{A} + \frac{\mathrm{K}-\mathrm{A}}{(1+\mathrm{Q}e^{-\mathrm{B}(\mathrm{x}-\mathrm{M})})^{1/\mathrm{v}}} \qquad (1)$$

where the lower asymptote (*A*) is set to 0, the upper asymptote (*K*) is set to 100, *v* is set to 1. Additionally, to compensate for a bias in the data (over- and under-representation of high and low scores, resp.) we added to the data set 1000 virtual points with DTW score 25 and human score 0 and 100 points with DTW score 75 and human score 100. Then values of the remaining parameters (*B*, *Q* and *M*) were determined using the Nelder-Mead simplex algorithm [181], to minimize the Root Mean Square Error (RMSE) over the data set specified. The optimal parameters for the model given by formula (1) are: *B* = 0.1034; *Q* = 1.2721; *M* = 44.4934. Figure 96 shows the optimal fit using this model plotted on top of the individual data points.

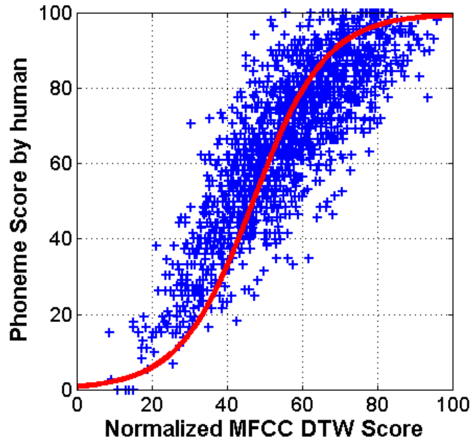


Figure 96: Model for mapping MFCC DTW scores to a 0 to 100% scale equivalent to human phoneme scoring.

Normative data obtained in 10 hearing listeners are depicted in Figure 97 as the median scores at the presented intensities. The median 50% speech reception threshold (SRT) for both the OWN words and the SAT words when scored by DTW was shown to be 21 dB SPL. When scored by a human the median SRT for the SAT words showed to be 19.5 dB SPL. The published normative data for the Brugse SAT specifies a 50% SRT of 20 dB SPL for monaural normal hearing listeners. It is remarkable to see that, even at higher stimulus levels, median DTW scores do not reach more than 85%.

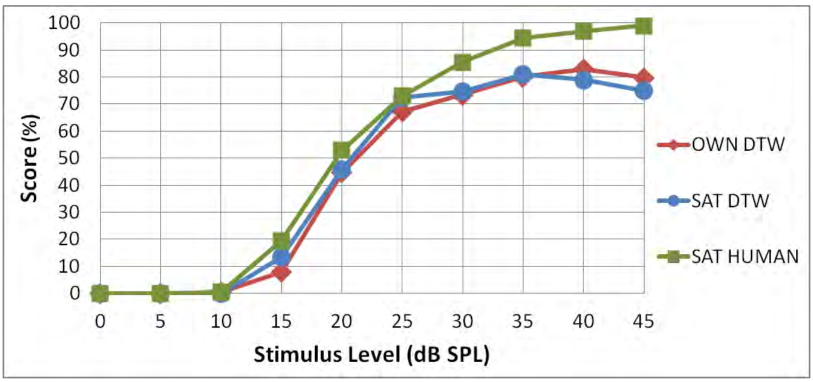


Figure 97: Normative data obtained in 10 normal hearing listeners. The figure shows the median scores at 5dB intervals between 0 and 45 dB SPL for the subject's own words (OWN) scored by DTW and for conventional SAT words scored by both DTW and human phoneme scoring (ref. Table 5).

3.6.6. DISCUSSION

The strong correlations found between human and DTW scores indicate that an automated scoring mechanism may be suitable to assess speech perception deficits. The correlation in the Dutch (Netherlands) is slightly lower, and the correlation in the Portuguese is markedly lower, than in the other languages/regions. The authors have no explanation for this observation. It may be attributed to the audiologist's day to day variance in phoneme error judgment. Another reason could be the variance in background noise when the initial recordings have been obtained in a different location than where the actual test session took place.

When comparing the median 50% SRTs in normal hearing listeners, it is clear that the use of short words extracted from the subject's daily reading is equivalent to using conventional SAT words, like the Brugse CVC words. Not only the SRTs resulting from the use of both types of stimuli are the same, but also below and above the 50% score point, results from both tests are very similar.

A clinically irrelevant 1.5 dB upward shift is observed in the median SRT when human scoring is replaced by DTW scores. This makes us believe that the results obtained by DTW scoring may be comparable to conventional speech audiometric test results and therefore clinically usable. The ceiling effect observed around 80% when using DTW scoring, may be attributed to the fact that in obtaining these normative data, the recordings took place in an office room, while the actual speech perception was assessed in a clinical sound treated room. The difference in ambient noise between the two rooms may have introduced a floor effect in the DTW distance, which presents itself as a ceiling effect in the speech perception scores, however further investigation is needed to confirm this hypothesis.

3.6.7. ACKNOWLEDGEMENTS

The reported study was conducted within a Seventh Framework Programme Project FP7 OPTI-FOX 262266, supported by the European Union.

B. Vaerenberg received a PhD grant for this work from the IWT (Agency for Innovation by Science and Technology, Baekeland-mandaat IWT090287).

3.7. COMBINED ELECTRIC AND ACOUSTIC HEARING PERFORMANCE

Combined electric and acoustic hearing performance with Zebra® speech processor: speech reception, place and temporal coding evaluation

Cochlear Implants Int. 2012 Nov 15.

B Vaerenberg[1,2], V Péan[3], G Lesbros[3], G De Ceulaer[1], K Schauwers[1], K Daemers[1], D Gnansia[3], PJ Govaerts[1]

[1] The Eargroup, Antwerp-Deurne, Belgium
[2] Laboratory of Biomedical Physics, University of Antwerp, Belgium
[3] MXM – Neurelec, Vallauris, France

Abstract

The objective of this study is to assess the auditory performance of Digisonic cochlear implant users with electric stimulation (ES) and electro-acoustic stimulation (EAS) with special attention to the processing of low frequency temporal fine structure. Six patients implanted with a Digisonic® SP implant and showing low-frequency residual hearing were fitted with the Zebra® speech processor providing both electric and acoustic stimulation. Assessment consisted of monosyllabic speech identification tests in quiet and in noise at different presentation levels, and a pitch discrimination task using harmonic and disharmonic intonating complex sounds [182]. These tests investigate place and time coding through pitch discrimination. All tasks were performed with ES only and with EAS. Speech results in noise showed significant improvement with EAS when compared to ES. Whereas EAS did not yield better results in the harmonic intonation test, the improvements in the disharmonic intonation test were remarkable, suggesting better coding of pitch cues requiring phase locking. These results suggest that patients with residual hearing in the low-frequency range still have good phase locking capacities, allowing them to process fine temporal information. ES relies mainly on place coding but provides poor low frequency temporal coding, whereas EAS also provides temporal coding in the low frequency range. Patients with residual phase locking capacities can make use of these cues.

3.7.1. INTRODUCTION

Whereas cochlear implants (CI) may provide good speech understanding in quiet in persons with severe and profound hearing loss, speech understanding in background noise and music listening still remain a challenge for most CI users. This is believed to be at least in part attributable to the current CI's limited ability to encode pitch [183] [184]. This relates to both impaired frequency selectivity (see Moore, 2007 [45] for a review) and impaired perception of temporal fine structure (TFS) cues (the rapid oscillations with a rate close to the centre frequency of the band [185] [186]

[155]; see [151] for a review. The frequency selectivity required for speech perception in noise is finer than for speech understanding in quiet [187]. Spectral selectivity is tonotopically coded in the cochlea. Most implant users however distinguish less than 10 channels of distinct ‘‘place–frequency’’ information across the entire spectral range [97], which does not suffice for good speech understanding in background noise. TFS cues are especially important for speech understanding in fluctuating noise and listening in noise valleys [188] [155]. In a normal-hearing auditory system, this fast temporal information is mainly coded by phase-locking mechanisms within an auditory channel. TFS cues however are not successfully transmitted by current CI processors. Even if CI algorithms would improve in temporal pitch coding, it remains questionable whether the CI users would benefit from it, since it has been shown that TFS coding by means of electrical stimulation has an upper limit of 300 Hz [189].

Recent improvements in CIs and soft surgical procedure now allow some preservation of residual acoustic hearing in the low-frequency range [190]. This is often realized by reducing the electrode array insertion depth, either by a partial insertion [191] [192] [193], or by using dedicated low-traumatic electrode arrays [194] [195]. In these patients, usable acoustic hearing is typically preserved up to frequencies of 500 to 1000 Hz. This allows acoustical stimulation in the low frequencies while the mid and high frequencies are stimulated electrically by means of the implant. Thus, these patients perceive sound via a combined electro-acoustic stimulation (EAS). It is reported that this combined stimulation improves the subjective sound quality and also the speech recognition in background noise [192] as well as pitch perception [196].

The assessment of speech understanding in quiet and in noise is common clinical practice, but assessing the coding of TFS cues, related to pitch perception and the underlying phase locking mechanism, is not. A§E (Auditory Speech Sound Evaluation [197] [107] [157]) is an audiological test suite that includes harmonic intonation (HI) and disharmonic intonation (DI) tests [182] for the assessment of TFS coding.

The goal of the present study is to evaluate the speech understanding in quiet and in noise, and the pitch perception using the HI/DI tests of A§E in six patients implanted with a Digisonic® SP device (Neurelec, Valauris, France) and a Zebra® speech processor providing EAS.

3.7.2. MATERIALS AND METHODS

3.7.2.1. PATIENTS

Six Digisonic® SP (Neurelec, France) users with preserved residual low-frequency hearing after implantation were identified (Table 8). The median age at implantation was 51 years (range 9-81 y). Until the moment of implant surgery, the subjects used different kinds of high-powered hearing aids that were adequately fitted and maintained. All subjects had residual low frequency hearing prior to implantation and this was at least in part and unintentionally preserved after implantation with full insertion of the electrodes. Figure 98 shows the pure-tone thresholds before and after surgery. Post surgical measures were performed on the EAS testing session day.

Table 8: Individual clinical history.

Patient #	Age at implantation (years)	Duration of cochlear implant use (years)	Etiology
1	71	1	Unknown
2	43	2	Progressive
3	81	2	Unknown
4	55	3	Congenital
5	9	5	Congenital
6	45	1	Progressive

The subjects initially received electrical stimulation (ES) only with the Digi SP processor which was programmed according to routine techniques, covering frequencies from 195 to 8008 Hz. Table 8 indicates ES use duration for each patient.

As soon as Neurelec was able to provide EAS by means of the new Zebra® processor, the subjects received additional acoustical amplification through an acoustic receiver in an individual ear mould. The median period of “ES only” was 2 years (range 2-4 y). Included patients had no experience of EAS stimulation prior to testing. Written informed consent was obtained from all patients.

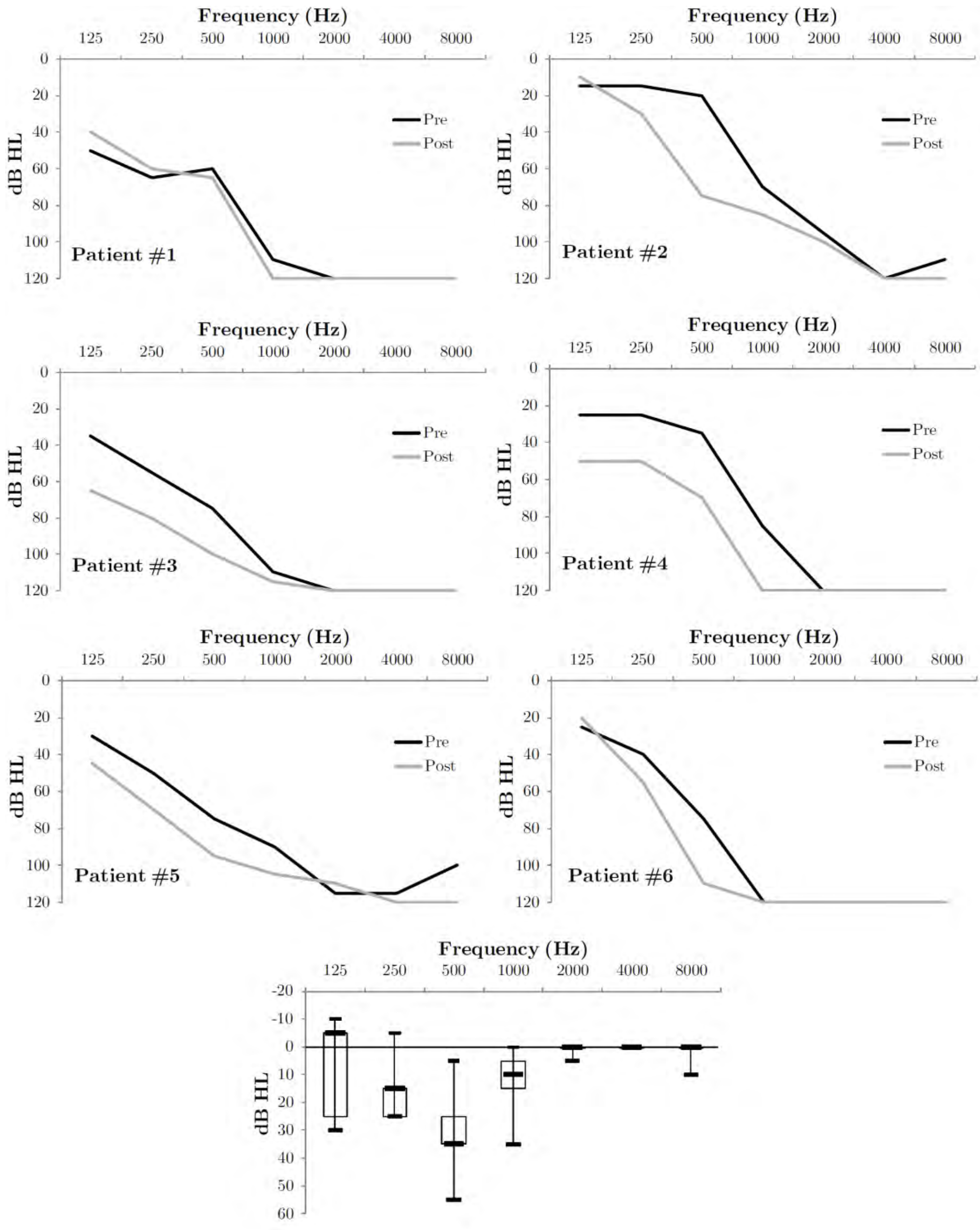


Figure 98: Individual pure tone thresholds before (Pre) and after surgery (Post). Pre- and post operative hearing thresholds differences are indicated in the bottom figure box plots showing median values, quartiles and ranges.

3.7.2.2. DEVICE

The Zebra® processor is a CI speech processor, integrating an acoustic output (see Figure 99). It is compatible with the Neurelec Digisonic® SP CI. Its shape is the same as the standard Digi SP and Saphyr® SP speech processors. The electrical signal is transmitted through a coil and magnet, as for all CIs, and the acoustical signal through a Sonion (Roskilde, Denmark) canal-receiver. The computation of both the electric and acoustic signals is performed in the same chipset: the incoming signal is analyzed in several frequency bands (Fast Fourier Transform analysis generating 64 frequency bands, linearly spaced between 190 and 8200 Hz), and all these bands are routed to the input of the electrical and acoustical processing software. This guarantees equality and synchronization of the acoustical and electrical input samples.

Figure 99: The Neurelec Zebra® speech processor.

The coding strategy for electrical stimulation is MPIS (Main Peaks Interleaved Sampling [198] [199]) as used in the classical Digisonic® SP implant with the Digi SP processor. This strategy is based on spectral multi-peak extraction, and interleaved stimulation. The number of transmitted peaks is a parameter that may be modified (default setting: 10 transmitted peaks out of 20 extracted peaks). Loudness coding is realized by varying pulse duration, and pulse amplitude remains constant over time (amplitude set at fitting). The stimulation rate may be set between 260 and 1000 pps per electrode. The default factory setting is 600 pps per electrode.

For the acoustical stimulation, all 64 frequency bands from 190 to 8200 Hz are processed, and gains from 0 to 42 dB can be applied for each band separately with a single band compression of which the parameters (compression rate, attack time and release time) are set in the fitting software.

3.7.2.3. FITTING

For the fitting of the acoustical gain, the half-gain rule was applied to determine the necessary amount of acoustic gain, in order to obtain aided thresholds of about 30 dB HL. However, because of the acoustic power limit of the speech processor and also to avoid distortions due to overamplification, the acoustical amplification applied was 30 dB at all frequencies between 195 and 8008 Hz, except for subject S2, who received 30 dB amplification from 195 to 719 Hz, tapering down to nil between 716 and 1572 Hz.

3.7.2.4. OUTCOME MEASURES

All six patients underwent audiological testing, both in the ES and the EAS stimulation mode. Pure tone thresholds were performed pre- and postoperatively. This was carried out in a sound treated audiometric room using a Madsen Aurical system (GN Otometrics) with free-field loudspeaker outputs calibrated to dB Hearing Level. The loudspeaker was positioned at 0^{o} azimuth, 1m from the subject's head. Thresholds to warble tones at octave frequencies between 125 and 8000 Hz were recorded using standard clinical audiometric methods.

Speech audiometry in quiet was performed with open set monosyllabic CVC-words (NVA-lists [200]), presented at 40, 55, 70 and 85 dB SPL, using the same room and equipment as above. Two lists of 12 words were used at each intensity level and phoneme scores were recorded. For speech audiometry in noise, open set monosyllabic CVC-words (Brugse-lists [180]) were presented at 10, 5, 0 and -5 dB signal-to-noise ratio (SNR) with speech-shaped noise at 65 dB SPL. One list of 17 words was used at each SNR and phoneme scores were recorded.

The coding of low frequency TFS was assessed using the HI and DI tests of A§E [107]. The details of these tests and normative data are described elsewhere [182]. Briefly, both tests use low frequency harmonic complexes to find the just noticeable difference (JND, also called difference limen or threshold) for pitch discrimination in individual subjects. In each trial of both the HI and DI test, 2 stimuli are presented consecutively, one of which contains an intonation, while the other one does not. The test is a same-different discrimination task. The non-intonating stimulus is a harmonic complex signal having a fundamental frequency (F0) of 200 Hz and 3 higher harmonics (with frequencies of 2F0, 3F0 and 4F0). The intensities of the harmonics decrease in comparison with F0 (-6 dB at 2F0, -12 dB at 3F0, -18 dB at 4F0). Both in HI and DI tests, this non-intonating sound is presented in contrast to an intonating sound. The intonating sounds used in the HI test feature a frequency sweep of all harmonics (including F0) from NF0 to N(F0+ΔF), with N=2, 3 and 4 respectively. In the DI test however, the intonating sounds feature a sweep of the fundamental frequency only (F0 to F0+ΔF), whereas the higher harmonics are kept fixed at their initial frequency. As a consequence the harmonic separation of partial tones is distorted by the sweep, hence a disharmonic (or dissonant) intonation. A JND is sought using an adaptive staircase procedure [156]. In the current study HI and DI tests were performed using an audio cable connected to the auxiliary input of the processor to deliver the stimuli directly to the implant.

3.7.2.5. DATA ANALYSIS AND STATISTICAL METHODS

Because of the limited number of included patients, non-parametric statistics were used for all variables. Box-and-Whisker plots are used for graphical representation. Wilcoxon tests for paired samples were conducted to compare the audiological results obtained with EAS to those obtained with ES. The cut-off level for statistical significance was set at 0.05.

3.7.3. RESULTS

3.7.3.1. SPEECH PERCEPTION IN QUIET

The phoneme scores for speech in quiet are presented in Figure 100. Gains between the two stimulation conditions in terms of intelligibility for each patient are also presented. In ES mode, the median phoneme scores ranged from 20 to 55% for presentation levels between 40 and 85 dB SPL. In EAS mode, patients showed correct identification of 27 to 63% of phonemes for the same presentation levels. Gains within patients between the two stimulation modes are shown in Figure 3b with median values ranging from 8% to 16%. None of the differences were statistically significant.

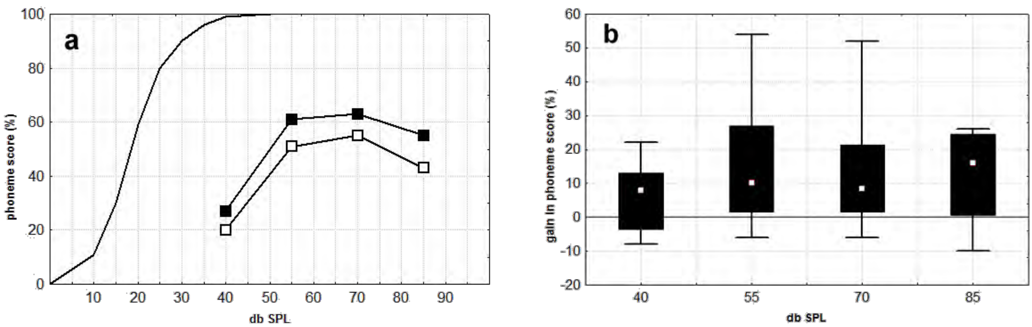


Figure 100: Monosyllabic speech audiometry in quiet. (a) Median phoneme scores at different presentation levels, in electric-only (E, white squares) and electro-acoustic (EA, black squares) conditions. The solid curve depicts normal-hearing performance. (b) Box and whisker plots showing the intra-individual differences between electric-only and electro-acoustic conditions (EA – E). Central point: median; Box: lower to upper quartile; Whiskers: lower to upper range.

3.7.3.2. SPEECH PERCEPTION IN NOISE

Results for speech perception in noise in ES and EAS modes are shown in Figure 101. Patients had median scores between 27 and 51% in ES mode, and between 27 and 59% in EAS, for SNRs between -5 and 10 dB. Median gains between the two stimulation modes were about 10% for all SNRs. Significant differences between the two stimulation modes were found at 10 dB SNR ($p<0.05$).

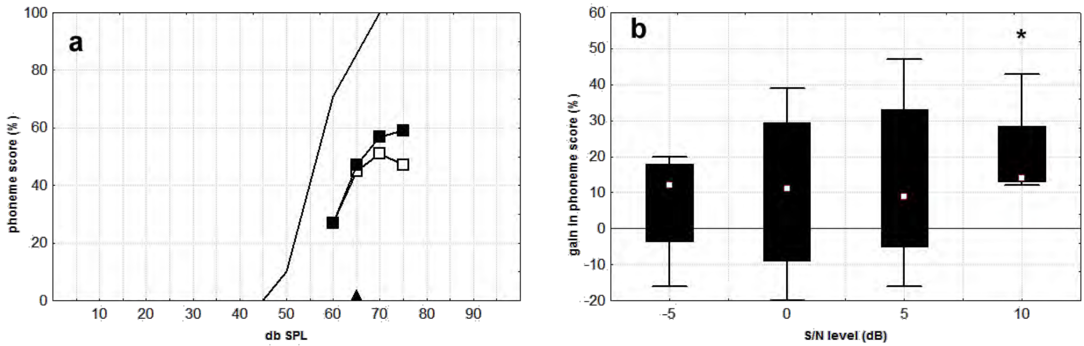


Figure 101: Monosyllabic speech audiometry in noise (65 dB SPL noise). (a) Median phoneme scores at different presentation levels, in electric-only (E, white squares) and electro-acoustic (EA, black squares) conditions. The solid curve depicts normal-hearing performance. (b) Box and whisker plots showing the intra-individual differences between electric-only and electro-acoustic conditions (EA – E). Asterisk: significant with p < 0.05.

3.7.3.3. HI AND DI TESTS

JNDs from HI and DI tests in ES and in EAS conditions are shown in Figure 102; standard scores obtained on hearing listeners are also given in the same plot. Gains between ES and EAS are also presented. For HI tests, JNDs measured were similar in ES and EAS, with the median value around 7 Hz. For DI tests, median values for JNDs were 44 Hz in electric-only mode and 12 Hz in electro-acoustic mode. Median values for gains between the two listening conditions were 0 Hz for the HI test and 24 Hz for the DI test. Comparing the two listening conditions, statistical analyses revealed that JNDs for the HI and DI tests were not significantly different (HI: p=.42; DI: p=.08).

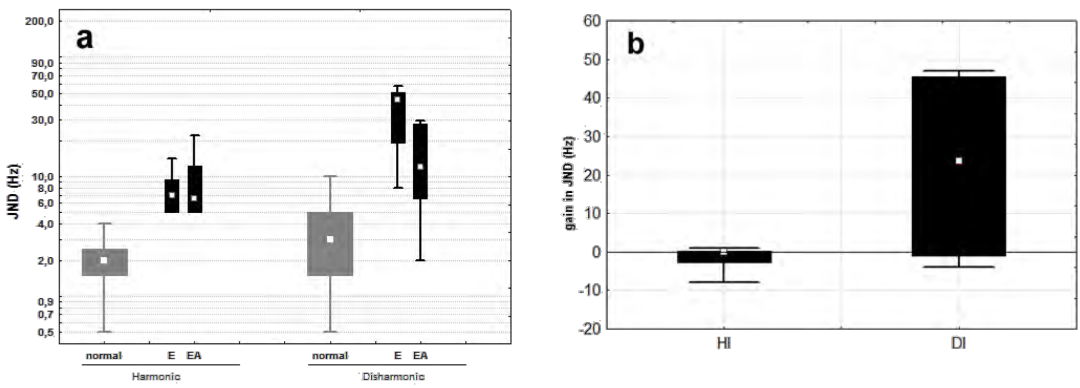


Figure 102: A§E pitch discrimination test results. (a) Box and whisker plots showing the JNDs (Hz) on harmonic and disharmonic tests, for hearing listeners and implanted subjects in electric-only (E) and electro-acoustic (EA) conditions. (b) Box and whisker plots showing the intra-individual differences between electric-only and electro-acoustic conditions (EA – E).

3.7.4. DISCUSSION

3.7.4.1. HEARING PRESERVATION

Cochlear implantation for patients with residual hearing has never been evaluated with the Digisonic® SP implant. The current study has investigated results in Digisonic CI users in whom the hearing was preserved unintentionally. Several studies have investigated hearing preservation using Med-El CIs (C40+ with flex EAS or Medium electrode) with 'long' electrode arrays and a soft surgery approach. For example Gstöettner, *et al.* [201] reported hearing to be preserved in 12 out of 18 patients with average threshold deteriorations ranging from 10 to 30 dB HL. In [202], it ranged from 10 to 25 dB HL in 8 out of 9 patients. Kiefer *et al.* [203] reported that at least partial preservation of hearing was accomplished in 11 out of 13 patients, and the mean threshold change for those 11 patients was approximately 15 dB at the lower frequencies, while the remaining two patients suffered essentially total losses. James *et al.* [204] reported a 25 dB loss in the lower frequencies for 12 patients implanted with a long electrode, including the data for two patients who suffered total losses. With the Nucleus CI with Hybrid L electrode-array, Lenarz *et al.* [190] reported median losses ranging from 10 to 15 dB HL in the low frequency range.

The current results suggest that hearing can be preserved with the standard Digisonic® SP implant with mean hearing loss induced by surgery ranging from 10 to 30 dB HL in the lower frequencies. This degree of hearing preservation seemed comparable to other devices. As said, this hearing preservation was unintentional and only occurred in a minority of cases.

3.7.4.2. SPEECH PERFORMANCE IN QUIET AND IN NOISE

The results of this study indicate an advantage of combined EAS compared to ES for speech understanding in quiet and in noise, which was statistically significant at 10 dB SNR. These results are consistent with those reported by others. For example, Kiefer *et al.* [203] performed speech recognition tests with monosyllabic words in quiet at 70 dB SPL in patients implanted with other EAS implants (Med-El, Combi 40/40+ and TEMPO+ processor). They reported a mean score of 54% with ES and 62% with EAS (compared to 50% and 64% in the present study using Digisonic SP implant and Zebra processor). Consistent with the present results, their differences were not statistically significant. Using another device from the same manufacturer (Med-El PulsarCI100) in similar test conditions, Prentiss *et al.* [205] found 38% correct identification in quiet for monosyllabic word identification with ES, and 47% EAS, with no statistical difference between these two conditions.

None of the previous studies on EAS performed monosyllabic word or phoneme recognition in noise in CI users as in the present study. However, speech identification in noise using sentences was always found to be significantly better with EAS, for SNRs at +5 or +10 dB [165] [206] [202] [204] [205]. In the present study, identification of monosyllabic words in noise was performed and a significant difference between EAS and ES was observed for 10 dB SNR, which is consistent with the other studies.

3.7.4.3. HI AND DI TESTS RESULTS

There was no statistically significant difference between ES and EAS on the HI test. Both with EAS and with ES only, the performance was poorer compared to hearing subjects and consistent with larger data sets on CI users with different devices (Eargroup, unpublished results). Nevertheless, the HI results are still fairly good, demonstrating reasonable pitch discrimination abilities in CI users when high frequency cues are available in the complex signal.

Table 9: Individual HI and DI test results (JND in Hz).

Patient#	Electric-only		Electro-acoustic	
	HI	DI	HI	DI
1	8	57	9	10
2	7	41	7	14
3	7	23	6	27
4	5	47	5	2
5	5	8	5	8
6	14	49	22	29

One could argue that both temporal envelope and temporal fine structure cues may have contributed to the pitch perception in some parts of the tests. However, we believe this to be very unlikely. Both HI and DI signals were constructed to feature temporal envelopes that are stable in time, except for the 30ms linear fade-in and fade-out, but these are identical for all stimuli. Stimulus envelopes by means of Hilbert transforms did not suggest any available cues in the envelope. There is some variation in the envelope due to the added white noise in the stimuli, but this is totally random and should therefore be unusable as a cue. It should be noted that temporal envelope cues may result from beating when the fundamental frequency in the DI stimulus approaches the 400Hz harmonic. This effect has been described in Vaerenberg et al. (2011) and comes into play only at rather high delta f (> 130 Hz (70 Hz beating) for normal hearing subjects and > 60Hz in CI users (140Hz beating) [159]). In the current study, the improved pitch perception is unlikely to be attributed to this phenomenon because all subjects obtain JNDs smaller than 60 Hz in all conditions and even below 30 Hz in the EAS condition.

Also loudness cues are unlikely to have played a role in the patient's abilities to discriminate the sounds, even when taking into consideration the fairly steep slopes in some unaided audiograms. The pitch perception was always assessed in the aided condition (either ES or EAS). Audiograms were recorded from all subjects in 3 aided conditions: ES only, AS only and EAS. In ES and EAS modes the audiograms were flat (+/- 10dB) over all frequencies for all listeners. If subjects were able to use loudness cues to obtain better results in EAS mode then they would originate from the acoustic stimulation, because electric maps were identical in both ES and EAS conditions. However, for the DI test, the relevant frequency range to explain the observed JNDs is 200 Hz to 250 Hz and audiograms obtained with AS only showed an average absolute difference between the thresholds at 125Hz and at 250Hz of 7.5 dB. The maximum difference between these thresholds, observed in S5, was 20 dB. S5 however showed no improvement by adding acoustic stimulation. For the other subjects it is also hard to imagine that they could have extracted a loudness cue from a sweep of around 10 Hz of the fundamental frequency, because the difference in absolute thresholds in the range of this sweep is likely to be less than a decibel. Therefore we believe it is reasonable to assume that all frequencies in the stimuli caused an equivalent loudness percept and the subject used pitch as a cue rather than loudness.

It might be interesting to consider the possibility that frequencies moved between electrodes as the stimulus changed. All maps used during the study featured a linear spacing of frequencies in the range 195 Hz to 977 Hz over the six most apical electrodes, yielding a band width of 130 Hz per electrode and an upper cutoff frequency of 326 Hz for the most apical channel. In the DI test only ΔF > 125 Hz should cause activation of the second electrode. All subjects have considerably lower JNDs, which makes us believe that place coded pitch by electrical stimulation is unlikely. However, when considering the possibility of spectral leakage by the FFT into adjacent channels, it could well be that some subjects were able to extract cues from the subtle increase/decrease of cross-channel leakage when frequencies shift up/down. This might explain why some subjects obtain JNDs as low as 8 Hz on the DI test using ES only.

In the HI test the 800 Hz harmonic (4F0) would move to the next channel for a ΔF as low as 12 Hz, resulting in a possible place cue for JNDs recorded above this ΔF. As ΔF becomes larger, more harmonics move to a next channel (2F0 at 28 Hz, 3F0 at 39 Hz, 4F0 again at 45 Hz, etc.). It is evident that the larger ΔF results in the more salient place pitch cue in the HI test.

The signal processing by the Zebra processor uses an 8 ms window as input for its FFT. When the MPIS strategy maps the amplitude spectrum to electrode activation the phase spectrum is lost. Therefore we assume that it is unlikely to have a temporal pitch cue within one channel. But as for the spectral leakage that may have caused subtle place cues, we cannot entirely exclude that temporal cues may have originated from small fluctuations in the channel's current level that result from artefacts of the signal processing in response to the shifting frequency (e.g. the segmentation of the signal in frames may cause a temporal modulation on the current level if the frame length is not aligned with the input signal's periodicity and the effect of the applied window is not able to compensate for this). This phenomenon may also have contributed to some subjects' small JNDs observed in the DI test with ES only.

One may disagree with the above reasoning and argue that these small JND's obtained with ES only indicate the DI test itself to be invalid. However, although we acknowledge the theoretical grounds on which such doubts are based, we believe they are unlikely to explain the above mentioned observations for reasons given before in this and previous papers.

As for the acoustical part of the stimulation, we find it very reasonable to believe that at moderate ΔFs (< 30 Hz), the subjects would have trouble extracting a place pitch cue, especially when considering their hearing losses and the resulting broadening of auditory filters. It seems more reasonable to us to attribute the gain resulting from adding acoustical stimulation, to a temporal pitch cue. For the acoustical processing, the Zebra uses the phase spectrum in its inverse FFT, after applying gains to the amplitude spectrum, such that TFS is restored in the acoustical output of the system, allowing for temporal pitch cues (and thus phase locking) in the processing of the acoustical signal by the subject's the auditory system.

We believe it is an important finding that EAS improved the DI test results substantially in patients #1, 2, 4 and 6 (Table 9). The group results were not significant though, which may be explained by the fact that patients #3 and 5 already showed good results with ES, which could not be improved by EAS. It is remarkable that DI results in all EAS users in this study are within or near the results obtained in hearing subjects. It is assumed that the DI test assesses the patient’s phase locking capacity, whereas HI may benefit from both phase locking and place coding [182]. This is because the DI test only provides low frequency TFS cues whereas the HI test provides both low and high frequency cues. Taken together, these results suggest that electric and acoustic stimulation may provide complimentary information. Certainly ES is used for place coding, yielding fairly good results (but still poorer than in hearing subjects). Acoustic stimulation may provide low-frequency TFS that will be processed by the remaining phase locking capacities in case of residual low frequency hearing. The present study tends to confirm the effect first discussed in the study by Turner *et al.* [192], in which the authors found release from masking between steady and fluctuating noise for EAS when compared to ES, which they interpreted as suggestive for better TFS processing with EAS.

3.8. PITCH PERCEPTION & SPEECH IN NOISE

Pitch perception & speech in noise with Med-El FS4 strategy

The 9th International Conference on Cholesteatoma and Ear Surgery, Nagasaki, Japan, 2012

Paul J Govaerts[1,2]; Bart Vaerenberg[1,2]; Geert De Ceulaer[1], Kristin Daemers[1], Karen Schauwers[1], Martine Coene[1,2,3]

[1] The Eargroup, Antwerp-Deurne, Belgium
[2] Laboratory of Biomedical Physics, University of Antwerp, Belgium
[3]Free University of Amsterdam

Abstract

Current cochlear implants have difficulties conveying low frequency temporal fine structure (TFS). TFS is important for pitch perception, spatial separation of multiple voices and speech understanding in noise. The MED-EL Maestro Cochlear Implant comes with the FS4 speech coding strategy aiming at an improved TFS coding at the apical electrodes. A§E harmonic intonation (HI) and disharmonic intonation (DI) tests are clinical psycho-acoustical tests to assess the coding of TFS. Five adult MED-EL CI users underwent A§E HI/DI and speech audiometry in temporally modulated noise with FSP and FS4 speech algorithms. There was a clear correlation between A§E DI and speech audiometric results with 3 CI users improving on both tests when shifting from FSP to FS4, and 2 users not showing any change. These results indicate that FS4 is capable of improving the coding of TFS in a number of CI-users.

3.8.1. INTRODUCTION

In daily life the relevant cues for voicing, melody, intonation and other musically and linguistically important percepts are conveyed by relatively low frequency pitch, relating mainly to the fundamental frequency (F0) [184] [183]. The way the cochlea codes spectral content of sound can be explained by two underlying mechanisms, place-coding and phase-locking. Both are complementary and overlapping. It is believed that for low frequency signals such as the fundamental frequencies of human voices, phase locking of the temporal pattern of nerve responses to the temporal fine structure of the signal is the more dominant cue for conveying pitch [207]. Poor low frequency pitch perception may play a role in a number of frequently encountered complaints by current CI-users, like poor music appreciation or poor spatial separation of multiple speakers [155] [153].

Many of the current CI strategies however, are designed to replace the tonotopical organization of the cochlea. They attempt at place coding the frequency spectrum of the input signal. But the limited number of electrodes and their relatively wide current spread result in a limited number of physical sites that can be targeted for stimulation. These limitations cause today's CI systems to have difficulties in conveying the fine structure of the signal. An attempt to improve over purely tonotopical (i.e., envelope based) strategies is provided by MED-EL's FSP strategy family (Figure 34). The basic idea behind these strategies is that they use the most apical electrodes for time coding, in order to improve TFS perception. They adapt the rate of stimulation in these channels to changes in the temporal fine structure of the input signal. This is achieved by so called Channel-Specific Sampling Sequences (CSSS), which are a series of pulses that are triggered by zero-crossings in the bandpass filter's output and that aim at transmitting temporal fine structure cues, such as fluctuations of the fundamental frequency of a signal.

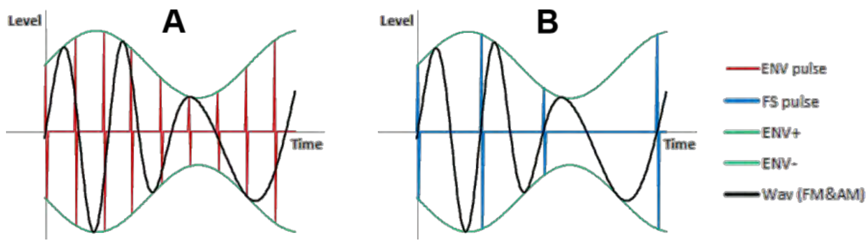


Figure 103: Illustration of MED-EL's FSP strategy (B) compared to a purely envelope based strategy (A). The temporal envelope (green lines) is used to modulate the amplitude of pulses (red and blue lines). In the FSP strategy, in addition to using the envelope, the temporal fine structure is used to determine the stimulation rate in a subset of apical channels. Pulses for these channels are triggered by zero-crossings in the bandpass filter's output (black lines). [adapted from MED-EL G.m.b.H.]

With its new FS4 strategy MED-EL aims at further improving on their existing FSP strategy. FSP uses 1 or 2 of these CSSS channels. FS4 uses 4 of them and stimulates at a higher rate, improving the accuracy of time-coding the zero crossings in the TFS of the input signal.

3.8.2. MATERIAL & METHODS

Seven adult MED-EL CI users participated after informed consent. A§E HI/DI was performed at 70 dB SPL presentation level. Details of these tests are explained in detail in “Clinical assessment of pitch perception”. The participants also underwent tests for speech understanding in temporally modulated noise (speech shaped noise with 8 Hz modulation frequency, 80 dB depth, 65 dB SPL presentation level). Monosyllabic Dutch words were presented at different presentation levels and phoneme scores were recorded. All tests were done with both the MED-EL FSP and FS4 algorithms.

3.8.3. RESULTS

Using FSP a median JND of 35 Hz was found on the harmonic intonation test and practically none of the subjects were able to perform the disharmonic intonation. For the FS4 strategy there may be a small improvement in the harmonic intonation, but the remarkable difference is observed in the disharmonic intonation. Although more than half of the subjects is still not able to perform the task, there are a number of individuals who obtain reasonable or even normal JND's. This indicates that their phase locking abilities have improved when switching to FS4.

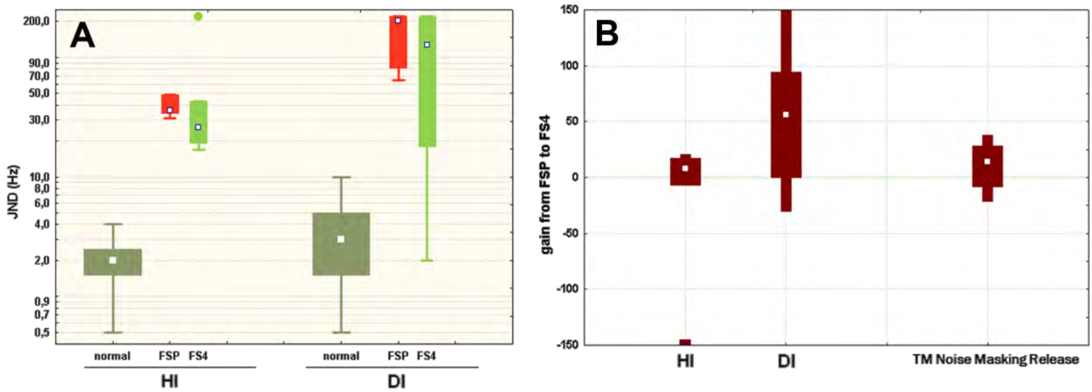


Figure 104: A: box and Whisker plots show the results (JND) on HI/DI tests in hearing controls (dark green) and in the 7 CI users using the FSP strategy (red) and FS4 strategy (green). B: gain in HI and DI JNDs and in the speech perception scores at a signal to noise ratio of 0 dB (i.e., masking release of the temporally modulated noise) when switching from FSP to FS4.

The median gain when switching from FSP to FS4 was found to be 8 Hz on the HI test, which can hardly be considered an improvement. On the DI test however, almost all subjects show an improvement, with a median gain of 56 Hz. This is a clear indication of the improved temporal processing of the FS4 strategy. The speech perception results in temporally modulated noise show a similar effect. At a signal to noise ratio of 0 dB, the median gain in speech understanding was 14% and some subjects even showed 40% improvement when switching to FS4 strategy.

3.8.4. CONCLUSION

The perception of low frequency pitch is closely related to temporal fine structure and therefore related to phase locking abilities. We have found pitch perception and the ability of phase locking to be clinically relevant. Using the A§E intonation tests this can easily be tested in clinical settings. A correlation was found between the results on A§E disharmonic intonation and release of masking in speech audiometry, which is far more difficult and time consuming to test. The results indicate that the choice of a speech coding strategy may have an impact on the coding of TFS, where the FS4 strategy improves the perception of fine structure in a number of CI users.

3.9. AN AUDIOMETRIC TEST BOX FOR HEARING ASSESSMENT IN CI RECIPIENTS

An audiometric test box for hearing assessment in CI recipients

European Symposium on Paediatric Cochlear Implantation 2013, Istanbul, Turkey, 2013

B Vaerenberg[1,2], H Van Oosten[2,3],
A Vreeswijk[3], E Schipperen[3], G De Ceulaer[1], PJ Govaerts[1]

[1] The Eargroup, Antwerp-Deurne, Belgium
[2] Laboratory of Biomedical Physics, University of Antwerp, Belgium
[3] Akoestische Bouw Projecten BV, Geertruidenberg, The Netherlands

Abstract

This report describes the research and development of Otocube and A§E 2012. This combination of hardware and software provides well insulated and calibrated test conditions for psychoacoustical testing in CI recipients. A desktop test room has been developed which provides acoustic insulation, calibrated liminal and supraliminal sound presentation, a flat frequency response and real-time monitoring of the sound delivered in the box. The technical specifications of Otocube are IEC 60645 compatible. A§E 2012 provides psychoacoustical tests for the assessment of pure tone detection thresholds, speech understanding, spectral discrimination and temporal resolution. Otocube with A§E2012 is a portable desktop test box to replace a fully equipped audiological test booth. It allows extensive outcome measuring with substantial reduction in resources and may be an intermediate step towards remote fitting.

3.9.1. INTRODUCTION

Typical clinical sound treated rooms are relatively expensive in terms of construction, required equipment (an audiometer with transducers), maintenance (regular calibration) and size (they occupy space in the clinic that could be put to other use). Yet they are essential in an outcome based fitting approach to ensure reliable measuring conditions. In many clinics sound treated rooms are over-occupied and the repeated testing for fitting CI recipients would put strain on the availability of these rooms even more. To address these issues, the idea of a portable desktop test box to facilitate psychoacoustical measures in CI recipients has been conceived. This section reports on the development and specifications of such a device.

3.9.2. OTOCUBE

In the context of a European research project (FP7-SME 262266 OPTI-FOX) and together with engineers from Akoestische Bouw Projecten BV, Geertruidenberg, The Netherlands, an acoustically insulated test chamber, called Otocube, was developed. Otocube is specifically designed for conduction psychoacoustical experiments in CI recipients. The basic idea is to position the recipient's speech processor at a fixed disctance from a loudspeaker built into the test room, and to connect the speech processor to the recipient's implant using a long headpiece cable running through the casing of the test room (Figure 105).

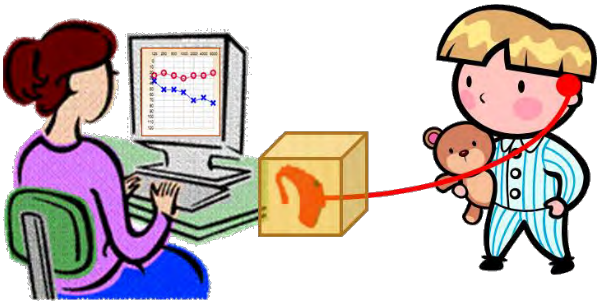


Figure 105: Illustration of the concept of Otocube, a sound insulated desktop test box for performing psychoacoustic measurements in CI recipients.While still connected to the recipient's implant system trough a long coil cable, the speech processor inside Otocube receives acoustic signals in controlled free field conditions.

When the box is closed measurements can be performed using the built-in audiometer, amplifier and loudspeaker. Otocube is small enough to fit on a typical desk. This desktop format and its limited weight (15 kg) make the box portable. It can be moved around within the clinical premises or even be taken on the road to remote locations.

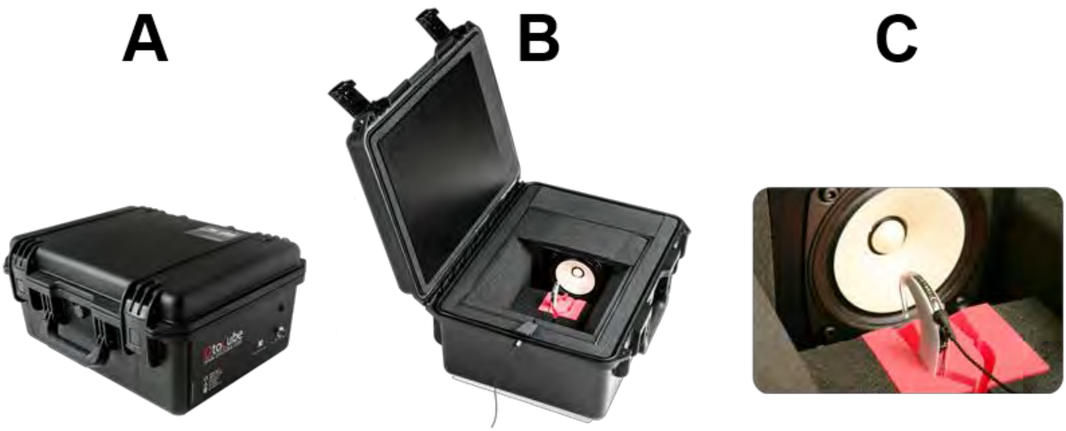


Figure 106: Pictures of the Otocube, showing the device with its lid closed (A) providing the necessary insulation during testing. Inside the box (B) near free field conditions are created at the position of the speech processor (C).

Figure 106 shows thefinal design of the Otocube. The Otocube comes with a built-in type-1 measurement microphone. This microphone is used for calibration and real-time monitoring. At the back side there are 2 connections: one to connect the USB cable going to the PC running A§E 2012 and one for power that comes from the medical grade power supply.

3.9.3. ELECTRONICS LAYOUT

Figure 107 shows a schematic of the Otocube's electronics layout. The device is separated into two distinct compartments: one for acoustic stimulation and one containing the electronics. The signals propagate along 2 paths: one for stimulation and one for monitoring. During stimulation, signals are sent digitally over the USB cable to the 24 bits digital to analog converter. The analog signals are amplified by a self-contained ultrahigh fidelity Class-D amplifier of 180 Watts. The speaker converts the analog signal to acoustical sound waves in the box. These are picked up by the speech processor's microphone, where they are processed and sent through the coil cable to the patient's receiver. The monitoring chain works in the opposite direction. Any sound in the box gets picked up by the built-in Otocube measurement microphone. These signals are fed trough a preamp and converted to a digital signal that is sent to the computer. An audiologist has the opportunity to monitor these signals to hear what's happening inside the box. There is also the possibility for the audiologist to talk through the Otocube speaker to communicate with the patient whose speech processor is inside the box.

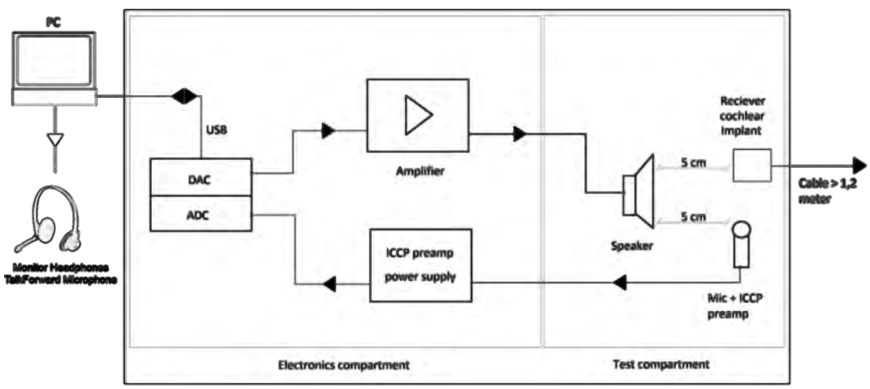


Figure 107: Schematic overview of the electronics layout of Otocube.

3.9.4. FREQUENCY RESPONSE

The Otocube's frequency response (depicted in Figure 108) complies with the requirements regarding the frequency response of clinical audiometric equipment (ISO 60645-2).

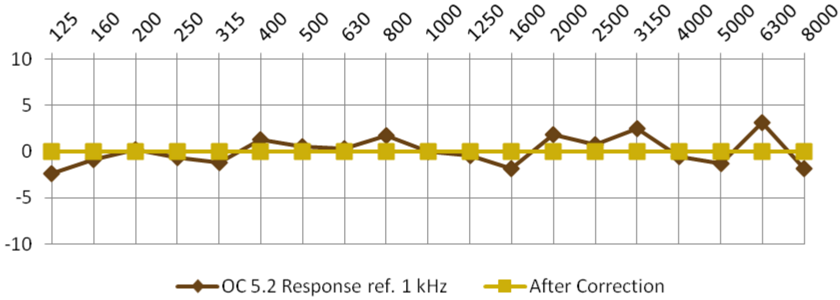


Figure 108: The Otocube's frequency response shown as the attenuation in dB (horizontal axis) for each frequency (vertical axis). The dark points show the frequency response as measured in third octave filterbands centered at the target frequencies. In A§E additional corrections are applied for tests that use narrow band signals (e.g., loudness scaling and tonal audiometry) resulting in flat frequency response (light points).

3.9.5. ACOUSTIC INSULATION

To comply with the standards for maximum permissible ambient noise levels for audiometric test environments (i.e., ANSI S3.6-2004), a lot of effort has been directed towards the Otocube's acoustic insulation characteristics. The results of the final insulations tests are depitced in Figure 109 and show that there is more than 40 dB of attenuation for all frequencies in the range that is clinically relevant, from 250 Hz to 8 kHz.

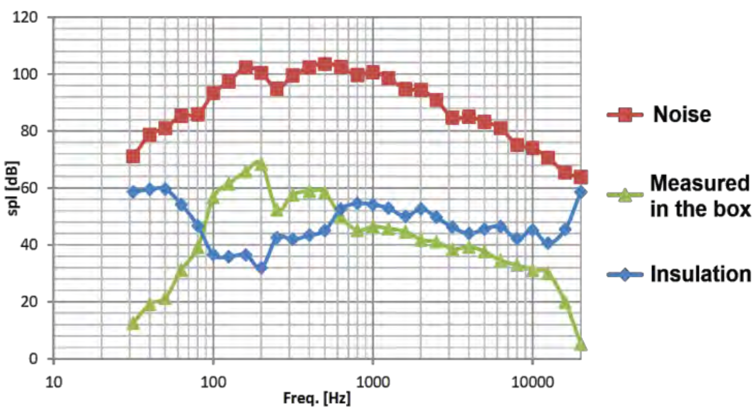


Figure 109: Insulation characteristics of the Otocube. The household spectrum shaped noise that was presented outside of the box during the measurement is plotted in red. The sound levels that were measured inside the box while this noise was present are shown in green. Shown in blue is the amount of insulation per frequency.

3.9.6. SOFTWARE

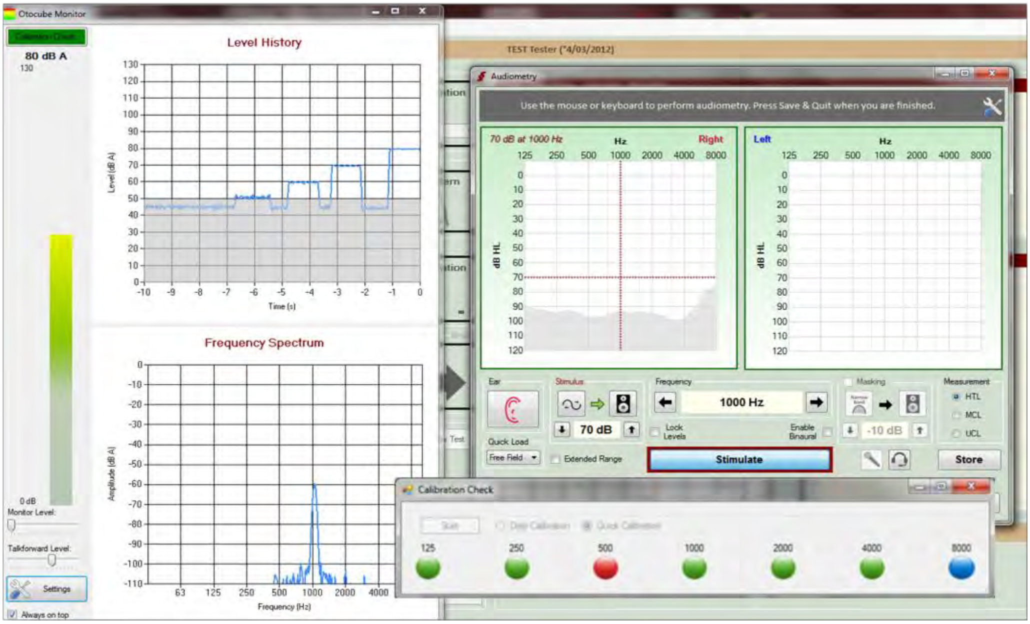


Figure 110: Screenshots of the Otocube software, showing the monitor (left side) with its sound level meter, amplitude spectrum and level history, the A§E audiometry module (right side) and the calibration verification tool (bottom).

Measurements with Otocube are performed through the A§E software, which contains modules for tonal audiometry and speech audiometry in quiet and in noise. Phoneme Discrimination, Loudness Scaling and tests for pitch perception can also be performed with A§E. Because of the built-in measurement microphone that acts as a sound level meter, it is possible to accurately measure the sound levels inside the box at any time. For this reason there is a monitoring application that is able to automatically verify the calibration of sound levels at different frequencies. This tool also allows visualizing any sound present in the box. It provides a real-time sound level meter, a history of sound levels and the real-time spectrum. It also allows communicating with the patient using the acoustical monitor and talk forward functions. Screenshots of the software are shown in Figure 110.

3.9.7. RESULTS

The Otocube has been evaluated in a small number of CI centres in Belgium and France. Tonal audiometry has been performed in 8 CI users, both using Otocube and using a classical sound treated room with calibrated audiometer. The results are depicted in Figure 111 A, which shows for each frequency the differences that were observed between the thresholds in the two test environments. Considering test-retest variability, no significant differences exist between the results. On almost all frequencies, there is a median difference of 5 dB or less. Only at 250Hz thresholds seem to be better in Otocube than in the sound treated room. This may be explained by the fact that Otocube's insulation is better at these low frequencies and that masking effects caused by ambient noise may have elevated the thresholds obtained in the classical sound treated room.

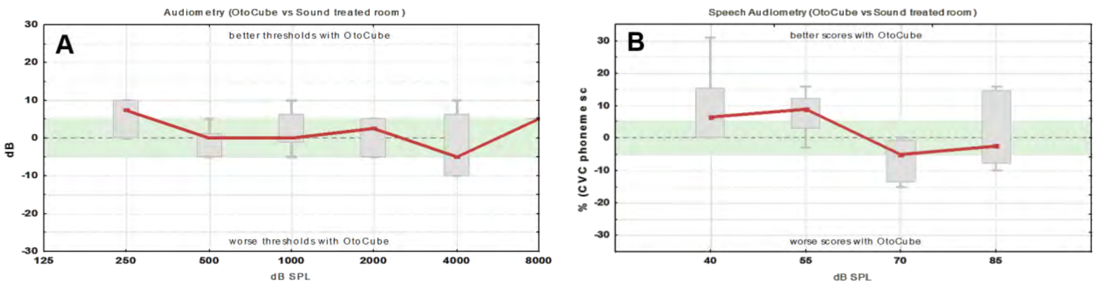


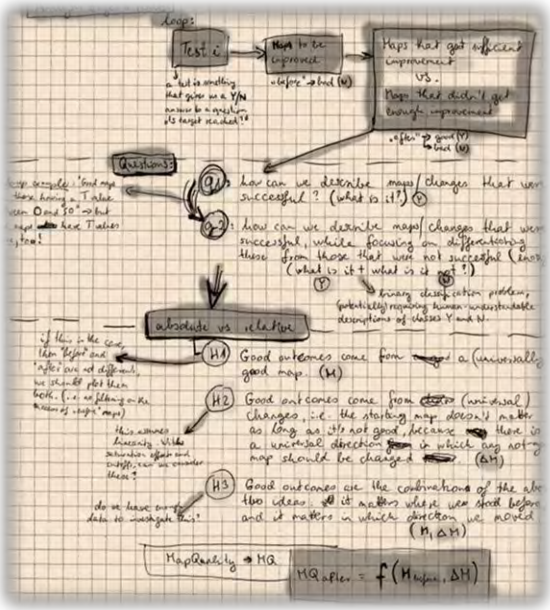
Figure 111: A: differences between detection thresholds for tone audiometry of 8 CI users measured in Otocube and in a clinical sound treated room. B: differences in speech perception scores at 4 intensities of 7 CI users between measurements in Otocube and in a clinical sound treated room.

The same comparison was done for speech audiometry in 7 CI users. From this limited number of results, the impression rises that scores obtained in the two environments are equivalent. There is a median difference of about 5% in speech intelligibility at most intensities. It appears that when presenting stimuli at low intensities, Otocube tends to measure better speech understanding. This might also be an effect of the superior insulation of Otocube; however more investigation is needed to come to conclusions on this matter.

3.9.8. CONCLUSIONS

Otocube is an acoustically insulated test chamber, for measuring outcome in CI users. It's an alternative for an entire sound treated room in clinics where these test rooms may be over-occupied. Efforts have been made to keep it small and portable, without compromising the insulation characteristics. The box is “plug and play” and comes with user-friendly software. The feedthrough provides space for multiple cables, allowing for "live" fitting, while the speech processor is inside the box.

CHAPTER 4: MODELLING THE IMPACT OF FITTING ON OUTCOME



Zoltán Szlávik's notes of a brainstorm session on optimization of the model through data mining.

4.1. INTRODUCTION

4.1.1. MOTIVATION

Cochlear hearing loss results in deficiencies in the 3 signal transduction mechanisms that are essential for normal hearing: (1) intensity coding, (2) tonotopy and (3) phase locking. The acoustical amplification provided by traditional hearing aids contributes to the restoration of intensity coding only, and even in that aspect the benefit is limited (mainly audibility is restored, while the dynamic range remains impaired). CIs on the other hand are able to bypass the limitations imposed by the damaged cochlea and therefore allow for a more extensive restoration of both loudness coding and tonotopy. Phase locking abilities remain largely unaddressed but in case there is residual low frequency phase locking capacity, acoustical stimulation (complementary to electrical stimulation, be it bimodal or hybrid (EAS)) can be used to improve in that area.

The tonotopical interface of the CI remains quite rigid, by design (i.e., a limited number of electrodes exhibiting fairly wide current spread) but also because the knowledge on how to effectively adjust the allocation of frequency bands to CI channels is lacking. The dropping of electrodes is a rough yet often usefull attempt at optimizing tonotopy. But other than that, CI fitting is largely about mapping loudness and the fundamentals of the models developed in this project also come down mainly to adjusting intensity. This however, may be an adequate approach for the time being, since deficits in spectral processing can often be reduced to loudness coding issues within specific frequency bands. As such these deficits can be alleviated through adjustments to intensity mapping in specific channels, causing certain sites within the neural population to receive more or less stimulation than others, hence altering the spectrum-place relation, i.e., tonotopy.

From the introduction of the CI 30 years ago onwards, the main challenge during programming has therefore been related to intensity coding and consisted out of determining an appropriate EDR for stimulating each of the recipient’s implant electrodes. In seeking these optimal ranges for electrical stimulation the imperative objectives were: (1) audibility, (2) no harm or discomfort and (3) optimal loudness growth and resolution. Most often the methods for reaching these objectives were based on behavioral measurements of loudness sensation in response to electrical stimulation of a single electrode. As a result, the field has come to define optimal EDRs in terms of this type of stimulation. The focus on determining such perfect EDRs for each recipient has remained up until today and the field invests heavily in finding ways to determine them in increasingly shorter times.

An essential shortcoming in this approach has been the lack of a methodology to confirm the psychoacoustical validity of such optimal EDRs. The stimulation of a single electrode is little related to sounds encountered in everyday environments. As a consequence, the effect on everyday hearing performance (expressed in the psychoacoustical domain) is insufficiently assessed as a means to evaluate and drive the programming of CIs. In general, psychoacoustical measurements, such as speech perception tests, are only performed for documentation purposes and to assure that changes to the recipient’s map do not adversely affect speech understanding. Over time this has evolved into a vicious circle in which determining the EDR has become a goal by itself rather than the means to an end.

Moreover, the people responsible for fitting (usually audiologists or engineers) having extensive expertise are rare [10]. Even for those "expert fitters" it seems almost impossible to master all programming parameters and their interactions, and more importantly to predict in a reliable manner the impact on a recipient's auditory performance of changing them [11] [12] [13]. In addition, the physiological activation patterns at the level of the brain stem and higher nerve paths are hardly known, which makes that no verifiable or measurable reference is available to objectively assess an intervention [208]. This lack of knowledge is even more explicit in the hearing impaired population as the nerve population at the cochlear level is usually pathologically organized [209] [210] [211].

The current practice of CI fitting therefore is constrained to the manipulation of a limited set of parameters (mainly channel EDRs) [14] [15] [16]. Today, fitting is a manual process in which parameter changes are justified through patient feedback. This is by definition a subjective approach, usually targeted at auditory comfort ("does this sound more pleasant or not?") and often not in line with a recipient’s auditory performance as it would be expressed through psychoacoustical measurements [17] [18]. In addition, such methods hardly take into account effects of adaptation and learning, for which it is known that "take-home" experience is required [99] [212] [213]. Finally, the elicitation of such feedback is only possible in adults and older children, disregarding the important population of infants [214] [12]. In summary, it can be stated that the fitting of CIs according to the current “state of the art” is suboptimal and subject to improvement.

Obviously, this observation is not new, and over the years CI fitting experts in the field have been feeling the need to systematize and optimize their fitting processes. A number of initiatives have led to different methods. For instance evolutionary, probabilistic or genetic algorithms [215] [216] [217], interpolation [218], principal component analysis [219] or techniques using historical data [220] have emerged and have been tried and/or used to improve the quality and/or time spent at programming implants. The use of objective measurements (impedance, ECAP NRT [19] [20] [21], EABR [124], ESRT [22], EART [99], medical imaging [221], etc.) to adjust the processor has gained popularity in recent years. However, the correlation of these measurements to the actual optimal settings was revealed to be limited. Moreover, these methods still remain focused only on finding appropriate EDRs and other fitting parameters are often neglected. The CI manufacturers themselves have also made efforts towards the optimization of fitting. This has led to advisory tools (Cochlear's Hearing Mentor™ for example) available in the fitting software. These approaches however are again based on the patient's subjective feedback of comfort and sound quality, not on measurable hearing performance.

4.1.2. TOWARDS FITTING FOR PERFORMANCE

The Eargroup has chosen, since many years, to adopt an outcome-driven approach, in which CI fitting is motivated by and tested for auditory performance, as measured by behavioural psychoacoustic tests. This has lead to the systematized evaluation of a number of psychoacoustical performance measures and the accumulation of mainly empirical/heuristic knowledge on the onderlying relationships to the map settings of the CI speech processor. Prior to this PhD project however, fitting at the Eargroup was done manually and pragmatically, based on years of their experience. In order to make this more evidence-based, it was necessary to investigate the relation between electrical stimulation (resulting from a specific map) and auditory performance (i.e., outcome) in a more fundamental way.

The relationship between a map and the outcome it yields is currently insufficiently known to allow for reliable prediction of the effects of map changes on outcome. Partial knowledge exists about the relations between sound features and outcome on the one hand and between sound features and processor parameters on the other hand. In order to gain insight into how the processing of sound by the human ear and the implanted cochlea in particular, leads to auditory performance, a better understanding of the relationship between physiology (the cochlea), the CI (processor) and measurements of hearing performance (outcome) is needed. This relationship can be described through the characteristics of sound (sound features) that compose human hearing: loudness, spectral content and temporal information [222] (Figure 112) and is deducted from the way sound features are coded by cochlear implants and how those features are perceived by the listener.

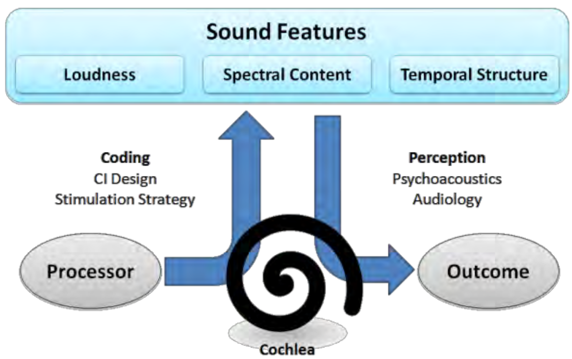


Figure 112: Schematic illustration of the relation between processor configuration and outcome. The design of CIs and their stimulation strategies determine how features of sound are processed (coded) and presented to the neural interface as auditory cues. The preservation of these cues is essential for adequate auditory performance. The field of audiology and psychoacoustics assesses this preservation through a number of outcome measures that relate to and give insight into possible deficits in the perception of sound features.

The **cochlea** is the receptor involved in converting the various characteristics of sound into a pattern of action potentials in the population of auditory neurons. The (electro)physiological coding of sound features, in particular intensity and spectral and temporal content, is highly complex and still subject of evolving insights [155] [223] [186] [224]. "The ear & cochlear hearing loss" provides a summary of physiological aspects of the cochlea which are a prerequisite for accurately describing the relationships between sound features and auditory performance on the one hand and the CI processor strategies on the other hand.

"Cochlear implants & fitting" provides an overview of current generation CI technology, and on the current state of the art in CI fitting. An optimal adjustment of the CI **processor** will approach the physiological sound treating processes most efficiently and result in the nearest to normal hearing possible. By comparing the operation of the processor to the cochlear function, it may become clear how a CI attempts to reconstruct these physiological processes and where the limitations of a CI come into play [225] [226]. The analysis of stimulation strategies and the parameters (map) available to adjust them, leads to the construction of functions that describe the effects of parameter changes on the coding of sound features [227].

"Measurements & outcome" provides an overview of the outcome measurements used in the context of CI fitting. To evaluate how the map affects auditory performance we investigated how the coding of these sound features is expressed in **outcome** measurements [228] [229] [230] [207] (for example, how the discrimination of phonemes /a/ and /u/ relates to transmission of cues in certain frequency bands [231]).

Based on those relations (processor – sound features and sound features – outcome), it would be possible to construct a working model for the relation between processor parameters ($m_{1\text{-}i}$) and outcome ($o_{1\text{-}j}$).

$$(o_1, o_2, \dots, o_j) = f(m_1, m_2, \dots, m_i) \tag{9}$$

An evident problem is the fact that the number of parameters is large, both at the input and at the output side of the equation (9). A CI has over one hundred configurable parameters, which are surely in part interdependent. The outcome is also described by a multitude of variables, a number that will increase as additional measurements will be developed, in the context of present and other research projects. Moreover, clinical experience suggests that this relationship is markedly sensitive to individual variations. This inter-subject heterogeneity makes it unlikely that a relationship between processing parameters and auditory performance can be described by universal formulas (i.e. the existence of a “perfect map” that maximizes performance in all individual recipients). On the other hand, clinical expertise also indicates that the relationship is more stable when expressed through changes in the variables at hand rather than through the absolute values of those variables. It would therefore be more efficient to represent the relationship through differentials, such that a modification of processor parameters (Δm) causes a change in performance (Δo).

$$(\Delta o_1, \Delta o_2, \dots, \Delta o_j) = f(\Delta m_1, \Delta m_2, \dots, \Delta m_i) \tag{10}$$

4.1.3. OBJECTIVE

The primary objective of this PhD project is to establish a useful model that provides deeper insight into the adjustment of cochlear implants. That model could then be used to optimize the the process of CI fitting by introducing a more systematic and verifiable approach. Three components are essential for process optimization: (1) defining targets for the system, (2) measuring the state of the system and (3) algorithms effective in moving the system's state towards target. When applied to CI fitting, this approach comes down to measuring a CI recipient's auditory performance and making targeted adjustments to the speech processor, in order to improve hearing performance, i.e., bringing it closer to the set targets.

These targets for hearing performance were defined with the auditory capabilities of the normal hearing population in mind. A detailed description is given in the manuscript "Setting and reaching targets with computer-assisted CI fitting" but in summary these targets comprise the following objectives:

- Pure tone audibility thresholds of 30 dBHL or better for frequencies between .5 and 8 kHz (35 dBHL or better for 250 Hz tones);
- Spectral discrimination of 18 out of the 20 A§E Phoneme Discrimination contrasts;
- Loudness growth of low, mid and high frequency narrow band noises corresponding to the 95% confidence interval in hearing subjects;
- Relatively constant word recognition scores across the range of intensities between 40 and 85 dBSPL.

In order to reach those targets in a systematized manner, a model that predicts the necessary changes to the map is needed. Given the extensive expertise in CI fitting driven by the above mentioned measures of psychoacoustical performance that is present at the Eargroup, a rule based deterministic model was constructed based on the Eargroup's clinical heuristics. The development of this model is explained in more detail in the next section "A deterministic model based on clinical heuristics".

Eventually this model gave rise to the intelligent agent "Fox" that is being used, already today, to optimize and automate the process of CI fitting. The software design of Fox is described later in this chapter and its development and application are described in the manuscript "Development of Fox", where it is argued that Fox allows a systematic approach focussing on outcome, reducing the fitting time and improving the quality of fitting. It introduces principles of artificial intelligence in the process of CI-fitting.

In a later stage it was also explored how probabilistic modelling could be applied to the problem domain. Given the elevated complexity of the relations between map and outcome and the biological variability across subjects, it is not unreasonable to assume that non-deterministic (stochastic) artificial intelligence (AI) has a similar or even stronger modelling power to handle the many uncertainties in the relationships. A first attempt to build a probabilistic model is described in the manuscript "A probabilistic graphical model". An intrinsic problem with probabilistic models is the exponential growth of computational requirements when increasing the number of parameters included in such a model. To handle the large number of parameters in the problem domain, a new canonical model was developed, aming at both reducing the number of probabilities (elicited from experts or learned from databases) to be explicitely specified and keeping the computational workload within limits. The details of this canonical model are given in “The tuning model”. Since the stochastic approach is still in the early stages of exploration, its applicability is subject to future research. It may well be that a combination of both deterministic and stochastic approaches will result in the most effective way to tackle the problem.

4.2. A DETERMINISTIC MODEL BASED ON CLINICAL HEURISTICS

4.2.1. INTRODUCTION

The construction of a model for tuning maps based on deficits in outcome measures was conducted in an agile bottom-up approach. At first relationships between a limited number of variables were defined based on theoretically derived dependencies between map changes (Δ map) and changes in outcome (Δ outcome). Those dependencies were subsequently tested against, refined by and extended with empirical knowledge elicited from CI fitting experts (Figure 113). For instance, from a theoretical point of view one may expect that an audiometric threshold which is markedly higher than the CI's microphone sensitivity can be improved by an increase in the EDR Minimum level(s) on the corresponding electrode(s). Clinical experience however showed that this is not always the case (presumably due to the presence of an internal noise in the CI system that is dependent on the EDR Minimum level). This bottom-up approach allowed for adding more and more variables to the model step by step, both on the input side (map changes) and at the output side (outcome changes). The resulting functions were validated case-wise on patients by the Eargroup's CI fitting team in an iterative review process, and tweaked by experts when need be.

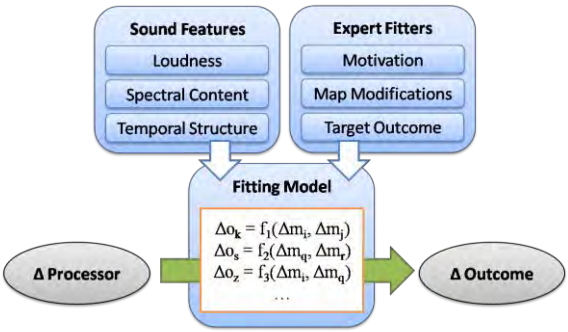


Figure 113: The initial model is based on a theoretical framework. This model reflects dependencies between changes to the processor's configuration (Δ map) and changes in auditory performance (Δ outcome). These dependencies were deducted from the synthesis of existing literature/knowledge (how sound features are processed by CI stimulation strategies and how the perception of those features is expressed in psychoacoustical measurements), combined with empirical knowledge of CI fitting experts.

In the mean time, a knowledge base was permanently fed with raw data consisting of outcome linked to maps. For every of the numerous (> 10.000) outcome test performed, all parameters values of the map used during the test are stored together with the test result. This coupling of outcome to the map that was used to obtain it, should allow refining the model by statistical analysis (e.g., principal component analysis, multiple regression analysis, etc.) of the collected data set (Figure 114 A). In addition, the fitting model was implemented in an engine that could predict the necessary map changes in function of desired outcome changes (Figure 114 B). This way the model was already deployable during early stages of its development such that it could be tested for its effectiveness in moving CI recipients closer to the auditory targets.

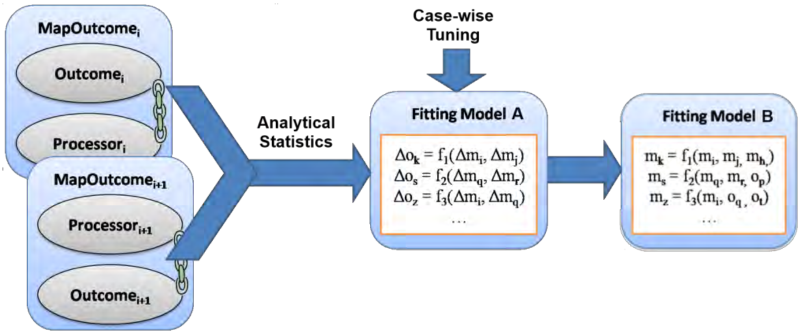


Figure 114: Refinement of the theoretically deducted model by statistical analysis of collected data sets and by case-wise tuning after validation on real subjects (model A). The model would be constructed in a way that it can be interrogated for the map most likely resulting in a desired outcome, given the map that was used in obtaining the observed outcome (model B).

After the initial validation and feasibility assessment, this deterministic rule-based model was made available to clinicians, also outside of the Eargroup. This was accomplished through a software tool called Fox (Fitting to Outcome eXpert), which is described in "Development", found later on in this chapter. In this section the development process of the deterministic rule set underlying this model is presented. That rule set is called the Eargroup's advice and has mainly been constructed and optimized for the Advanced Bionics (AB) HiRes90k implant system. Care has been taken however to keep it as generic as possible for the sake of future portability and applicability to other CI systems.

4.2.2. RULE SET DEVELOPMENT

A dedicated Integrated Development Environment (IDE) was built to enable the progressive construction and iterative tuning of rules. These iterations were based on the previous experiences, both successes and failures, in resolving outcome deficits in individual recipients. This IDE is called the Fox Advice Designer of which a screenshot is shown in Figure 115. The IDE uses its own object model and programming language with a highly simplified syntax which enables rules to be defined through straightforward statements using specific high level operations that act on easily accessible map and outcome variables (e.g., “add 20 units to the EDR Minima of low frequency channels when the audiometric threshold at 250 Hz is greater than 40 dB HL”). This allows clinicians to write, review and modify the rule set without the need for them to have any specific programming skills.

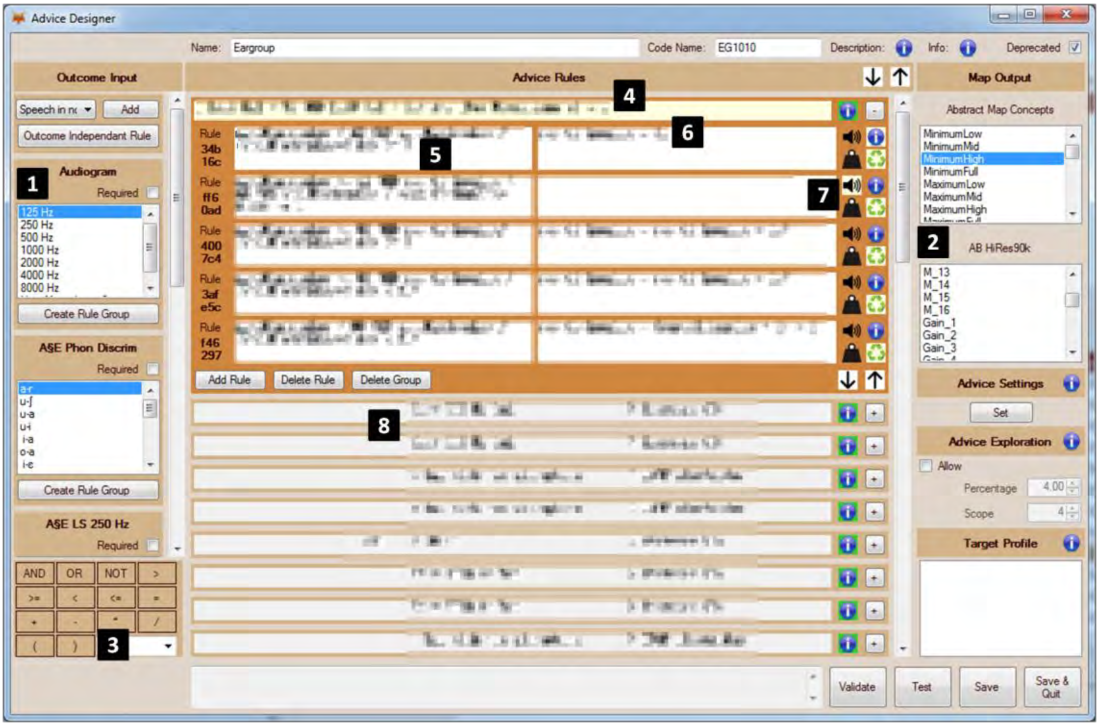


Figure 115: The Fox Advice Designer, allowing clinicians to define a set of rules consisting of Outcome Parameters (1), Map Parameters (2) and Operators (3). Rules are grouped by a common Outcome Condition (4) and every rule has its own Map Condition (5) and Map Effect (6), which is executed when both Outcome and Map Condition are met. In addition to a Map Effect (changing values of Map Parameters) a rule may also be attributed with Additional Effects (7): a message to the audiologist, the request to perform an additional outcome measurement and a weight that determines the rule's importance when Map Effects of different rules are in conflict. When Rule Groups are collapsed, a textual description of the group is shown (8). ***Due to reasons of intellectual property parts of this figure are blurred in the published version of this dissertation.***

The choice of a dedicated IDE to develop and maintain the rule set has a number of advantages. Firstly, it enforces a structured process of defining rules. This process consists of defining the Outcome Condition (OC), the Map Condition (MC) and Map Effect (changing values of Map Parameters) (ME) for each rule. Rules having the same OC are grouped together. Every rule has its own MC and ME, which is executed when both OC and MC are met. In addition to a ME a rule may also be attributed with Additional Effects (AE): a message to the audiologist, the request to perform an additional outcome measurement and a weight that determines the rule's importance when MEs of different rules are in conflict.

The IDE treats every rule as a separate entity, for which a complete revision history (including all previous versions) is maintained. Figure 116 shows a plot of the cumulative number of modifications to rules through time.

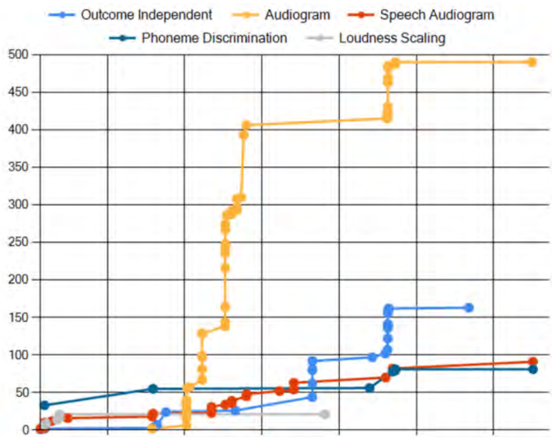


Figure 116: Development of the Eargroup's advice. The plot shows the cumulative number of rule set modifications per type of outcome. Each dot is a modification of a rule at a specific point in time.

The isolation of rules into separate entities also enables transaction logging at the rule level, meaning that for every rule a history is maintained of which map parameters it manipulates every time the advice is solicited for a particular map and associated set of outcomes. In this transaction log the input value of every map variable is recorded, together with its output value after application of the rule. This allows for instance the evaluation of a rule's application rate (how often it is executed) and application scope (which map parameters it affects). Figure 117 shows such an analysis for the Eargroup's advice at the level of the different outcome measures.

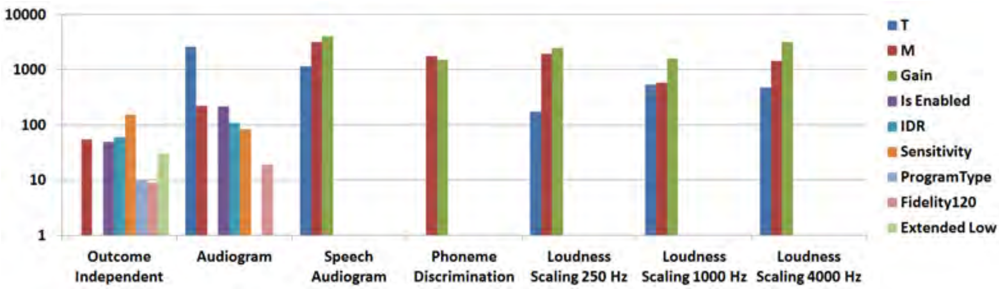


Figure 117: Map modifications resulting from the Eargroup's advice. For each type of outcome, the bars indicate the number of map variables that have been adjusted by application of rules that are based on that outcome.

Because every rule is an isolated entity, rules are executed in parallel upon solicitation of the advice. This means that there is no chronological or hierarchical dependency between rules. Whenever multiple rules attempt to manipulate the same map parameter, the final value of the map parameter is determined by averaging the rules' separate effects, considering their relative weights.

Rules sets are maintained separately from the software that executes them. This means that rule sets can be modified without the need to recompile or otherwise modify the software to solicit advices.

The rule set is enforced to act within certain constraints that relate to patient safety. For every map parameter 2 ranges are defined: warning range and absolute limits. Although these absolute limits are often more stringent than the range of valid values as defined in the manufacturer’s fitting software, the execution of a rule cannot assign a value that is not within this range. The warning range can be exceeded, but in these cases the audiologist’s attention is drawn to these parameters by a visual highlight in the Fox software. A more detailed description on medical safety and risk management is given in the section “Fox software design”.

4.2.3. CLINICAL HEURISTICS OF THE EARGROUP ADVICE

This section has been removed due to reasons of intellectual property.

4.3. FOX SOFTWARE DESIGN

This section has been removed due to reasons of intellectual property.

4.4. DEVELOPMENT OF FOX

Development of a software tool using deterministic logic for the optimization of cochlear implant processor programming

Otol Neurotol. 2010 Aug;31(6):908-18

P J Govaerts[1,2], B Vaerenberg, [1,3], G De Ceulaer[1], K Daemers[1], C De Beukelaer[1]

[1]The Eargroup, Antwerp-Deurne, Belgium
[2]CNTS and [3]Laboratory of Biomedical Physics, University of Antwerp, Belgium

Abstract

An intelligent agent "FOX" was developed to optimize and automate CI programming. The current paper describes the rationale, development and features of this tool. CI fitting is a time-consuming procedure to define the value of a subset of the available electrical parameters based on behavioural responses. It is comfort driven with high intra- and inter-individual variability. Its validity in terms of process control can be questioned. Good clinical practice would require an outcome-driven approach. An intelligent agent may help to solve the complexity of addressing more electrical parameters based on a range of outcome measures. FOX is a knowledge-based software application which consists of deterministic rules that analyze the map-settings in the processor together with psycho-acoustic test results (audiogram, A§E phoneme discrimination, A§E loudness scaling, speech audiogram) obtained with that map. The data transfer to and from this agent is either manual or through seamless digital communication with the CI-fitting database and the psychoacoustic test suite. It recommends and executes modifications to the map settings to improve the outcome. FOX is an operational intelligent agent for which a CE class I medical device mark was obtained. The principles of this agent are described, its development and modes of operation, and a case example is given. FOX is capable to improve the measured outcome. It is argued that this novel tool allows a systematic approach focussing on outcome, reducing the fitting time and improving the quality of fitting. It introduces principles of artificial intelligence in the process of CI-fitting.

4.4.1. INTRODUCTION

Cochlear implantation is now widely accepted as an effective treatment for profound deafness [232] [233]. Several commercial devices are currently available, but all share many common features, such as the basic combination of an externally worn sound processor which delivers power and coded signal to an implanted receiver package, via a transcutaneous RF transmission link, which in turn delivers a sequence of electrical pulses to an array of electrodes surgically placed into the scala tympani of the cochlea. There are also considerable similarities between the various coding strategies employed in different devices, which define the pattern of electrical pulses delivered to the cochlea in response to acoustic input to the processor.

Following surgical implantation, the external sound processor must be appropriately programmed and customized for the individual. The aim of this is to set a number of electrical parameters to ensure that the electrical pattern generated by the internal device in response to sound stimulation, yields an optimal auditory percept [84]. Several electrical parameters are available and all their values together are commonly called the MAP. Finding and programming the optimal values for an individual is commonly called the act of fitting. It is achieved using proprietary software and a hardware interface connected to the processor, and depends on behavioural responses from the CI user.

From the early years onwards, many electrical parameters have been set at default values which are mostly left untouched during the fitting process. Fitting is usually restricted to setting the threshold of audibility for electric stimulation, plus dynamic range, for each electrode separately. Both levels may vary considerably among individuals and among different electrodes along the array within individuals. For this reason the initial task is for the audiologist to measure the threshold and some measure of upper loudness tolerance (such as “most comfortable level”) for each electrode, in order to define a range of outputs that provides a comfortable percept when the resultant MAP is activated.

After the initial “fitting” and activation of the processor, several review sessions are normally required to re-measure these levels in order to accommodate the increase in dynamic range that typically occurs as the user becomes accustomed to the electrical stimulation over the first few months of device use [98]. The need for follow-up sessions is particularly important for young children as it is generally very difficult to assess sensitivity to electrical stimulation in this population due to their cognitive status and lack of experience of auditory sensations. Following stabilization of electrical dynamic range fitting sessions are usually limited to periodical checks, typically annually, as long as progress is normal.

While threshold and upper loudness levels are the main parameters commonly used for the generation of an appropriate MAP, there are many others that can be adjusted within the fitting software. The most common additional adjustment is the de-activation of individual electrodes if deemed necessary, usually if they show high thresholds, small dynamic ranges or produce non-auditory stimulation. However, the majority of available parameters are rarely adjusted in normal clinical practice; these include parameters such as band pass filter boundaries, gain, microphone sensitivity, output compressive function, inter pulse interval, stimulation rate and so on.

It is important to notice that the main criterion used is the patient’s behavioural response. This reflects detection at low intensities to set the lower stimulation level and some appreciation of comfort, maximal comfort or discomfort to set the upper level.

Once behavioural fitting parameters are stable it is usually assumed that the MAP is optimally adjusted. Occasionally, the user may complain about the subjective quality or tone of the auditory percept but generally MAP adjustments are not based on formal outcome measures. If a user is performing at a lower level than might be expected then fitting measures may be repeated, but if these appear reliable then it will be accepted that performance is probably optimal for that particular user, as it is well known that outcomes vary considerably even within a relatively homogeneous group of CI users [234].

Repeated fitting sessions, even when they merely address the limited number of electrical parameters described earlier, are very time-consuming for a CI centre, and there is therefore a perceived need to make this process as efficient as possible. However, apart from time considerations, the efficiency of the process is clearly also affected by how much benefit is gained by very accurate processor fitting. There exists a school of thought that the central auditory system is able to accommodate to a fairly wide range of inputs from the cochlea, such that as long as speech sounds are audible then the language processing centres of the auditory system can satisfactorily adapt through neural plasticity. To this end, several studies have shown that processor fittings can be simplified to a certain degree without significant detriment [235] [16].

While this line of thinking may have useful connotations for certain clinical situations, it remains an accepted fact that accurately adjusted processor MAPs do generally result in better outcomes in terms of speech understanding [236] [237]. The practical question, however, is how to achieve this without spending excessive amounts of clinical time.

Several ways to reduce the fitting time have been developed over the years. They can be summarized by two strategies: 1) to introduce objective measures that serve to predict the optimal MAP values and 2) to set MAP values on a group of electrodes rather than on individual electrodes.

Objective measures are often performed during surgery, although they can also be performed at any moment after surgery. They include measurements of the electrically-evoked compound action potential (eCAP) using back-telemetry [238], electrically-evoked auditory brainstem recordings [239] and electrically-evoked stapedius reflex thresholds [240]. These have been shown to identify stimulation levels within the behavioural dynamic range, but show considerable variability and do not accurately indicate the limits (threshold and MCL) of the dynamic range [241] [22]. They are mainly used as a starting point for user MAPs, where behavioural measures are still important in order to fine tune the processor fitting. Much work has been carried out in order to optimize the correlation between these objective measures and their behavioural equivalents, with most effort in recent years directed towards eCAP measures, with the hope that they might satisfactorily be used as a means of “automated” fitting, dispensing with the need for behavioural measures altogether [242].

Changing the MAP values for a group of electrodes is facilitated in the fitting software of the different devices. For instance, several electrodes can be selected together and their MAP values can be modified group-wise. Shift- and tilt-functions allow changing the profile of the lower or upper stimulation levels of the entire electrode-array [128], etc.

A limitation of traditional processor fitting is that it depends on the experience and knowledge of the audiologists or other personnel performing the measurements and adjustments. Behavioural responses, especially when obtained from patients with no or little auditory experience, may vary according to the methodology employed, instructions to the patient, and so on. Training in fitting is usually provided primarily by the CI manufacturers, but there exists no standardized methodology, which makes it difficult to verify the quality of this aspect of the fitting process. Anecdotal reports from clinical specialists working with CI manufacturers suggest that patients with grossly inappropriate MAPs are occasionally encountered, even in centres where the usual amount of training has been provided. One can argue that after more than 20 years of cochlear implantation, the act of fitting is still a matter of craftsmanship where much time is invested to set merely a partial number of the electrical parameters based on behavioural responses relating to a level of detection and some level of comfort and of which the reliability can be questioned.

One of the basic tenets of the system developed here is that it is the cochlea that is the main site of dysfunction in the typical CI user. Therefore, no matter to what extent the central auditory system is able to “compensate” for an imperfect signal from the cochlea, it will inevitably be able to function better if the output from the cochlea can itself be optimized. Furthermore, the cochlea is clearly the level of the auditory system to which we have access during CI programming.

Outcome measures that reflect cochlear function are limited, at least in terms of tests that can be readily performed in a routine clinical setting. Audiometry can assess detection, but speech recognition testing involves higher level linguistic processing and so only indirectly relates to cochlear function.

Largely to address this problem, we developed a test battery known as the “Auditory Speech Sounds Evaluation” or “A§E” [107] [243]. This is a psycho-acoustical test suite attempting to assess these cochlear functions in more detail. The core module is a discrimination test based around 20 pairs of speech sounds, which are presented in an oddity paradigm and which can provide a clinical indication of the frequency resolving power of the cochlea. More recently, we have added a loudness scaling module that indicates loudness growth at 250 Hz, 1 kHz and 4 kHz.

To date audiometry, A§E phoneme discrimination (20 phoneme pairs), A§E loudness scaling (with narrow band noise centered at 250, 1000 and 4000 Hz) and speech audiometry (open set monosyllables presented at 40, 55, 70 and 85 dB SPL) are routinely used in our centre to measure the quality of the fitting. Strategies have been developed to feed back this information to MAP-changes in order to improve the measured outcome. This approach however has faced us with the complex relationships, correlations and interdependencies between the many variables, electrical and psycho-acoustical, at both sides of the equation. For any professional, even the very experienced one, it becomes difficult to master all these functional relationships. For that matter we have made a first attempt to introduce artificial intelligence in this process.

Artificial intelligence (AI) is a relatively new science with many theoretical applications, one of which is the making of rational decisions to maximize outcome in complex systems. It not only attempts to understand but also to build intelligent entities [244]. An intelligent agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators. For our purpose, the psycho-acoustical tests serve as sensors and the MAP (together with the fitting software) as actuator. Internally the agent function is implemented by an agent program. It is beyond the scope of the present paper to elaborate in detail on AI. Briefly, the program is based on knowledge, logic and learning skills. The core consists of logic, which can be either deterministic or non-deterministic (also called stochastic or probabilistic). Deterministic logic is typically rule-based. Typical forms of non-deterministic logic are neural networks, genetic algorithms, etc. A comprehensive state of the art of AI can be found in Russel & Norvig [245].

Over the past 10 years we have been developing an intelligent software system or intelligent agent which is designed to optimize CI processor MAPs. In its actual state it uses the psycho-acoustical outcome measures mentioned earlier, although it is conceived to handle other measures, like electrophysiological test results or questionnaires, as well. It analyzes the actual MAP settings together with the outcome obtained with it. Its primary aim is then to provide recommendations for mapping adjustments to optimize the electrical signal presented to the cochlea without the need for conventional behavioural fitting measures, which are subject to the limitations outlined above.

This software tool is termed the “Fitting to Outcomes eXpert” or FOX, and has a CE Class I Medical Device mark. The aim of this report is to outline the principles behind its development, describe its main features and to demonstrate its function through some case examples.

4.4.2. PRINCIPLES BEHIND THE DEVELOPMENT OF FOX

FOX (Registered with Interdeposit Digital Number BE.010.0112303.000.R.P.2008.035.31230) is based on a set of programming rules which have been established from analysis of clinical MAPs and outcomes over several years’ experience with over 600 CI users at our centre. The system, which is written using .net technology, currently contains a large number of determinist “rules” which link a range of outcome measures to the most important parameters that can be adjusted within the CI fitting software. This particular set of rules constitutes the Eargroup “advice”, but additional “advices” can be developed and added to FOX from other sources (e.g. other clinical experts, CI manufacturers, etc.) and a user-friendly interface allows the input of additional rules by professionals without the need for knowledge of programming languages. Separate advices (each made up of a set of rules) are available for different situations, such as different CI devices, types of processor or the type of fitting session, as any particular rule may operate differently under different situations.

FOX can be used as a stand-alone software package, but is also able to interface directly with proprietary outcome data sources and CI fitting software through direct synchronization. In this report, we demonstrate how it operates together with the SoundWave fitting system from Advanced Bionics, but it can potentially interface with fitting software from other CI manufacturers. FOX works as an iterative process and can be run several times. The basic mode of operation is illustrated in Figure 118, which shows options for independent function and when interfaced with the Soundwave software ("CI tables") and the Audiqueen database containing outcome data. Thus, FOX takes an existing CI MAP and analyses the outcome data associated with that MAP. Using deterministic logic based on its set of rules, it then recommends changes to the CI MAP that are expected to improve outcomes. Following these changes, outcome measures can be repeated and fed back to FOX, which may suggest further changes or confirm an optimal fitting.

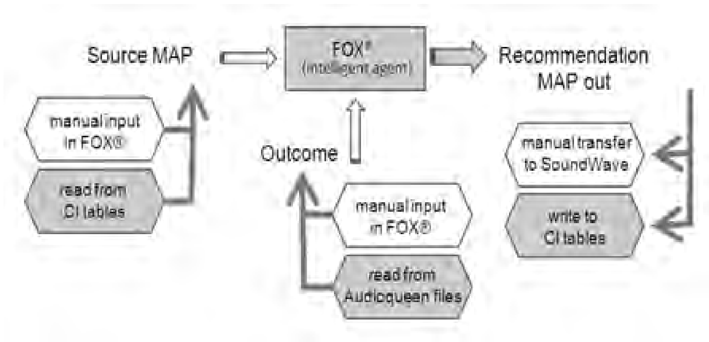


Figure 118: FOX working principle. An initial program as well as various psycho-acoustical test results may be input into FOX , which then delivers fitting recommendations as output. Shaded boxes illustrate function when interfaced with proprietary outcome and CI fitting software, while unfilled boxes denote standalone function.

Table 10 illustrates the operation of a typical rule. The top row shows the outcome condition that elicits the execution of the rule.

Table 10: Typical rule featuring in the Eargroup's advice. See text for details.

[z-s] + [a-r] < 2		
Rule 4d7 186	map.Maximum_0_600 < 330	map.Maximum_0_600 = map.Minimum_0_600 + (map.Maximum_0_600 – map.Minimum_0_600) * (100 + (2 – [z-s] – [a-r]) * 10)/100

In this case it translates as: "IF the listener fails to discriminate the contrasts /z-s/ or /a-r/ of the A§E phoneme discrimination, THEN execute the rule". The left column shows the breakdown of the possible effects of the rule based on additional criteria that consider the actual MAP settings. In this case the first additional criterion (Rule 4d7 186) reads: "IF the average M-level of the electrodes coding the acoustic frequencies between 0 and 600 Hz is lower than 330 clinical units, THEN execute what follows". The right column shows the effect produced by the execution of the rule. In this case, the first effect reads: “increase the dynamic range of the electrodes coding the frequencies 0-600 Hz by 20% if both contrasts were not discriminated OR by 10% if only one contrast was not discriminated”.

Outcome measures that may be input to FOX include the following:

- Acoustic (free field) thresholds from 250 Hz to 8 kHz
- Loudness growth function for 250 Hz, 1 kHz and 4 kHz
- Auditory Speech Sound Evaluation (A§E) discrimination of 20 phoneme contrasts at 70 dB HL [re. 1kHz narrow-band noise]
- Speech audiogram (scores at 40, 55, 70 and 85 dB SPL)

Additional outcome measures can potentially be incorporated into FOX, following development of appropriate rules. These could potentially include other behavioural test data, objective test data (eCAP measures, stapedius reflex thresholds etc.), questionnaire data or other performance measures.

Mapping parameters currently incorporated into FOX include the following:

- Electrical thresholds (T levels) and upper loudness limits (M levels)
- Input dynamic range
- Gain
- Electrode activation / de-activation
- Processing strategy (HiRes, HiRes 120 etc.)
- Pulse rate
- Bandpass Filter boundaries
- Automatic Gain Control
- Sensitivity
- Volume

In the future, rules for additional parameters may be developed. However, we consider that the parameter adjustments currently available are adequate for MAP optimization in the large majority of cases.

It should also be noted that a number of safety measures are available to control the risk of errors and of overstimulation. These provide warnings or constraints to the MAP settings or changes that are allowed, restricting the operating freedom within stricter and safer limits than those the manufacturer’s fitting software allows. Some are based on clinical expertise and intuition. For example, increases in the maximal or most comfortable stimulation level (M-level for AB-devices) are restricted to 80 clinical units per iteration. Other safety measures are based on statistical analysis of all MAPs that have ever been given to CI-users. For example, the distribution of all these setting is defined by an average value μ and a standard deviation SD for each MAP-parameter.

Whenever an advice attempts to modify the value of an electrical parameter to beyond the interval $\mu \pm 2$ SD, this attempt is highlighted to alert the audiologist. Whenever an advice attempts to modify the value to beyond the interval $\mu \pm 5$ SD, FOX will block the modification, alert the audiologist and provide the option to take over programming outside the control of FOX. These additional safety measures are particularly important when fitting is being performed by relatively inexperienced clinicians.

4.4.3. FEATURES AND OPERATION OF THE FOX SYSTEM

FOX is able to provide MAPs for the initial switch-on sessions, based on demographic data, and this “automap” function is described below. In addition, the system can be used in order to optimize MAPs that have been originally generated from standard behavioural fitting sessions.

A user-friendly graphical interface presents a list of available MAPs for a given patient. These MAPs are available to FOX by means of synchronization between the proprietary fitting software and its own database. An individual MAP is selected and read into FOX.

A specific “advice” is then selected from a list, according to, for example, the particular CI device or type of fitting session (such as initial switch-on). Figure 119 shows the advice selection screen together with the list of outcomes that can be entered for that particular advice. The audiologist can enter whatever outcomes have been obtained from the patient using the MAP being analysed. These test results can be entered manually or they can be imported seamlessly from the A§E test suite or an Audiqueen (Otoconsult, Belgium) export file.

Figure 119: The "advice" selection screen. In this case it is a follow-up session with an Advanced Bionic's device. On the left are listed the outcome measures that can be entered for this particular advice.

Once data entry is complete FOX analyses the MAP settings together with the outcome and formulates its feedback. The response from FOX is in two forms: messages and MAP changes. Figure 120 shows a screen shot of the software showing a typical FOX response. Outcome measures that were entered are listed in the left panel. The main panel contains several "messages", in this case highlighting that changes to the MAP are suggested, plus a prompt requesting additional outcome data.

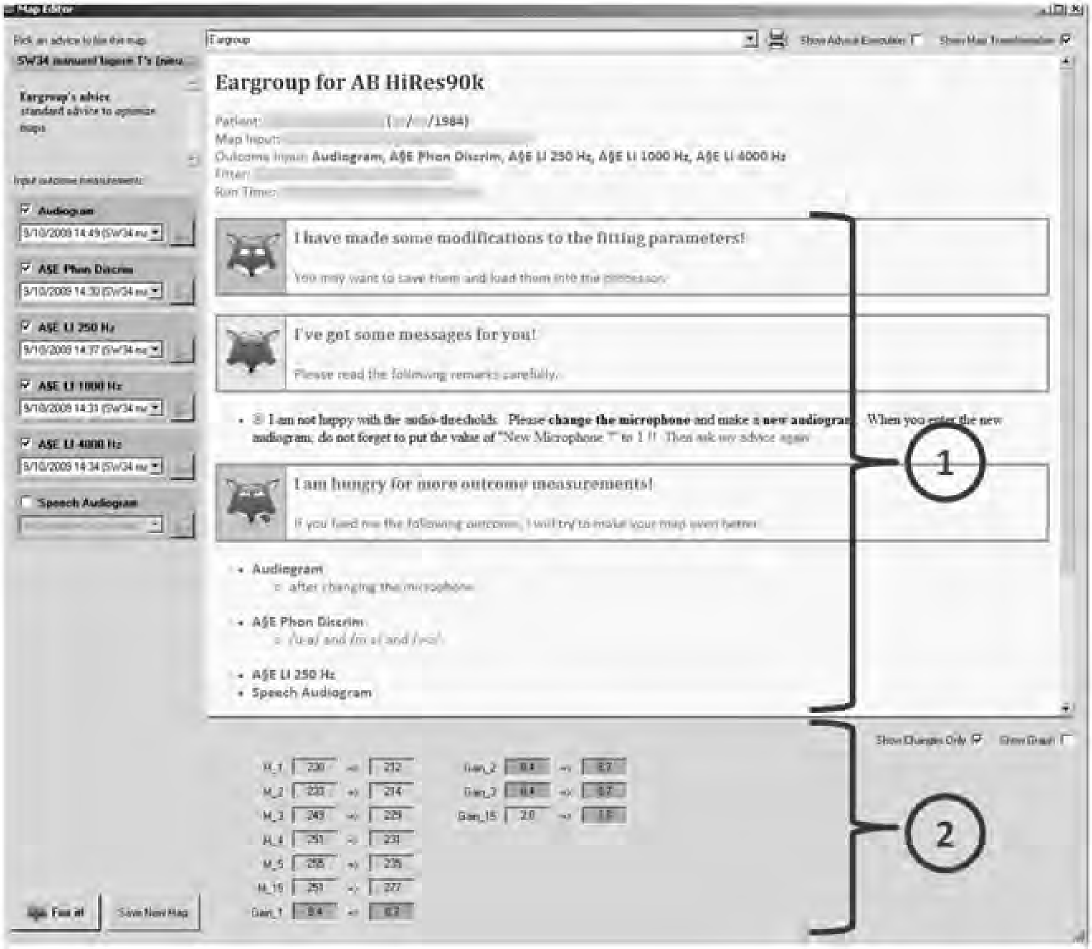


Figure 120: A typical response from FOX, following input of outcome data, is illustrated in this screen shot. The output consists of messages (1) and suggested modifications of the MAP settings (2). See text for further details.

At the bottom of the screen are recommendations to adjust M levels for 6 electrodes and gains for 4 electrodes. These fitting parameter suggestions may be executed manually by the audiologist or automatically by direct communication with the fitting software.

Outcome measures using the modified MAP can then be made and entered into FOX, such that more than one iteration may be performed at a fitting session, whereby FOX will assess the new outcome measures with reference to the new MAP and then possibly suggest further MAP changes. Alternatively, depending on the type of session or time after initial activation, the patient may be advised to use the new MAP until the next fitting session, when outcome measures may then be performed. If no programming changes are required following analysis by FOX, then an appropriate message is returned by the system.

4.4.4. THE "AUTOMAP" FUNCTION

The current version of FOX contains an advice for the production of MAPs in the absence of any pre-existing behavioural fitting measures or outcome measures. These "automaps" are generated based on a statistical analysis of all available MAPs from the CI population that yielded good outcomes (where FOX judged that no further attempts to improve the outcome could be made), and would typically be used at the initial "switch-on" session, when an incremental series of up to 10 automaps can be generated in order to accommodate early increase in dynamic range (loudness tolerance). This makes the initial fitting process more systematic and can save a lot of clinical time. As soon as the CI user has a level of acceptance to electrical stimulation, and the first outcome measures are available, FOX can then be used to individualize and optimize these MAPs. The case example below provides further details of the automap function.

In the future, we plan to develop rules so that FOX can generate automaps for specific subgroups of patients based on their medical history, age and duration of deafness, audiological and other data.

4.4.5. CASE EXAMPLE

A 22 year old lady requested a cochlear implant when she was about to finish university studies. She had been diagnosed with a 60 dB sensorineural hearing loss of unknown aetiology at the age of 3. She received hearing aids immediately and entered mainstream education. Her hearing thresholds had further deteriorated to 90 dB HL by the age of 12.

Imaging suggested normal cochlear morphology, and surgery was uneventful. An Advanced Bionics HiRes90k device was implanted with full insertion of the electrode array, and first fitting took place three weeks later. A series of 10 automaps, with incrementally increasing stimulation levels, was created (from quietest to loudest these are known as the switch-on automap, Silver 1, 2 and 3, Gold 1, 2 and 3 and Ivory 1, 2 and 3). The switch-on MAP was used for the duration of the switch-on session. At the end of this session, the silver MAPs were programmed in one speech processor and the gold MAPs in a second processor and the patient received both processors to take home. She was instructed to start with the MAP Silver 1 and to switch to an incrementally louder MAP every day as long as the sound percept remained tolerable.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
T	83	72	68	67	65	67	62	65	65	68	72	71	68	72	72	70
M	201	203	211	232	244	252	254	256	251	251	245	235	241	240	241	233
Gain	3.5	3.3	3.3	3	1.6	0.6	0.1	0.1	-0.3	-0.7	-0.6	-0.9	-1.3	-1.5	-2.1	-1.5
Is Enabled	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Filter Band Low Cutoff	250	416	494	587	697	828	983	1168	1387	1648	1958	2326	2762	3281	3898	4630
Filter Band High Cutoff	416	494	587	697	828	983	1168	1387	1648	1958	2326	2762	3281	3898	4630	8700
Clipping Level	541	541	541	541	541	541	541	541	541	541	541	541	541	541	541	541
Impedance	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

PulseWidth	20.7	ExtendedLow	Extended	PulseWidthSetting	Manual	ThresholdSetting	Manual
IDR	80	Volume Range Lower	5	ProcessorType	Harmony		
AGC	Enabled	Volume Range Upper	5	ProgramType	HiRes-S		
Sensitivity	-10	AudioMixType	Aux Only	Fidelity120	Disabled		

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
T	83	72	68	67	65	67	62	65	65	68	72	71	68	72	72	70
M	234	258	243	259	274	269	274	279	279	288	292	290	293	291	290	284
Gain	4.6	4.3	4.1	3.5	2.8	1.1	0.6	0.3	0	0	0	0	-1	-1	-1	-1
Is Enabled	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Filter Band Low Cutoff	250	416	494	587	697	828	983	1168	1387	1648	1958	2326	2762	3281	3898	4630
Filter Band High Cutoff	416	494	587	697	828	983	1168	1387	1648	1958	2326	2762	3281	3898	4630	8700
Clipping Level	541	541	541	541	541	541	541	541	541	541	541	541	541	541	541	541
Impedance	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5

PulseWidth	20.7	ExtendedLow	Extended	PulseWidthSetting	Manual	ThresholdSetting	Manual
IDR	80	Volume Range Lower	5	ProcessorType	Harmony		
AGC	Enabled	Volume Range Upper	5	ProgramType	HiRes-S		
Sensitivity	-10	AudioMixType	Aux Only	Fidelity120	Disabled		

Figure 121: MAPs at session 2 (Gold 2, top) and session 3 (Ivory 1, bottom)

The second session was one week after switch-on. The patient had increased up to MAP Gold 2 (Figure 121 top) and outcomes were measured using this MAP. According to the routine follow-up protocol in the Eargroup, audiometry and A§E phoneme discrimination were assessed, and the results are given in Figure 122. Both the MAP settings and the outcomes were entered into FOX, which recommended leaving the MAP unchanged. The speech processor was loaded with MAP Gold 2 plus two higher automaps (Gold 3 and Ivory 1) and the patient was instructed to try the higher automaps occasionally to see whether they were comfortable.

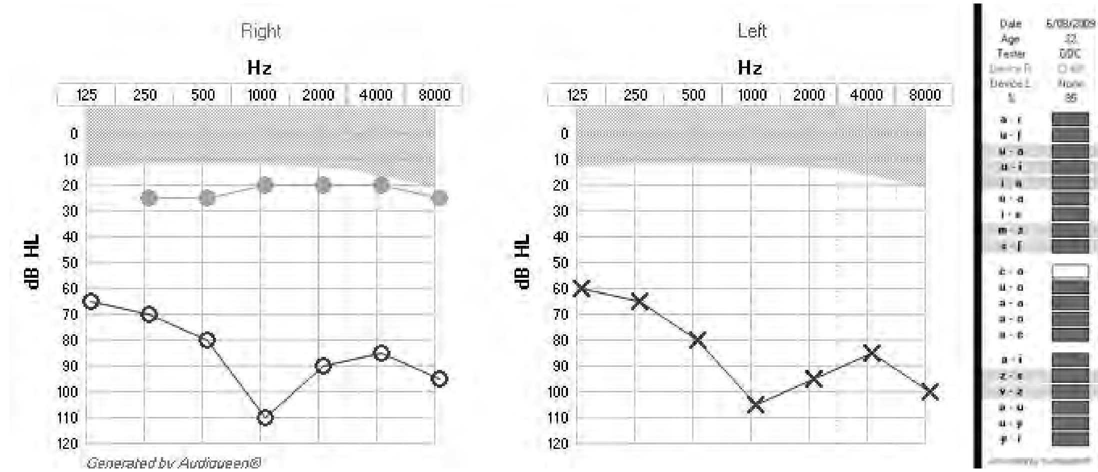


Figure 122: Outcomes at session 2. The audiogram (left) shows the unaided results with headphones before implantation (pure tone average of 93 dB HL on both sides) and the results with the CI in free field after implantation (PTA average of 22 dB HL). The A§E discrimination (right) shows that 19 of the phoneme contrasts were well discriminated (grey fields) and that 1 contrast was not (white field).

The third postoperative session was scheduled for two months later, i.e. 10 weeks after switch-on. At this time the patient had moved up two more automap levels (to Ivory 1, Figure 121 bottom) and outcomes were measured using this MAP. A§E loudness scaling and speech audiometry (open set CVC list at 40, 55, 70 and 85 dB SPL) were assessed, and the results are given in Figure 123. Both the MAP settings and the outcomes were entered into FOX.

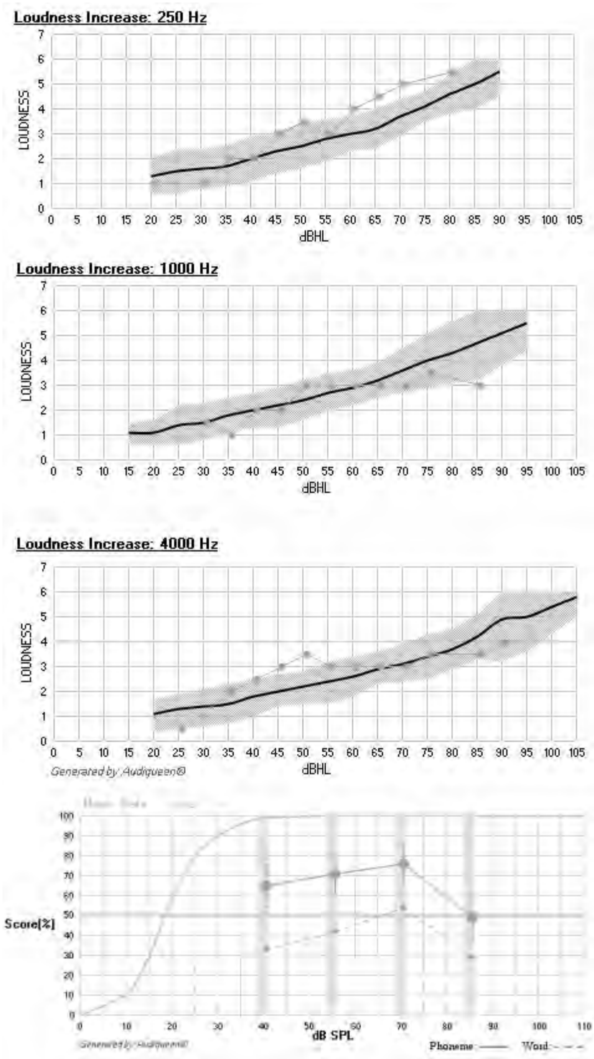


Figure 123: Outcomes at session 3. The loudness scaling at 250, 1000 and 4000 Hz are plotted (dots connected by solid line) on the top three graphs representing the perceived loudness on the vertical axis (ranging from 0 = inaudible to 6 = too loud) as a function of the presented intensity. The thick black line and the grey zone represent the average score and 95% confidence interval respectively in normally hearing listeners. The speech audiogram (bottom graph) shows the phoneme and word scores of open set monosyllable lists presented at 40, 55, 70 and 85 dB SPL.

On this iteration, FOX proposed some MAP changes and to repeat speech audiometry and A§E loudness scaling at 250 Hz (Figure 124). It can be seen that loudness percepts at 250 Hz were louder than ideal and that the speech audiometry shows some rollover at 85 dB SPL (Figure 123). The suggested MAP changes were an overall slight decrease of the M-level, an increase of the gain on 5 most basal electrodes and a slight increase of the pulse width. Figure 125 shows the repeated outcome measures after these were implemented.

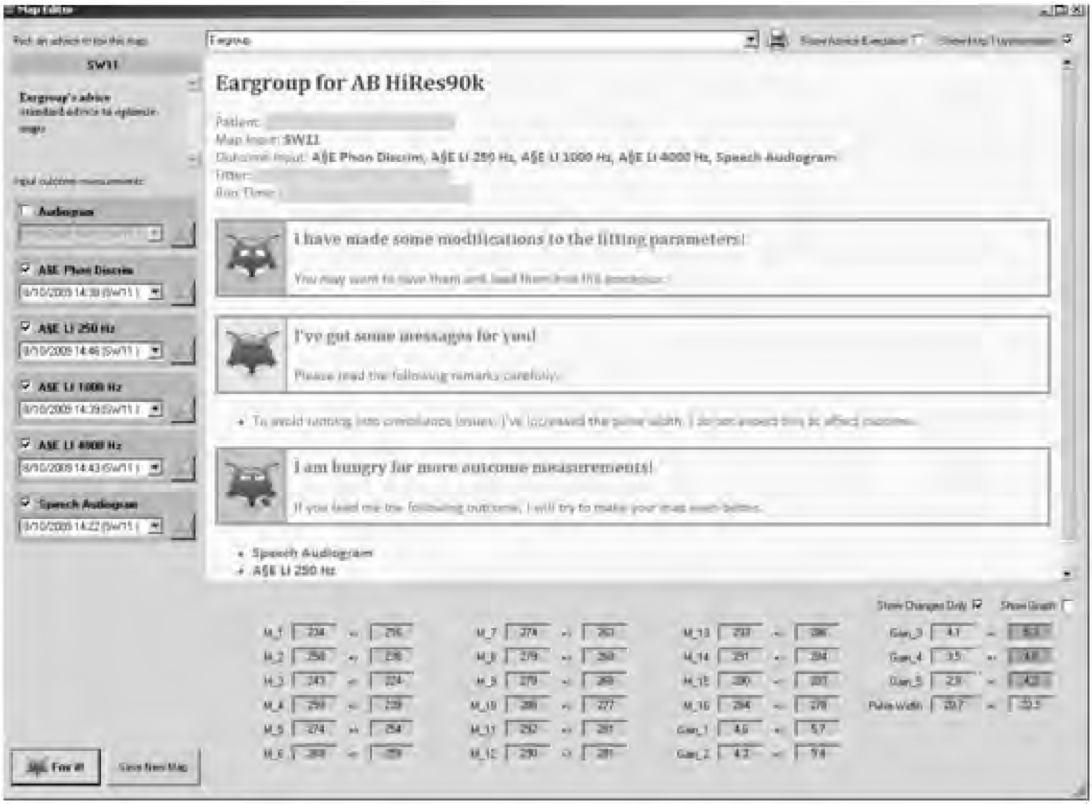


Figure 124: FOX advice at session 3. FOX proposed some MAP changes and to repeat speech audiometry and A§E loudness scaling at 250 Hz.

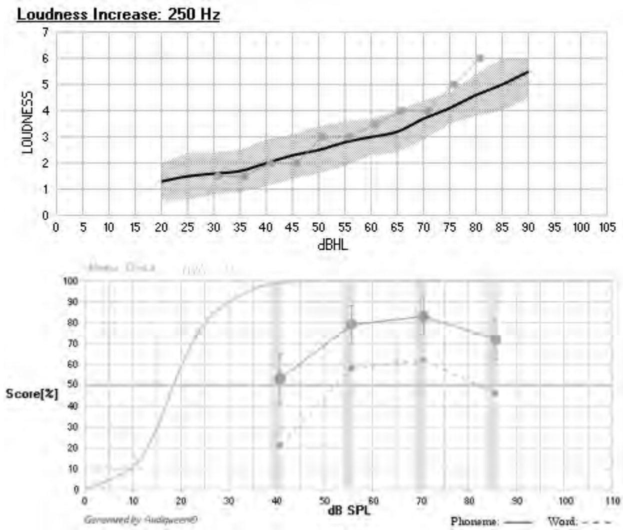


Figure 125: Repeated outcomes at session 3, after implementing the MAP changes proposed by FOX. The loudness scaling at 250 Hz shows values that are more within the normal zone than previously. Speech audiometry shows better scores at 55, 70 and 85 dB SPL and less rollover.

A further iteration of FOX was then run, using the new MAP parameters and the repeated outcome measures. On this occasion FOX proposed a few minor MAP changes (further lower the M-level and increase the gain on the 5 most basal electrodes) but did not request any new outcome measures (Figure 126). These changes were implemented and the patient returned home.

Figure 126: FOX response at session 3 based on the repeated outcome measures.

4.4.6. DISCUSSION

The traditional approach to CI programming has remained essentially unchanged since the introduction of commercial devices some 20 years ago. Generally, the fitting process (as it is performed in usual clinical practice) can be considered to be “comfort driven”, in that the primary goal is to provide electrical stimulation within the dynamic range (from threshold up to most comfortable level). This applies to individual electrodes and to active MAPs when multiple electrodes may be active. Sometimes, CI users may report their auditory percept to be too soft or too loud, or to have an undesirable tonal quality (too boomy, for example). From these reports, adjustments to the stimulation limits are normally made to optimize loudness levels (usually adjustment of M levels). For tonal adjustments, M level and/or gain adjustments would be typical. However, the process of making the percept as comfortable as possible may not necessarily be desirable in terms of long term benefit from the device. What is immediately most comfortable may not provide the best speech understanding. This point can perhaps be clearly illustrated by the situation of fitting hearing aids to patients suffering from presbycusis. Such patients have often suffered from long term high frequency hearing loss and tend to dislike amplified high frequencies initially, even though these are critical for speech understanding.

Another difficulty is that users become adapted to a particular stimulation pattern (MAP), so that any parameter changes tend to result in an initial decrement in perceived sound quality. Due to this, patients may resist potentially beneficial modifications or may be asked to trial new MAPs for periods of days or weeks, so that the process of MAP optimization can take considerable time and sometimes numerous clinical appointments.

This traditional approach to fitting has been a legitimate one as more sophisticated methods have not been available (apart from the incorporation of objective measurements that generally aim to achieve the same goals as behavioural measures). However, in this time clinicians working in this field have developed considerable theoretical, empirical and heuristic knowledge, such that a more systematic approach, such as is offered by FOX, may represent a significant improvement in fitting methodology and, hence, produce better outcomes. As a way of achieving this, a fundamental principle of FOX is to make parameter adjustments that are based on outcomes, rather than comfort. Indeed, it seems irresponsible to adjust such a highly technical device for such an important sensory function without any measurable outcome as feedback. As outlined earlier, FOX can potentially utilise rules based on a wide range of outcome measures, including subjective questionnaires. However, the central focus is on optimization of the signal delivered to the cochlea. This is the level of the auditory pathway where fitting parameters will have their most direct impact. The cochlea is responsible for detection and for fundamental acoustic discrimination, and optimization of these processes will result in optimized identification and recognition at the higher levels of the auditory pathway that serve language processing.

Audiometry and speech discrimination tests have been used as outcome measures for many years and so it is probably a reasonable first step to use these measures where possible. However, we now have additional outcomes that can be used. Thus, at our centre we also place considerable emphasis on the use of phoneme discrimination and loudness growth, two modules of the A§E psycho-acoustical test suite, to gain additional information on cochlear function (as outlined in the introduction).

As mentioned in the introduction to this report, standardization of the fitting process is another important issue, especially in view of the wide range of skills and experience of the clinicians performing these tasks. There is a large number of fitting parameters available to the audiologist, some of which interact with each other, and it is therefore likely that certain potential adjustments are sometimes overlooked by inexperienced audiologists. Even when reliable outcome measures are available, there can be another difficulty in that the relationships between the many patient-related factors, outcome variables and fitting parameters is very complex, making it difficult for an audiologist, in the typical clinical situation, to make systematic judgements on which parameters to adjust in order to gain the best outcomes.

The introduction of a system such as FOX as an “intelligent agent” using deterministic logic provides an opportunity to cope with this complexity. It is a first step towards the introduction of artificial intelligence in the fitting of cochlear implants. At this stage we have opted for a deterministic approach and heuristic rules, in contrast to nondeterministic approaches (such as with neural networks, genetic algorithms etc), mainly because the latter require instantaneous feedback of large amounts of outcome data.

This is manageable in systems such as gaming, labyrinth-tasks, pattern-recognition, etc, but not in the human being where each outcome measure takes of the order of 10 minutes.

With the traditional approach to fitting, the initial stages (the period from “switch-on” until the electrical dynamic range is stable) usually take up a lot of clinical time – perhaps 5-10 sessions for post-lingually deafened adults and more for pre-lingually deaf children. It is questionable whether this time is well spent. Many CI users have no experience or no recollection of normal hearing and so they are often unable to reliably make the judgements required for the audiologist to set fitting parameters.

Furthermore, even if a MAP can be generated with apparently reliable estimates of the lower and upper loudness limits there are usually large changes in these over the initial days and weeks following activation [98], meaning that measurements are often repeated at each visit and the MAP modified accordingly.

Automation, such as may be provided by the use of FOX, may save a lot of time in these early stages. The “quality” of CI fitting (i.e. outcomes) is inevitably dependant on the time spent, whatever approach is used. Figure 127 provides a hypothetical relationship where a certain amount of fitting time is needed to obtain satisfactory outcomes (solid line), but spending ever increasing fitting time will result in diminishing returns. Naturally, we tend to favour a time input that provides the best compromise between time spent and quality of the outcome. From this starting point, the introduction of FOX can thus provide two options; either to spend the same amount of time on fitting as before and thus obtain better outcomes, or to spend less time to obtain similar outcomes to before (dashed line). Again, the choice of which option to follow will typically depend on available resources, financial factors and so on.

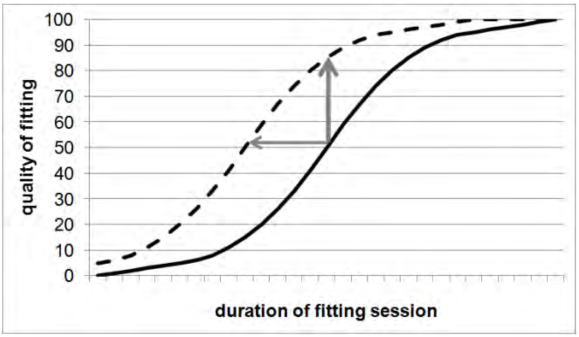


Figure 127: Hypothetical relationship between time spent on CI fitting (abscissa) and the quality of the fitting obtained (ordinate). The solid line represents the relationship using the "traditional" approach to fitting, while the dashed line represents the relationship using FOX. From a starting point on the solid line, the incorporation of FOX can be used to either (i) spend the same amount of fitting time to achieve better outcomes, or (ii) spend less time to achieve the same outcome.

In addition to its application in routine clinical fitting, it is possible that the systematic approach provided by FOX may have uses in other related situations. One such application may be in clinical research, where it is conceivable to design advices to conduct clinical trials, for example to try to find out whether the individual setting of stimulation rate can optimize results. If one designs a rule that uses an outcome to set the stimulation rate, then this can be used for driving a variety of related studies. The systematic approach will not only improve the robustness of the study design, it will also allow to diligently explore other MAP parameters than the ones commonly used to date.

A further advantage of the use of an intelligent agent lies in the possibility to equip it with learning skills, allowing an almost continuous improvement of the rules based on the permanently monitored effects. This could be either "case-wise", e.g. where negative results in a single case can be analyzed and contribute towards rule modifications, or "group-wise", based on the statistical analysis of group data, which will allow us to expand our rules using such data both from our own centre and from many others. At present, rules are only modified after the intervention of and approval by an expert team consisting of at least one audiologist, ENT-specialist and software engineer.

Future developments will include automatic self-learning capacities to become part of FOX.

This report demonstrates that FOX is currently a useful and user-friendly clinical tool. It may be the first step in a new approach to CI fitting, and one which leads the way towards the use of automated expert systems. Several further developments and refinements are currently under consideration, in many cases the main task being the collection and analysis of additional fitting-related data in order to establish the required new rules.

These further developments include the following:

1. Refinement of current rules through analysis of additional clinical data
2. Currently available rules are based on the expertise of our own centre. In terms of AI this is known as a "local optimum". The operation of the intelligent agent in other areas with other local optimums may expand the zone and dimensions of operation. For instance, we have already noticed that we tend to work with relatively large electrical dynamic ranges. Other centres often program much narrower EDRs, which may have consequences on MAP modifications in order to obtain a desired outcome effect.
3. We would like to conduct clinical trials in which we address a number of the currently used fitting parameters in a more systematic way. For instance, does it make a difference to systematically set the T-level at 10% of the M-level, to set it at higher levels or possibly to even set it to 0 clinical units?
4. We would like to include additional fitting parameters that have not yet been addressed at this stage. These could include stimulation rate, the choice of sequential versus simultaneous stimulation, the frequency band limits for each electrode, etc.
5. We are also keen to introduce new outcome measures, such as the results of electrophysiological tests or questionnaires, using the expertise of clinicians experienced in interpretation of such data in order to create rules based on these outcomes.

4.4.7. ACKNOWLEDGEMENTS

The authors acknowledge Dr Michael Boedts who performed the intervention in the case example.

4.5. A PROBABILISTIC GRAPHICAL MODEL FOR CI FITTING

A probabilistic graphical model for tuning cochlear implants

Artificial Intelligence in Medicine, Lecture Notes in Computer Science Volume 7885, 2013, pp 150-155

Iñigo Bermejo[1], Francisco Javier Diez[1], Paul J Govaerts[2], Bart Vaerenberg[2,3]

[1] ETSI Informatica, UNED, Madrid, Spain
[2] The Eargroup, Antwerp-Deurne, Belgium
[3] Laboratory of Biomedical Physics, University of Antwerp, Belgium

Abstract

Severe and profound hearing losses can be treated with cochlear implants (CI). Given that a CI may have up to 150 tunable parameters, adjusting them is a highly complex task. For this reason, we decided to build a decision support system based on a new type of probabilistic graphical model (PGM) that we call tuning networks. Given the results of a set of audiological tests and the current status of the parameter set, the system looks for the set of changes in the parameters of the CI that will lead to the biggest improvement in the user's hearing ability. Because of the high number of variables involved in the problem we have used an object-oriented approach to build the network. The prototype has been informally evaluated comparing its advice with those of the expert and of a previous decision support system based on deterministic rules. Tuning networks can be used to adjust other electrical or mechanical devices, not only in medicine.

4.5.1. INTRODUCTION

Cochlear implants (CI) are being successfully applied to treat severe and profound hearing losses. A CI consists of a speech processor that analyzes the sound and an array of electrodes placed into the cochlea, which pass an electrical signal directly to the auditory nerve. After implantation, CIs need to be programmed or “fitted” to optimize the user's hearing capability. This is usually a challenging and time-consuming task that is typically performed by highly trained audiologists or medical doctors. CI centres and manufacturers have developed their own heuristics, usually in the form of “if-then" rules applied in a very flexible but individual and often uncontrollable way. Aiming at improving the process of CI fitting, an application called FOX [101], was developed by Otoconsult and the Eagroup. It is being used in several centres across Europe. FOX is based on parameterized deterministic rules, which entails some limitations, such as the difficulty to maintain the knowledge base when the number of rules increases and the inability to learn from data. The Opti-FOX project was conceived to overcome these limitations. In the beginning, an approach with supervised classification algorithms such as the k-NN classifier was attempted, but failed to progress due to the complexity of the problem and the small number of records available to learn from. In order to improve the results of FOX, the most promising approach seemed to build a probabilistic graphical model (PGM) because this type of model can combine expert knowledge, the power of probabilistic reasoning, and the ability to learn from data. This paper describes briefly a new type of PGM especially tailored for tuning programmable devices and how it has been used to build a decision support system for fitting CIs.

4.5.2. TUNING NETWORKS

A tuning network consists of an acyclic directed graph (ADG) containing chance, decision and utility nodes, and a probability distribution. As in other types of PGMs, a decision node represents a variable that is under the direct control of the decision maker, while chance nodes represent features of the system over which the decision maker has no direct control, and utility nodes represent the decision maker's preferences, measured on a numerical scale. In tuning networks, each property of the system is modelled by a relative-value node that represents a change in its value and, optionally, by an absolute-value node. In the case of a tuneable parameter (for example, the sensitivity of the microphone), the relative-value node is a decision node because the programmer of the CI can increase, decrease, or keep the value of the parameter; the absolute-value node is represented as a chance node for which we have evidence, because the value of a tuneable parameter is always known. We may also have evidence about the absolute-value nodes that represent measurements, such as the result of a test. Utility nodes are always relative-value nodes, as they represent the increase or decrease in the user's performance as a consequence of tuning some parameters. An important component of tuning networks is the tuning model, a new canonical model based on the property of independence of causal interaction (ICI) [246] [247] [248]. Canonical models represent how a variable is probabilistically influenced by a set of parent variables [249], in general assuming a pattern of causal interaction. Their main advantage is that the number of parameters (conditional probabilities) is proportional to the number of parents, while in the general case it grows exponentially. ICI models assume that each parent produces the effect with a certain probability, independently of the values of the other parents, and the global effect is determined by a function, specific of each type of ICI model, that combines the individual effects; for example, in the noisy OR the effect is present when at least one of the causes has produced the effect. A unique feature of the tuning model is that it assumes that every variable involved has exactly three values: increased, decreased, and not-changed, while other ICI models, such as the noisy OR and the noisy AND, assume that all variables are boolean, and other models, such as the noisy MIN and the noisy MAX impose no restriction about the number of values of each variable [248]. The tuning model assumes that a change in one of the parents causes a change in the child variable with a certain probability. When some of the parents induce an increase and others cause a decrease, the global effect depends on whether there are more increases than decreases, or vice versa, or there is a tie. It is therefore similar to a majority voting function. Tuning networks differ from influence diagrams [250] in that they do not have a total ordering of the decisions because the order in which the parameters are tuned does not affect the result. Additionally, all the evidence is available before making the decisions, as it is in each session when programming a CI, while in influence diagrams some decisions provide evidence that can be used in subsequent decisions. As a consequence, the algorithms for evaluating these two types of models are very different, see Section "Inference".

4.5.3. CONSTRUCTION OF THE MODEL

4.5.3.1. MODEL CONSTRUCTION

VARIABLES IN THE MODEL

In the tuning network, the tuneable parameters are those of the CI; as mentioned above, each one is represented by an absolute value chance node and a relative-value decision node. Each electrode has several tuneable parameters, e.g., the T level (the softest electrical input level detectable by the user), the M level (the electrical input level perceived as loud but comfortable), etc. Besides, the CI has a set of tuneable parameters that are electrode-independent, i.e., global to the implant, such as the volume of the microphone. The model also represents the results of a battery of different tests, such as audiometries, phoneme discrimination and speech recognition tests. Each measurement of a test is modelled with a chance node representing the current value of the test (this node receives evidence when performing the test), a chance node representing the expected change in the result of the tests given the changes in the tuneable parameters, and a utility variable defining the utility function based on the other two. Other nodes represent internal properties of the device, such as the amount of energy in the auditory nerve, which depend on the tuneable parameters and in turn affect the results of the tests. The global utility of the model is the sum of the results of all tests; therefore maximizing this utility is the same as optimizing the user's hearing ability. The resulting model contains 202 nodes and 664 links.

ELICITATION OF NUMERICAL PARAMETERS

The probabilities and utilities have been assessed by the expert: the probabilities are subjective estimates based on his expertise while the utilities have been estimated by roughly assigning monetary value to positive and negative changes in the results of tests.

OBJECT-ORIENTED PROBABILISTIC NETWORKS

The network, containing sets of repeated structures (such as electrodes, frequency bands and tests) was modelled following the object-oriented paradigm for PGMs as proposed by [251] [252]. A class defines a structure consisting of a set of attributes and their probabilistic relations and is connected with other classes through their input parameters, namely instances of other classes. An OOPN consists of a set of instances and their causal relations.

4.5.3.2. INFERENCE

Inference in a tuning network consists in looking for the optimal strategy, i.e., the set of changes in the tunable parameters that maximizes the global expected utility. As an exhaustive search would be computationally unaffordable, we have implemented a greedy search and score algorithm that examines myopically the space of possible strategies. The search is initialized by setting all policies for all decision nodes to “no change”. It then iteratively looks for the single change in the strategy, i.e. a change in a decision node's policy that maximizes the global utility function. The score for each strategy, namely the global expected utility given the strategy, is computed using an inference algorithm. Given the high number of variables in the model and its high connectivity, the cost of running exact inference algorithms is unaffordable. For that reason, we decided to use an approximate inference algorithm, namely a likelihood weighting method [253] adapted to networks with utility nodes, whose spatial and temporal complexities grow linearly with the number of nodes instead of exponentially. The main drawback of likelihood weighting is that its accuracy decreases with extremely unlikely evidence, but it still fits our needs as the observed nodes usually have no extreme probabilities. We compared the results of this greedy algorithm, in simplified versions of the model with those of an exact inference algorithm (variable elimination) and both returned the same optimal strategy under different evidence scenarios. The execution time of the greedy algorithm, which has been implemented to run in parallel taking advantage of multiple core processors, depends on the number of changes proposed by the optimal strategy, but in a regular desktop computer (Intel Core i5-2500 @ 3.30GHz and 8GBs of RAM) is usually under a minute.

4.5.3.3. EVALUATION OF THE MODEL

We have initially built a prototype for the low-frequency electrodes, i.e., those in the range from 250 to 1000 Hz. This model has been tested on a set of cases taken from a database of real CI users. The recommendations output by the probabilistic model have been compared with those of FOX, the expert system based on deterministic rules, having expert CI fitters as judges. In many cases, the recommendations of the probabilistic model agreed with both FOX and the experts. There were, however, some cases in which the probabilistic model recommended some interventions that surprised the experts, but they never deemed them nonsensical. On the contrary, they described them as “intelligent”, “smart” and “worth trying”.

Figure 128: Screenshot of the prototype network.

On July 31, 2012 a patient at Otoconsult had a poor performance in the speech understanding tests, in spite of having an audiometry in the range of normality. The audiologists using their expertise and FOX's support, were not able to improve her ability to understand spoken words. However, when her implant was fitted using the advice of our prototype network, her performance increased to the level of normality. Of course, this isolated result does not prove that the probabilistic model outperforms FOX or the audiologists in general, but it is a promising result.

4.5.4. CONCLUSIONS AND FUTURE WORK

In the context of the European project Opti-FOX, we have built a PGM for programming CIs. The development of tuning networks and our framework for OOPNs has been motivated by the needs encountered in this project, but they can be applied to adjust other electrical or mechanical devices, not only in medicine. The advantages of the probabilistic model with respect to FOX, the rule based system, are that it is capable of complex reasoning whereas FOX mostly concatenates rules, that FOX is deterministic while the probabilistic model handles uncertainty, and that the probabilistic model will be fine-tuned by learning from data. However, FOX is still a more mature project that has been evaluated extensively and includes features that the probabilistic model still lacks, such as the ability to determine the quantity by which the value of a parameter should be changed. The most obvious next step in the project is to test the developed prototype on real CI users. Besides, we are currently working on learning the conditional probabilities from a database, in order to fine-tune the probabilities elicited by the expert. Given that the probabilistic model contains unobservable variables, the usual parametric learning algorithms cannot be applied. Instead, we are using the Expectation Maximization (EM) algorithm, applied to the learning of Bayesian networks as proposed by Lauritzen [254]. Another aspect with room for improvement is the granularity of the variables. Relative-value variables were discretized into three intervals (increase, decrease, no change) to reduce the complexity of the problem. This oversimplification prevents the probabilistic model from accurately predicting the effect of small changes in the parameters of the CI. Finally, the programming of a CI has a temporal aspect: it usually involves several sessions and the history of each patient is relevant. Unfortunately, the current model only considers the current values of the parameters. Turning our system into a partially-observable Markov decision process (POMDP) would allow us to model that temporal evolution and determine the optimal sequence of tests and parameter adjustments.

4.6. THE TUNING MODEL

The tuning model

In preparation

Francisco Javier Diez[1], Paul J Govaerts[2], Bart Vaerenberg[2,3], Iñigo Bermejo[1]

[1] ETSI Informatica, UNED, Juan del Rosal 16, Madrid, Spain
[2] The Eargroup, Antwerp-Deurne, Belgium
[3] Laboratory of Biomedical Physics, University of Antwerp, Belgium

Abstract

This paper describes a new type of canonical model: the tuning model, which represents how a change in some variables (the parents) affects another variable (the child). It belongs to the family of models based on the property of independence of causal interaction. The motivation for this work was the need to model the interactions between variables in a Bayesian network for programming cochlear implants.

4.6.1. INTRODUCTION

Construction of directed graphical probabilistic models, such as Bayesian networks [249] and influence diagrams [250], requires specification of many conditional probability distributions of the form P(y|x), where X = {X_1 , . . . , X_n} is the set of parents of a node Y in the network, see Figure 129. The set {Y} ∪ X is called a family. When all the variables in a family are discrete, P(y|x) can be expressed in form of a conditional probability table (CPT), whose size grows exponentially with the number of nodes. In general, the numerical parameters are obtained from databases or assessed by human experts and, for this reason, it is usually difficult to build a CPT for a family having more than three or four parents.

One way of reducing the complexity of elicitation of numerical probabilities is to rely on so-called canonical models, which allow for building probability distributions from a small number of parameters. The term “canonical” is used because such models are elementary units used in the construction of more complicated models [249]. Different canonical models may coexist in any probabilistic network. For instance, in causal Bayesian networks that model real-world domains, it is not uncommon that a significant number of families interact through OR/MAX-models, a few through AND-models and the rest of the families do not correspond to any canonical model, which implies that their CPT must be explicitly given.

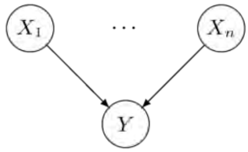


Figure 129: A family consisting of a child node Y and its parents, X = {X1, . . . , Xn }.

Canonical models are useful not only because they simplify the construction of probabilistic models (knowledge engineering), but also because they save storage space and computation time [248] and because they correspond to causal patterns that can be exploited to generate user explanations [255].

In this paper we present a new canonical model, called the tuning model, that represents how a change in some variables (the parents) affects another variable (the child). The motivation for this work is that, when building a Bayesian network for programming cochlear implants, we needed a model that could represent how a change in some of the parameters of the physical device affect other properties of the system, which in turn may affect the subject's ability to detect and recognize the sound. Given that none of the existing models fitted our needs, we devised a new model based on the property of independence of causal interaction (ICI), which is discussed at length in Diez and Druzdzel, 2006 [248].

The rest of the paper is structured as follows: Section "Background" introduces the notation and the fundamentals of ICI models, Section "The tuning model" defines the tuning model, Section "Application of the tuning model in practice" explains how to apply it to model real-world problems, Section "Discussion" discusses the main features of the new model, and Section "Conclusions and future work" presents the conclusions and proposes some lines for future research.

4.6.2. BACKGROUND

4.6.2.1. NOTATION

We will use capital letters to represent variables and lower case letters to represent their values. For instance, v will represent a possible value of variable V. In the same way, V will denote a set of variables $\{V_1, \ldots, V_n\}$, and v a certain n-tuple $(v_1, \ldots, v_n)$, where v_i represents a value taken by variable V_i .

4.6.2.2. ICI MODELS

One kind of probabilistic model is that based on the assumption of independence of causal influence (ICI).

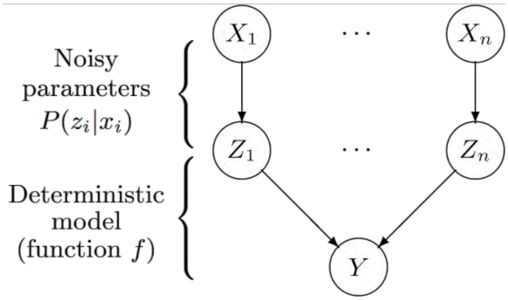


Figure 130: Auxiliary structure for the derivation of a noisy ICI model.

Noisy ICI models can be defined by introducing n auxiliary variables $\{Z_1, \ldots, Z_n\}$, as shown in Figure 130, such that Y is a deterministic function of the Z_is, y = f(z), and the value of each Z_i depends probabilistically on X_i, as captured by the CPT $P(z_i|x_i)$. In most ICI models, the Z_is have a causal interpretation. However, we can just see them as auxiliary variables that are used for deriving the equations and are not part of the model. The conditional probability P(y|x) is obtained by marginalizing out the Z_is:

$$P(y|\mathbf{x}) = \sum_{z} P(y|\mathbf{z}) * P(\mathbf{z}|\mathbf{x}) \tag{11}$$

where

$$P(y|\mathbf{z}) = \begin{cases} 1, & \text{if } y = f(\mathbf{z}) \\ 0, & \text{otherwise} \end{cases} \tag{12}$$

Therefore,

$$P(y|\mathbf{x}) = \sum_{\mathbf{z}|f(\mathbf{z})=y} P(\mathbf{z}|\mathbf{x}) \tag{13}$$

Independence of causal influence (ICI) means that there are no interactions among the causal mechanisms by which the X_is affect the value of Y. Given the graph in Figure 2, this property is equivalent to the absence of links $X_i \rightarrow Z_j$ and $Z_i \rightarrow Z_j$ for all $i \neq j$, which means that

$$P(\mathbf{z}|\mathbf{x}) = \prod_{i} P(z_i|x_i) \tag{14}$$

and, consequently,

$$P(y|\mathbf{x}) = \sum_{\mathbf{z}|f(\mathbf{z})=y} \prod_{i} P(z_i|x_i) \tag{15}$$

Each parameter $P(z_i|x_i)$ of a canonical model is associated with a particular link $X_i \rightarrow Y$, while each parameter $P(y|x)$ in a CPT corresponds to a certain configuration x made up by all the parents of Y , and cannot be associated with any particular link. This property, stemming from the ICI assumption, entails two advantages from the point of view of knowledge engineering. The first is a significant reduction in the number of parameters required to specify a model, from O(exp(n)) in a general model to O(n) in a canonical model. This can amount to a substantial reduction of the elicitation effort. For example, a binary node with 10 binary parents will have a CPT consisting of 2^{11} = 2,048 numerical parameters. Adding one more node doubles this number to 2^{12} = 4,096 parameters. In contrast, a noisy OR model would require only 10 and 11 parameters, respectively. The second advantage is that the parameters in canonical models lend themselves to fairly intuitive interpretations, which facilitates the task of eliciting them from human experts. As mentioned above, canonical models not only require fewer parameters than ordinary CPTs, but also their parameters are more intuitive and easier to estimate.

It is possible to define leaky ICI models, which only differ from their noisy counterparts in the addition of another auxiliary variable, Z_L, which represents the effect of the variables not explicitly represented in the model [248].

4.6.3. THE TUNING MODEL

4.6.3.1. MATHEMATICAL DEFINITION OF THE TUNING MODEL

A noisy ICI model is defined by three elements: the domains of the variables, the function f, and some constraints on the values of $P(z_i|x_i)$.

In the tuning model, all the variables have the same domain, { − , 0 , + }, where − represents a decrease in the value of the variable, + represents an increase, and 0 means “no change”. If the variable is denoted by V, we will sometimes write v^- / v^0 / v^+ instead of − / 0 / + to make it more clear what variable we are speaking of.

The function of the tuning model is defined as follows:

$$f_{tuning}(\mathbf{z}) = \begin{cases} y^+, \text{if } n^+(\mathbf{z}) > 0 \\ y^0, \text{if } n^+(\mathbf{z}) = 0 \\ y^-, \text{if } n^+(\mathbf{z}) < 0 \end{cases} \tag{16}$$

where $n^+(\mathbf{z})$ is a function that returns the number of variables that take the value + in configuration z minus the number of those that take the value −. For example, $n^+(z_1^+, z_2^+, z_3^+) = 3$, $n^+(z_1^+, z_2^-, z_3^0) = 0$, and $n^+(z_1^+, z_2^-, z_3^-) = -1$. Therefore, $f_{tuning}(z_1^+, z_2^+, z_3^+) = y^+$, $f_{tuning}(z_1^+, z_2^-, z_3^0) = y^0$, and $f_{tuning}(z_1^+, z_2^-, z_3^-) = y^-$

A constraint that we impose on $P(z_i|x_i)$ is that

$$P(z_i^0 |\, x_i^0) = 1 \qquad (17)$$

which implies that $P(z_i^+ |\, x_i^0) = P(z_i^- \,|\, x_i^0) = 0$. Therefore, if we introduce four parameters for each link, c_i^{++} , c_i^{+-} , c_i^{-+} and c_i^{--} , the CPT for that link has the form shown in Table 11Table 11: Conditional probability table for link X_i → Y in the tuning model.. It is possible to prove from Equation (15) that when all the Xs take the value 0, then Y takes the value 0 with absolute certainty. When X_i takes the value + and the other Xs take the value 0, then Y takes the value + with probability c^{++} and the value − with probability c^{+-}. Similarly, when X_i takes the value − and the other Xs take the value 0, then Y takes the value + with probability c^{-+} and the value − with probability c^{--}. Therefore, the four c-parameters quantify the individual impact of X_i on Y.

Table 11: Conditional probability table for link X_i → Y in the tuning model.

$P(z_i \| x_i)$	x_i^-	x_i^0	x_i^+
z_i^+	c_i^{-+}	0	c_i^{++}
z_i^0	$1 - c_i^{-+} - c_i^{--}$	1	$1 - c_i^{++} - c_i^{+-}$
z_i^-	c_i^{--}	0	c_i^{+-}

4.6.3.2. CLASSES OF INTERACTIONS

We have seen that the general form of the conditional probability table associated with link X_i → Y is as shown in Table 11. However, it is possible to impose a second constraint: for each variable X_i and each value of this variable, P ($z_i^+ | x_i$) = 0 or P ($z_i^- | x_i$) = 0; put another way:

$$(c_i^{-+} = 0 \lor c_i^{--} = 0) \land (c_i^{+-} = 0 \lor c_i^{++} = 0) \qquad (18)$$

Therefore, when this constraint holds for a link X_i → Y , only two parameters are different from 0 — in contrast with the general case, which requires four independent parameters — and that link must belong to one of four classes: direct, inverse, always increasing, and always decreasing.

The direct class is shown in Table 12. The values in the x_i^0 column are imposed by the first constraint of the tuning model (Eq. (17)).

The x_i^- column implies that a decrease in X_i causes a decrease in Y with a probability c_i^{--}, such that c_i^{--} > 0. It may occur, with probability 1 − c_i^{--}, that a decrease in X_i fails to cause a change in Y, but that decrease can never cause an increase in Y. Similarly, the x_i^+ column means that an increase in X_i causes an increase in Y with a probability c_i^{++}. Therefore, this class represents a positive influence of X_i on Y [256] and a positive correlation between both variables.

Similarly, the inverse class is characterized by $c^{++} = c^{--} = 0$, $c^{-+} > 0$, and $c^{+-} > 0$, which implies that a decrease in X_i causes an increase in Y, and vice versa, thus leading to a negative correlation between both variables.

Table 12: Conditional probability table for a link $X_i \rightarrow Y$ of the direct class.

$P(z_i\|x_i)$	x_i^-	x_i^0	x_i^+
z_i^+	0	0	c_i^{++}
z_i^0	$1-c_i^{--}$	1	$1-c_i^{++}$
z_i^-	c_i^{--}	0	0

The relations that define the always decreasing class are: $c^{-+} = c^{++} = 0$, $c^{--} > 0$, and $c^{+-} > 0$. Therefore, any change in X_i will cause a decrease in Y. The properties of the always increasing class are analogous.

We may impose a third constraint: the symmetry of the influence. In the case of a direct interaction, it implies that $c_i^{++} = c_i^{--}$, i.e., the probability that an increase in X_i causes an increase in Y is the same as the probability that a decrease in X_i causes a decrease in Y. In the case of an always decreasing interaction, the probability of a decrease in Y is the same for an increase in X_i as for a decrease: $c_i^{+-} = c_i^{--}$. A link satisfying the condition of symmetry requires only one parameter.

Several kinds of interaction may coexist within the same family. For example, in a family with four parents, the link $X_1 \rightarrow Y$ might be general (i.e., free from the second and third constraints, as shown in Table 11), $X_2 \rightarrow Y$ might be a direct interaction, $X_3 \rightarrow Y$ might be direct and symmetric, and X_4 might be always decreasing. The total number of parameters for this model would be 4 + 2 + 1 + 2 = 9.

If the interaction of this family did not use any canonical model, its conditional probability table would require $3^5 = 243$ parameters, but given that there are $3^4 = 81$ constraints among them, this family would require 243 − 81 = 162 independent parameters. Obtaining those parameters from a database is unreliable, unless in the case of a huge database, because many of the configurations of the Xs will not be represented. Obtaining those parameters from an expert would be impossible in practice not only for the amount of time required, but mainly because estimating the probability of Y for each configuration of the Xs exceeds by far the cognitive capabilities of the human mind.

4.6.4. APPLICATION OF THE TUNING MODEL IN PRACTICE

4.6.4.1. CAUSAL INTERPRETATION OF THE TUNING MODEL

In the tuning model Y represents a parameter of a system whose value depends on the values taken on by other parameters, $\{X_1, \ldots, X_n\}$. Each auxiliary variable Z_i, associated with link $X_i \rightarrow Y$, as shown in Figure 130, indicates whether a change in X_i (from x_i^0 to x_i^+ or x_i^-) has induced a change in Y : z_i^+ indicates an increase (from y^0 to y^+) while z_i^- indicates a decrease (from y^0 to y^-).

The first constraint, given by Equation (17), means that when there is no change in X_i, then there is no change in Y.

The second constraint, given by Equation (18), means that a change in X_i may cause either an increase or a decrease in Y, but not both; this assumption seems reasonable for some domains, but there might be others in which an increase (or a decrease) in X_i sometimes produces an increase in Y and sometimes a decrease.

The function $f_{tuning}(z)$, given by Equation (16), means that when some of the X_is induce an increase in Y and others cause a decrease, the global effect depends on whether there are more increases than decreases, or vice versa, or there is a tie.

The tuning model is used to build Bayesian networks in which each variable represents a property of the system. In these networks, a node without parents represents a physical parameter that can be adjusted by the user, while a node Y with parents $\{X_1, \ldots, X_n\}$ represents a parameter or a property of the system whose value depends on other parameters (its parents). There are two different interpretations of such a Bayesian network.

The diagnostic interpretation assumes that the optimal tuning is unique, i.e., there is only one configuration of the parameters (the inputs) that makes the system perform optimally. The three values of each variable are interpreted as {decreased, optimal, increased}. The current output of the system is introduced as evidence into the Bayesian network and the goal of inference is to “diagnose” which parameters are not properly tuned. Therefore, inference proceeds down-up, i.e., from the observed outputs to the inputs. This method has two advantages: first, it does not require a global gain function, and second, inference is more efficient, because it evaluates the network only once, while the variational interpretation needs to evaluate the network once for each change in one of the parameters.

The variational interpretation tries to predict the impact that a change in the value of the parents (the causes or the inputs) will have on the children (the effects or the outputs). The three values of each variable, {−, 0, +} are interpreted as {decrease, no-change, increase}. Initially all the variables take on the value no-change by definition. The process of inference consists of computing the posterior probability of each output variable for each change that the user may impose on the input variables. Therefore, in this interpretation, inference proceeds top-down, i.e., from (the possible adjustments of) the inputs to the outputs, using predictive reasoning. The changes that lead to an improved performance will be applied. When an improvement in some of the output variables comes together with a worsening in others, it is necessary to have a utility function that measures the global gain in performance.

4.6.4.2. CONSTRUCTION OF A TUNING BAYESIAN NETWORK

The construction of a Bayesian network for the tuning of a physical system begins by selecting the variables. Some of them will represent variations in the parameters of the system. The domain of each of these variables will be {−, 0, +}, which implies a discretization of a continuous variable. The value 0 might indicate that the value of the parameter has not changed at all, and −/+ might represent any increase/decrease in its value, no matter how small it might be. However, in practice it is better that 0 indicates "no significant change", − represents a significant decrease and + represents a significant increase. It is the knowledge engineer, in collaboration with human experts, who must determine what constitutes a significant change. For instance, the threshold might be ±5% of the absolute value of the parameter represented by the variable; for a different variable tuned with higher or lower precision the threshold might be ±2% or ±10%, respectively. This threshold might be different for each variable in the Bayesian network, but it must be very clearly defined, because it will affect the elicitation of the conditional probabilities.

The second step in the construction of a Bayesian network is to draw causal links between the variables, which is usually the easiest task in the construction of the network.

The third step is to analyze for each family in the Bayesian network the possibility of applying a canonical model. The conditions for applying an OR, a MAX, a MIN, or an XOR model are discussed in [248]. The first condition for applying a tuning model is that all the variables involved in the family have the same domain: {−, 0, +}. The second condition is that the effects of the parents can be combined by applying the function f_{tuning} defined in Equation (16), which basically states that each change in one of the Xs may produce an increase or a decrease in Y, and the resulting value of Y depends on whether there are more increases than decreases, or vice versa, or there is a tie. The third condition is that the increase or decrease produced by each X_i only depends on the value taken by this variable, not on the values of the other Xs; this condition seems difficult to assess for a human expert because in fact the individual effects are combined by the function f_{tuning} and consequently it is difficult to think of the individual effects "before" being combined. Therefore, it is reasonable to give the third condition for granted and assume that the tuning model can be applied whenever the first two conditions hold.

The fourth step is to obtain the numerical parameters, i.e., the conditional probabilities for each family in the Bayesian network. In the tuning model, each link $X_i \rightarrow Y$ must be analyzed independently of the others. The first question is: "Does the second constraint, given by Equation (18), hold for this link, or is it possible that the same change in X_i sometimes causes an increase in Y and other times a decrease?" In the latter case it will be necessary to obtain four parameters: c^{++}, c^{+-}, c^{-+}, and c^{--}. However, some of the parameters might coincide; for example, using causal knowledge we might state that $c^{++} = c^{--}$ and $c^{-+} = c^{+-}$ (assumption of symmetry). This would reduce the number of independent parameters to be estimated.

On the contrary, if the second constraint holds, the next question to be asked is: "What class of interaction is this: direct, inverse, always increasing, or always decreasing?". The last question is about symmetry. For example, in the case of a direct interaction, the question is: "Does a decrease in X_i cause a decrease in Y with the same probability that an increase in X_i causes an increase in Y?". If there is symmetry, we only need to elicit one parameter; otherwise, we need two.

Then, we have to estimate the numerical value(s) of the parameter(s) of each link. The question to be asked for each parameter can be derived from the mathematical definition of the tuning model. For example, the question for parameter c^{++} is: "What is the probability that an increase in X_i causes an increase in Y when there is no change in the other parents of Y?" The questions for the other parameters are analogous.

Finally, we must consider for that family whether a noisy tuning model suffices or it is necessary to apply a leaky tuning model [248]. The question is: "Is it possible that a change in some of the physical parameters not explicitly represented in the Bayesian network causes a change in Y ?" If the answer is affirmative, the question: "What is the probability that they cause an increase in Y (when none of the explicit parents change)?" will give us an estimate for the leak parameter c_L^+ , while the question: "What is the probability that they cause a decrease in increase in Y ?" will yield c_L^-.

4.6.4.3. CONDITIONED INTERACTIONS

It may occur in practice that the effect of a certain variable - say X_1 - on Y depends on the value of a third variable, C. For example, X_1 may represent a change in the value of a physical parameter (an increase or a decrease) while C represents the absolute value of that parameter. In this situation we can apply a modelling trick consisting of adding an auxiliary variable A_1, as shown in Figure 131.

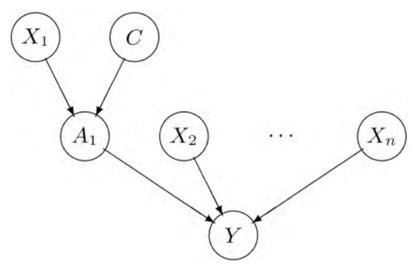


Figure 131: Conditioned interaction: the effect of X1 on Y depends on the value of variable C . The combined effect of X1 and C is modelled by the auxiliary variable A1.

The interaction between A_1 and Y is given by the identity matrix: $P(a_1|y) = \delta_{a1,y}$, where δ is Kronecker's delta function; put another way, the link $A_1 \rightarrow Y$ is a deterministic symmetric direct interaction (see Table 12) with $c_i^{--} = c_i^{++} = 1$.

Table 13: Conditional probability table for the auxiliary variable A_1.

	$C = low$			$C = medium$			$C = high$		
$P(a_i \mid x_i)$	x_i^-	x_i^0	x_i^+	x_i^-	x_i^0	x_i^+	x_i^-	x_i^0	x_i^+
a_i^+	0	0	0.9	0	0	0.4	0.04	0	0.26
a_i^0	0.1	1	0.1	0.4	1	0.6	0.68	1	0.67
a_i^-	0.9	0	0	0.6	0	0	0.28	0	0.07

Table 13 shows with a hypothetical example how the effect of X_1 on Y may depend on a conditioning variable, C. When the value of C is low, the interaction is bottom-up and symmetric: in 90% of cases, a decrease in X_1 causes a decrease in Y, and vice versa. When C = medium, the interaction is also bottom-up and symmetric, but the effect is qualitatively smaller, i.e., it occurs in a lower proportion of cases. When C = high, the effect of a change in X_1 is unpredictable: it may cause an increase in Y but it may also cause a decrease, and the variability is asymmetric: it is higher for an increase in X_1 than for a decrease.

4.6.5. DISCUSSION

The tuning model arose from a need encountered when building a Bayesian network for a real-world problem: the programming of cochlear implants, which is an electronic device that allows deaf people to hear almost normally. In this domain of application, it soon became clear that the diagnostic approach was inappropriate, hence we decided to use the predictive approach, which seems to be leading us to positive results. In this project we are using OpenMarkov, an open-source tool for editing, learning, and doing inference with probabilistic graphical models. We have implemented in this tool the tuning model, which can be applied to any family in which every variable can take on three values.

The fact that the tuning model arose from a real problem is a difference with some of the canonical problems proposed in the literature, which came out from mathematical speculation and have never been implemented in a software tool nor used in practice.

4.6.6. CONCLUSIONS AND FUTURE WORK

In this paper we have presented a new canonical model that is now in the toolbox of knowledge engineers building probabilistic graphical models. In particular, it has been implemented in OpenMarkov, an open-source tool and used to build a model for a real-world application: the programming of cochlear implants.

A possible line for future research is to explore the behavior of the tuning model when using different combination functions; thus, instead of the "democratic" function f_{tuning}, defined in Equation (16), which assigns the same weight to each parent, we might have a function in which each parent "votes" with a different weight. However, such a function would require more parameters, and the increase in the complexity of the model, instead of improving its accuracy, might be counterproductive.

Another line of research is to improve the integration of the tuning model with exact and/or stochastic algorithms, in order to improve their efficiency.

4.6.7. ACKNOWLEDGEMENTS

This work has been supported by the project OptiFox (grant 262266 of the 7th Framework Programme of the European Union). I.B. has received a predoctoral fellowship from the Universidad Nacional de Educaci´on a Distancia (UNED).

CHAPTER 5: CLINICAL IMPLEMENTATION

Audiological tests are performed on a child, with aid of Accolado, Otoconsult's VRA Reward System.

5.1. INTRODUCTION

To validate the new fitting concept, which we describe as 'target driven, computer assisted CI fitting' or 'FOX-fitting', we have set out to conduct a number of studies investigating its optimization power. As said before, the approach that has been taken consists of 2 distinct processes: (1) Automaps to let the recipient get accustomed to increasing levels of electrical stimulation during the first few weeks after switch-on and (2) Tuning the map by measuring outcome and adjusting the map accordingly.

These processes are described in detail in the manuscript "Experiences of the use of Fox in new users". This report outlines the fitting protocol that is typically followed at the Eargroup for post-lingually deafened adult CI recipients using the Fox system from switch-on onwards. The timing of 4 sessions in the first six months was found to be adequate to optimize the subjects' maps in the great majority of cases. Across these 4 sessions the total time spent is of the order of 2.5 hours, which includes all "audiological" issues, i.e. technical explanations, device programming and performance measures. This compares favourably to fitting times reported by traditional methods. This report demonstrates that good results can already be obtained with a relatively small clinical workload and that a systematic outcome-driven approach, with the assistance of an intelligent agent like Fox, is capable of selectively improving test results. Adding another session of 30 minutes at 9 months keeps the total time spent during the first year following the proposed protocol well below the median of 7 hours spent by the clinics taking part in our global survey [83].

The manuscript "Evaluation of Fox with established cochlear implant users" reports on a study to evaluate whether Fox is able to complement standard clinical procedures in clinics other than the Eargroup. Ten adult post-lingually deafened and unilateral long term CI users underwent speech perception assessment with their current clinical program. One iteration of Fox optimization was performed and the program adjusted accordingly. After a month of take home experience a second iteration of Fox optimization was performed. Following this, the assessments were repeated without further acclimatization. Sound field aided thresholds were found to be significantly better for the Fox than for the clinical program. Group speech scores in noise were not significantly different between the two programs while three individual subjects had improved speech scores with the Fox map, two had worse speech scores and five were the same. This means that 2 iterations of FOX fitting were able to improve the performance in 30% of CI users who untill then had been conventionally fitted as well as possible by experienced audiologists.

"Multicentre assessment of Fox in new cochlear implant users" reports on a controlled, randomised, clinical study conducted in CI centres in Germany, United Kingdom and France. The aim was to compare the overall fitting time and the overall speech perception performance, between Fox and standard clinical fitting procedures (Control group). The results showed a significant improvement in word scores in quiet (35%, $p = 0.02$) and sentences in +5dB signal to noise (23%, $p=0.04$) for the Fox group compared to the Control group at six months. The fitting time for Fox was also significantly reduced at 14 weeks ($p<0.001$) and equivalent over the six month period. There was much less overall variance in the Fox results. From this it is concluded that the use of Fox produced results that were at least equivalent to conventional fitting methods for all the outcome measures tested. Despite including more testing of outcomes during fitting and the adjustment of a greater range of parameters, Fox does not add to the fitting time. Fox appears highly efficient and effective in providing an optimal map.

The paper "Setting and reaching targets with computer-assisted CI fitting" contains a retrospective data analysis on 255 adults and children in 14 participating centres. The paper aims to demonstrate the feasibility of defining a substantial set of psychoacoustic outcome measures with preset targets and to adopt a systematic methodology for reaching these targets. For each patient, 66 measurable psychoacoustical outcomes were recorded several times after cochlear implantation: free field audiometry (6 measures), speech audiometry (4), spectral discrimination (20) and loudness growth (36), defined from the A§E test battery. These outcomes were reduced to 22 summary variables. The initial results were compared with the latest results. Results showed that the use of Fox significantly improved the proportion of the 22 variables on target. When recipients used the automated maps provided at switch-on, more than half (57%) of the 22 targets were already achieved before any further optimisation took place. Once the Fox system was applied there was a significant 24% ($p < 0.001$) increase in the number of targets achieved.

This study demonstrates that it is feasible to set targets and to report on the effectiveness of a fitting strategy in terms of these targets. Fox provides an effective tool for achieving a systematic approach to programming, allowing for better optimisation of recipients' maps. The setting of well defined outcome targets, allowed a range of different centres to successfully apply a systematic methodology to monitoring the quality of the programming provided.

5.2. EXPERIENCES OF THE USE OF FOX IN NEW USERS

Experiences of the use of FOX, an intelligent agent, for programming cochlear implant sound processors in new users

Int J Audiol. 2011 Jan; 50(1):50-8

Bart Vaerenberg[1,2], Paul J Govaerts[1,3], Geert De Ceulaer[1], Kristin Daemers[1], Karen Schauwers[1]

[1] *The Eargroup, Antwerp-Deurne, Belgium*
[2] *Laboratory of Biomedical Physics, University of Antwerp, Belgium*
[3] *CLiPS Research Center, University of Antwerp, Belgium*

Abstract

This report describes the application of the software tool "Fitting to Outcomes eXpert" (FOX) in programming the cochlear implant (CI) processor in new users. FOX is an intelligent agent to assist in the programming of CI processors. The concept of FOX is to modify maps on the basis of specific outcome measures, achieved using heuristic logic and based on a set of deterministic "rules". A prospective study was conducted on eight consecutive CI-users with a follow-up of three months. Eight adult subjects with postlingual deafness were implanted with the Advanced Bionics HiRes90k device. The implants were programmed using FOX, running a set of rules known as Eargroup's EG0910 advice, which features a set of "automaps". The protocol employed for the initial 3 months is presented, with description of the map modifications generated by FOX and the corresponding psychoacoustic test results. The 3 month median results show 25 dB HL as PTA, 77% (55 dB SPL) and 71% (70 dB SPL) phoneme score at speech audiometry and loudness scaling in or near to the normal zone at different frequencies. It is concluded that this approach is feasible to start up CI fitting and yields good outcome.

5.2.1. INTRODUCTION

Currently available commercial cochlear implant (CI) systems share many common features. All consist of an external (usually ear-level) sound processor which processes the incoming microphone signal and converts this to a command stream which is delivered to an internal receiver via a radio link through the intact skin. The internal receiver is surgically implanted in the mastoid bone and is connected to a silastic electrode array which is inserted into the scala tympani of the cochlea, at or near the round window. The array supports several electrode contacts which are designed to stimulate individual populations of spiral ganglion cells along the cochlea. Low frequency information is directed to electrodes placed apically and high frequency information to basal electrodes, thus preserving the natural tonotopicity of the cochlea.

Following surgical placement of the internal components, the external sound processor must be adjusted ("programmed") by the audiologist so that the characteristics of the stimulating current match the requirements of the individual. There are many processing parameters that can be adjusted, but the most commonly used are the output limits of the stimulating current. This is because the minimum current required for eliciting an auditory percept, and the current where the percept becomes uncomfortably loud, are known to vary between individuals, and between electrodes in a given individual, due to several factors including local neural survival and the exact position of the electrode contacts.

Therefore, before the CI can be used to deliver a signal in response to microphone input, these current limits must be programmed for each electrode, so that stimulation is always within the comfortable range for the user. This is usually performed by presenting short current pulse trains at varying current levels in order to identify the threshold and the maximum comfortable level psychophysically for each electrode. The initial "fitting" session usually involves setting these levels and possibly also checking whether any individual electrodes need to be de-activated, usually due to high thresholds or production of non-auditory sensations. This process is straightforward, though time consuming, in adults, but is far more difficult with young children, and it sometimes takes several fitting sessions to identify these current levels for all electrodes.

Once the set of programming parameters ("map") has been defined and downloaded to the sound processor, the user can start to use the CI. However, a process of adaptation to the electrical signal usually occurs over the first few weeks or months of device use, such that initially loud sounds become perceptibly quieter as the user becomes accustomed to the new signal [98]. As a result, the current limits usually have to be increased gradually in order to accommodate to this change. This can be achieved, in part, by volume control adjustment, or by having several maps loaded into the processor with different current limits. However, additional fitting sessions are usually required so that the audiologist can repeat psychophysical measurements and optimize the user's everyday map.

Thus, adjustment to the parameters of the map is usually based on comfort, with the assumption that the most comfortable map will also be optimal in terms of performance. Performance is normally monitored periodically (primarily by speech recognition measures), but performance outcomes do not usually result in review of the map parameters unless the CI user is performing considerably poorer than would normally be expected. Expectations, however, tend to be rather imprecise, as it is well known that the performance of CI users varies greatly, even among relatively homogeneous subject groups [234]. Unfortunately, the most comfortable map is not necessarily the one that provides the best performance, a finding which is well known and documented in hearing aids [257]. Furthermore, a revised map may well result in a decrement of performance initially, so that many protracted trials may be required before map parameters are optimized.

A particular concern with respect to the usual fitting procedure is its validation. Electrical stimulation at a single electrode or even a group of electrodes produces an electrical field which does not correspond to any physiological acoustic stimulation of the system. T and C levels may substantially differ depending on which procedure was used to set them. It can be questioned whether the minimal and maximal levels identified in this way truly represent the optimal stimulation zone of the subject once the full array is active [258]. This is especially the case in subjects who may have never heard before, who have been deprived of hearing for a long period of time or in children.

With these considerations in mind, we have developed an alternative approach to processor adjustment, which is based on specific outcomes, rather than comfort. Our method has involved the development of an intelligent agent known as the “Fitting to Outcomes eXpert” or FOX. This report describes some early experiences and outcomes of using the FOX software tool in routine fitting of post-lingually deafened recipients of the Advanced Bionics CI system.

The principles and mode of operation of FOX are described in detail by Govaerts et al [101]. Briefly, FOX considers results of several specific performance measures that reflect cochlear function and resolution and assesses whether the parameters of the map in use can be adjusted to improve these measured outcomes. The output of FOX consists of recommendations for any map modifications that it considers are required. The decision process employs heuristic logic and is based on a set of deterministic “rules” derived from theory and experience (often trial and error) which is called an “advice”. To date the only existing advice is the Eargroup’s advice, which has been developed by analysing maps and performance measures from over 600 CI users implanted at our centre (Eargroup) over several years. The set of rules currently in use constitute Eargroup’s EG0910 advice, but FOX is able to work with other sets of advice rules that may be developed in the future or by other centres. An advice would typically contain hundreds of conditional rules and rule sets. The detailed structure of the rule set of an advice is not disclosed as it is subject to intellectual property.

The performance measures currently utilized by FOX are (i) free field audiometry (250 Hz to 8 kHz), (ii) A§E phoneme discrimination [231], (iii) A§E loudness scaling at 250 Hz, 1 kHz and 4 kHz, (iv) speech audiogram, using monosyllabic words at intensities from 40 to 85 dB SPL. Further details of these tests are provided in the methods section below. Map parameters considered by FOX are not restricted to threshold (T) and maximum comfortable level (M), but also include input dynamic range (the minimal and maximal sound levels between which the speech processor processes sound), electrode de-activation, gain (post-processing amplification applied to the signal), processing strategy, pulse rate and bandpass filter boundaries.

One key feature of FOX is the availability of a set of “automaps”, which are designed to be used for the initial CI activation (“switch on”), before outcome measures are available. The parameters of these automaps are based on features of a large number of “Green” maps that have yielded outcomes that FOX considers optimal. There is a growing set of Green maps and the statistics of this set form the basis for the parameters of the automaps. Based on these Green maps, an incremental series of 10 automaps is created that may be used over the first few months, within a protocol such as that described in the methods section below. As the CI user progresses through the series, the T and M levels are incrementally increased as a proportion of those levels used in all available Green maps. Other processing parameters remain constant throughout the automap series. The concept of an incremental series of preset maps was already proposed by others (e.g. [259]). But their preset fittings were based on the profile of the ECAP thresholds. Smoorenburg and colleagues argued however that the relation between ECAP thresholds and behavioural responses is not strong enough to allow for an accurate prediction of behavioural T and C levels in individual CI users [258]. This correlation appeared to be stronger in children and also in more recent publications, but it can be argued that this may result from a circular procedure, where the map-levels are first set based on ECAP thresholds resulting in a stronger correlation between both.

The aim of the present study was to use the FOX system to program the sound processor in a group of new users of the Advanced Bionics (AB) HiRes90k device over the first 3 months, focusing on the use of the automap feature, and to document performance outcomes and the map modifications recommended by FOX.

5.2.2. METHODOLOGY

5.2.2.1. SUBJECTS

Eight consecutive subjects who received an AB HiRes90k device between June and December 2009 entered the study. They were all post-lingually deafened, showed good speech production prior to implantation with Speech Intelligibility Ratings (SIR) of 1 or 2 [260] and all but one used a hearing aid in at least one ear. No re-implantations were included and all subjects had full electrode insertion according to the surgical report.

Key demographics for each subject are provided in Table 14. Implantation was performed at an average subject age of 59 years (range 13-76 years). Surgery was performed by 4 different surgeons. Three subjects received the implant in the right ear, five in the left ear.

Table 14: Subjects demographics.

Subject	Birth date	Etiology	Preop PTA CI-ear	Preop Imaging	Preop Hear Aid
1	4/06/1938	Menière's	102	normal	contralateral
2	10/10/1969	Sudden idiopathic	120	normal	contralateral
3	26/01/1987	Progressive idiopathic	95	normal	bilateral
4	2/04/1996	Early acquired idiopathic	100	normal	bilateral
5	13/09/1933	Otosclerosis	118	otosclerosis	bilateral
6	2/03/1941	Otosclerosis	120	otosclerosis	ipsilateral

7	3/06/1961	Menière's	80	normal	No
8	11/10/1933	Progressive idiopathic	93	normal	bilateral

5.2.2.2. CHARACTERISTICS OF "AUTOMAPS"

For this study FOX 1.1 was used with Eargroup's EG0910 advice (further referred to as FOX1.1$^{(EG0910)}$). The 10 automaps are called "Switch on", then "Silver 1, 2, 3", "Gold 1, 2, 3" and "Ivory 1, 2, 3". The switch-on map has T- and M-levels set to 20 and 90 current units respectively. The statistical basis for the incremental increase in T and M levels for the other automaps is outlined in Table 15. Essentially, there is a gradual increase in these values as a percentage of the "ideal" parameters as defined by those identified from our Green maps. Subjects were randomly assigned to one of two processing strategies, HiRes or HiRes 120. This approach is part of a further study comparing these two strategies, but does not impact on the process or outcomes of the present study.

Table 15: Statistical basis for T and M levels used in the automap series.

Name	Statistical basis
Switch-on	Flat map with T-levels at 20 CU and M-levels at 90 CU
Silver 1	All variables are set between their value in the switch-on and in the Gold 1 map, at 1/4th of the interval
Silver 2	All variables are set between their value in the switch-on and in the Gold 1 map, at 2/4th of the interval
Silver 3	All variables are set between their value in the switch-on and in the Gold 1 map, at 3/4th of the interval
Gold 1	P25: All map-variables have values corresponding to the 25th percentile of the population of green maps
Gold 2	P50: All map-variables have values corresponding to the 50th percentile of the population of green maps
Gold 3	All variables are set between their value in the Gold 2 and in the Ivory 1 map, halfway the interval
Ivory 1	**P75**: All map-variables have values corresponding to the 75th percentile of the population of green maps
Ivory 2	All variables are set between their value in the Ivory 1 and in the Ivory 3 map, halfway the interval
Ivory 3	**P97**: All map-variables have values corresponding to the 97th percentile of the population of green maps

5.2.2.3. PERFORMANCE MEASURES

FREE-FIELD AUDIOMETRY

This was carried out in a sound treated audiometric room using a Madsen Aurical system (GN Otometrics) with free-field loudspeaker outputs calibrated to dB Hearing Level. The loudspeaker was positioned at 0^{o} azimuth, 1m from the subject's head. Thresholds to warble tones at 125, 250, 500, 1000, 2000, 4000, and 8000 Hz were recorded using standard clinical audiometric methods.

A§E PHONEME DISCRIMINATION

The A§E test suite is loaded onto the same PC as that running the Aurical system. Output is fed to the AUX input of the Aurical. The phoneme discrimination module is a discrimination test based around 20 pairs of vowels and consonants, which can provide a clinical indication of the frequency discriminating power of the auditory system. Discrimination of all 20 phoneme contrasts of the "eargroup's list" was measured at 70 dB SPL. Full details of calibration and test procedure are provided by Govaerts et al [107].

A§E LOUDNESS SCALING

This is a loudness scaling procedure where narrow band noise of 250, 1000 or 4000 Hz is presented at different intensities (5 dB steps presented at random between 2 limits). The limits are set during a training session to just below the lowest audible level and just below the level which is too loud for the subject (typically about 20 and 90 dB SPL). The 1876 ms stimulus is presented at least twice at each presentation level, and the subject is required to indicate loudness using a 7-point visual-analogue scale, ranging from "inaudible" to "too loud". The median score at each presentation level is recorded at the end of the test. A "loudness index" was calculated for each test, which is the RMS-value (root mean square) of the scores compared to the average score at the same intensity in normally hearing listeners. A sign (positive or negative) was applied to this score, according to the sign of the sum of all differences between each of the subject's score and the corresponding average in hearing listeners. The average RMS in hearing listeners is 0 with a 95% confidence interval of -0.8 to +0.8 (from our own unpublished data). A RMS value of -1.1, for example, indicates an abnormal loudness scaling with more scores lower than the average in hearing listeners.

SPEECH AUDIOGRAM

Open set monosyllabic CVC-words (NVA-lists, [177]) were presented at 40, 55, 70 and 85 dB SPL, using the same room and equipment as above. Two lists of 12 words were used at each intensity level and phoneme scores recorded.

5.2.2.4. FOX IMPLEMENTATION

The FOX software is installed on several computers within a local area network, as is the Soundwave fitting software and the A§E and Audiqueen software. FOX is able to interface seamlessly between these modules in order to read outcome measures and implement required map modifications.

Automaps are automatically generated when a new CI-subject is entered into FOX. If accepted by the audiologist, they are permanently saved in the Soundwave fitting software and can be loaded to the sound processor as with any maps generated by other means. When output measures are available, then FOX generates recommendations for map modifications which can be either accepted or rejected by the audiologist. If accepted, then FOX automatically implements the required modifications and activates the new map. Full details are provided by Govaerts et al [101].

5.2.2.5. FITTING AND ASSESSMENT PROTOCOL

All subjects were fitted with the AB Harmony sound processor and were randomly allocated to either the HiRes or the HiRes coding strategy. The procedure was the same for all subjects and made use of the incremental series of automaps. This has been the routine clinical procedure for all CI users in our centre for several years and was not modified for this study. The following provides a step-by-step sequence of the procedures carried out:

1ST SESSION (S1)

- The first ("switch on") automap was activated in "live mode". As long as this was tolerated by the subject then the rest of the session was spent counselling the subject regarding operation of the external hardware and aspects of early device use
- At the end of the session the subject received two sound processors. One was a loan processor containing maps Silver 1-2-3 and the other was the subject's own processor containing maps Gold 1-2-3. The subject left the clinic with map Silver 1 active and was instructed to change to the next map every 2-3 days as long as the auditory percept was comfortable.
- This session typically lasted 30-60 minutes, most of which was spent on counselling and familiarization. No performance testing was carried out.

2ND SESSION (S2), TYPICALLY 2 WEEKS AFTER SWITCH-ON

- The aim of the second session was to identify any electrodes that may require deactivation. FOX can efficiently perform this task using the results of free field audiometry, but it is important to involve a competent audiologist to make judgments on any electrodes that produce non-auditory stimulation, usually involving the facial nerve [261].
- The audiogram was performed and the results entered into FOX. Impedance telemetry measures were also performed at this point. FOX decided whether or not any electrodes require deactivation and provided appropriate suggestions.
- At the end of the session the subject was given the same map he/she came in with (with or without deactivated electrodes) with either one lower and one higher, or with two higher automaps in the 3 memory slots of the processor (depending on discussions with the

subject) and the subject was instructed to try to assess the relative comfort of these maps over the following 2 weeks. The aim in this period was mainly to assess the most comfortable map, rather than trying to increase the levels.

- This session typically lasted 15-20 minutes.

3RD SESSION (S3): TYPICALLY 4 WEEKS AFTER SWITCH-ON

- The primary aim of this session was to optimize the subject's preferred automap using the audiogram and A§E phoneme discrimination performance measures. These tests are detection and discrimination tasks, which we consider do not exhibit significant learning effects
- Free-field audiometry and the A§E phoneme discrimination test were conducted as described above, and the results were input into FOX, which analyzed the parameters of the map being used and formulated recommendations to modify the map in an attempt to improve the test results if appropriate. If map modifications were requested by FOX, then the performance tests were usually repeated. If FOX does not recommend map changes (either initially or after map modifications) it outputs a message suggesting that the fitting is "optimal".
- When FOX had no further recommendations, the subject was sent home with the optimized map. The previous map was also provided as a back-up, but the subject was strongly encouraged to use the new map as much as possible. This session typically lasted 30 minutes.

4TH SESSION (S4): TYPICALLY 2½ – 3 MONTHS AFTER SWITCH-ON

- The aim of this session was to modify the subject's everyday map based on results from A§E loudness scaling and speech audiometry. The latter two tests are identification tasks, which we believe are subject to learning effects and changes over time. If the S3 results had not been optimal, FOX would have requested to also repeat free field audiometry or A§E phoneme discrimination if indicated.
- As in the former session, FOX analyzed the map parameters and the test results and formulated recommendations for map modifications, if indicated, until no further testing was requested ("optimal" map assessed by FOX). If the session ended without having obtained "optimal" results according to FOX, then the latest modifications were saved into the processor and the pending outcome requests (PORs) were retained for the next session, which was typically scheduled after another 3 months.
- This session typically lasted 60 minutes.

5.2.3. RESULTS

The median interval from surgery to switch-on (S1) was 21 days (range 17 – 22), and the intervals from switch-on to S2, S3 and S4 were 11 (7 – 16), 28 (21 - 42) and 78 (46 – 111) days respectively.

Table 16 shows the progression of maps that was in use by each subject at the start of fitting sessions 2, 3 and 4, plus the final map at the end of session S4. Even by the start of session S2 chosen maps were already at an advanced stage, ranging from Silver 3 to Ivory 1. Over the remaining sessions there was an overall gradual progression, though several subjects did not change much between S2 and S4. Subjects 1 and 2 initially set themselves automaps that turned out to be slightly too high and later dropped back slightly by the last session. The syntax “Ivory 2#1” denotes automap Ivory 2 which has been modified through one iteration of FOX using outcome measures.

Table 16: Automaps in use by each subject at the start of sessions S2, S3 and S4, plus the final map programmed at the end of S4. (#1) and (#2) denote modifications implemented by FOX following consideration of outcome measures.

Case	Strategy	S2	S3	S4	End S4
1	HiRes	Gold 3	Ivory 1	Gold 1	Gold 1#1
2	HiRes	Ivory 1	Ivory 1	Gold 3	Gold 3#1
3	HiRes	Gold 2	Gold 2	Ivory 2	Ivory 2#1
4	HiRes 120	Ivory 1	Ivory 2	Ivory 2#1	Ivory 2#2
5	HiRes 120	Silver 3	Gold 1	Ivory 1	Ivory 1#1
6	HiRes 120	Gold 1	Gold 1	Gold 1	Gold 1
7	HiRes 120	Gold 3	Ivory 1#1	Ivory 2#1	Ivory 2#2
8	HiRes 120	Gold 3	Gold 3#1	Gold 1#1	Gold 1#2

5.2.3.1. MODIFICATIONS AT SESSION S2

Following impedance telemetry and free-field audiometry, FOX deactivated electrodes 15 and 16 (the most basal) in two subjects (7 and 8) who showed poor thresholds at 6 and 8 kHz. Figure 132 shows the audiograms obtained before and after electrode de-activation in these subjects. All other subjects had satisfactory thresholds across the frequency range examined. In this group of subjects no electrodes were deactivated due to non-auditory stimulation.

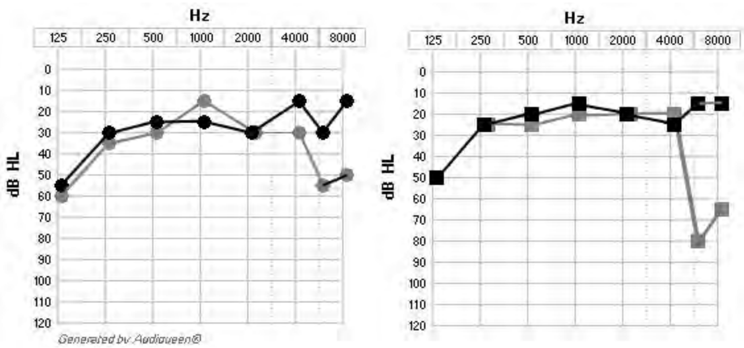


Figure 132: Audiograms of Subjects 7 and 8 at Session 2, thresholds before (light) and after (dark) inactivation of electrodes 15 and 16 by FOX.

5.2.3.2. MODIFICATIONS AT SESSION S3

All initial audiometric thresholds were judged satisfactory, with median thresholds of 21 dB HL (range 13-28 dB HL). Group results are shown in Figure 133.

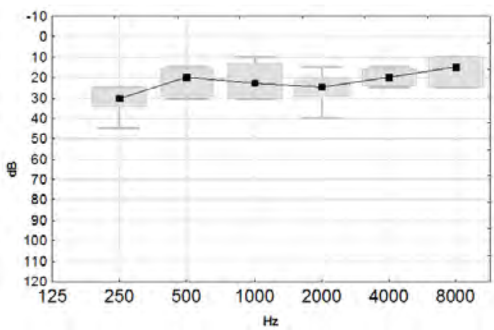


Figure 133: Free-field thresholds obtained for the group at the start of session S3. Central points indicate the median values, boxes the 25th and 75th percentiles and whiskers the limits.

A§E phoneme discrimination was also good in all cases. Four subjects discriminated 19 out of 20 contrasts and the other four 20 out of 20. FOX did not suggest any map modifications as a result but modified the pulse width in 1 case to avoid possible compliance problems based on the measured impedances. The results of A§E loudness scaling and speech audiometry are shown in Figure 134 A and B respectively.

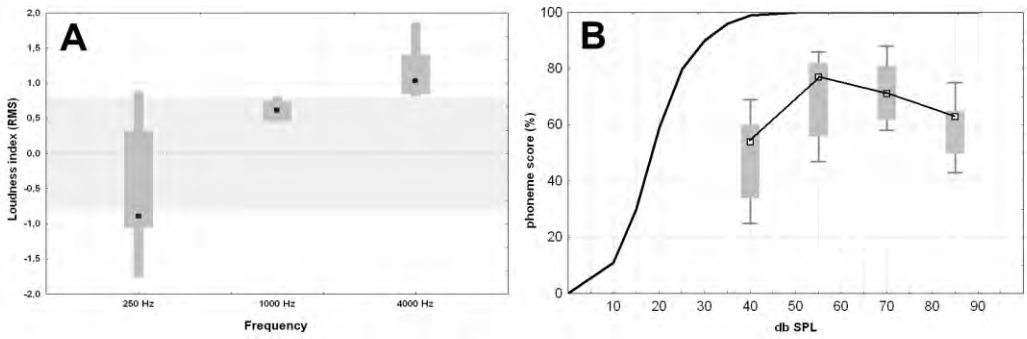


Figure 134: A: loudness scaling results for the group at session S3 for 250, 1000 and 4000 Hz. Central points indicate the median values, boxes the 25th and 75th percentiles and whiskers the limits. B: Speech audiometry group results for the session S3. Central points indicate the median values, boxes the 25th and 75th percentiles and whiskers the limits. The solid curve represents the median of scores from normally hearing individuals.

5.2.3.3. MODIFICATIONS AT SESSION S4

Table 17 summarizes the modifications and final outcomes of the 4th session for the 8 subjects. For each subject, the outcome measures are shown that prompted FOX to make modifications, as well as the re-measured outcome measures after the modifications. For example, subject 4 had a score of 0.9 on the loudness scaling at 250 Hz, which became 0.8 after the modification. The columns on the right list the parameters modified by FOX as well as the final outcome. For example, FOX changed the T-levels, the M-levels and the gains in subject 4 and the final outcome of session 4 was “optimal”, meaning that FOX had no further recommendations. “POR” denotes “pending outcome requests”, meaning that FOX still had recommendations for further map changes which would be addressed at the next follow-up session.

Table 17: Outcome measures resulting in modifications to maps during the final fitting session S4. Left columns (Audio, Speech LS 250, 1000, 4000) show the outcome measures (initial and final) that required map parameter modifications. The right columns show the modifications that were recommended by FOX, plus the final result. POR = pending outcome request.

Case	Audio	Speech	LS 250	LS 1000	LS 4000	Map modifications	Final result
1					1.3 > 0.4	T, gain	Optimal
2			- 1.8			Manual inactivation of apical electrodes	POR LS 250
3		Roll-over > no roll-over	0.7 > 0.6			M, gain, pulse width	Optimal
4			0.9 > 0.8		1.0 > 0.8	T, M, gain	Optimal
5	2000 Hz: 40 > 20 dB		-1.0 > -0.8		1.9 > 1.6	T, M, gain, drop electrodes and change to HiRes	POR LS 4000
6							Optimal
7					1.1 > 1.1		POR LS 4000
8					1.4 > 1.3		POR LS 4000

Subject 2 had manual deactivation of apical electrodes in an attempt to improve loudness scaling at 250 Hz. Based on a 40 dB HL threshold at 2 kHz, FOX recommended the deactivation of electrodes 9, 10 and 11 (with frequency bands centred at 1387 Hz, 1648 Hz and 1958 Hz respectively) in subject 5 and to change the strategy from HiRes120 to HiRes. The audiologist decided to only deactivate electrode 11 and to change strategy on this occasion, which resulted in a threshold of 20 dB HL.

Subject 5 had poor free field audiometric thresholds at 2 kHz. FOX decided to deactivate electrode 11, which represents this frequency. This subject was using the HiRes120 strategy, where two adjacent electrodes are always stimulated simultaneously (in order to achieve current steering). Inactivating one electrode would cause a gap in the array sequence since current steering by means of 2 non-adjacent electrodes (the electrode before and the one after the inactivated electrode) was not possible by the fitting software. To avoid possible problems related to this, FOX therefore also changed from HiRes120 to HiRes strategy. The HiRes strategy is a monopolar strategy where only electrode is stimulated at the time. Inactivating an electrode causes a physical gap in the stimulation pattern, which is intentional, but it does not jeopardize the strategy as such.

5.2.4. DISCUSSION

This report outlines the fitting protocol that we typically follow for post-lingually deafened adult CI recipients using the FOX system running the Eargroup’s advice. This approach has several key features. Firstly, the switch-on session uses an automap generated by FOX, so that the majority of the session is spent counselling the subject rather than focusing on technical programming. We explain to the subject that we only expect him to hear a comfortable auditory percept initially, and that optimization will follow at subsequent sessions. This tends to reduce possible anxiety relating to the belief that sound clarity is dependent on the subject’s psychophysical responses. It also postpones the “fine tuning” of a program to a time when the subject has already habituated to the electrical signal. Secondly, we send the subject home from the first fitting session with two processors containing a total of 6 incremental automaps. This enables him to gradually increase stimulation levels over the ensuing two weeks, and often we find the subjects do not require much current increase after this period (see Table 16). Thirdly, we find the timing of 4 sessions in the first six months to be adequate to optimize the subjects’ maps in the great majority of cases. Across these 4 sessions the total time spent is of the order of 2.5 hours, which includes all “audiological” issues, i.e. technical explanations, device programming and performance measures. To the best of our knowledge, no publications exist reporting the time which is usually spent at fitting according to other procedures. However, it is our impression that our reported 2,5 hours in the first six months compare favourably with fitting times reported by traditional methods. Finally, the programming is outcome-driven where outcome is defined as psychoacoustical performance at the level of detection (audiogram), discrimination (A§E phoneme discrimination) and identification (A§E loudness scaling and speech audiogram).

The series of automaps is based on maps that have been proven to yield good outcomes ("green maps") in children and adults who were able to undergo all the tests. One consequence of this is that these automaps will change over time with the growing number of such green maps. Another consequence is that they can also be used in young children who are not yet able to undergo the psychoacoustical tests. The statistical approach used to generate these maps and their systematic use in all new CI-users provides confidence that they may also be suitable for the young child. With young children, the audiologist obviously needs to provide careful guidance to the parents as to how often the incremental maps should be changed and which signs of possible discomfort or intolerance to look for. It is our experience that the individual course is not significantly different in our paediatric compared to our adult subjects.

In the set of subjects reported here the free field audiometry results were satisfactory in the majority of cases, without the need for any modifications. This is to be expected, as audiometric thresholds are chiefly dependant on processor parameters, rather than subject-specific factors (Boyd 2006). However, the examples shown in Figure 132 demonstrate that FOX1.1[*(EG0910)*] was able to improve abnormally poor high frequency thresholds by deactivating basal electrodes. Without electrode-specific psychophysical measures it is not possible to state whether the electrodes involved were defective, were associated with high electrical thresholds or were, perhaps, outside the cochlea. In the case of electrodes with high electrical thresholds one could argue that alternative parameter adjustments, such as increase in M levels and/or pulse width, could make an electrode useable without the need for deactivation. However, this can have a negative impact on the loudness scaling and can often only be done at the expense of a decreased pulse rate, and often auditory percepts are less clear from electrodes that have significantly different electrical dynamic range characteristics than other electrodes along the array. Deactivation of electrodes also changes other parameters, such as bandpass filter boundaries, and it can be a complex task to weigh up the relative advantages and disadvantages of electrode de-activation. FOX is able to take these considerations into account through reference to large numbers of existing subject maps and outcomes, and is able to verify recommendations immediately through repetition of the outcome measure that initiated the recommendation. In this context it is relevant to consider that the use of FOX in such a decision-making process can be particularly beneficial when the audiologist is relatively inexperienced in CI programming.

The A§E phoneme discrimination module is one of the key tools in the function of FOX as it reflects the spectral discrimination abilities of the cochlea, which is the level at which programming changes are effective, rather than at higher levels of the auditory pathways that are important for speech discrimination and language processing. In the cases reported here, A§E phoneme discrimination was perfect in all subjects by the end of session 4, and in most cases did not require any programming modifications. This is an illustration of the ceiling effect that is often encountered with this test. Although the results are always less than perfect prior to implantation, even with well-fitted hearing aids (this is one of our selection criteria), they usually become "normal" very soon after implantation. This test contributed to the fine tuning of the device in only the minority of cases, but identification of poor phoneme discrimination is considered vital and testing the 20 contrasts in an adult subject typically takes only 10-15 minutes, so does not increase the clinical workload significantly.

Based on this and our previous experience, we feel that A§E phoneme discrimination needs only to be assessed once in most cases, probably fairly soon after device activation (e.g. during the 3rd session). Adopting such a scheme would mean that during the first 6 months after surgery, no more than 2 – 2 ½ hours need to be spent for each subject, spread over 4 sessions.

A§E loudness scaling showed slightly abnormal loudness ratings at 250 Hz (too soft) and 4 kHz (too loud) in several subjects at session 3 (Figure 134 A). We presume that this can be explained by the fact that all were used to the sound of hearing aids and that it takes more time to accommodate to the new perception. There was a marked inter-individual variation at 250 Hz. The 4 subjects with unaided thresholds of worse than 100 dB prior to implantation (Subjects 1, 2, 5 and 6), were the ones who scored the 250 Hz sounds as softer than the other subjects. These were the ones who were used to the strongest amplification with hearing aids. In session 4, FOX1.1$^{(EG0910)}$ recommended modifications to reduce the 4 kHz loudness percept in five cases, but this was only partially successful, resulting in pending outcome requests (LS 4000) at the end of session 4 in three of the subjects. On the other hand, FOX1.1$^{(EG0910)}$ improved the loudness scaling at 250 Hz in four subjects and was successful in three of these. The remaining subject could not be retested due to unavailability relating to a separate severe medical condition. She reported a distorted sound percept when narrow band noise of 250 Hz was presented. The audiologist decided to deactivate the most apical electrode manually which corrected the distorted percept and LS 250 remained as a POR for checking at the next session.

Speech audiometry can be important, particularly in order to identify excessive roll-over at high intensities. Some roll-over is inevitable as the highest intensity speech components are subject to output limitation inherent in the processor function, but often roll-over can be increased due to subject-specific factors such as electrode compliance limits or abnormal loudness growth. Subject 3 showed a high degree of roll-over (Figure 135), which was successfully corrected by FOX1.1$^{(EG0910)}$ through modifications to M levels, gain and pulse width (Table 17).

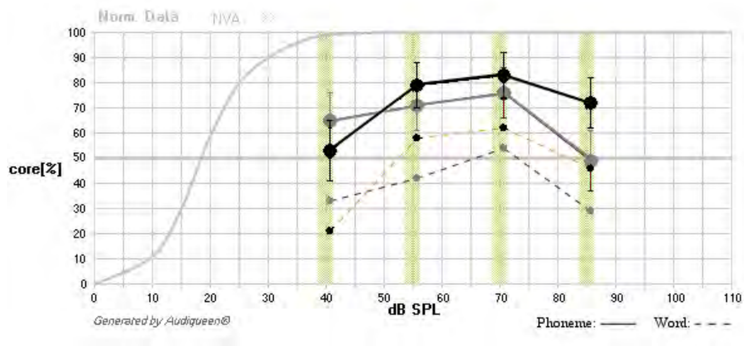


Figure 135: Speech audiometry for Subject 3, before (light symbols) and after (dark symbols) modification of M levels, gain and pulse width by FOX. Solid lines indicate phoneme scores and dashed lines indicate word scores.

FOX is able to manipulate more variables than those routinely modified by most audiologists, including T-levels, M-levels, gains, pulse width, filter boundaries, the activation state of electrodes and even changing the stimulation strategy. Many of these parameters interact with each other, such that efficient programming requires a comprehensive understanding of these issues by the audiologist. In four out of the eight subjects reported in this study the fitting, based on the outcome results at the end of session 4, was considered "optimal" by FOX1.1*(EG0910)*. The other four cases had only minor remaining issues, relating to loudness scaling at a single frequency, resulting in new map modifications with pending outcome requests that would be addressed at the next fitting session.

While FOX is able to efficiently manage most aspects of programming, it is perhaps worthwhile pointing out that an experienced audiologist is still an important component of the fitting process. Reliable outcome measures are critical for optimal use of FOX® and the role of the audiologist here should not be underestimated. There are also some programming issues that FOX is less able to assess accurately, such as non-auditory stimulation, and there may be subject-specific factors relating to lifestyle (music appreciation, for example) which might impact on programming preferences. When FOX® makes recommendations for programming parameter changes these may be accepted or rejected by the audiologist and such decisions require good understanding of the fitting process. In this way, FOX® becomes a useful tool for the experienced audiologist in the fitting process.

On the other hand, FOX might be expected to be especially useful when an experienced audiologist is not available. From anecdotal experience, grossly inappropriate maps are occasionally encountered that have been generated by inexperienced audiologists, a situation which would never occur if FOX is used as an assistant for programming. A key aim of FOX®, therefore, is to provide a systematic approach to programming which can standardize fitting across different centres.

Thus, our initial fitting protocol, using the FOX1.1*(EG0910)* software application, is fundamentally different from traditional methods in that it starts "blindly" with preprogrammed processors. This is a “one size fits all” approach at the start with the “tailoring” of the program to the individual subject at a later stage. It may seem weird to use a “one fits all” approach with preset maps coming from other CI-users. However, previous studies using principal-components analysis (PCA) showed that both the profiles of ECAP thresholds and the conventional T and C levels across the full electrode array are governed by two factors, the major being the overall level (termed shift), and accounting for 90% of the variance [262]. Our switch-on approach incorporates this factor by offering an incremental series of automaps taking care of this shift-effect. The tailoring to the individual profile of the CI-user can be based on electrophysiological measures (like ECAP thresholds), but as outlined in the introduction, these ECAP thresholds only weakly correlate to the behaviourally obtained map-levels. Our tailoring is done with a strong emphasis on outcome measurements. At this stage all recipients have access to the same series of start-up automaps, so the only individual variability lies in the level ultimately tolerated. Future automaps may be different for different subgroups of CI-users depending on factors still to be defined, such as age at implantation, duration or cause of deafness, etc. This report demonstrates that good results can already be obtained with a relatively small clinical workload and that a systematic approach, with the assistance of an intelligent agent like FOX, is capable of selectively improving test results. It is likely that further improvements can be expected with increasing experience and data analysis.

It can be argued that huge differences exist in CI-programming strategy between different centres and even between different audiologists from one single centre and that all strategies seem to yield equally good results. However it is our feeling that hardly any outcome is ever measured or presented. Most papers report on correlations between map-levels based on ECAP-measures to those obtained behaviourally [258] [263]. If psycho-acoustic outcome is presented, this is almost always word or phoneme scores on speech lists presented at one or two presentation levels (typically 60-70 dB SPL). These results depend not only on the cochlear functioning but also on the central processing of the signals and as a consequence on the cognitive functioning of the CI-user, the duration of deafness and many other factors. The inter-individual variability is very high which makes it statistically almost impossible to demonstrate differences between different programming strategies. For an individual patient, we believe that it is justified to try to optimize the detection threshold and the coding of loudness and spectral content by modifying the fitting parameters and we speculate that this results in better speech understanding ultimately.

To date, the set of rules we have worked with are derived from mapping data and outcomes recorded in our centre, i.e. the “Eargroup advice”. FOX is keeping track of all the MAP data with their corresponding outcome and also of the changes made and the measured effects of these changes. This growing database is now analyzed on a regular basis and if possible, the rules are modified to further optimize the advice. Future developments will include automating this analysis and rule optimization such that FOX will include a self-learning engine. As outlined above, the Eargroup’s advice targets the optimization of psychoacoustical outcomes. FOX, however, should not be associated solely with the Eargroup’s advice. It should be emphasized that other experienced groups are able to develop their own “set of advice rules” which can use the same or other outcomes. It is perfectly conceivable that other outcomes may be used, such as electrophysiological test results or even subjective questionnaires. FOX incorporates a user-friendly interface which allows the input of additional rules by professionals without the need for knowledge of programming languages. There may be several advantages for audiologists to become involved in this process, as (i) it encourages the expert to critically analyze his way of working and turn it into a systematic set of rules, (ii) it makes the individual's expertise available to peers and (iii) it systematizes the fitting procedures, making it more easy to share skills with others and to provide a standardized procedural approach.

5.2.5. CONCLUSION

It is concluded that the introduction of the intelligent agent FOX in the programming of cochlear implants is feasible and yields good results as measured by means of psycho-acoustic tests. It represents the introduction of artificial intelligence in this domain. It is anticipated that this will systematize CI programming, reduce the fitting time and optimize the results. Future developments include multi-center trials with FOX, further improvements of the Eargroup's set of rules, the introduction of other outcome measures, the creation of rules that address even more electrical parameters and the development of other sets of rules reflecting the procedures used by other experts in the field. The incorporation of a self-learning engine will allow a continuous improvement of the rules based on the experience in real CI-users.

5.2.6. ACKNOWLEDGEMENTS

Support: BV received a PhD grant for this work from the IWT (Agency for Innovation by Science and Technology, Baekeland-mandaat IWT090287).

5.3. EVALUATION OF FOX WITH ESTABLISHED COCHLEAR IMPLANT USERS

Evaluation of the "Fitting to Outcomes Expert" FOX with established cochlear implant users

Cochlear Implants International, in press

A Buechner[1], B Vaerenberg[2,3], D Gazibegovic[4], M Brendel[4], G De Ceulaer[2], PJ Govaerts[2], T Lenarz[1]

[1] Medical University of Hannover, Department of Otolaryngology, Hannover
[2] The Eargroup, Antwerp-Deurne, Belgium
[3] Laboratory of Biomedical Physics, University of Antwerp, Belgium
[4] Advanced Bionics European Research Center GmbH, Hannover, Germany

Abstract

The objective of this study is to evaluate the possible impact of "Fitting to Outcomes eXpert (FOX)" on cochlear implant (CI) fitting in a clinic with extensive experience of fitting a range of cochlear implant systems, as a way to assess whether a software tool such as FOX is able to complement standard clinical procedures. Ten adult post-lingually deafened and unilateral long term users of the Advanced Bionics cochlear implant system (Clarion CII or HiRes 90K™) underwent speech perception assessment with their current clinical program. One cycle "iteration" of FOX optimization was performed and the program adjusted accordingly. After a month of using both, clinical and FOX programs a second iteration of FOX optimization was performed. Following this, the assessments were repeated without further acclimatization. Sound field aided thresholds were significantly lower for the FOX than the clinical program. Group speech scores in noise were not significantly different between the two programs but three individual subjects had improved speech scores with the FOX MAP, two had worse speech scores and five were the same. It was observed that in this group of subjects FOX prescribed programming modifications in all subjects. Improvement in soundfield aided thresholds was the only measure which showed an overall statistically significant improvement. However, FOX showed improved speech perception in noise for some individuals. For this group of well-fitted patients, FOX did improve outcomes in some individuals. Significant improvements were made for the group in sound field aided thresholds, but overall speech perception scores in noise remained unchanged.

Summary

This manuscript was not yet published at the time of publication of this dissertation. For reasons of copyright, the details of this manuscript have been removed. Only the figures showing the most important results and the general conclusion are preserved in the published version of this dissertation.

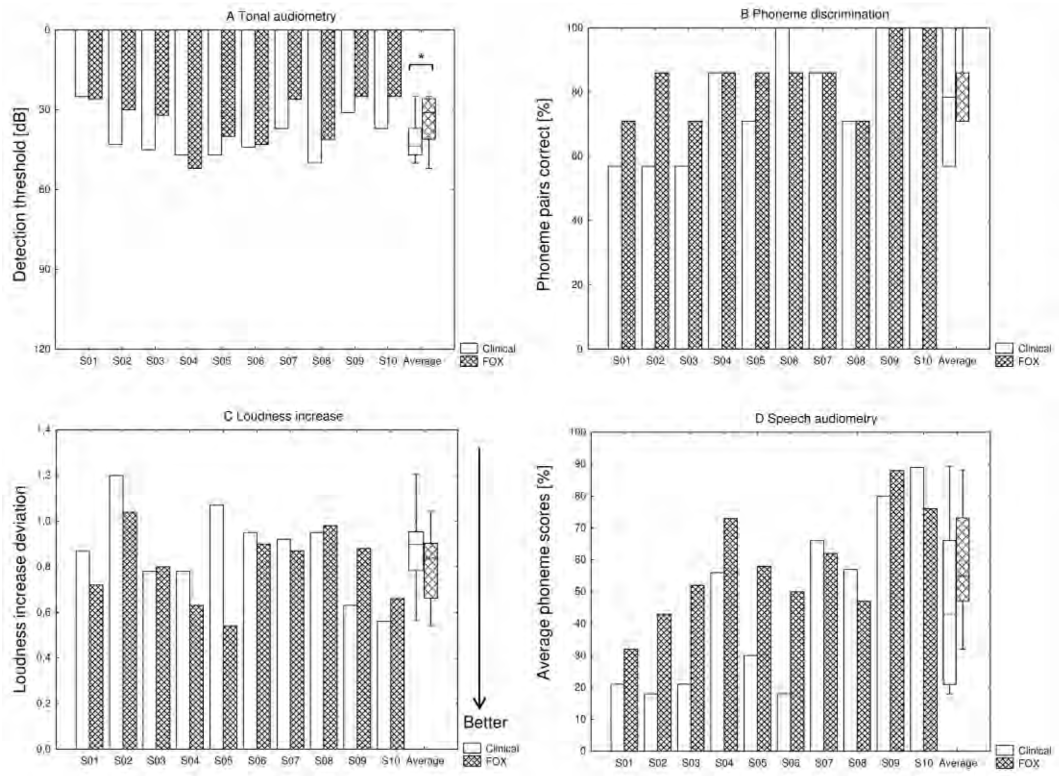


Figure 136: Average scores are shown for the individual subjects, for each of the four FOX outcome measures. The box plots show the median, upper and lower quartiles for the group for each measure. Lower loudness increase scores indicate lower average deviations from the normative data (normal hearing). Marked differences are significant at level $p < 0.05$.

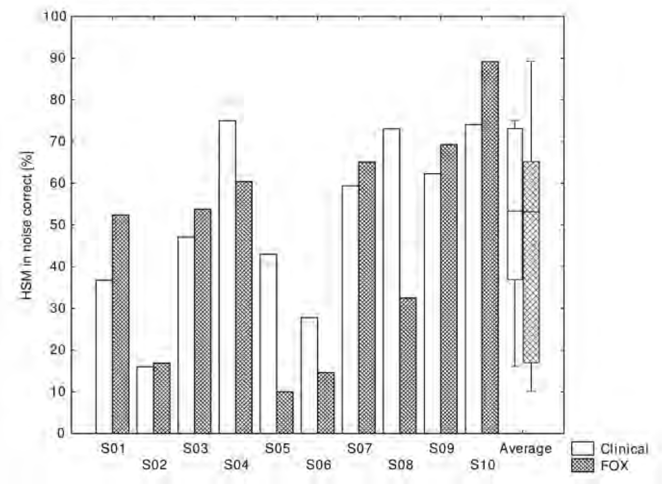


Figure 137: Individual and group median scores for HSM sentences in noise for the clinical and FOX optimized programs.

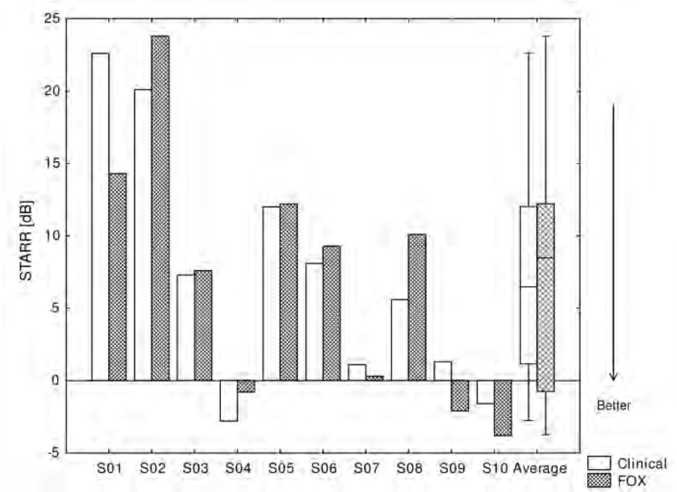


Figure 138: Individual and group median S/N ratios for the clinical and FOX programs for the adaptive STARR test. Lower scores indicate better results. The critical difference for individual test results is 2.2dB.

In conclusion, FOX provides a standardized approach to fitting based on outcome measures rather than comfort, and is perhaps particularly useful in clinics without extensive specialist experience. The results indicated that for this group of well-fitted patients, FOX did improve outcomes in some individuals, particularly on measures that are not typically assessed in normal clinical practice, such as speech recognition over a range of presentation levels. Significant improvements were made for the group in sound field aided thresholds, but overall speech perception scores in noise remained unchanged.

5.4. MULTICENTRE ASSESSMENT OF FOX IN NEW COCHLEAR IMPLANT USERS

Assessment of "Fitting to Outcomes Expert" FOX with new cochlear implant users in a multicentric study

Cochlear Implants International, in preparation

Rolf-Dieter Battmer, Stephanie Borel, Martina Brendel, Anzel Britz, Andreas Büchner, Huw Cooper, Claire Fielden, Dzemal Gazibegovic, Romy Goetze, Paul Govaerts, Thomas Lenarz, Isabelle Mosinier, Joanne Muff, Terry Nunn, Vaerenberg Bart, Zebunissa Vanat

UKB, Berlin, Germany
Beaujon, Paris, France
Advanced Bionics, Clinical Research International
Guy's and St.Thomas', London, UK
MHH, Hannover, Germany
University hospital, Birmingham, UK
Eargroup, Antwerp, Belgium
Adenbrookes, Cambridge, UK

Abstract

The objective of this study is to compare the overall fitting time and the overall speech perception performance, during the first six months after initial stimulation, of computer-assisted fitting with the Fitting to Outcome eXpert" (FOX^{TM}) and a standard clinical fitting procedure. The study is designed as a controlled, randomised, clinical trial. The study sample contained 27 newly implanted recipients of the Advanced Bionics HiRes 90K cochlear implant taken from tertiary referral centres in Germany, United Kingdom and France. There was a significant improvement in word scores in quiet (35%, p = 0.02) and sentences in +5dB signal to noise (23%, p=0.04) for the FOX group compared to the Control group at six months. The fitting time for FOX was also significantly reduced at 14 weeks (p<0.001) and equivalent over the six month period. There was much less overall variance in the FOX results. It can be concluded that the use of FOX assisted fitting produced results that were at least equivalent to conventional fitting methods for all the outcome measures tested. Despite including more testing of outcomes during fitting and the adjustment of a greater range of parameters, FOX does not add to the fitting time. Computer assisted fitting appears highly efficient and effective in providing an optimal MAP.

Summary

This manuscript was not yet published at the time of publication of this dissertation. For reasons of copyright, the details of this manuscript have been removed. Only the figures showing the most

important results and the general conclusion are preserved in the published version of this dissertation.

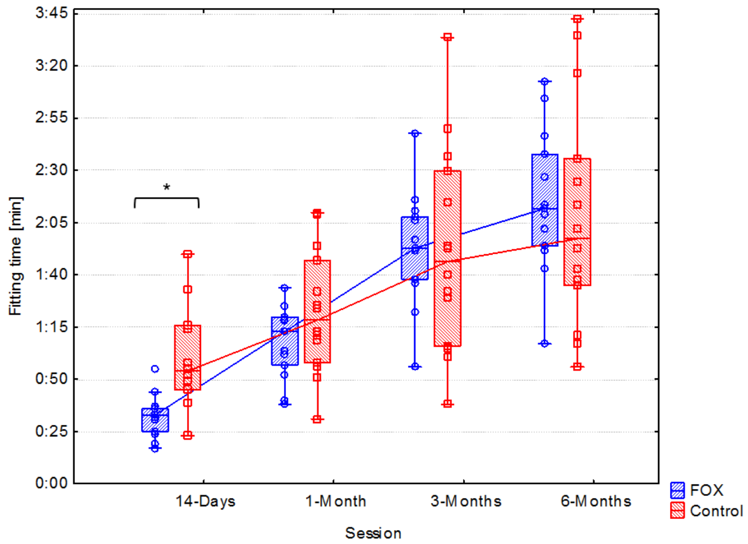


Figure 139: Cumulative fitting time by session for each group over the 6 month assessment period. Starred brackets indicate a statistically significant difference. N=27 at 14-days, 1-month, 3-months and n=25 at 6-months.

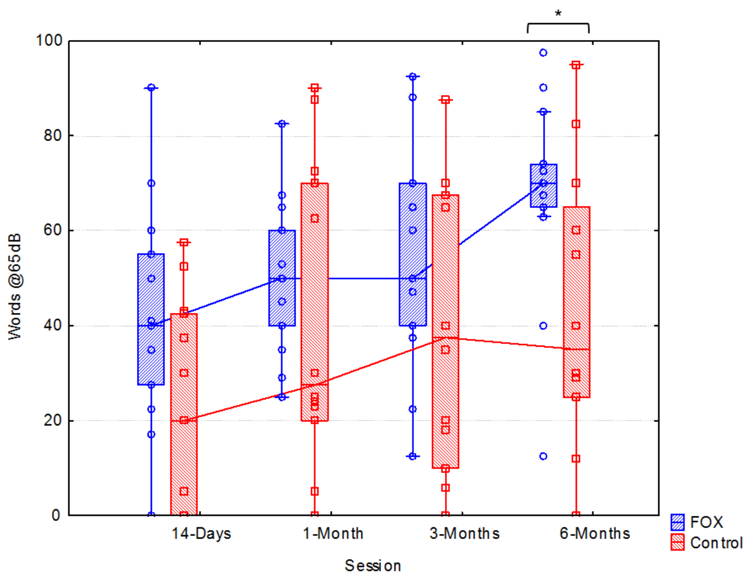


Figure 140: Words in quiet by session for each group over the 6 month assessment period. Starred brackets indicate a statistically significant difference. N=13 for the FOX group and N = 14 for the control group. At 6-months only, n=12 for the control group.

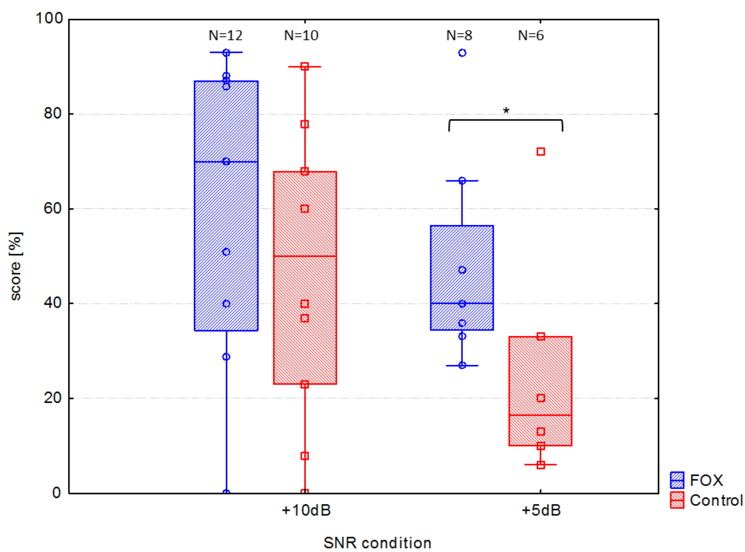


Figure 141: 6-month scores for sentences in noise. Starred brackets indicate a statistically significant difference.

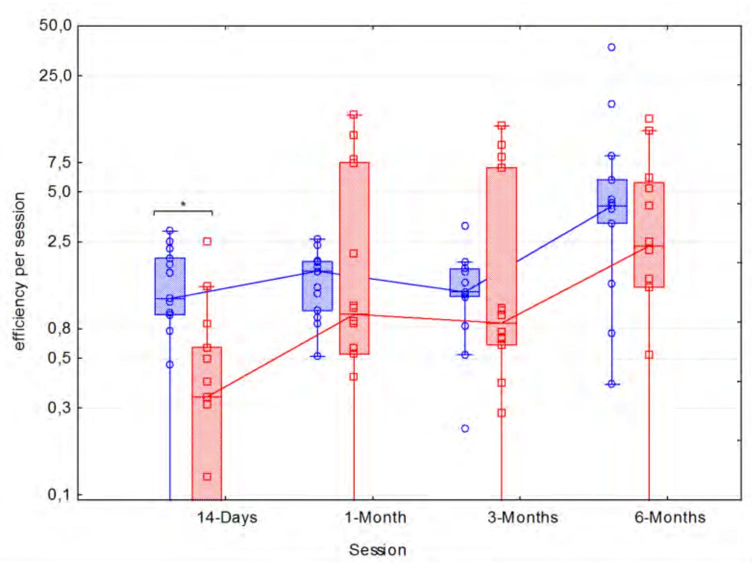


Figure 142: Efficiency index for each session, calculated as the mean word score divided by the overall fitting time. Starred brackets indicate a statistically significant difference. N=27.

In conclusion, the use of FOX assisted fitting produced results that were at least equivalent to conventional fitting methods for all the outcome measures tested. There was a significant improvement in word scores in quiet and sentences in noise for the FOX group compared to the control group at 6 months and the efficiency index and fitting time results indicate that, despite including more testing of outcomes during fitting and the adjustment of a greater range of parameters, FOX does not add to the fitting time. Based on these results, it appears that computer assisted fitting is possible to establish in different clinical environments and leads to improved results and less variation in fitting time and outcomes over time. Particularly at the initial stage of fitting, the FOX method appears highly efficient and effective in providing an optimal MAP.

5.5. SETTING AND REACHING TARGETS WITH COMPUTER-ASSISTED CI FITTING

Setting and reaching targets with computer-assisted cochlear implant fitting

The Scientific World Journal, 2014

Bart Vaerenberg[1,2], Geert De Ceulaer[1], Zoltán Szlávik[3], Patrizia Mancini[4], Andreas Buechner[5], Paul J Govaerts[1]

[1] *The Eargroup, Antwerp, Belgium*
[2] *Laboratory of Biomedical Physics, University of Antwerp, Belgium*
[3] *Dept of Computer Science, VU University Amsterdam, The Netherlands*
[4] *Dept of Sense Organs, University Sapienza, Rome, Italy*
[5] *Dept Otolaryngol, Medizinische Hochschule Hannover, Germany*

Abstract

The paper aims to demonstrate the feasibility of defining a substantial set of psychoacoustic outcome measures with preset targets and to adopt a systematic methodology for reaching these targets in a large group of subjects, by more than one clinical centre. The study is designed as a retrospective data analysis in a multicentre setting with 14 participating centres. 255 adults and children using the Advanced Bionics HiRes90k cochlear implant underwent target driven fitting with the Fitting to Outcomes eXpert (FOX) system. For each patient, 66 measurable psychoacoustical outcomes were recorded several times after cochlear implantation: Free field audiometry (6 measures) and speech audiometry (4), spectral discrimination (20) and loudness growth (36), defined from the A§E test battery. These outcomes were reduced to 22 summary variables. The initial results were compared with the latest results. The state of the fitting process could be well monitored by means of the measured variables. The use of the FOX computer assisted CI-programming significantly improved the proportion of the 22 variables on target. When recipients used the automated MAPs provided at switch-on, more than half (57%) of the 22 targets were already achieved before any further optimisation took place. Once the FOX system was applied there was a significant 24% ($p < 0.001$) increase in the number of targets achieved. This study demonstrates that it is feasible to set targets and to report on the effectiveness of a fitting strategy in terms of these targets. FOX provides an effective tool for achieving a systematic approach to programming, allowing for better optimisation of recipients' MAPs. The setting of well defined outcome targets, allowed a range of different centres to successfully apply a systematic methodology to monitoring the quality of the programming provided.

5.5.1. INTRODUCTION

Cochlear Implants (CI) have become the standard treatment for bilateral severe to profound hearing loss with over 30,000 recipients implanted per year worldwide. Cochlear implant (CI) processors must be appropriately programmed and customized for the recipient [84] [85]. The aim of this is to set a number of parameters to ensure that the electrical pattern generated by the device in response to sound, yields optimal speech intelligibility. Several electrical parameters are available and all their values together are commonly called the MAP. Finding and programming the optimal values for a recipient is commonly called the act of fitting. It is achieved using proprietary software and a hardware interface connected to the processor, and depends on behavioral responses from the CI recipient.

We've recently conducted a global survey to make an inventory of the current practice in CI fitting worldwide [83]. Data were obtained from 47 centres from 17 different countries and 5 different continents. The analysis was based on a written questionnaire, a cross-sectional analysis of 5 consecutive fitting sessions for each centre, a 2-day group debate, and a 2 hour individual oral interview with each centre. It was concluded that current clinical practice in most centres could be defined as setting global profiles of maximum current levels and to a lesser extent of minimum current levels, mainly based on subjective loudness perception by the CI user. Other MAP parameters were rarely modified. It was also shown that measurable targets were only defined for pure tone audiometry. Huge variation appeared to exist across centres in virtually all aspects of CI fitting. The authors concluded that in the absence of targets or well defined outcome measures, it is impossible to compare all these differences or to judge whether some yield better results or are more efficient than others.

Hence, the authors believe that optimizing the process of CI fitting requires defining outcome measures and targets and adopting systematic approaches and algorithms to reach target. At present, there are no agreed standards or targets for both what should be adjusted, or the outcomes expected. Subjective loudness or other comfort measures are relevant, but it should be taken for granted that professionals in the field are aware of this and take care of this. Comfort as such can hardly suffice as target for such an intrusive and costly intervention as cochlear implantation. Placing an implant in the cochlea aims at taking over the function of this sensory organ and it seems obvious that any target should relate to a functional aspect of this organ. This function is the coding of sound and many features of this are well known. Psychoacoustic tests aim at testing the coding of these features in the clinic. Sound field audiograms provide a measure for the correct setting of MAP parameters and targets of 30 dB HL are used by many centers [237] [268] [83]. But audiometric thresholds only partially reflect cochlear performance. The core function of the cochlea is to code for the differences in intensity and spectral content. Assessing this requires supraliminal tests. Speech perception measures are often used but results don't depend on a good cochlear functioning alone, but also on central processing of sound and cognitive capacities. Irrespective of the speech material used, results on speech perception tests in the CI population typically range between 0 and 100% and the factors identified so far merely explain a few percentages of the variation between CI recipients [278]. Therefore it is very difficult to define preset speech audiometrical targets for individual CI recipients.

The Eargroup decided many years ago to use a fixed set of outcome measures to assess the “state of the aided cochlea” after implantation; this set of tests consists of tonal audiometry, speech audiometry and two tests of the A§E psycho-acoustical test battery (Otoconsult, Antwerp, Belgium), namely the spectral discrimination and loudness scaling tests [107]). This provides a method of continuously monitoring the “auditory state” of a CI recipient over time and goes beyond the level of subjective feedback alone. The use of the test battery also provides a set of measurable targets, which assess the auditory system at psychoacoustic level and can be compared to normal values. For each of the measured points in this test battery we defined targets for the performance level considered acceptable (see material and methods (Table 19) and discussion). These targets are near to the normal values as found in hearing subjects. If the target is not reached, then performance is considered suboptimal and changes to the MAP may be indicated.

The software application Fitting to Outcomes eXpert (FOX) system, described in previous papers, introduced a systematic methodology to make adjustments to the MAP, based on the target outcomes from the A§E test battery [101] [102]. FOX is a software tool that uses a deterministic logic, based on a set of pre programmed rules, to recommend changes to a MAP to improve outcome. The recommendations are presented to the audiologist who remains in charge and has the option to either accept or overrule the advice. The outcome measures are then repeated and used to determine if a parameter change has been effective in improving performance. A particular feature of FOX is the use of 10 incremental auto MAPs for the initial period of adaptation after switch on. This approach to predefined MAP settings in the early stages has also been used by others, but based, for instance, on eCAP measures recorded intraoperatively [259]. The FOX MAPs however, are based on statistical analysis of all the ideal or ‘green’ MAPs on the database, defined as MAPs where recipients have reached the target outcomes [102].

The purpose of this study is to demonstrate the concept and feasibility of process optimisation by setting targets in a substantial set of psychoacoustic outcome measures and adopting a systematic methodology for reaching these preset targets in a large group of subjects, by more than one clinical centre.

5.5.2. METHODS

A retrospective study was conducted to assess the results of computer assisted CI fitting in terms of a set of psychoacoustic outcome measures.

5.5.2.1. SUBJECTS

The data for 255 consecutive subjects fitted, almost all (N=228) from switch-on, with the FOX programming system from January 2008 were retrospectively extracted from the FOX database. All subjects used an Advanced Bionics (AB) HiRes90k device (Advanced Bionics LLC, USA), as FOX was until recently only set up for use with AB software. The CI recipients came from 14 different centres all of whom followed the same procedure. Most came from the Eargroup in Antwerp, Belgium (152), four centres contributed at least 10 subjects (21 each from the University Sapienza in Rome, Italy and from the MHH University in Hannover, Germany, 17 from the Yorkshire Cochlear Implant Service in Bradford, U.K. and 10 from the University Hospital in Nijmegen, the Netherlands) and 10 centres (see acknowledgements) from France, India, Italy, Lebanon, Morocco and U.K. contributed between one and nine CI recipients each.

5.5.2.2. FITTING PROCEDURE

All the CI recipients were fitted by an experienced audiologist who was assisted by FOX according to the procedures outlined in Govaerts et al [101] and Vaerenberg et al [102]. Briefly, the recipient received the first statistically derived auto MAP at switch-on, with T and M levels set to approximately 20 and 90 clinical units respectively, T-mic only selected and volume range set to ±5%. The recipient was then instructed to move stepwise to each of the next maps every second or third day and to try and move up to auto MAP 5 or higher, but to stop as soon as it becomes uncomfortable. This typically took two weeks. Once this level was reached, the fine-tuning of the MAP assisted by FOX began. This was done in a staged procedure comprising three sessions over three months (Table 18). Targets were defined for all tests from the psycho-acoustic test battery and are listed below. The initial focus was on detection and discrimination of the acoustic signal, using audiometry and A§E phoneme discrimination as outcome measures. Thereafter identification was optimised using loudness scaling and speech audiometry. If the measured outcome was within the target range defined, the audiologist (assisted by FOX) did not undertake any modifications. If the outcome was not within target, FOX made recommendations for modifying the MAP in an attempt to bring the outcome closer to target. In most of the cases, the audiologist accepted the recommendations made, although he/she had the option to overrule them. The same outcome was then measured again and if still out of target FOX made further suggestions, changing the MAP several times before resting its case.

Table 18: Overview of the fitting procedure.

Session	Programming	Outcome measure
Switch-on	Auto maps loaded	None
Session 2 (2 weeks)	Electrode deactivation (if required)	Impedance Telemetry, Free Field Audiometry

Session 3 (4 weeks)	MAP optimization as recommended by FOX, but only if targets not reached	Free Field Audiometry, Phoneme Discrimination
Session 4 (10-12 weeks)	MAP optimization as recommended by FOX, but only if targets not reached	Loudness Scaling, Speech Audiometry

5.5.2.3. OUTCOME MEASURES AND TARGETS

The following outcome measures were used to assess the results:

- Free field audiometry (6 raw data points): Thresholds determined in Free Field with loudspeaker positioned at 1m from the subject and warble tones presented at 250, 500, 1000, 2000, 4000 and 8000 Hz.
- Spectral discrimination (20 raw data points): A§E phoneme discrimination using 20 speech sound contrasts (a-r, u-ʃ, u-a, u-i, i-a, o-a, i-ɛ, m-z, s-ʃ, ɛ-a, u-o, ə-a, ə-o, ə-ɛ, ə-l, z-s, v-z, ə-u, u-y, y-i) presented at 70dB SPL in an oddity paradigm, 1m from the subject (see Govaerts 2006 [107] for test details). A result of yes or no was recorded for the discrimination of each contrast, yielding 20 results, which were grouped to one variable representing the cumulative score on 20.
- Loudness growth function (36 raw data points): A§E loudness scaling test using 1/3rd octave narrow band noises, centred at 250, 1000 and 4000 Hz. A 1876ms stimulus was presented twice at each level and scored on a visual analogue scale ranging from 0 (inaudible) to 6 (too loud). Levels were randomly presented at 5 dB increments between 30 and 80 dB HL. This yielded 36 values. Scores were pooled for four different levels (30-35-40 dB HL, 45-50-55 dB HL, 60-65-70 dB HL and 75-80-85 dB HL), leading to 12 variables for further analysis.
- Speech audiometry (4 raw data points): Monosyllabic CVC word lists with phoneme scoring presented at 40, 55, 70 and 85 dBSPL, 1m from the subject. The slope between two neighbouring points was then calculated, yielding 3 variables for further analysis.

This yielded 66 raw data points, some of which were grouped such that the final number was reduced to 22 outcome variables (listed in Table 19.) for further analysis. Audiometry was performed in all subjects, but the other tests were not performed in all because, due to age or cognitive ability, this was not always possible. Table 19 shows how many patients underwent each outcome measure at least twice during their follow up.

Table 19: Overview of outcome variables with value definitions for 'target' and 'close to target'. N: number of included records; Outcome variable: see text for more information; Target dimensions: for Audiometry: dB HL; for Spectral discrimination: score on 20; for Loudness scaling: average score on visual-analog scale; for speech audiometry: difference in phoneme score between 2 presentation levels (see text for details). (*) for loudness scaling, the target values correspond to the 95% confidence interval in hearing subjects.

Audiological Test	N	Outcome variable	Target	almost on target	% on target at first	% on target at last	% almost on target at last
Audiometry	255	250 Hz	35 dBHL	40 dBHL	56	80	88
	255	500 Hz	30 dBHL	40 dBHL	71	84	92
	255	1000 Hz	30 dBHL	40 dBHL	69	84	89
	255	2000 Hz	30 dBHL	40 dBHL	64	85	90
	255	4000 Hz	30 dBHL	40 dBHL	55	81	90
	255	8000 Hz	30 dBHL	40 dBHL	55	77	89
Spectral Discrimination	102	set of 20 contrasts	18/20	17/20	82	97	99
Loudness Scaling*	177	250 Hz (30-40 dBSPL)	1.1 - 2.8	0.8 - 3.1	47	71	76
	178	250 Hz (45-55 dBSPL)	1.9 - 3.6	1.6 - 3.9	62	82	88
	180	250 Hz (60-70 dBSPL)	2.9 - 4.4	2.6 - 4.7	59	82	91
	182	250 Hz (75-85 dBSPL)	4.1 - 5.8	3.8 - 6.1	42	70	90
	180	1000 Hz (30-40 dBSPL)	1.2 - 2.3	0.9 - 2.6	58	76	81
	180	1000 Hz (45-55 dBSPL)	1.9 - 2.9	1.6 - 3.2	49	73	87
	181	1000 Hz (60-70 dBSPL)	2.7 - 3.7	2.4 - 4.0	45	67	83
	182	1000 Hz (75-85 dBSPL)	3.4 - 5.1	3.1 - 5.4	75	88	90
	178	4000 Hz (30-40 dBSPL)	0.6 - 2.1	0.3 - 2.4	71	90	94
	180	4000 Hz (45-55 dBSPL)	1.3 - 2.4	1.0 - 2.7	41	67	80
	137	4000 Hz (60-70 dBSPL)	1.9 - 3.4	1.6 - 3.7	37	56	68
	178	4000 Hz (75-85 dBSPL)	2.6 - 4.2	2.3 - 4.5	46	60	80
Speech Audiometry	58	differential scores at 40 vs 55 dBSPL	-15 - 15 %	-20 - 20 %	19	34	40
	92	differential scores at 55 vs 70 dBSPL	-15 - 15 %	-20 - 20 %	65	87	91
	89	differential scores at 70 vs 85 dBSPL	-15 - 15 %	-20 - 20 %	81	94	96

For each of these 22 outcome variables, a target and near target for acceptable performance was defined as shown in Table 19. The rationale for these targets is addressed in the discussion section. Briefly, the targets for audiometry were 30 dB, which corresponds to the lower limits of the device microphone and the front end technology. The targets for spectral discrimination were set at 90%, since this is a prerequisite for good speech understanding. The targets for loudness scaling were the 95% confidence interval in hearing subjects and the targets for the speech audiometric slopes were set empirically at ±15%. We calculated two measures for success: (1) the target hit rate (THR) for each outcome variable and (2) the subject's hit rate (SHR) for each CI recipient. This was done at two moments, namely after switch-on ('initially') and when the optimisation was considered to be completed by FOX ('finally'). The THR was calculated for each outcome measure as the percentage of subjects who had reached the target. If the target was not reached we looked at the interval between the initial and the final measurement. A small interval indicates that the fitting process may not yet have been finished and that further optimisation might still be possible if additional programming sessions would be undertaken. The THR in that case might be an underestimation of the real success rate. The SHR was calculated for each subject as the percentage of the 22 targets which was reached by the subject. In addition for both THR and SHR we also calculated the percentages with results within the 'almost on target' range according to the definitions of Table 19. These will be referred to as tolerant THR (tTHR) and tolerant SHR (tSHR) henceforth.

Descriptive statistics were used to present the results as histograms for THRs and box and whisker plots for SHRs. Nonparametric statistics were used to compare the initial and final THRs and SHRs (Wilcoxon paired rank tests) with a cut-off level of significance set at 0.05.

5.5.3. RESULTS

66 psychoacoustic points were measured to monitor the fitting in 255 consecutive CI recipients. Some results were grouped such that a total of 22 outcome variables were obtained to describe the 'state' of the CI fitting process. For all variables a target was defined in a strict sense ('on target') and a more tolerant sense ('almost on target'). Hence the state of the process was measured at two moments, marked as initial and final. The initial state refers to the first time that the outcome was measured, which is typically after the automated switch-on procedure. It therefore reflects the success rate of this start-up procedure. The final state is the last time the outcome was measured. Since all CI-recipients were fitted for target (by the audiologist assisted by FOX), this final state reflects the success rate of this fitting approach.

The THRs and tTHRs of all 22 outcome variables individually are shown in Figure 143 A and Figure 144 A and also in Table 19. After the initial switch-on procedure, the THR ranged from 19% for the variable [Speech Audiometry 40-55 dB] to 82% for the variable [Spectral Discrimination]. After this switch-on procedure the computer (FOX) assisted fitting improved the THR for all of the individual outcome variables (median improvement = 21%; $p < 0.001$). As displayed by the figures, the THR was substantially different across outcome variables. For instance, the loudness scaling results show that the coding of soft sounds at 4000 Hz already reached target at initial evaluation in 71% of the subjects and that this improved to 90% at the final evaluation (Figure 144 A). This is in contrast to the speech audiometry at soft presentation levels where the slope between the 40 and 55 dB SPL presentation levels only reached target in 19% of the subjects at the initial evaluation and in 34% at the final evaluation (Figure 143 A). In both cases, the subjects who did not reach target had been evaluated more than a year after the initial stage (375 days for the loudness scaling and 524 days for the speech audiometry, Figure 144 A and Figure 143 A respectively). From this it can be inferred that sufficient time had passed to try optimizing these outcomes and that it would be unlikely that they would further improve.

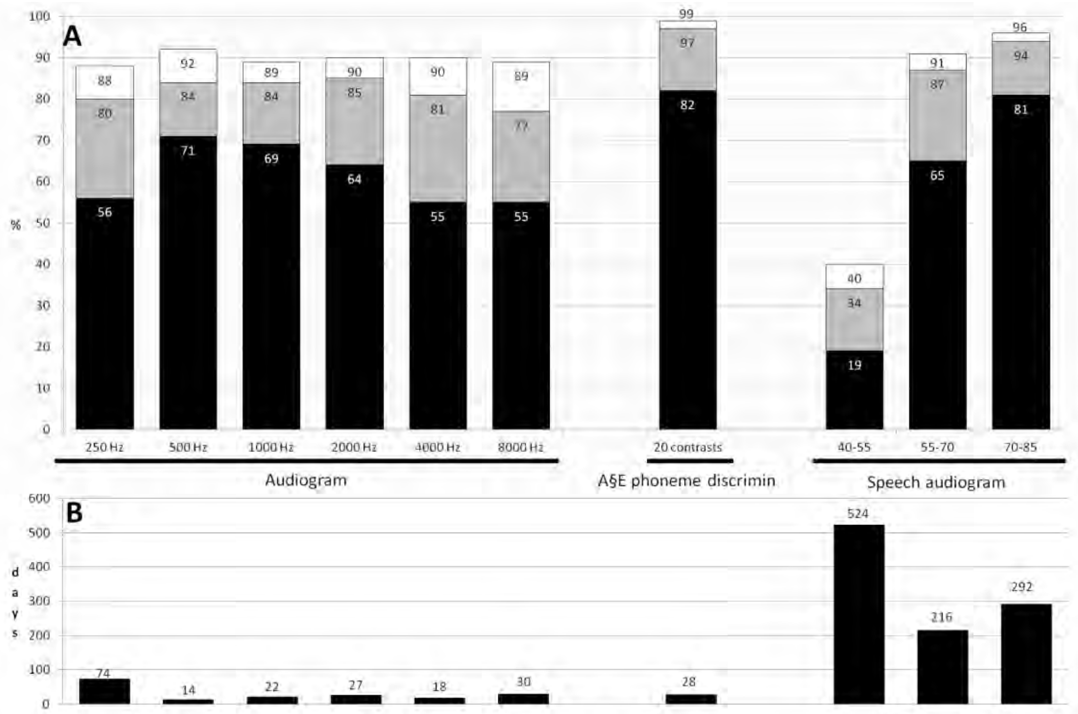


Figure 143: A: the percentages of CI-users who performed on target (THR) at initial testing (black), final testing (gray) and almost on target (tTHR) at final testing (white) on Audiometry, A§E phoneme discrimination (A§E phoneme discrimin) and Speech Audiometry. B: the interval between the initial and final measurement for those CI-users who did not reach target at the latest measurement.

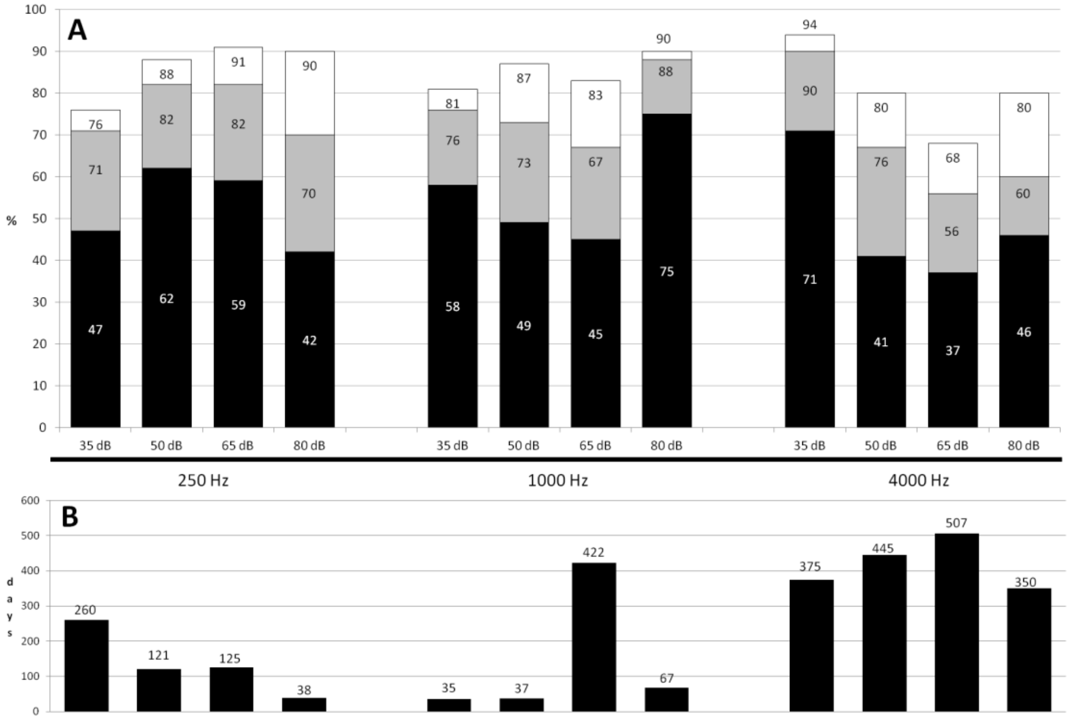


Figure 144: the results for Loudness Scaling at 250, 1000 and 4000 Hz. See Figure 143 for interpretation.

The SHR results are shown in Figure 145. This representation allows analyzing how close CI recipients come to target when all 22 measures are taken into consideration. For instance, it shows an average SHR of 57% after switch-on by means of the automated MAPs and before any further optimisation took place. This means that the average CI recipient is on target for almost 13 of the 22 outcome variables. The computer-assisted fitting yielded a significant improvement of 24% in SHR, from 57% to 81% ($p < 0.001$). A further significant improvement of 8% ($p < 0.001$) was seen when the almost on target values were applied.

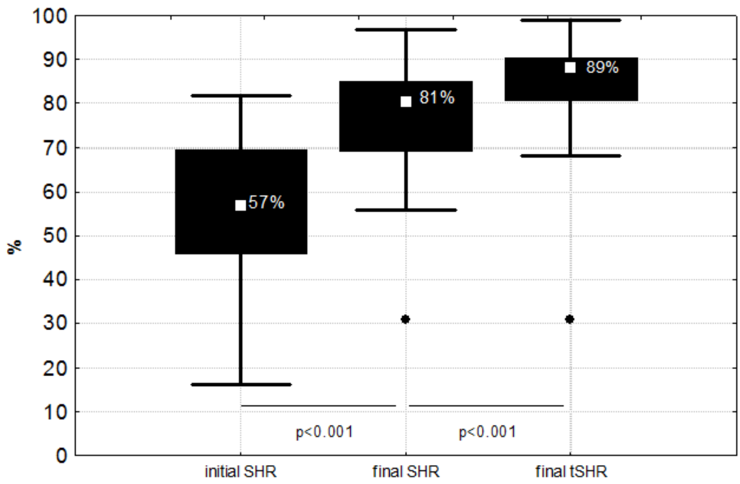


Figure 145: the distribution of the success rates of the 22 outcome variables (SHR) with the median value (central dot), quartile range (box) and range (whiskers).

5.5.4. DISCUSSION

Optimizing any process requires (1) a number of parameters to be adjusted within specific constraints; (2) quantitative or objective performance measures that need to reach predefined targets and (3) a systematic approach with methods and algorithms rather than trial and error. When applying this to the process of CI fitting, the first requirement refers to the MAP parameters which can be modified by means of the CI programming software. The next 2 requirements are not obvious, as revealed by the global survey which has recently been conducted [83]. At present CI fitting is performed by experts in the field who have an idea of what the expected level of performance for an individual recipient should be and who make MAP adjustments if this target is not reached. Assessing the success of changing a parameter usually relies solely on patient feedback. There is no universal set of quantitative measures which is commonly used to quantify the "auditory state of the aided cochlea", and for which well defined targets are commonly accepted. Also the basis for adjusting the MAP parameters is often heuristics or trial and error. Systematic approaches are lacking both in textbooks [88] [85] and as revealed by the global survey [83].

This report shows that the setting of well defined outcome targets did allow a range of different centres to apply a systematic methodology to monitoring the quality of the programming provided. In an age where good clinical practice requires an evidence based approach, it is essential to have the ability to objectively monitor and audit the success of the treatment provided. The use of clear targets enables audiologists to define what is meant by an optimised MAP and provides consistency across different professionals and centres.

For outcomes to be effective they must be measureable in most clinical settings and reliably repeatable. They must also provide an accurate assessment of auditory performance and preferably be independent of the language spoken. The auditory system is complex and therefore requires a complex assessment system; one single measure is unlikely to be sufficient to provide all the information required. Like any other sensory organ, the cochlea is responsible for detecting its particular signal, sound, and for discriminating two sounds which differ in one of their components. In the absence of a global consensus on such targets, we have chosen the psycho-acoustical targets as used in this study because we believe that, within the context of programming, they reflect well the state of the aided cochlea. They combine measures at the level of detection (audiometry), spectral discrimination (A§E phoneme discrimination) and identification (A§E loudness scaling and speech audiometry). They cover the coding of the sound features intensity and spectral content and most can be used for both adults and children, as they do not require a high cognitive or language level and are easy to implement across a wide range of centres. One can argue the choice of outcome variables and with this paper we do not intend to state that the variables chosen here compose the best selection. We do intend to open the debate and to make the point that a consensus would be very helpful in moving forward the discussion on the quality of CI fitting. The targets and 'tolerant' targets set for each outcome were empirical though educated choices. The audiometrical targets were set at 30 db HL since this is close to the technological limit of the current CI devices, which is defined by a combination of the microphone sensitivity, the front end preprocessing and filter bank steps and the internal noise of the electronic circuitry [279]. The target for spectral discrimination was set at 85%, which means that 17 out of the 20 contrasts presented are well discriminated by the CI recipient. Previous unpublished results of our team have shown that good speech intelligibility (>= 60% phoneme score on monosyllabic speech lists) is only obtained in listeners with at least 85% score on the phoneme discrimination test. The target for loudness scaling was set to be the 95% confidence interval in hearing listeners. And for speech audiometry, we have chosen not to use absolute scores as target because there is no such value which is valid for all CI recipients. This is because speech audiometry results depend on much more than just a good replacement of the cochlear function, which explains the huge variability in this outcome across CI users [278]. On the other hand, a valid target can be to maintain the best score for a given CI user across a wide range of presentation levels. Therefore we use to present speech not at one single presentation level but rather at different levels (40-55-70-85 dB SPL) and our target is to have scores at these levels which are as close as possible to the best of all four scores. This is reflected in the slopes of the three lines connecting the four scores being as close as possible to zero; hence we set slopes of 0 ± 15 as empirical target.

Once these targets were set, we introduced a systematic approach to change the MAPs based on the outcome obtained. The FOX computer assisted programming system provides such a systematic approach across centres. All the centres included were able to use the FOX system effectively and to perform the tests required. In this study, the use of the FOX system significantly improved the audiologists' ability to achieve the target outcomes set at the beginning of the study. The initial switch-on MAPs provided were solely based on the statistical derivation of T, M and Gain levels, with incremental increases in T and M levels applied as the MAP number used was increased. With these MAPs, more than half (57%) of the 22 targets were already achieved before any further optimisation took place. Once the FOX system was applied and optimisation began, there was a significant 24% increase in the number of target achieved, as measured at the last fitting session. Also, the spread of SHR across subjects decreased from 66% initially (range 16-82%) to 41% finally (range 56-97%) or even 31% if near to target scores are also tolerated (tSHR range 68-99%). This indicated that the approach under study is capable of delivering robust results across different CI recipients from different CI centres.

The approach of systematic fitting for target also allows looking at and interpreting the individual results for each outcome measure (THR). For instance, FOX was able to improve the THR for free field audiometry outcomes by a minimum of 13% and by as much as 26% at 4000Hz and 24% at 250Hz. Although this measure merely reflects the front-end technological capacity of the device, it still requires customized programming to achieve good results in every individual. The results for phoneme discrimination were good without any optimisation, with 82% achieving target. This is in line with previous reports on the use of FOX [102]. It is not unexpected, because cochlear implants are conceived primarily to restore the tonotopical organisation and also because the contrasts used in the A§E phoneme discrimination task are rather easy, thus causing ceiling effects. Nevertheless it is important to achieve good results on this task because spectral discrimination is one of the core tasks of the cochlea and a prerequisite for good central processing and identification. At the same time they are not sufficient to guarantee good supraliminal identification, which is reflected by the fact that for the identification tasks (loudness scaling and speech intelligibility), the THR with the Auto MAPs is considerably lower, namely between 37 % and 71%.

For speech audiometry equivalent performance across presentation levels is considered to be an area where correct fitting of the device can directly impact performance. For the un-optimised Auto MAPs, at what could be considered to be the key intensity comparison of 55 and 70dB levels, the THR was 65%. This was improved with FOX optimisation to 87%. However, the THR for the 40 and 55dB comparison remained low at 40%, even after optimisation. This is in line with previous reports showing that speech intelligibility at these quiet levels is very challenging [280].

The loudness scaling targets were harder to achieve with the Auto MAPs, with the 4000Hz frequencies being the most difficult. This was again also shown in the smaller sample reported by Vaerenberg et al [102] and it was assumed that the difficulty in measuring the loudness outcomes relates to the distorted loudness picture that some recipients have become used to during long periods of hearing aid use [102]. However, despite the difficulties, once FOX optimisation was applied, 9 out of the 12 measures showed a THR between 70 – 90% which represented an improvement of between 13% and 28% over the un-optimised MAPs. Again, once the ‘almost on target’ outcomes were applied, two out of the other three measures below the 70% on target value, increased to at least 80%.

Results were based on the last available measurement and not the measurement when the fitting was considered to be optimal. Typically, a reasonable interval to achieve optimum would be around one month for audiometry and spectral discrimination and six months for loudness scaling and speech audiometry. Therefore for some measures, the interval between the initial and final measurements was much less than that ideally required and if optimisation was continued, then further improvements in the percentage of subjects achieving target could be expected.

A final consideration is on the validity of the outcome variables chosen. Although the authors feel that the main justification for the current selection of outcome variables lies in the fundamentals of sound coding by the cochlea, as argued above, the ultimate proof of their validity will come from better speech understanding in quiet and in noise. It is beyond the scope of this paper to further address this issue, but two studies have been conducted in other centres than the Eargroup where speech understanding in quiet and in noise has been analyzed after conventional fitting compared to computer assisted and target driven fitting. The first study was conducted at MHH (Hannover, Germany), where 10 long term CI users who had always been fitted in the conventional way entered a single FOX iteration based on the 66 measured outcome points [281]. Speech audiometry with monosyllabic words in quiet improved instantaneously in 7 CI recipients and deteriorated in 3. Speech audiometry with sentences in noise improved instantaneously in 6 CI recipients and deteriorated in 4. Another Sentence Test with Adaptive Randomized Roving levels (STARR test) [272] [273] yielded better results in 4 CI recipients and worse results in 6. Although not statistically significant, this shows that a single iteration with the computer assisted and target driven approach can further improve the speech understanding in more than 50% of CI users who have been fitted throughout many sessions by expert hands. The second study was a multicenter study conducted in 6 CI centres in Germany, France and the UK [282]. New CI recipients were randomized to enter either a conventional fitting arm or a computer assisted, target driven arm during 3 months starting at switch-on. The computer assisted, target driven approach used the 66 outcome measures as mentioned before and the FOX application to assist the audiologist. Speech audiometry was assessed in quiet and in noise at 6 months and it showed significantly better results in the computer assisted, target driven arm compared to the conventional arm both in quiet and in noise. These results, although preliminary and in small numbers, indicate that the target driven systematic approach may be considered promising.

5.5.5. CONCLUSIONS

This study demonstrates that it is feasible to set targets and to report on the effectiveness of a fitting strategy in terms of these targets. This is demonstrated with the FOX-assisted strategy as example. The psycho-acoustical measures chosen here were selected because they measure the behavioural response to acoustic stimulation, both in terms of loudness and frequency, and provide the building blocks for eventual speech perception and language development. This study also demonstrates that the application of the FOX system provides an effective tool for achieving a systematic approach to programming, allowing for better optimisation of the MAPs, when measured by the set targets. When recipients used the automated MAPs provided at switch-on, more than half (57%) of the 22 targets were already achieved before any further optimisation took place. Once the FOX system was applied there was a significant 24% increase in the number of targets achieved. The setting of well defined outcome targets allowed a range of different centres to successfully apply a systematic methodology to monitoring the quality of the programming provided.

5.5.6. ACKNOWLEDGEMENTS

The authors gratefully acknowledge the following CI centres who contributed data to the FOX database: University La Sapienza (Rome, Italy), Medical University Hannover (Hannover, Germany), Yorkshire CI Service (Bradford, UK), UMS Radboud (Nijmegen, The Netherlands), Guys and St Thomas Hospital (London, UK), Hôpital Beaujon (Paris, France), RIFMED (Casablanca, Maroc), Hôpital Sacré Coeur (Beirut, Lebanon), Hôpital Avicenne (Paris, France), University Hospital Federico II (Naples, Italy), Midlands Adult CI Program (Birmingham, UK), Hörtherapiezentrum (Potsdam, Germany), MERF (Chennai, India), Mesiarc (Calicut, India).

BV received a PhD grant for this work from the IWT (Agency for Innovation by Science and Technology, Baekeland-mandaat IWT090287). Part of the work is supported by a 7th Framework for SME grant (“Opti-Fox”) from the European Commission.

REFERENCES

[1] Y. Tong, R. Black, G. Clark, I. Forster, J. Millar, B. O'Loughlin and J. Patrick, “A preliminary report on a multiple-channel cochlear implant operation.,” *J Laryngol Otol,* pp. 679-695, 1979.

[2] G. Clark, Sounds from Silence: the bionic ear story., St Leonards NSW, Australia: Allen & Unwin, 2000.

[3] K. Schauwers, S. Gillis and P. Govaerts, “The characteristics of prelexical babbling in hearing-impaired children after cochlear implantation between 5 and 20 months of age,” *Ear Hear,* pp. 29(4): 627-637, 2008.

[4] S. Archbold, M. Harris, G. O'Donoghue, T. Nikolopoulos, A. White and H. Richmond, “Reading abilities after cochlear implantation: the effect of age at implantation on outcomes at 5 and 7 years after implantation.,” *Int J Pediatr Otorhinolaryngol,* pp. 72(10):1471-8, 2008.

[5] G. Barton, P. Stacey, H. Fortnum and S. AQ, “Hearing-impaired children in the United Kingdom, IV: cost-effectiveness of pediatric cochlear implantation,” *Ear Hear,* pp. 27(5):575-88, 2006.

[6] P. Govaerts, C. De Beukelaer, K. Daemers, G. De Ceulaer, M. Yperman, T. Somers, I. Schatteman and F. Offeciers, “Outcome of cochlear implantation at different ages from 0 to 6 years.,” *Otol Neurotol,* pp. 23(6): 885-90, 2002.

[7] H. Francis, N. Chee, J. Yeagle, A. Cheng and J. Niparko, “Impact of cochlear implants on the functional health status of older adults.,” *Laryngoscope,* pp. 112(8 Pt 1):1482-8, 2002.

[8] G. O’Donoghue, “Hearing without ears: do cochlear implants work in children?,” *Britisch Medical Journal,* pp. 318(7176):72-3, 1999.

[9] S. Connell and T. Balkany, “Cochlear implants,” *Clin Geriatr Med,* pp. Aug;22(3):677-86, 2006.

[10] A. Ramos, C. Rodriguez, P. Martinez-Beneyto, D. Perez, A. Gault and J. Carlos Falcon, “Use of telemedicine in the remote programming of cochlear implants.,” *Acta Otolaryngol,* pp. Jul 22:1-8, 2008.

[11] M. Skinner, L. Holden and T. Holden, “Parameter selection to optimize speech recognition with the Nucleus implant.,” *Otolaryngol Head Neck Surg,* pp. Sep;117(3 Pt 1):188-95, 1997.

[12] T. Zwolan and B. Griffin, “How we do it: tips for programming the speech processor of an 18-month-old.,” *Cochlear Implants Int.,* pp. Dec;6(4):169-77, 2005.

[13] T. Braunschweig, P. Schelhorn-Neise, F. Biedermann and P. Weisser, “Examination to

possibilities of a physiological fitting of cochlear implants," *Laryngorhinootologie,* pp. Jun;83(6):387-90, 2004.

[14] K. Plant, M. Law, L. Whitford, M. Knight, S. Tari, J. Leigh, K. Pedley and E. Nel, "Evaluation of streamlined programming procedures for the Nucleus cochlear implant with the Contour electrode array.," *Ear Hear,* pp. Dec;26(6):651-68, 2005.

[15] M. C. Zimmerman-Phillips S, "Programming features of the CLARION Multi-Strategy Cochlear Implant," *Ann Otol Rhinol Laryngol Suppl.,* pp. Apr;177:17-21., 1999.

[16] P. J. Boyd, "Effects of programming threshold and maplaw settings on acoustic thresholds and speech discrimination with the MED-EL COMBI 40+ cochlear implant," *Ear and hearing,* vol. 27, no. 6, pp. 608-618, 2006.

[17] J. Shallop and K. Ash, "Relationships among comfort levels determined by cochlear implant patient's self-programming, audiologist's programming, and electrical stapedius reflex thresholds.," *Ann Otol Rhinol Laryngol Suppl.,* pp. Sep;166:175-6, 1995.

[18] L. Spivak and P. Chute, "The relationship between electrical acoustic reflex thresholds and behavioral comfort levels in children and adult cochlear implant patients.," *Ear and hearing,* pp. vol. 15, no2, pp. 184-192, 1994.

[19] K. Seyle and C. Brown, "Speech perception using maps based on neural response telemetry measures.," *Ear Hear.,* pp. Feb;23(1 Suppl):72S-79S, 2002.

[20] G. Caner, L. Olgun, G. Gültekin and M. Balaban, "Optimizing fitting in children using objective measures such as neural response imaging and electrically evoked stapedius reflex threshold.," *Otol Neurotol,* pp. Aug;28(5):637-40, 2007.

[21] K. Gordon, B. Papsin and R. Harrison, "Programming cochlear implant stimulation levels in infants and children with a combination of objective measures.," *Int J Audiol,* pp. Dec;43 Suppl 1:S28-32, 2004.

[22] J. Allum, R. Greisiger and R. Probst, "Relationship of intraoperative electrically evoked stapedius reflex thresholds to maximum comfortable loudness levels of children with cochlear implants.," *Int J Audiol,* pp. 41(2):93-9, 2002.

[23] A. Ramos, C. Rodriguez, P. Martinez-Beneyto, D. Perez, A. Gault and J. Carlos Falcon, "Use of telemedicine in the remote programming of cochlear implants.," *Acta Otolaryngol,* pp. Jul 22:1-8, 2008.

[24] M. Skinner, L. Holden and T. Holden, "Parameter selection to optimize speech recognition with the Nucleus implant.," *Otolaryngol Head Neck Surg,* pp. Sep;117(3 Pt 1):188-95, 1997.

[25] T. Zwolan and B. Griffin, “How we do it: tips for programming the speech processor of an 18-month-old.,” *Cochlear Implants Int.,* pp. Dec;6(4):169-77, 2005.

[26] T. Braunschweig, P. Schelhorn-Neise, F. Biedermann and P. Weisser, “Examination to possibilities of a physiological fitting of cochlear implants,” *Laryngorhinootologie,* pp. Jun;83(6):387-90, 2004.

[27] K. Plant, M.-A. Law, L. Whitford, M. Knight, S. Tari, J. Leigh, K. Pedley and E. Nel, “Evaluation of streamlined programming procedures for the Nucleus cochlear implant with the Contour electrode array.,” *Ear Hear,* pp. Dec;26(6):651-68, 2005.

[28] P. Boyd, “Effects of programming threshold and maplaw settings on acoustic thresholds and speech discrimination with the MED-EL COMBI 40+ cochlear implant.,” *Ear Hear,* pp. Dec;27(6):608-18, 2006.

[29] J. Shallop and K. Ash, “Relationships among comfort levels determined by cochlear implant patient's self-programming, audiologist's programming, and electrical stapedius reflex thresholds.,” *Ann Otol Rhinol Laryngol Suppl.,* pp. Sep;166:175-6, 1995.

[30] L. Spivak and P. Chute, “The relationship between electrical acoustic reflex thresholds and behavioral comfort levels in children and adult cochlear implant patients.,” *Ear and hearing,* pp. vol. 15, no2, pp. 184-192, 1994.

[31] K. Seyle and C. Brown, “Speech perception using maps based on neural response telemetry measures.,” *Ear Hear.,* pp. Feb;23(1 Suppl):72S-79S, 2002.

[32] G. Caner, L. Olgun, G. Gültekin and M. Balaban, “Optimizing fitting in children using objective measures such as neural response imaging and electrically evoked stapedius reflex threshold.,” *Otol Neurotol,* pp. Aug;28(5):637-40, 2007.

[33] K. Gordon, B. Papsin and R. Harrison, “Programming cochlear implant stimulation levels in infants and children with a combination of objective measures.,” *Int J Audiol,* pp. Dec;43 Suppl 1:S28-32, 2004.

[34] J. Allum, R. Greisiger and R. Probst, “Relationship of intraoperative electrically evoked stapedius reflex thresholds to maximum comfortable loudness levels of children with cochlear implants.,” *Int J Audiol,* pp. 41(2):93-9, 2002.

[35] E. Shaw, “Transformation of sound pressure level from the free field to the eardrum in the horizontal plane,” *The Journal of the Acoustical Society of America,* vol. 56, p. 1848, 1974.

[36] D. W. Batteau, “The role of the pinna in human localization,” *Proceedings of the Royal Society of London. Series B. Biological Sciences,* vol. 168, no. 1011, pp. 158-180, 1967.

[37] B. R. Glasberg and B. C. Moore, “Prediction of absolute thresholds and equal-loudness contours using a modified loudness model,” *The Journal of the Acoustical Society of America,* vol. 120, p. 585, 2006.

[38] B. Johnstone, R. Patuzzi and G. Yates, “Basilar membrane measurements and the travelling wave,” *Hearing Research,* vol. 22, no. 1, pp. 147--153, 1986.

[39] E. Borg, “A quantitative study of the effect of the acoustic stapedius reflex on sound transmission through the middle ear of man,” *Acta oto-laryngologica,* vol. 66, no. 1-6, pp. 461--472, 1968.

[40] I. Klockhoff and H. Anderson, “Reflex activity in the tensor tympani muscle recorded in man: Preliminary report,” *Acta oto-laryngologica,* vol. 51, no. 1-2, pp. 184--188, 1960.

[41] R. W. Baloh and K. Kerber, Baloh and Honrubia's Clinical Neurophysiology of the Vestibular System, Oxford university press, 2010.

[42] U. Friberg, J. Stahle and A. Svedberg, “The natural course of Meniere's disease,” *Acta Oto-Laryngologica,* vol. 96, no. S406, pp. 72-77, 1983.

[43] M. Torres and F. Giraldez, “The development of the vertebrate inner ear,” *Mechanisms of development,* vol. 71, no. 1, pp. 5--21, 1998.

[44] P. Dallos and R. Fay, The Cochlea, Springer London, Limited, 2012.

[45] B. C. Moore, Cochlear hearing loss: physiological, psychological and technical issues, Wiley. com, 2007.

[46] P. Dallas, “The active cochlea,” *The journal of neuroscience,* vol. 2, no. 12, pp. 4575-4585, 1992.

[47] L. Nordang, “Glutamate is the afferent neurotransmitter in the human cochlea,” *Acta oto-laryngologica,* vol. 120, no. 3, pp. 359--362, 2000.

[48] D. Kemp, “Otoacoustic emissions, travelling waves and cochlear mechanisms,” *Hearing research,* vol. 22, no. 1, pp. 95--104, 1986.

[49] A. Crawford and R. Fettiplace, “The frequency selectivity of auditory nerve fibres and hair cells in the cochlea of the turtle,” *The Journal of physiology,* vol. 306, no. 1, pp. 79-125, 1980.

[50] R. Galambos, “Suppression of auditory nerve activity by stimulation of efferent fibers to cochlea,” *Journal of Neurophysiology,* 1956.

[51] T. Konishi, P. E. Hamrick and P. J. Walsh, “Ion transport in guinea pig cochlea: I. Potassium and

sodium transport," *Acta oto-laryngologica,* vol. 86, no. 1-6, pp. 22-34, 1978.

[52] D. Purves, G. Augustine and D. Fitzpatrick, Neuroscience. 2nd edition, Sunderland (MA): Sinauer Associates, 2001.

[53] R. Lewis and A. Hudspeth, "Voltage-and ion-dependent conductances in solitary vertebrate hair cells," *Nature,* vol. 304, pp. 538-541, 1983.

[54] A. Palmer and I. Russell, "Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells," *Hearing research,* vol. 24, no. 1, pp. 1-15, 1986.

[55] H. Davis, "Some principles of sensory receptor action," *Physiological reviews,* vol. 41, pp. 391--416, 1961.

[56] E. Goldstein, Sensation and Perception, Cengage Learning, 2013.

[57] B. Nayagam, M. Muniak and D. Ryugo, "The spiral ganglion: connecting the peripheral and central auditory systems.," *Hearing research,* vol. 278, no. 1, pp. 2--20, 2011.

[58] A. Siegel and H. Sapru, Essential Neuroscience. 2nd edition, Lippincott Williams & Wilkins, 2010.

[59] W. Shofner and E. Young, "Excitatory/inhibitory response types in the cochlear nucleus: relationships to discharge patterns and responses to electrical stimulation of the auditory nerve," *Journal of neurophysiology,* vol. 54, no. 4, pp. 917--939, 1985.

[60] B. R. Glasberg and B. C. Moore, "A model of loudness applicable to time-varying sounds," *Journal of the Audio Engineering Society,* vol. 50, no. 5, pp. 331--342, 2002.

[61] P. Roland, W. Meyerhoff and B. Marple, Hearing Loss, Thieme, 1997.

[62] L. Robles and M. A. Ruggero, "Mechanics of the mammalian cochlea," *Physiological reviews,* vol. 81, no. 3, pp. 1305-1352, 2001.

[63] D. H. Johnson, "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," *The Journal of the Acoustical Society of America,* vol. 68, no. 4, pp. 1115--1122, 1980.

[64] E. Selkirk, "Sentence prosody: Intonation, stress, and phrasing," *The Handbook of Phonological Theory,* 1996.

[65] B. H. Repp and H. B. Lin, "Integration of segmental and tonal information in speech perception: a cross-linguistic study (A)," *Journal of the Acoustical Society of America,* pp. 87(S1), S46, 1990.

[66] P. F. Assmann, "Fundamental frequency and the intelligibility of competing voices.," 1999.

[67] A. E. Carney and D. A. Nelson, “An analysis of psychophysical tuning curves in normal and pathological ears,” *The Journal of the Acoustical Society of America,* vol. 73, no. 1, pp. 268--278, 1983.

[68] H. Fletcher, “Auditory patterns,” *Reviews of modern physics,* vol. 12, no. 1, p. 47, 1940.

[69] B. C. Moore and B. R. Glasberg, “Suggested formulae for calculating auditory-filter bandwidths and excitation patterns,” *The Journal of the Acoustical Society of America,* vol. 74, p. 750, 1983.

[70] R. Drullman, “Temporal envelope and fine structure cues for speech intelligibility,” *Journal of the Acoustical Society of America,* vol. 97, no. 1, pp. 585--592, 1995.

[71] K. R. Bachmann and J. C. Arvedson, “Early identification and intervention for children who are hearing impaired,” *Pediatrics in Review,* vol. 19, no. 5, pp. 155-165, 1998.

[72] A. Davis, J. Bamford, I. Wilson, T. Ramkalawan, M. Forshaw, S. Wright and others, “A critical review of the role of neonatal hearing screening in the detection of congenital hearing impairment.,” *Health technology assessment (Winchester, England),* vol. 1, no. 10, pp. i--iv, 1996.

[73] C. Kennedy and D. McCann, “Universal neonatal hearing screening moving from evidence to practice,” *Archives of Disease in Childhood-Fetal and Neonatal Edition,* vol. 89, no. 5, pp. F378--F383, 2004.

[74] H. M. Fortnum, A. Q. Summerfield, D. H. Marshall, A. C. Davis, J. M. Bamford, A. Davis, C. Yoshinaga-Itano and S. Hind, “Prevalence of permanent childhood hearing impairment in the United Kingdom and implications for universal neonatal hearing screening: questionnaire based ascertainment studyCommentary: Universal newborn hearing screening: implications for coordinating and developing services for deaf and hearing impaired children,” *Bmj,* vol. 323, no. 7312, p. 536, 2001.

[75] J. G. Clark, “Uses and abuses of hearing loss classification.,” *Asha,* vol. 23, no. 7, p. 493, 1981.

[76] P. Sellick, R. Patuzzi and B. Johnstone, “Measurement of basilar membrane motion in the guinea pig using the M{\"o}ssbauer technique,” *The journal of the acoustical society of America,* vol. 72, p. 131, 1982.

[77] H. J. McDermott and C. M. McKay, “Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea,” *The Journal of the Acoustical Society of America,* vol. 96, p. 155, 1994.

[78] P. C. Loizou, O. Poroy and M. Dorman, “The effect of parametric variations of cochlear implant

processors on speech understanding," *The Journal of the Acoustical Society of America,* vol. 108, p. 790, 2000.

[79] K. Arora, P. Dawson, R. Dowell and A. Vandali, "Electrical stimulation rate effects on speech perception in cochlear implants," *International journal of audiology,* vol. 48, no. 8, pp. 561-567, 2009.

[80] T. Balkany, A. Hodges, C. Menapace, L. Hazard, C. Driscoll, B. Gantz, D. Kelsall, W. Luxford, S. McMenomy, J. G. Neely and others, "Nucleus freedom North American clinical trial," *Otolaryngology--Head and Neck Surgery,* vol. 136, no. 5, pp. 757-762, 2007.

[81] A. Büchner, C. Frohne-Büchner, R. D Battmer and T. Lenarz, "Two years of experience using stimulation rates between 800 and 5000 pps with the clarion CII implant," in *International Congress Series*, 2004.

[82] A. E. Vandali, L. A. Whitford, K. L. Plant, G. M. Clark and others, "Speech perception as a function of electrical stimulation rate: using the Nucleus 24 cochlear implant system," *Ear and hearing,* vol. 21, no. 6, pp. 608-624, 2000.

[83] B. Vaerenberg, C. Smits, G. De Ceulaer, E. Zir, S. Harman, N. Jaspers, Y. Tam, M. Dillon, T. Wesarg, D. Martin-Bonniot, L. Gaertner, S. Cozma, J. Kosaner, S. Prentiss, P. Sasidharan and J. J. Briaire, "Cochlear Implant Programming: A Global Survey on the State of the Art," *The Scientific World Journal,* vol. 2014, 2014.

[84] Y. Cope and C. Totten, "Cochlear implants for young children. Second Edition.," S. A. Barry McCormick, Ed., London: Whurr, 2003, pp. 217-56.

[85] W. H. Shapiro and T. S. Bradham, "Cochlear implant programming," *Otolaryngologic Clinics of North America,* vol. 45, no. 1, pp. 111-127, 2012.

[86] R. Carhart and J. F. Jerger, "Preferred method for clinical determination of pure-tone thresholds," *Journal of Speech and Hearing Disorders,* vol. 24, no. 4, p. 330, 1959.

[87] B. S. Wilson, C. C. Finley, D. T. Lawson, R. D. Wolford, D. K. Eddington and W. M. Rabinowitz, "Better speech recognition with cochlear implants," *Nature,* vol. 352, no. 6332, pp. 236-238, 1991.

[88] J. Wolfe and E. C. Schafer, Programming cochlear implants, Plural Publishing, 2010.

[89] Cochlear, "Cochlear Ltd. Custom Sound 3.2 Help contents," 2012.

[90] MED-EL, "MED-EL m.b.H. Maestro 4.0 Help contents," 2012.

[91] AB, "Advanced Bionics LLC. SoundWave 2.1 Help contents," 2012.

[92] Neurelec, “Neurelec SA. Digimap 3.4 Help contents,” 2012.

[93] I. Hochmair, P. Nopp, C. Jolly, M. Schmidt, H. Schosser, C. Garnham and I. Anderson, “MED-EL cochlear implants: state of the art and a glimpse into the future,” *Trends in amplification,* vol. 10, no. 4, pp. 201-219, 2006.

[94] D. Koch, M. Downing, M. Osberger and L. Litvak, “Using current steering to increase spectral resolution in CII and HiRes 90K users.,” *Ear and Hearing,* vol. 28.2, pp. 38S-41S, 2007.

[95] B. Stobich, C. Zierhofer and E. Hochmair, “Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy.,” *Ear and hearing,* vol. 20.2, no. 0196-0202 (Linking), pp. 104-116, 1999.

[96] F.-G. Zeng, G. Grant, J. Niparko, J. Galvin, R. Shannon, J. Opie and P. Segel, “Speech dynamic range and its effect on cochlear implant performance,” *The Journal of the Acoustical Society of America,* vol. 111, p. 377, 2002.

[97] L. M. Friesen, R. V. Shannon, D. Baskent and X. Wang, “Speech recognition in noise as a function of the number of spectral channels: comparison of acoustic hearing and cochlear implants,” *The Journal of the Acoustical Society of America,* vol. 110, p. 1150, 2001.

[98] E. Walravens, D. Mawman and M. O'driscoll, “Changes in psychophysical parameters during the first month of programming the Nucleus Contour and Contour Advance cochlear implants,” *Cochlear Implants International,* vol. 7, no. 1, pp. 15-32, 2006.

[99] Y. Henkin, R. Kaplan-Neeman, C. Muchnik, J. Kronenberg and M. Hildesheimer, “Changes over time in the psycho-electric parameters in children with cochlear implants,” *Int J Audiol.,* pp. Jul;42(5):274-8, 2003.

[100] Y. Henkin, R. Kaplan-Neeman, J. Kronenberg, L. Migirov, M. Hildesheimer and C. Muchnik, “A longitudinal study of electrical stimulation levels and electrode impedance in children using the Clarion cochlear implant,” *Acta oto-laryngologica,* vol. 126, no. 6, pp. 581-586, 2006.

[101] P. J. Govaerts, B. Vaerenberg, G. De Ceulaer, K. Daemers, C. De Beukelaer and K. Schauwers, “Development of a software tool using deterministic logic for the optimization of cochlear implant processor programming,” *Otology \& Neurotology,* vol. 31, no. 6, pp. 908-918, 2010.

[102] B. Vaerenberg, P. J. Govaerts, G. De Ceulaer, K. Daemers and K. Schauwers, “Experiences of the use of FOX, an intelligent agent, for programming cochlear implant sound processors in new users,” *International journal of audiology,* vol. 50, no. 1, pp. 50-58, 2011.

[103] J. H. Frijns, W. M. C. Klop, R. M. Bonnet and J. J. Briaire, “Optimizing the number of electrodes

with high-rate stimulation of the clarion CII cochlear implant," *Acta oto-laryngologica,* vol. 123, no. 2, pp. 138-142, 2003.

[104] J. K. Shallop, "Objective measurements and the audiological management of cochlear implant patients," *Advances in Oto Rhino Laryngology,* vol. 53, pp. 85-111, 1997.

[105] N. Dillier, W. Lai, E. von Wallenberg, B. van Dijk, F. Akdas, M. Aksit, C. Batman, A. Beynon, S. Burdo, J.-M. Chanal and others, "Normative findings of electrically evoked compound action potential measurements using the neural response telemetry of the Nucleus CI24M cochlear implant system," *Audiology and Neurotology,* vol. 10, no. 2, pp. 105-116, 2005.

[106] L. G. Spivak, P. M. Chute, A. L. Popp and S. C. Parisier, "Programming the cochlear implant based on electrical acoustic reflex thresholds: patient performance," *The Laryngoscope,* vol. 104, no. 10, pp. 1225-1230, 1994.

[107] P. Govaerts, K. Daemers, M. Yperman, C. De Beukelaer, G. De Saegher and G. De Ceulaer, "Auditory speech sounds evaluation A§E: a new test to assess detection, discrimination and identification in hearing impairment," *Cochlear implants international,* vol. 7(2), pp. 92-106, 2006.

[108] G. De Ceulaer, S. Johnson, M. Yperman, K. Daemers, F. E. Offeciers, G. M. O'Donoghue and P. J. Govaerts, "Long-term evaluation of the effect of intracochlear steroid deposition on electrode impedance in cochlear implant patients," *Otology \& neurotology,* vol. 24, no. 5, pp. 769-774, 2003.

[109] J. L. Vargas, M. Sainz, C. Roldan, I. Alvarez and A. de la Torre, "Long-term evolution of the electrical stimulation levels for cochlear implant patients," *Clinical and experimental otorhinolaryngology,* vol. 5, no. 4, pp. 194-200, 2012.

[110] K. Plant, M.-A. Law, L. Whitford, M. Knight, S. Tari, J. Leigh, K. Pedley and E. Nel, "Evaluation of streamlined programming procedures for the Nucleus cochlear implant with the Contour electrode array," *Ear and hearing,* vol. 26, no. 6, pp. 651-668, 2005.

[111] B. Vaerenberg, G. De Ceulaer, Z. Szlávik, P. Mancini, A. Buechner and P. Govaerts, "Setting and reaching targets with computer-assisted cochlear implant fitting," *The Scientific World Journal,* 2014 (in press).

[112] B. McCormick, S. Archbold and S. Sheppard, Cochlear implants for young children, Whurr, 2002.

[113] S. Waltzman and J. Roland, Cochlear implants, Thieme, 2006.

[114] J. K. Niparko, Cochlear implants: Principles and practices, Wolters Kluwer Health, 2009.

[115] A. Starr and D. Brackmann, "Brain stem potentials evoked by electrical stimulation of the cochlea in human subjects.," *The Annals of otology, rhinology, and laryngology,* vol. 88, no. 4 Pt 1, pp. 550-556, 1978.

[116] J. W. Hall, New Handbook for Auditory Evoked Responses, Pearson Allyn and Bacon, 2007.

[117] P. R. Kileny, S. Zimmerman-Phillips, J. L. Kemink and S. Schmaltz, "Effects of preoperative electrical stimulability and historical factors on performance with multichannel cochlear implant.," *The Annals of otology, rhinology, and laryngology,* vol. 100, no. 7, p. 563, 1991.

[118] J. K. Shallop, L. VanDyke, D. W. Goin and R. E. Mischke, "Prediction of behavioral threshold and comfort values for Nucleus 22-channel implant patients from electrical auditory brain stem response test results," *Ann Otol Rhinol Laryngol,* vol. 100(11):, pp. 896-8, 1991.

[119] S. M. Mason, S. Sheppard, C. W. Garnham, M. E. Lutman, G. M. O'Donoghue and K. P. Gibbin, "Application of intraoperative recordings of electrically evoked ABRs in a paediatric cochlear implant programme," *Adv. Otorhinolaryngol.,* vol. 48, pp. 136-141, 1993.

[120] A. Hodges, S. Butts and J. King, "Cochlear Implants, Objective Measures," C. H, Ed., Whurr Publishers, London, Philadelphia, 2003, pp. 81-93.

[121] A. V. Hodges, T. J. Balkany, R. A. Ruth, P. R. Lambert, S. Dolan-Ash and J. J. Schloffman, "Electrical middle ear muscle reflex: use in cochlear implant programming," *Otolaryngology--Head and Neck Surgery,* vol. 117, no. 3, pp. 255-261, 1997.

[122] C. D. Brown, T. Antognelli and W. P. Gibson, "Auditory brain stem response evoked by electrical stimulation with a cochlear implant," *Adv. Otorhinolaryngol.,* vol. 48, pp. 125-129, 1993.

[123] J. King, "Objective Measures and Programming Cochlear Implants," *Perspectives on Hearing and Hearing Disorders in Childhood,* vol. 19, no. 2, pp. 54-62, 2009.

[124] C. J. Brown, M. L. Hughes, B. Luk, P. J. Abbas, A. Wolaver and J. Gervais, "The relationship between EAP and EABR thresholds and levels used to program the Nucleus 24 speech processor: data from adults," *Ear and hearing,* vol. 21, no. 2, pp. 151-163, 2000.

[125] M. L. Hughes, K. R. Vander Werff, C. J. Brown, P. J. Abbas, D. M. Kelsay, H. F. Teagle and M. W. Lowder, "A longitudinal study of electrode impedance, the electrically evoked compound action potential, and behavioral measures in nucleus 24 cochlear implant users," *Ear and hearing,* vol. 22, no. 6, pp. 471-486, 2001.

[126] K. H. Franck and S. J. Norton, "Estimation of psychophysical levels using the electrically evoked compound action potential measured with the neural response telemetry capabilities of

Cochlear Corporation's CI24M device," *Ear and hearing,* vol. 22, no. 4, pp. 289-299, 2001.

[127] P. J. Abbas, C. Brown, M. Hughes, B. Gantz, A. Wolaver, J. Gervais and S. Hong, "Electrically evoked compound action potentials recorded from subjects who use the nucleus CI24M device.," *The Annals of otology, rhinology \& laryngology. Supplement,* vol. 185, p. 6, 2000.

[128] G. Smoorenburg, C. Willeboer and J. van Dijk, "Speech perception in nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds," *Audiol Neurootol.,* pp. Nov-Dec;7(6):335-47, 2002.

[129] J. L. Smullen, M. Polak, A. V. Hodges, S. B. Payne, F. F. Telischi and T. J. Balkany, "Facial nerve stimulation after cochlear implantation," *The Laryngoscope,* vol. 115, no. 6, pp. 977-982, 2005.

[130] K. B. Agung, S. C. Purdy and C. Kitamura, "The Ling sound test revisited," *Australian and New Zealand Journal of Audiology, The,* vol. 27, no. 1, p. 33, 2005.

[131] N. Tye-Murray and W. Clark, Foundations of aural rehabilitation: Children, adults, and their family members, Singular Publishing Group, 1998.

[132] P. Yantis, "Handbook of Clinical Audiology. 4th Ed.," J. Katz, W. L. Gabbay and D. S. Ungerleider, Eds., Baltimore, Williams & Wilkins, 1994, pp. 107-108.

[133] A. Boothroyd, "Developments in speech audiometry," *British Journal of Audiology,* vol. 2, no. 1, pp. 3-10, 1968.

[134] G. v. Békésy, "A new audiometer," *Acta Oto-Laryngologica,* vol. 35, no. 5-6, pp. 411-422, 1947.

[135] R. H. Margolis, B. R. Glasberg, S. Creeke and B. C. Moore, "AMTAS®: Automated method for testing auditory sensitivity: Validation studies," *International journal of audiology,* vol. 49, no. 3, pp. 185-194, 2010.

[136] M. Taylor and C. D. Creelman, "PEST: Efficient estimates on probability functions," *The Journal of the Acoustical Society of America,* vol. 41, p. 782, 1967.

[137] H. Levitt, "Transformed up-down methods in psychoacoustics," *The Journal of the Acoustical society of America,* vol. 49, p. 467, 1971.

[138] D. M. Green, "A maximum-likelihood method for estimating thresholds in a yes--no task," *The Journal of the Acoustical Society of America,* vol. 93, p. 2096, 1993.

[139] M. Grassi and A. Soranzo, "MLP: A MATLAB toolbox for rapid and reliable auditory threshold estimation," *Behavior Research Methods,* vol. 41, no. 1, pp. 20-28, 2009.

[140] B. Treutwein, "Adaptive psychophysical procedures," *Vision research,* vol. 35, no. 17, pp. 2503-2522, 1995.

[141] M. R. Leek, “Adaptive procedures in psychophysical research,” *Perception \& Psychophysics,* vol. 63, no. 8, pp. 1279-1292, 2001.

[142] S. A. Klein, “Measuring, estimating, and understanding the psychometric function: A commentary,” *Perception \& psychophysics,* vol. 63, no. 8, pp. 1421-1455, 2001.

[143] S. Amitay, A. Irwin, D. J. Hawkey, J. A. Cowan and D. R. Moore, “A comparison of adaptive procedures for rapid and reliable threshold assessment and training in naive listeners,” *The Journal of the Acoustical Society of America,* vol. 119, p. 1616, 2006.

[144] K. Hopkins and B. C. Moore, “Development of a fast method for measuring sensitivity to temporal fine structure information at low frequencies,” *International journal of audiology,* vol. 49, no. 12, pp. 940-946, 2010.

[145] C. Kaernbach, “Adaptive threshold estimation with unforced-choice tasks,” *Perception \& Psychophysics,* vol. 63, no. 8, pp. 1377-1388, 2001.

[146] A. Sek and B. C. Moore, “Implementation of a fast method for measuring psychophysical tuning curves,” *International Journal of Audiology,* vol. 50, no. 4, pp. 237-242, 2011.

[147] G. Wetherill, “Sequential estimation of quantal response curves,” *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 1-48, 1963.

[148] P. Boersma, “Praat, a system for doing phonetics by computer,” *Glot international,* vol. 5, no. 9/10, pp. 341-345, 2002.

[149] D. M. Green, “Stimulus selection in adaptive psychophysical procedures,” *The Journal of the Acoustical Society of America,* vol. 87, p. 2662, 1990.

[150] J. H. McDermott and A. J. Oxenham, “Music perception, pitch, and the auditory system,” *Current opinion in neurobiology,* vol. 18, no. 4, pp. 452-463, 2008.

[151] B. C. Moore, “The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people,” *Journal of the Association for Research in Otolaryngology,* vol. 9, no. 4, pp. 399-406, 2008.

[152] B. H. Repp and H.-B. Lin, “Integration of segmental and tonal information in speech perception: A cross-linguistic study,” *The Journal of the Acoustical Society of America,* vol. 87, p. S46, 1990.

[153] G. S. Stickney, P. F. Assmann, J. Chang and F.-G. Zeng, “Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences,” *The Journal of the Acoustical Society of America,* vol. 122, p. 1069, 2007.

[154] B. C. Moore, “The role of temporal fine structure processing in pitch perception, masking, and speech perception for normal-hearing and hearing-impaired people,” *Journal of the Association for Research in Otolaryngology,* vol. 9, no. 4, pp. 399-406, 2008.

[155] C. Lorenzi, G. Gilbert, H. Carn, S. Garnier and B. C. Moore, “Speech perception problems of the hearing impaired reflect inability to use temporal fine structure,” *Proceedings of the National Academy of Sciences,* vol. 103, no. 49, pp. 18866-18869, 2006.

[156] B. a. H. W. a. G. P. J. Vaerenberg, “Managed estimation of psychophysical thresholds,” *Journal of Hearing Science,* vol. 3, no. 3, pp. 19--31, 2013.

[157] W. Heeren, M. Coene, B. Vaerenberg, A. Avram, A. Cardinaletti, L. del Bo, A. Pascu, F. Volpato and P. Govaerts, “Development of the A§E test battery for assessment of pitch perception in speech,” *Cochlear Implants International,* vol. 13, no. 4, pp. 206-219, 2012.

[158] E. Buss, J. W. Hall III and J. H. Grose, “Temporal fine-structure cues to speech and pure tone modulation in observers with sensorineural hearing loss,” *Ear and hearing,* vol. 25, no. 3, pp. 242-250, 2004.

[159] R. V. Shannon, “Temporal modulation transfer functions in patients with cochlear implants,” *The Journal of the Acoustical Society of America,* vol. 91, p. 2156, 1992.

[160] A. van Katwijk, Accentuation in Dutch: an experimental linguistic study, Van Gorcum, 1974.

[161] P. M. Bertinetto, “The perception of stress by Italian speakers.,” *Journal of Phonetics,* vol. 8, pp. 385-95, 1980.

[162] M. Chatterjee and S.-C. Peng, “Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition,” *Hearing research,* vol. 235, no. 1, pp. 143-156, 2008.

[163] R. J. Zatorre, “Pitch perception of complex tones and human temporal-lobe function,” *The Journal of the Acoustical Society of America,* vol. 84, p. 566, 1988.

[164] R. Meddis and L. O'Mard, “A unitary model of pitch perception,” *The Journal of the Acoustical Society of America,* vol. 102, p. 1811, 1997.

[165] M. F. Dorman, R. H. Gifford, A. J. Spahr and S. A. McKarns, “The benefits of combining acoustic and electric stimulation for the recognition of speech, voice and melodies,” *Audiology and Neurotology,* vol. 13, no. 2, pp. 105-112, 2007.

[166] J. Rubinstein, W. Parkinson, R. Tyler and B. Gantz, “Residual speech recognition and cochlear implant performance: effects of implantation criteria,” *Otology \& Neurotology,* vol. 20, no. 4, pp. 445-452, 1999.

[167] M. Martin, Speech Audiometry, 2nd Edition, Wiley, 2006.

[168] J. Jerger, C. Speaks and J. L. Trammell, “A new approach to speech audiometry,” *Journal of Speech and hearing disorders,* vol. 33, no. 4, p. 318, 1968.

[169] I. J. Hirsh, H. Davis, S. R. Silverman, E. G. Reynolds, E. Eldert and R. W. Benson, “Development of materials for speech audiometry,” *Journal of Speech and Hearing Disorders,* vol. 17, no. 3, p. 321, 1952.

[170] J. F. Werker and R. C. Tees, “Phonemic and phonetic factors in adult cross-language speech perception,” *The Journal of the Acoustical Society of America,* vol. 75, p. 1866, 1984.

[171] J. N. Holmes, Speech synthesis and recognition, Taylor \& Francis, 2001.

[172] P. Hubbard, “A general introduction to computer assisted language learning,” *Computer Assisted Language Learning: Critical Concepts in Linguistics,* vol. 1, pp. 1-20, 2009.

[173] H. Sakoe and S. Chiba, “Dynamic programming algorithm optimization for spoken word recognition,” *Acoustics, Speech and Signal Processing, IEEE Transactions on,* vol. 26, no. 1, pp. 43-49, 1978.

[174] S. Davis and P. Mermelstein, “Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences,” *Acoustics, Speech and Signal Processing, IEEE Transactions on,* vol. 28, no. 4, pp. 357-366, 1980.

[175] M. C. Allen, T. P. Nikolopoulos and G. M. O'Donoghue, “Speech Intelligibility in Children After Cochlear Implanation,” *Otology \& Neurotology,* vol. 19, no. 6, pp. 742-746, 1998.

[176] A. Bosman and G. Smoorenburg, “Woordenlijst voor spraakaudiometrie,” *CD onder auspiciën van de Nederlandse Vereniging voor Audiologie,* vol. N/A, p. N/A, 1992.

[177] J. Wouters, W. Damman and A. Bosman, “Vlaamse opname van woordenlijsten voor spraakaudiometrie,” *Logopedie: informatiemedium van de Vlaamse vereniging voor logopedisten,* vol. 7, no. 6, pp. 28-34, 1994.

[178] K.-H. Hahlbrock, “Über Sprachaudiometrie und neue Wörterteste,” *European Archives of Oto-Rhino-Laryngology,* vol. 162, no. 5, pp. 394-431, 1953.

[179] J. L. Reis, Surdez Diagnóstico e Reabilitação, Gazela Artes Gráfica, 2002.

[180] W. Damman, “De Brugse lijsten voor spraakaudiometrie-versie 1989,” *Tijdschrift voor Logopedie en Audiologie,* vol. 20, p. 21, 1990.

[181] J. A. Nelder and R. Mead, “A simplex method for function minimization,” *The computer*

journal, vol. 7, no. 4, pp. 308-313, 1965.

[182] B. Vaerenberg, A. Pascu, L. Del Bo, K. Schauwers, G. De Ceulaer, K. Daemers, M. Coene and P. J. Govaerts, “Clinical assessment of pitch perception,” *Otology \& Neurotology,* vol. 32, no. 5, pp. 736-741, 2011.

[183] K. Gfeller, C. Turner, J. Oleson, X. Zhang, B. Gantz, R. Froman and C. Olszewski, “Accuracy of cochlear implant recipients on pitch perception, melody recognition, and speech reception in noise,” *Ear and hearing,* vol. 28, no. 3, pp. 412-423, 2007.

[184] K. E. Fishman, R. V. Shannon and W. H. Slattery, “Speech recognition as a function of the number of electrodes used in the SPEAK cochlear implant speech processor,” *Journal of Speech, Language and Hearing Research,* vol. 40, no. 5, p. 1201, 1997.

[185] K. Hopkins, B. C. Moore and M. A. Stone, “Effects of moderate cochlear hearing loss on the ability to benefit from temporal fine structure information in speech,” *The Journal of the Acoustical Society of America,* vol. 123, p. 1140, 2008.

[186] C. Lorenzi, L. Debruille, S. Garnier, P. Fleuriot and B. C. Moore, “Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal,” *The Journal of the Acoustical Society of America,* vol. 125, p. 27, 2009.

[187] Q.-J. Fu, R. V. Shannon and X. Wang, “Effects of noise and spectral resolution on vowel and consonant recognition: Acoustic and electric hearing,” *The Journal of the Acoustical Society of America,* vol. 104, p. 3586, 1998.

[188] D. Gnansia, V. Péan, B. Meyer and C. Lorenzi, “Effects of spectral smearing and temporal fine structure degradation on speech masking release,” *The Journal of the Acoustical Society of America,* vol. 125, p. 4023, 2009.

[189] F.-G. Zeng, “Temporal pitch in electric hearing,” *Hearing research,* vol. 174, no. 1, pp. 101-106, 2002.

[190] T. Lenarz, “Electro-Acoustic Stimulation of the Cochlea. Editorial.,” *Audiol Neurootol,* vol. 14 Suppl 1, p. 1, 2009.

[191] B. J. Gantz and C. W. Turner, “Combining acoustic and electrical hearing,” *The Laryngoscope,* vol. 113, no. 10, pp. 1726-1730, 2003.

[192] C. W. Turner, B. J. Gantz, C. Vidal, A. Behrens and B. A. Henry, “Speech recognition in noise for cochlear implant listeners: benefits of residual acoustic hearing,” *The Journal of the Acoustical Society of America,* vol. 115, p. 1729, 2004.

[193] C. Von Ilberg, J. Kiefer, J. Tillein, T. Pfenningdorff, R. Hartmann, E. Sturzebecher and R. Klinke,

"Electric-acoustic stimulation of the auditory system. New technology for severe hearing loss.," *ORL; Journal for oto-rhino-laryngology and its related specialties,* vol. 61, no. 6, p. 334, 1999.

[194] S. Helbig, U. Baumann, M. Helbig, N. von Malsen-Waldkirch and W. Gstoettner, "A new combined speech processor for electric and acoustic stimulation--eight months experience," *ORL,* vol. 70, no. 6, pp. 359-365, 2008.

[195] T. Lenarz, T. St{\"o}ver, A. Buechner, A. Lesinski-Schiedat, J. Patrick and J. Pesch, "Hearing conservation surgery using the hybrid-L electrode," *Audiology and Neurotology,* vol. 14, no. Suppl. 1, pp. 22-31, 2009.

[196] S. Brockmeier, M. Peterreins, A. Lorens, K. Vermeire, S. Helbig, I. Anderson, H. Skarzynski, P. Van De Heyning, W. Gstoettner and J. Kiefer, "Music perception in electric acoustic stimulation users as assessed by the Mu. SIC test," *Adv Otorhinolaryngol,* vol. 67, pp. 70-80, 2010.

[197] K. Daemers, M. Yperman, C. De Beukelaer, G. De Saegher, G. De Ceulaer and P. Govaerts, "Normative data of the A§E discrimination and identification tests in preverbal children," *Cochlear implants international,* vol. 7, pp. 107-116, 2006.

[198] F. Di Lella, A. Bacciu, E. Pasanisi, V. Vincenti, M. Guida and S. Bacciu, "Main peak interleaved sampling (MPIS) strategy: effect of stimulation rate variations on speech perception in adult cochlear implant recipients using the Digisonic SP cochlear implant," *Acta oto-laryngologica,* vol. 130, no. 1, pp. 102-107, 2010.

[199] D. S. Lazard, P. Bordure, G. Lina-Granade, J. Magnan, R. Meller, B. Meyer, E. Radafy, P.-E. Roux, D. Gnansia, V. Pean and others, "Speech perception performance for 100 post-lingually deaf adults fitted with Neurelec cochlear implants: comparison between Digisonic Convex and Digisonic SP devices after a 1-year follow-up," *Acta oto-laryngologica,* vol. 130, no. 11, pp. 1267-1273, 2010.

[200] J. Wouters, W. Damman and A. Bosman, "Evaluation of Flemish-Dutch materials for speech audiometry," *status: published,* vol. N/A, p. N/A, 1995.

[201] W. K. Gstöettner, P. Van de Heyning, A. Fitzgerald O'Connor, C. Morera, M. Sainz, K. Vermeire, S. McDonald, L. Cavalle, S. Helbig, J. Garcia Valdecasas and others, "Electric acoustic stimulation of the auditory system: results of a multi-centre investigation," *Acta oto-laryngologica,* vol. 128, no. 9, pp. 968-975, 2008.

[202] W. Gstöettner, S. Helbig, C. Settevendemie, U. Baumann, J. Wagenblast, C. Arnoldner and others, "A new electrode for residual hearing preservation in cochlear implantation: first clinical results," *Acta oto-laryngologica,* vol. 129, no. 4, pp. 372-379, 2009.

[203] J. Kiefer, S. Hohl, E. Stürzebecher, T. Pfennigdorff and W. Gstöettner, “Comparison of Speech Recognition with Different Speech Coding Strategies (SPEAK, CIS, and ACE) and Their Relationship to Telemetric Measures of Compound Action Potentials in the Nucleus CI 24M Cochlear Implant System,” *International Journal of Audiology,* vol. 40, no. 1, pp. 32-42, 2001.

[204] C. J. James, B. Fraysse, O. Deguine, T. Lenarz, D. Mawman, A. Ramos, R. Ramsden and O. Sterkers, “Combined electroacoustic stimulation in conventional candidates for cochlear implantation,” *Audiology and Neurotology,* vol. 11, no. Suppl. 1, pp. 57-62, 2006.

[205] S. Prentiss, K. Sykes and H. Staecker, “Partial deafness cochlear implantation at the University of Kansas: techniques and outcomes,” *Journal of the American Academy of Audiology,* vol. 21, no. 3, pp. 197-203, 2010.

[206] B. J. Gantz, C. Turner and K. E. Gfeller, “Acoustic plus electric speech processing: preliminary results of a multicenter clinical trial of the Iowa/Nucleus Hybrid implant,” *Audiology and Neurotology,* vol. 11, no. Suppl. 1, pp. 63-68, 2006.

[207] B. Moore, “Basic auditory processes involved in the analysis of speech sounds.,” *Philos Trans R Soc Lond B Biol Sci.,* pp. Mar 12;363(1493):947-63., 2008.

[208] M. Sainz, A. de la Torre and C. Roldán, “Application of statistical analysis of the programming maps for fitting cochlear implants in children,” *Cochlear Implants Int.,* pp. Dec;4 Suppl 1:59-60, 2003.

[209] C. Finley, T. Holden, L. Holden, B. Whiting, R. Chole, J. Neely, T. Hullar and M. Skinner, “Role of Electrode Placement as a Contributor to Variability in Cochlear Implant Outcomes.,” *Otology & Neurotology,* pp. 27:920-8, 2008.

[210] M. Skinner, T. Holden, B. Whiting, A. Voie, B. Brunsden, J. Neely, E. Saxon, T. Hullar and C. Finley, “In vivo estimates of the position of advanced bionics electrode arrays in the human cochlea.,” *Ann Otol Rhinol Laryngol Suppl.,* pp. Apr;197:2-24, 2007.

[211] D. O. Kim and C. E. Molnar, “A population study of cochlear nerve fibers: comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones,” *J Neurophysiol,* pp. 16-30, 1979.

[212] H. Thai-Van, J. Chanal, C. Coudert, E. Veuillet, E. Truy and L. Collet, “Relationship between NRT measurements and behavioral levels in children with the Nucleus 24 cochlear implant may change over time: preliminary report.,” *Int J Pediatr Otorhinolaryngol.,* pp. Apr 27;58(2):153-62, 2001.

[213] G. F. Smoorenburg, C. Willeboer and J. E. van Dijk, “Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action

potential thresholds," *Audiology and Neurotology,* vol. 7, no. 6, pp. 335-347, 2002.

[214] A. Beiter, S. Staller and l. R. Dowel, "Evaluation and device programming in children," *Ear Hear.,* pp. Aug;12(4 Suppl):25S-33S., 1991.

[215] J. Rubinstein and R. Hong, "Signal coding in cochlear implants: exploiting stochastic effects of electrical stimulation," *Ann Otol Rhinol Laryngol Suppl.,* pp. Sep;191:14-9, 2003.

[216] D. Baskent, C. Eiler and B. Edwards, "Using genetic algorithms with subjective input from human subjects: implications for fitting hearing aids and cochlear implants.," *Ear Hear.,* pp. Jun;28(3):370-80., 2007.

[217] D. Baskent and B. Edwards, "Simulating listener errors in using genetic algorithms for perceptual optimization," *J Acoust Soc Am.,* pp. Jun;121(6):EL238-43, 2007.

[218] C. Willeboer and G. Smoorenburg, "Comparing cochlear implant users' speech performance with processor fittings based on conventionally determined T and C levels or on compound action potential thresholds and live-voice speech in a prospective balanced crossover study.," *Ear Hear.,* pp. Dec;27(6):789-98, 2006.

[219] D. Cafarelli Deesa, N. Dilliere, W. Lai, E. von Wallenberg, B. van Dijk, F. Akdas, M. Aksit, C. Batman, A. Beynon, S. Burdo, J. Chanal, L. Collet, M. Conway, C. Coudert, L. Craddock, H. Cullington, N. Deggouj, B. Fraysse and others, "Normative Findings of Electrically Evoked Compound Action Potential Measurements Using the Neural Response Telemetry of the Nucleus CI24M Cochlear Implant System," *Audiol Neuraudiol,* pp. Vol. 10, No. 2, 2005.

[220] P. Govaerts.Europe Patent EP1338301, 2003.

[221] B. Whiting, T. Holden, B. Brunsden, C. Finley and M. Skinner, "Use of Computed Tomography Scans for Cochlear Implants.," *J Digit Imaging,* pp. 21(3):323-8, 2008.

[222] R. deCharms, D. Blake and M. Merzenich, "Optimizing sound features for cortical neurons.," *Science,* 1998.

[223] F.-G. Zeng, K. Nie, S. Liu, G. Stickney, E. Del Rio, Y.-Y. Kong and H. Chen, "On the dichotomy in auditory perception between temporal envelope and fine structure cues," *J. Acoust. Soc. Am.,* pp. Volume 116, Issue 3, pp. 1351-1354, 2004.

[224] H. Gockel, E. Hafter and B. Moore, "Pitch discrimination interference: the role of ear of entry and of octave similarity.," *J Acoust Soc Am.,* pp. Jan;125(1):324-7, 2009.

[225] S. Singh, Y. Kong and F. Zeng, "Cochlear Implant Melody Recognition as a Function of Melody Frequency Range, Harmonicity, and Number of Electrodes.," *Ear Hear.,* vol. 30, no. 2, p. 160, 2009.

[226] G. Stickney, P. Assmann, J. Chang and F. Zeng, “Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences.,” *J Acoust Soc Am.,* pp. Aug;122(2):1069-78, 2007.

[227] J. Carroll and F. Zeng, “Fundamental frequency discrimination and speech perception in noise in cochlear implant simulations.,” *Hear Res.,* pp. Sep;231(1-2):42-53, 2007.

[228] H. Cullington and F. Zeng, “Speech recognition with varying numbers and types of competing talkers by normal-hearing, cochlear-implant, and implant simulation subjects.,” *J Acoust Soc Am.,* pp. Jan;123(1):450-61., 2008.

[229] A. Bhattacharya and F. Zeng, “ompanding to improve cochlear-implant speech recognition in speech-shaped noise.,” *J Acoust Soc Am.,* pp. Aug;122(2):1079-89, 2007.

[230] S. Sheft, M. Ardoint and C. Lorenzi, “Speech identification based on temporal fine structure cues.,” *J Acoust Soc Am.,* pp. Jul;124(1):562-75, 2008.

[231] P. Govaerts, K. Daemers, M. Yperman, C. De Beukelaer, G. De Saegher and G. De Ceulaer, “Auditory speech sounds evaluation (A§E®): a new test to assess detection, discrimination and identification in hearing impairment.,” *Cochlear Implants International,* pp. 7(2): 97-106, 2006.

[232] A. Summerfield, D. Marshall and A. Davis, “Cochlear implantation: demand, costs, and utility.,” *The Annals of otology, rhinology \& laryngology. Supplement,* vol. 166, pp. 245-248, 1995.

[233] “NIH Consensus Statement.,” *Cochlear implants in adults and children,* vol. 13, no. 2, pp. 1-30, 1995.

[234] R. E. Roditi, S. F. Poissant, E. M. Bero and D. J. Lee, “A predictive model of cochlear implant performance in postlingually deafened adults,” *Otology \& Neurotology,* vol. 30, no. 4, pp. 449-454, 2009.

[235] A. J. Spahr and M. F. Dorman, “Effects of minimum stimulation settings for the Med El Tempo+ speech processor on speech understanding,” *Ear and hearing,* vol. 26, no. 4, pp. 2S--6S, 2005.

[236] P. Dawson, M. Skok and G. Clark, “The effect of loudness imbalance between electrodes in cochlear implant users.,” *Ear and hearing,* vol. 18, no. 2, pp. 156-165, 1997.

[237] M. Sainz, A. de la Torre, C. Roldan, J. M. Ruiz and J. L. Vargas, “Analysis of programming maps and its application for balancing multichannel cochlear implants,” *International journal of audiology,* vol. 42, no. 1, pp. 43-51, 2003.

[238] P. J. Abbas, C. J. Brown, J. K. Shallop, J. B. Firszt, M. L. Hughes, S. H. Hong and S. J. Staller, “Summary of results using the Nucleus CI24M implant to record the electrically evoked compound action potential,” *Ear and Hearing,* vol. 20, no. 1, pp. 45-59, 1999.

[239] C. J. Brown, P. J. Abbas, H. Fryauf-Bertschy, D. Kelsay, B. J. Gantz and others, “Intraoperative and postoperative electrically evoked auditory brain stem responses in nucleus cochlear implant users: implications for the fitting process,” *Ear and hearing,* vol. 15, no. 2, pp. 168-176, 1994.

[240] A. V. Hodges, T. J. Balkany, R. A. Ruth, P. R. Lambert, S. Dolan-Ash and J. J. Schloffman, “Electrical middle ear muscle reflex: use in cochlear implant programming,” *Otolaryngol Head Neck Surg,* vol. 117, no. 3 Pt 1, pp. 255-261, Sep 1997.

[241] S. Mason, K. Gibbin, C. Garnham, G. O'donoghue and T. Twomey, “Intraoperative electrophysiological and objective tests after cochlear re-implantation in a young child,” *British journal of audiology,* vol. 30, no. 2, pp. 67-70, 1996.

[242] C. M. McKay, L. Fewster and P. Dawson, “A different approach to using neural response telemetry for automated cochlear implant processor programming,” *Ear and hearing,* vol. 26, no. 4, pp. 38S--44S, 2005.

[243] K. Daemers, M. Yperman, C. De Beukelaer, G. De Saegher, G. De Ceulaer and P. Govaerts, “Normative data of the A§E® discrimination and identification tests in preverbal children.,” *Cochlear Implants International,* pp. 7(2): 107-116, 2006.

[244] D. Poole, A. Mackworth and R. Goebel, Computational Intelligence, Oxford University Press Oxford, 1998.

[245] S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik and D. D. Edwards, Artificial intelligence: a modern approach, vol. 74, Prentice hall Englewood Cliffs, 1995.

[246] D. Heckerman, “Causal independence for knowledge acquisition and inference,” in *Proceedings of the Ninth international conference on Uncertainty in artificial intelligence*, 1993.

[247] D. Heckerman and J. S. Breese, “Causal independence for probability assessment and inference using Bayesian networks,” *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on,* vol. 26, no. 6, pp. 826-831, 1996.

[248] F. J. Díez and M. J. Druzdzel, “Canonical probabilistic models for knowledge engineering,” 2006.

[249] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference, Morgan Kaufmann, 1988.

[250] R. A. Howard and J. E. Matheson, Readings on the principles and applications of decision analysis, Strategic Decisions Group, 1984.

[251] D. Koller and A. Pfeffer, “Object-oriented Bayesian networks,” in *Proceedings of the Thirteenth*

conference on Uncertainty in artificial intelligence, 1997.

[252] O. Bangsø and P.-H. Wuillemin, “Top-down construction and repetitive structures representation in Bayesian networks,” in *IN PROCEEDINGS OF THE THIRTEENTH INTERNATIONAL FLORIDA ARTIFICIAL INTELLIGENCE RESEARCH SOCIETY CONFERENCE*, 2000.

[253] R. D. Shachter and M. A. Peot, “Simulation approaches to general probabilistic inference on belief networks,” *arXiv preprint arXiv:1304.1526,* vol. N/A, p. N/A, 2013.

[254] S. L. Lauritzen, “The EM algorithm for graphical association models with missing data,” *Computational Statistics \& Data Analysis,* vol. 19, no. 2, pp. 191-201, 1995.

[255] C. Lacave and F. Diez, “A review of explanation methods for Bayesian networks,” *Knowledge Engineering Review,* vol. 17, pp. 107-127, 2002.

[256] C. Lacave, A. Onisko and F. Diez, “Use of Elvira's explanation facilities for debugging probabilistic expert systems,” *Knowledge-Based Systems,* vol. 19, pp. 730-738, 2006.

[257] A. Schaub, “Solving the trade-off between speech understanding and listening comfort,” *The Hearing Journal,* vol. 63, no. 7, pp. 26-28, 2010.

[258] C. Willeboer and G. F. Smoorenburg, “Comparing cochlear implant users' speech performance with processor fittings based on conventionally determined T and C levels or on compound action potential thresholds and live-voice speech in a prospective balanced crossover study,” *Ear and hearing,* vol. 27, no. 6, pp. 789-798, 2006.

[259] B. Almqvist, “The Almqvist/Lund procedure for adjusting a speech processor for a congenitally deaf child using objective measurements.,” *Cochlear Nucleus Report,* p. 3–4, 2004.

[260] R. M. Cox and D. M. McDaniel, “Development of the Speech Intelligibility Rating (SIR) test for hearing aid comparisons,” *Journal of Speech, Language and Hearing Research,* vol. 32, no. 2, p. 347, 1989.

[261] J. K. Niparko, D. L. Oviatt, N. J. Coker, L. Sutton, S. B. Waltzman and N. L. Cohen, “Facial nerve stimulation with cochlear implantation,” *Otolaryngology--Head and Neck Surgery,* vol. 104, no. 6, pp. 826-830, 1991.

[262] G. F. Smoorenburg, C. Willeboer and J. E. van Dijk, “Speech perception in Nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds,” *Audiology and Neurotology,* vol. 7, no. 6, pp. 335-347, 2002.

[263] P. J. Abbas, C. J. Brown and C. P. Etler, “Electrophysiology and device telemetry,” in *Cochlear implants*, Thieme, 2006, pp. 96-109.

[264] J. Baudhuin, J. Cadieux, J. B. Firszt, R. M. Reeder and J. L. Maxson, “Optimization of programming parameters in children with the advanced bionics cochlear implant,” *Journal of the American Academy of Audiology,* vol. 23, no. 5, p. 302, 2012.

[265] L. K. Holden, R. M. Reeder, J. B. Firszt and C. C. Finley, “Optimizing the perception of soft speech and speech in noise with the Advanced Bionics cochlear implant system,” *International journal of audiology,* vol. 50, no. 4, pp. 255-269, 2011.

[266] K. Plant, L. Holden, M. Skinner, J. Arcaroli, L. Whitford, M.-A. Law and E. Nel, “Clinical evaluation of higher stimulation rates in the nucleus research platform 8 system,” *Ear and hearing,* vol. 28, no. 3, pp. 381-393, 2007.

[267] M. W. Skinner, P. L. Arndt and S. J. Staller, “Nucleus (R) 24 advanced encoder conversion study: performance versus preference,” *Ear and Hearing,* vol. 23, no. 1, pp. 2S--17S, 2002.

[268] A. J. Spahr, M. F. Dorman and L. H. Loiselle, “Performance of patients using different cochlear implant systems: Effects of input dynamic range,” *Ear and Hearing,* vol. 28, no. 2, pp. 260-275, 2007.

[269] L. Craddock, “Cochlear Implants: A Practical Guide. Second Edition,” L. C. Huw Cooper, Ed., Wiley, 2006, pp. 274-298.

[270] A. Büchner, T. Lenarz, P.-P. Boermans, J. H. Frijns, P. Mancini, R. Filipo, C. Fielden, H. Cooper, M. Eklöf, A. Freijd and others, “Benefits of the HiRes 120 coding strategy combined with the Harmony processor in an adult European multicentre study,” *Acta oto-laryngologica,* vol. 132, no. 2, pp. 179-187, 2012.

[271] I. Hochmair-Desoyer, E. Schulz, L. Moser and M. Schmidt, “The HSM sentence test as a tool for evaluating the speech understanding in noise of cochlear implant users,” *Otology \& Neurotology,* vol. 18, no. 6, pp. S83--S86, 1997.

[272] P. J. Boyle, T. B. Nunn, A. F. O'Connor and B. C. Moore, “STARR: A Speech Test for Evaluation of the Effectiveness of Auditory Prostheses Under Realistic Conditions,” *Ear and hearing,* vol. 34, no. 2, pp. 203-212, 2013.

[273] S. Haumann, T. Lenarz and A. Buechner, “Speech perception with cochlear implants as measured using a roving-level adaptive test method,” *ORL,* vol. 72, no. 6, pp. 312-318, 2010.

[274] D. S. Lazard, C. Vincent, F. Venail, P. Van de Heyning, E. Truy, O. Sterkers, P. H. Skarzynski, H. Skarzynski, K. Schauwers, S. O'Leary, D. Mawman, B. Maat, A. Kleine-Punte, A. M. Huber, K. Green, P. J. Govaerts, B. Fraysse, R. Dowell, N. Dillier, E. Burke, A. Beynon, F. Bergeron, D. Baskent, F. Artares and P. J. Blamey, “Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model

over time.," *PLoS ONE,* vol. 7, no. 11, p. e48739, 11 2012.

[275] A. E. Holmes, R. Shrivastav, L. Krause, H. W. Siburt and E. Schwartz, "Speech based optimization of cochlear implants," *International Journal of Audiology,* vol. 51, no. 11, pp. 806-816, 2012.

[276] N. Courtade, "The Lafon test and speech retardation].," *Revue de laryngologie-otologie-rhinologie,* vol. 87, no. 3, p. 181, 1966.

[277] J. Bench, A. Kowal and J. Bamford, "The BKB (Bamford-Kowal-Bench) sentence lists for partially-hearing children," *British Journal of Audiology,* vol. 13, no. 3, pp. 108-112, 1979.

[278] P. Blamey, F. Artieres, D. Baskent, F. Bergeron, A. Beynon, E. Burke, N. Dillier, R. Dowell, B. Fraysse, S. Galego and others, "Factors affecting auditory performance of postlinguistically deaf adults using cochlear implants: an update with 2251 patients," *Audiology and Neurotology,* vol. 18, no. 1, pp. 36-47, 2012.

[279] B. Vaerenberg, P. Govaerts, T. Stainsby, P. Nopp, A. Gault and D. Gnansia, "A uniform Graphical Representation of Intensity Coding in Current Generation Cochlear Implant Systems," *Ear and Hearing,* p. (in press), 2014.

[280] J. B. Firszt, L. K. Holden, M. W. Skinner, E. A. Tobey, A. Peterson, W. Gaggl, C. L. Runge-Samuelson and P. A. Wackym, "Recognition of speech presented at soft to loud levels by adult cochlear implant recipients of three cochlear implant systems," *Ear and hearing,* vol. 25, no. 4, pp. 375-387, 2004.

[281] A. Buechner, B. Vaerenberg, D. Gazibegovic, M. Brendel, C. De, P. Govaerts and T. Lenarz, "Evaluation of the "Fitting to Outcomes Expert" FOX with Established Cochlear Implant Users," *Cochlear Implants Int.,* p. (submitted), 2014.

[282] R.-D. Battmer, S. Borel, M. Brendel, A. Britz, A. Büchner, H. Cooper, C. Fielden, D. Gazibegovic, R. Goetze, P. Govaerts, T. Lenarz, I. Mosinier, J. Muff, T. Nunn, B. Vaerenberg and Z. Vanat, "Assessment of "fitting to outcomes expert" fox with new cochlear implant users in a multicentric study," *Cochlear Implants Int.,* p. (submitted), 2014.

[283] McDermott, "Music perception, pitch, and the auditory system," *Current Opinion in Neurobiology,* p. 18:1–12, 2008.

[284] C.-R. Shyu, J. M. Green, D. P. Lun, T. Kazic, M. Schaeffer and E. Coe, "Image Analysis for Mapping Immeasurable Phenotypes in Maize," *Signal Processing Magazine, IEEE,* pp. 116--119, 2007.

[285] R. S. Tyler, P. Abbas, N. Tye-Murray, B. J. Gantz, J. F. Knutson, B. F. Mccabe, C. Lansing, C. Brown, G. Woodworth, J. Hinrichs and others, "Evaluation of five different cochlear implant

designs: audiologic assessment and predictors of performance," *The Laryngoscope,* vol. 98, no. 10, pp. 1100-1106, 1988.

[286] E. Selkirk, "The Handbook of Phonological Theory," Blackwell Publishers, 1995.

[287] MED-EL, "FocusOnFineHearing™ Technology," 2012.

[288] C. Kaernbach, "Simple adaptive testing with the weighted up-down method," *Attention, Perception, \& Psychophysics,* vol. 49, no. 3, pp. 227-229, 1991.

[289] Y. Henkin, R. Kaplan-neeman, C. Muchnik, J. Kronenberg and M. Hildesheimer, "Changes over time in the psycho-electric parameters in children with cochlear implants: Cambios en el tiempo, de los par{\'a}metros psico-el{\'e}ctricos en ni{\~n}os con implante coclear," *International journal of audiology,* vol. 42, no. 5, pp. 274-278, 2003.

[290] G. A. Gates, K. Daly, W. J. Dichtel, R. J. Dooling, A. J. Gulya, J. W. Hall, S. W. Jerger, J. E. Jones, M. H. Mayer, M. Pierschalla and others, "Cochlear implants in adults and children," *JAMA: The Journal of the American Medical Association,* vol. 274, no. 24, pp. 1955-1961, 1995.

[291] Cochlear, "Cochlear™ Clinical Guidance Document, N33595F ISS3 FEB10; 2012.," 2012.

[292] G. Clark, Cochlear implants: fundamentals and applications, Springer, 2003.

[293] M. Brendel, A. Buechner, B. Krueger, C. Frohne-Buechner and T. Lenarz, "Evaluation of the Harmony soundprocessor in combination with the speech coding strategy HiRes 120," *Otology \& Neurotology,* vol. 29, no. 2, pp. 199-202, 2008.

[294] P. J. Boyle, A. Büchner, M. A. Stone, T. Lenarz and B. C. Moore, "Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants," *International Journal of Audiology,* vol. 48, no. 4, pp. 211-221, 2009.

[295] S. Archbold, M. Harris, G. O'Donoghue, T. Nikolopoulos, A. White and H. Richmond, "Reading abilities after cochlear implantation: the effect of age at implantation on outcomes at 5 and 7 years after implantation.," *Int J Pediatr Otorhinolaryngol,* vol. 72, no. 10, pp. 1471-8, 2008.

[296] i. C. Lorenz, L. Debruille, S. Garnier, P. Fleuriot and B. Moore, "Abnormal processing of temporal fine structure in speech for frequencies where absolute thresholds are normal.," *J Acoust Soc Am.,* pp. Jan;125(1):27-30, 2009.

[297] K. Schauwers, S. Gillis and P. Govaerts, "The characteristics of prelexical babbling in hearing-impaired children after cochlear implantation between 5 and 20 months of age," *Ear Hear,* pp. 29(4): 627-637, 2008.

[298] S. B. Waltzman, N. L. Cohen, R. H. Gomolin and W. H. Shapiro, "Long-Term Results of Early

Cochlear Implantation in Congenitally and Prelingually Deafened Children," *American Journal of Otology,* p. 9, 1994.

[299] Y. Tong, R. Black, G. Clark, I. Forster, J. Millar, B. O'Loughlin and J. Patrick, "A preliminary report on a multiple-channel cochlear implant operation.," *J Laryngol Otol,* pp. 679-695, 1979.

[300] M. J. A. Makhdoum, A. F. M. Snik and P. van den Broek, "Cochlear implantation: a review of the literature and the Nijmegen results.," *The Journal of Laryngology & Otology,* pp. 1008-1017, 1997.

[301] P. Govaerts, C. De Beukelaer, K. Daemers, G. De Ceulaer, M. Yperman, T. Somers, I. Schatteman and F. Offeciers, "Outcome of cochlear implantation at different ages from 0 to 6 years.," *Otol Neurotol,* pp. 23(6): 885-90, 2002.

[302] H. Francis, N. Chee, J. Yeagle, A. Cheng and J. Niparko, "Impact of cochlear implants on the functional health status of older adults.," *Laryngoscope,* pp. 112(8 Pt 1):1482-8, 2002.

[303] S. Connell and T. Balkany, "Cochlear implants," *Clin Geriatr Med,* pp. Aug;22(3):677-86, 2006.

[304] C. Lorenzi, G. Gilbert, H. Carn, S. Garnier and B. C. J. Moore, "Speech perception problems of the hearing impaired reflect inability to use temporal fine structure," *PNAS,* pp. vol. 103 no. 49 18866-18869, 2006.

[305] G. Barton, P. Stacey, H. Fortnum and S. AQ, "Hearing-impaired children in the United Kingdom, IV: cost-effectiveness of pediatric cochlear implantation," *Ear Hear,* pp. 27(5):575-88, 2006.

[306] C. J. Brown, "The Relationship Between EAP and EABR Thresholds and Levels Used to Program the Nucleus 24 Speech Processor: Data from Adults," *Ear and Hearing,* pp. Volume 21(2)April 2000pp 151-163, 2000.

[307] M. Sainz, A. de la Torre and C. Roldán, "Application of statistical analysis of the programming maps for fitting cochlear implants in children," *Cochlear Implants Int.,* pp. Dec;4 Suppl 1:59-60, 2003.

[308] C. Finley, T. Holden, L. Holden, B. Whiting, R. Chole, J. Neely, T. Hullar and M. Skinner, "Role of Electrode Placement as a Contributor to Variability in Cochlear Implant Outcomes.," *Otology & Neurotology,* pp. 27:920-8, 2008.

[309] M. Skinner, T. Holden, B. Whiting, A. Voie, B. Brunsden, J. Neely, E. Saxon, T. Hullar and C. Finley, "In vivo estimates of the position of advanced bionics electrode arrays in the human cochlea.," *Ann Otol Rhinol Laryngol Suppl.,* pp. Apr;197:2-24, 2007.

[310] D. O. Kim and C. E. Molnar, "A population study of cochlear nerve fibers: comparison of spatial distributions of average-rate and phase-locking measures of responses to single tones," *J*

Neurophysiol, pp. 16-30, 1979.

[311] Y. Henkin, R. Kaplan-neeman, C. Muchnik, J. Kronenberg and M. Hildesheimer, “Changes over time in the psycho-electric parameters in children with cochlear implants,” *Int J Audiol.,* pp. Jul;42(5):274-8, 2003.

[312] H. Thai-Van, J. Chanal, C. Coudert, E. Veuillet, E. Truy and L. Collet, “Relationship between NRT measurements and behavioral levels in children with the Nucleus 24 cochlear implant may change over time: preliminary report.,” *Int J Pediatr Otorhinolaryngol.,* pp. Apr 27;58(2):153-62, 2001.

[313] G. F. Smoorenburg, C. Willeboer and J. E. van Dijk, “Speech perception in nucleus CI24M cochlear implant users with processor settings based on electrically evoked compound action potential thresholds,” *Audiol Neurootol.,* pp. Nov-Dec;7(6):335-47, 2002.

[314] A. Beiter, S. Staller and l. R. Dowel, “Evaluation and device programming in children,” *Ear Hear.,* pp. Aug;12(4 Suppl):25S-33S., 1991.

[315] J. Rubinstein and R. Hong, “Signal coding in cochlear implants: exploiting stochastic effects of electrical stimulation,” *Ann Otol Rhinol Laryngol Suppl.,* pp. Sep;191:14-9, 2003.

[316] D. Baskent and B. Edwards, “Simulating listener errors in using genetic algorithms for perceptual optimization,” *J Acoust Soc Am.,* pp. Jun;121(6):EL238-43, 2007.

[317] D. Baskent, C. Eiler and B. Edwards, “Using genetic algorithms with subjective input from human subjects: implications for fitting hearing aids and cochlear implants.,” *Ear Hear.,* pp. Jun;28(3):370-80., 2007.

[318] C. Willeboer and G. Smoorenburg, “Comparing cochlear implant users' speech performance with processor fittings based on conventionally determined T and C levels or on compound action potential thresholds and live-voice speech in a prospective balanced crossover study.,” *Ear Hear.,* pp. Dec;27(6):789-98, 2006.

[319] D. Cafarelli Deesa, N. Dilliere, W. Lai, E. von Wallenberg, B. van Dijk, F. Akdas, M. Aksit, C. Batman, A. Beynon, S. Burdo, J. Chanal, L. Collet, M. Conway, C. Coudert, L. Craddock, H. Cullington, N. Deggouj, B. Fraysse and others, “Normative Findings of Electrically Evoked Compound Action Potential Measurements Using the Neural Response Telemetry of the Nucleus CI24M Cochlear Implant System,” *Audiol Neuraudiol,* pp. Vol. 10, No. 2, 2005.

[320] B. Whiting, T. Holden, B. Brunsden, C. Finley and M. Skinner, “Use of Computed Tomography Scans for Cochlear Implants.,” *J Digit Imaging,* pp. 21(3):323-8, 2008.

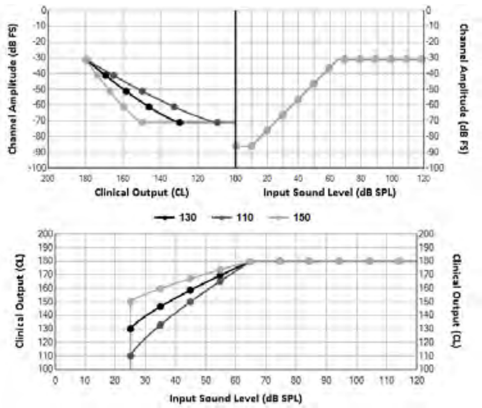
APPENDICES

Illustration of man and paper.

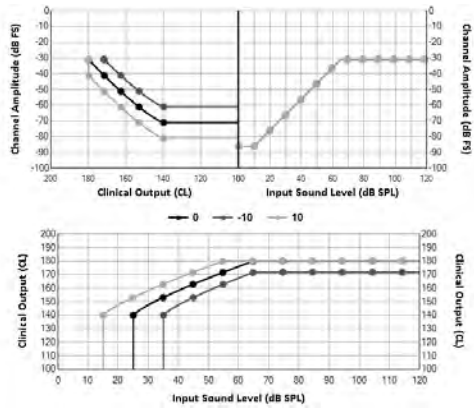
APPENDIX A: ICF PLOTS FOR MAP PARAMETERS

COCHLEAR

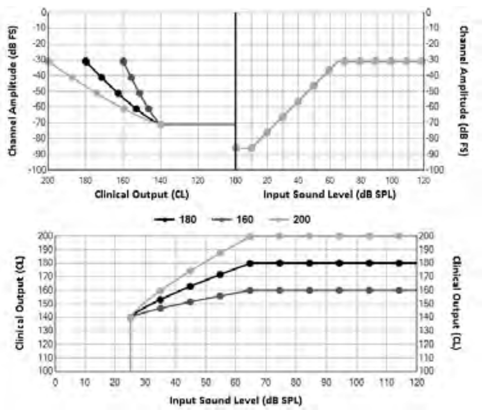
T



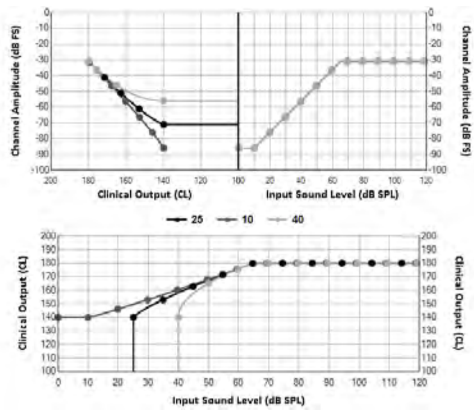
GAIN



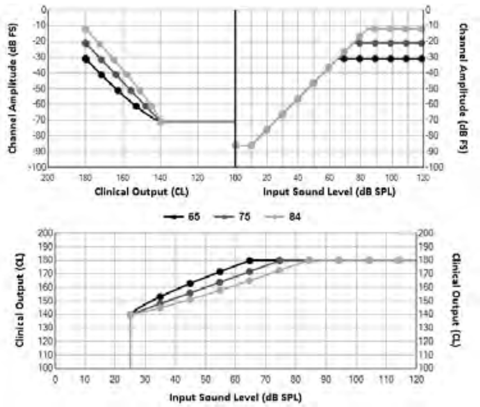
C



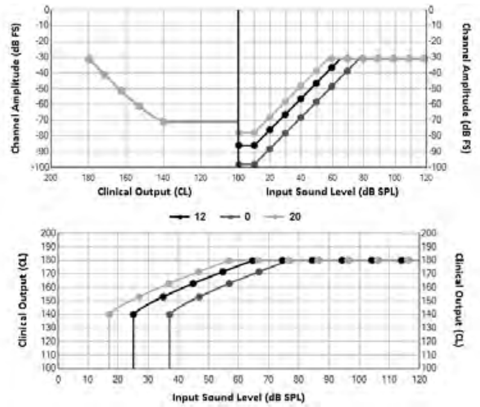
T-SPL



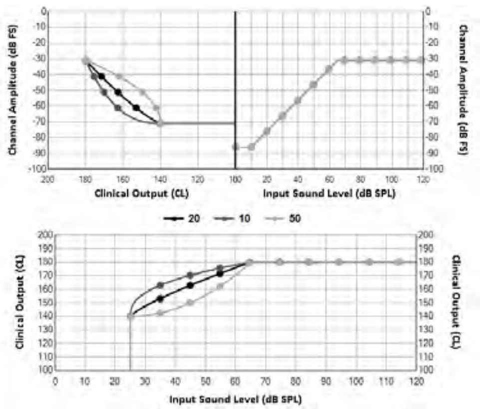
C-SPL



SENSITIVITY

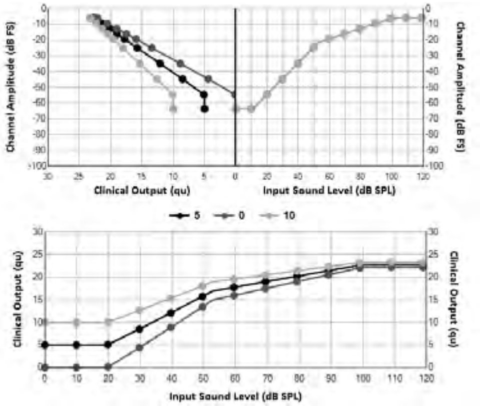


LOUDNESS GROWTH

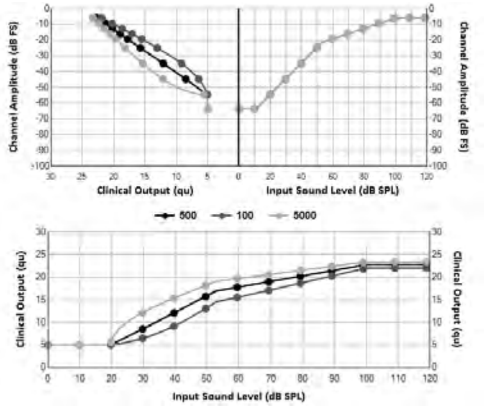


MED-EL

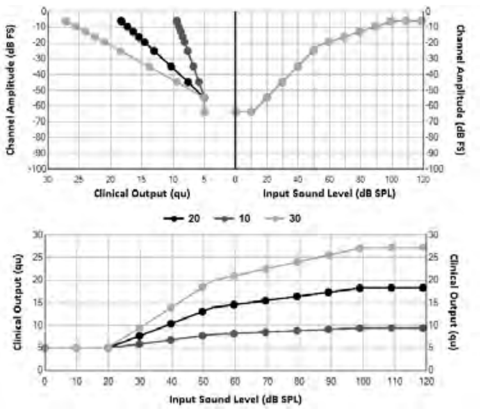
THR



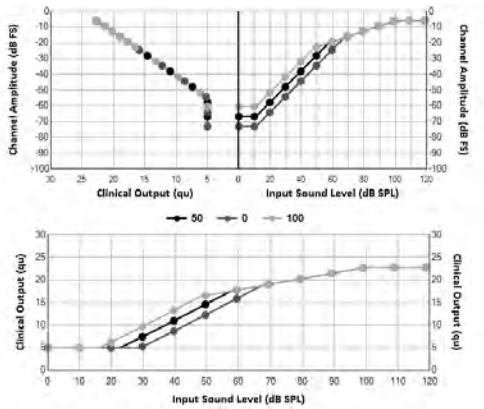
MAP LAW



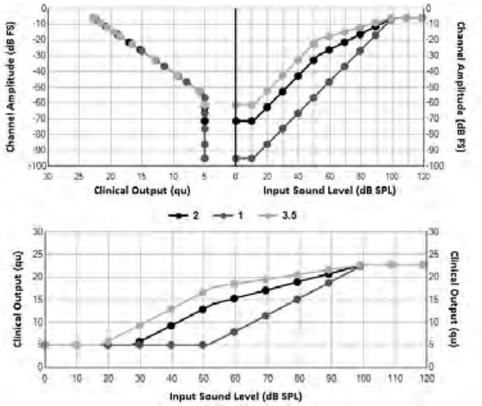
MCL



AGC SENSITIVITY

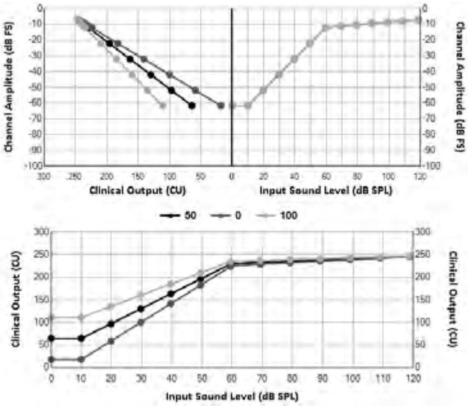


AGC COMPRESSION RATIO

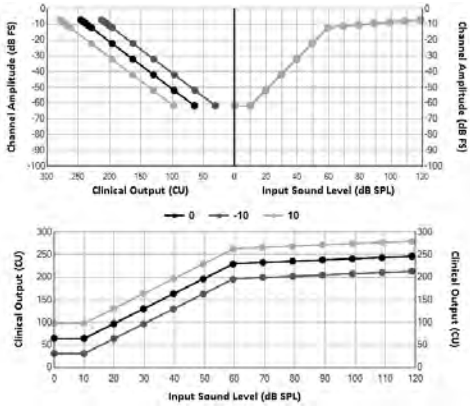


ADVANCED BIONICS

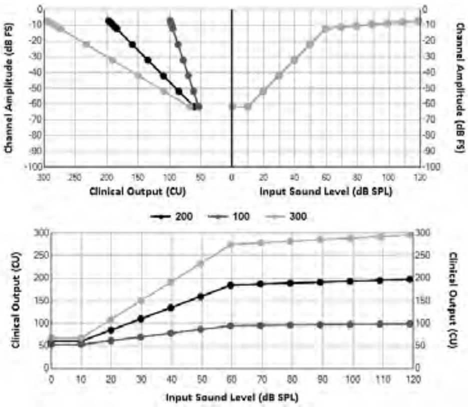
T



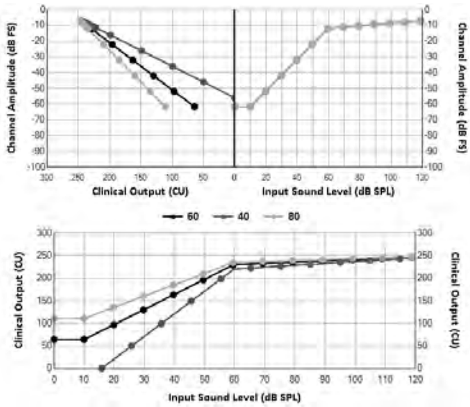
GAIN



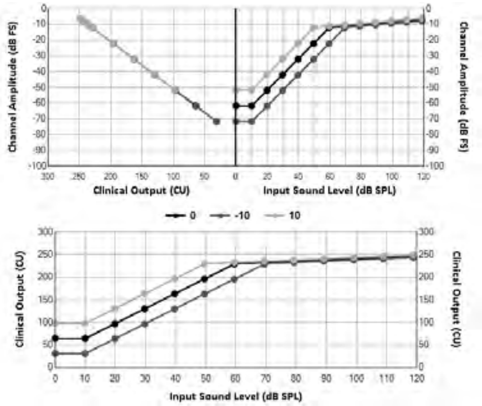
M



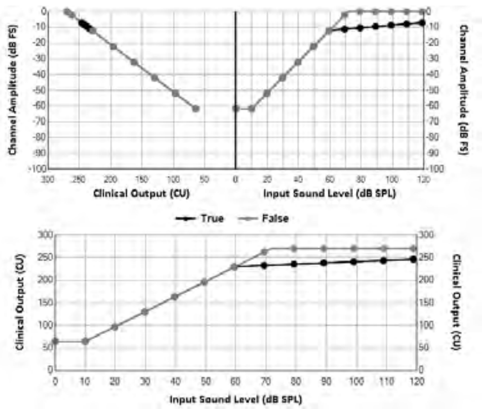
IDR



SENSITIVITY

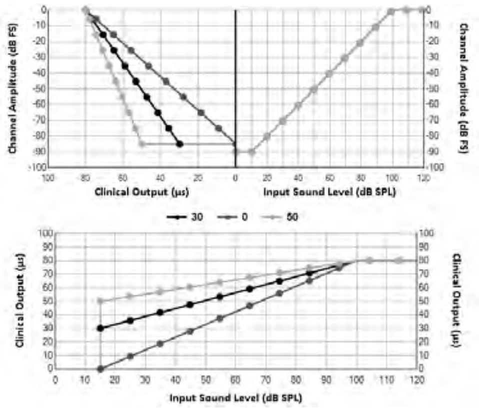


AGC

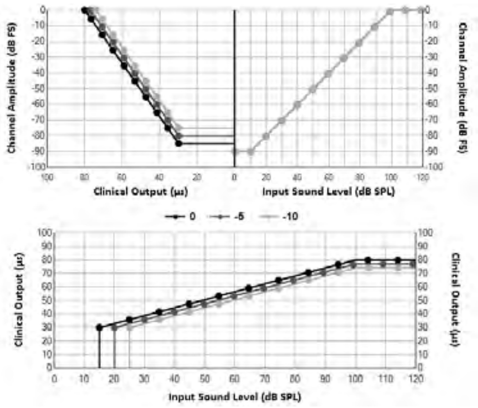


NEURELEC

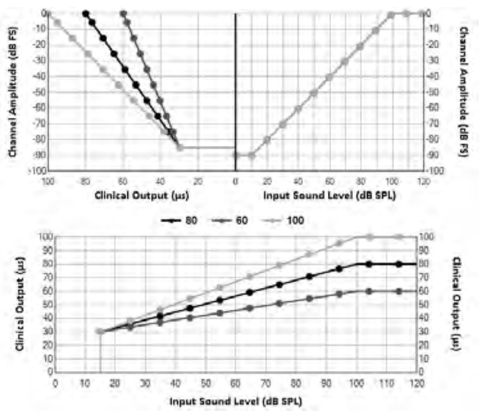
MIN



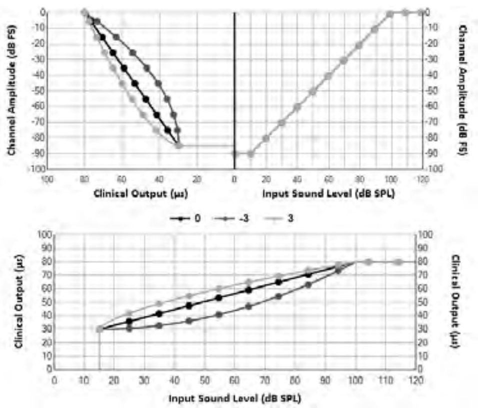
GAIN



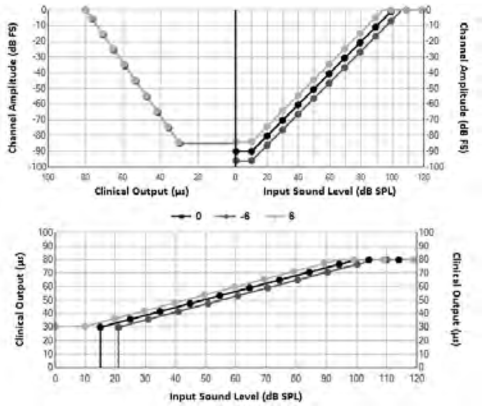
MAX



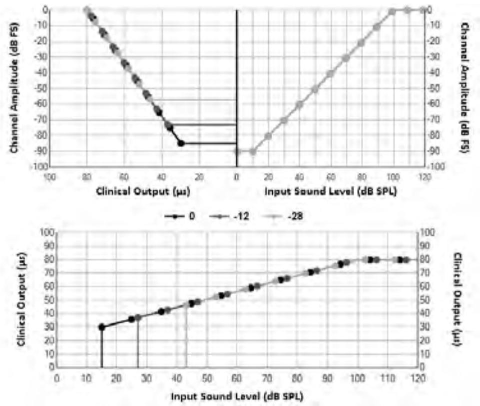
VOLUME



ANALOG GAIN



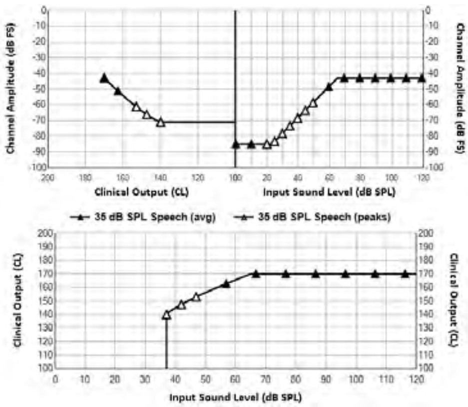
DENOISING



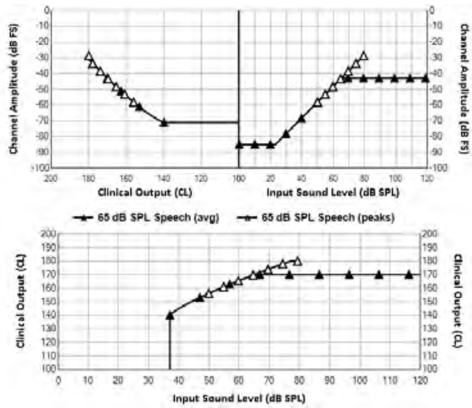
APPENDIX B: ICF PLOTS WITH FAST RESPONSES

COCHLEAR

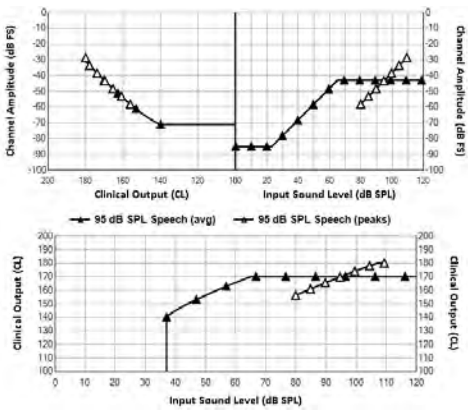
35 DB SPL WITH +/- 15 DB PEAKS



65 DB SPL WITH +/- 15 DB PEAKS

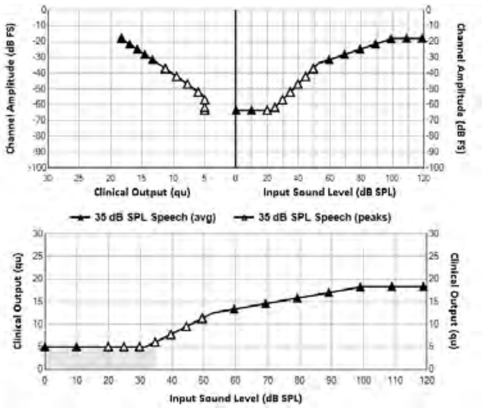


95 DB SPL WITH +/- 15 DB PEAKS

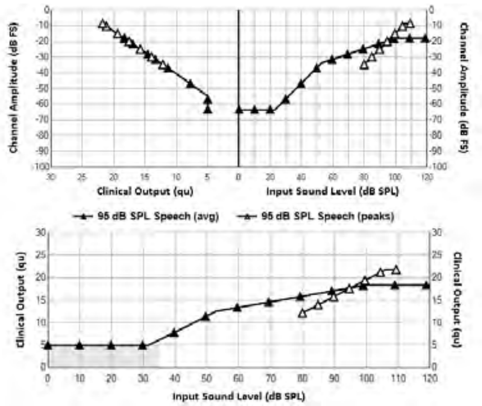


MED-EL

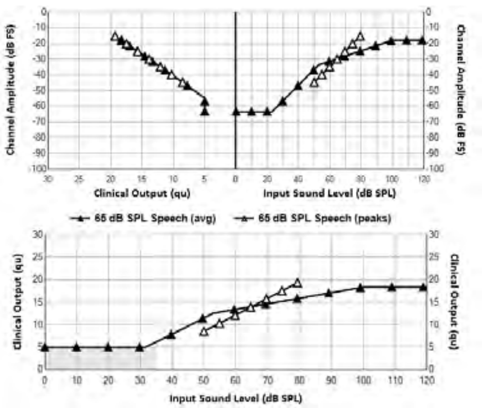
35 DB SPL WITH +/- 15 DB PEAKS



95 DB SPL WITH +/- 15 DB PEAKS

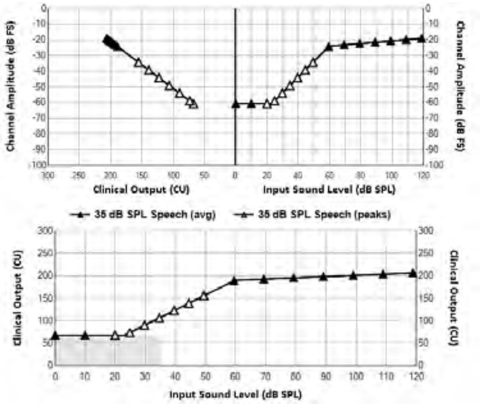


65 DB SPL WITH +/- 15 DB PEAKS

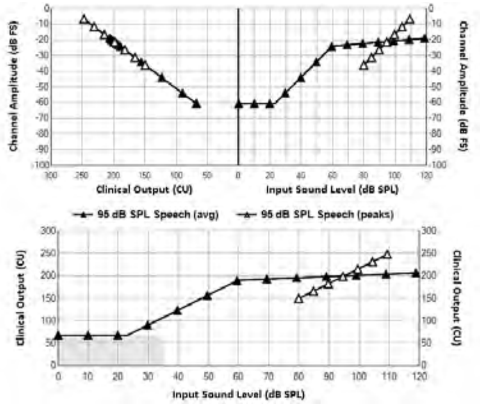


ADVANCED BIONICS

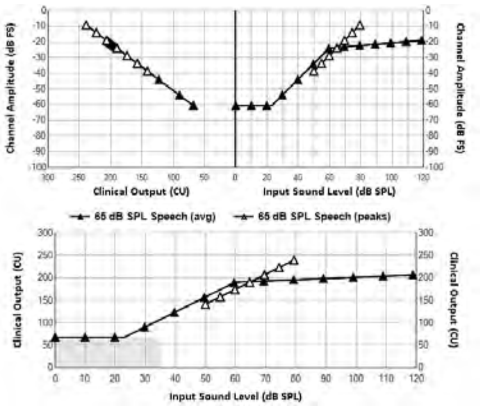
35 DB SPL WITH +/- 15 DB PEAKS



95 DB SPL WITH +/- 15 DB PEAKS

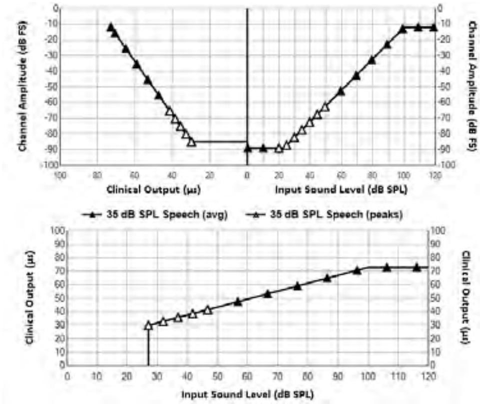


65 DB SPL WITH +/- 15 DB PEAKS

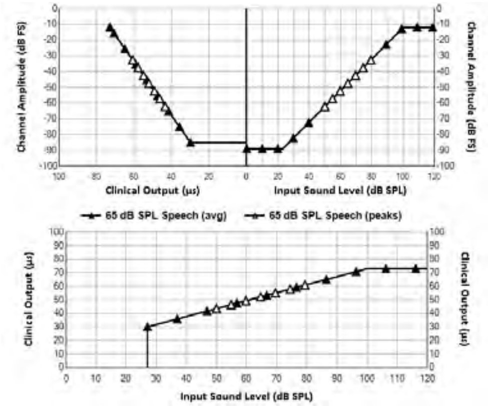


NEURELEC

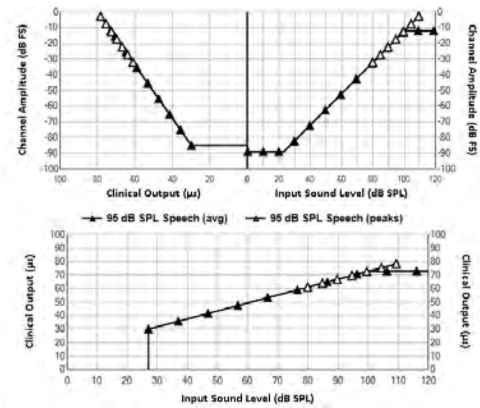
35 DB SPL WITH +/- 15 DB PEAKS



65 DB SPL WITH +/- 15 DB PEAKS



95 DB SPL WITH +/- 15 DB PEAKS



APPENDIX C: THE EARGOUP ADVICE

This appendix has been removed due to reasons of intellectual property.

APPENDIX D: TEMA MONTE CARLO SIMULATIONS

Table 20: Results of the Monte Carlo simulations. For both algorithms (TEMA and REF) each configuration was simulated 1000 times. Displayed are: Rejection ratio (REJECTED), the percentage of experiments where no threshold was found; threshold (THR) mean μ and standard deviation σ; number of trials (TRIALS) mean μ and standard deviation σ. Shaded fields are statistically significant (see text for details).

SIMULATION	CATEGORY	CONFIGURATION	TEST	TEMA					REF				
				REJECTED	THR		TRIALS		REJECTED	THR		TRIALS	
					μ	σ	μ	σ		μ	σ	μ	σ
Rejection Performance													
A1a	Gambler	Uses all options	WSP	99.5%	137.7	39.2	18.1	8.2	61.4%	152.2	46.2	16.7	7.3
A1b	Gambler	Uses all options	SI	88.0%	62.6	50.1	33.3	19.8	6.7%	83.0	53.9	20.4	5.2
A2a	Gambler	Thinks stimulus is always present	WSP	100.0%	N/A	N/A	19.2	7.2	33.0%	135.9	49.2	18.8	6.4
A2b	Gambler	Thinks stimulus is always present	SI	100.0%	N/A	N/A	13.0	0.0	100.0%	N/A	N/A	7.0	0.0
B1a	Cheater	reversalrate 1;delay 0	WSP	77.5%	78.5	29.7	26.9	7.1	0.0%	105.5	0.0	11.0	0.0
B1b	Cheater	reversalrate 1;delay 0	SI	78.8%	78.5	33.1	33.6	14.6	0.0%	105.5	0.0	11.0	0.0
B2a	Cheater	reversalrate 1;delay 3	WSP	70.6%	4.8	5.4	21.5	7.1	0.0%	9.0	0.0	14.0	0.0
B2b	Cheater	reversalrate 1 delay 3	SI	73.7%	6.7	6.2	22.2	7.7	0.0%	9.0	0.0	14.0	0.0
B3a	Cheater	reversalrate 2;delay 0	WSP	77.2%	111.4	33.1	27.2	8.0	0.0%	151.0	0.0	21.0	0.0
B3b	Cheater	reversalrate 2 delay 0	SI	80.1%	116.6	37.1	32.8	17.8	0.0%	151.0	0.0	21.0	0.0
B4a	Cheater	reversalrate 4;delay 0	WSP	86.2%	96.9	35.1	25.9	11.6	0.0%	113.0	0.0	41.0	0.0
B4b	Cheater	reversalrate 4;delay 0	SI	92.9%	103.6	32.9	30.1	24.4	0.0%	113.0	0.0	41.0	0.0
Threshold Accuracy													
C1a	Perfect	threshold 30	WSP	0.0%	27.5	0.0	21.3	4.3	0.0%	27.5	0.0	13.0	0.0
C1b	Perfect	threshold 30	SI	0.0%	27.5	0.0	21.8	4.8	0.0%	27.5	0.0	13.0	0.0
C2a	Perfect	threshold 2	WSP	0.0%	3.0	0.0	14.4	2.3	0.0%	3.0	0.0	14.0	0.0
C2b	Perfect	threshold 2	SI	0.0%	3.0	0.0	14.3	2.5	0.0%	3.0	0.0	14.0	0.0
C3a	Perfect	threshold 175	WSP	0.0%	176.5	0.0	23.6	7.4	0.0%	176.5	0.0	12.0	0.0
C3b	Perfect	threshold 175	SI	0.0%	176.5	0.0	23.7	7.2	0.0%	176.5	0.0	12.0	0.0
D1a	Normal	threshold 50;stdev 10	WSP	0.0%	49.8	4.7	23.1	4.7	0.0%	50.1	4.5	17.2	3.2
D1b	Normal	threshold 50 stdev 10	SI	0.0%	50.3	4.7	23.9	5.2	0.0%	50.0	4.4	17.1	3.1
D2a	Normal	threshold 15;stdev 10	WSP	0.3%	14.3	4.9	23.1	4.8	0.0%	15.1	4.2	18.2	3.1
D2b	Normal	threshold 15;stdev 10	SI	0.3%	15.0	4.9	23.6	4.9	0.0%	15.3	4.3	18.3	3.1
D3a	Normal	threshold 150;stdev 70	WSP	2.1%	145.3	24.9	27.5	6.6	0.2%	145.1	21.4	20.1	4.0
D3b	Normal	threshold 150;stdev 70	SI	4.1%	153.2	22.2	37.6	16.3	0.2%	146.0	21.2	20.1	3.8
D4a	Normal	threshold 50;stdev 150	WSP	49.3%	57.4	30.3	32.0	10.0	0.6%	59.1	29.4	20.8	4.2
D4b	Normal	threshold 50;stdev 150	SI	49.0%	58.9	31.3	43.4	21.4	0.4%	60.2	28.6	20.9	4.3
E1a	Drift	initial 100;target 20;stdev 20;step -8;delay 10	WSP	0.6%	23.7	14.9	40.7	7.0	0.0%	61.4	24.0	24.0	5.6
E1b	Drift	initial 100;target 20;stdev 20;step -8;delay 10	SI	1.0%	24.3	14.0	41.8	8.1	0.0%	61.9	23.5	24.1	5.2
E2a	Drift	initial 20 target 3 stdev 5 step -2 delay 5	WSP	9.8%	6.0	2.4	20.5	4.1	2.0%	5.5	2.7	18.0	2.4
E2b	Drift	initial 20;target 3;stdev 5;step -2;delay 5	SI	13.8%	6.7	3.1	21.7	5.1	1.9%	5.5	2.7	18.0	2.4
E3a	Drift	initial 50 target 70 stdev 30 step 5 delay 15	WSP	0.5%	61.3	13.5	29.3	7.5	0.0%	54.7	10.6	20.1	4.0
E3b	Drift	initial 50;target 70;stdev 30;step 5;delay 15	SI	0.1%	64.8	12.3	34.3	11.7	0.0%	54.7	11.0	20.2	4.0
E4a	Drift	initial 50 target 20 stdev 10 step -5 delay 0	WSP	0.0%	20.5	5.2	25.3	4.6	0.0%	20.4	4.4	19.2	3.0
E4b	Drift	initial 50;target 20;stdev 10;step -5;delay 0	SI	0.1%	20.9	5.1	25.9	5.2	0.0%	20.6	4.2	19.1	2.9
E5a	Drift	initial 50;target 10;stdev 5;step -2;delay 0	WSP	0.0%	14.4	7.4	29.4	4.8	0.0%	24.8	5.3	16.4	2.4
E5b	Drift	initial 50;target 10;stdev 5;step -2;delay 0	SI	0.0%	13.8	6.6	30.2	4.8	0.0%	25.2	5.0	16.1	2.2
E6a	Drift	initial 50;target 10;stdev 5;step -1;delay 0	WSP	0.0%	35.5	6.4	23.0	5.5	0.0%	38.4	3.7	14.9	2.4
E6b	Drift	initial 50;target 10;stdev 5;step -1;delay 0	SI	0.0%	34.8	8.0	24.9	9.2	0.0%	38.4	3.8	15.1	2.4

APPENDIX E: CURRICULUM VITAE

PERSONAL INFORMATION

Bart Vaerenberg
Herentalsebaan 10
2140 Antwerp, Belgium

vaerenberg@gmail.com

+32 485 628 718

Belgian, male, born September 21st 1982

PROFESSIONAL EXPERIENCE

2005 – 2007: Software Developer at the Eargroup, Antwerp

2007 – 2013: Software Engineer at Otoconsult NV, Antwerp

EDUCATION

2002 – 2005: Bachelor in Applied Computer Sciences, KdG, Antwerp

2005 – 2007: Master in Industrial Sciences, Artesis, Antwerp

2009 – 2013: PhD Student at the University of Antwerp, Antwerp

PUBLICATIONS

1. Govaerts PJ, **Vaerenberg B**, De Ceulaer G, Daemers K, Schwauwers K. Development of a software tool using deterministic logic for the optimization of cochlear implant processor programming. **Otol Neurotol 2010.**

2. Heeren W, **Vaerenberg B**, Coene M, Daemers K, Govaerts PJ, Avram A, Cardinaletti A, De Ceulaer G, Del Bo, L, Gillis S, Pascu A, Rooryck J, Schauwers K, van Heuven V, Volpato F. Linguistic assessment tools for the Digisonic® Dual electric-acoustic speech processor. **Cochlear Implants Int 2010.**

3. **Vaerenberg B**, Govaerts PJ, De Ceulaer G, Daemers K, Schwauwers K. Experiences of the use of FOX, an intelligent agent, for programming cochlear implant sound processors in new users. **Int J Audiol 2011.**

4. Heeren W, Coene M, **Vaerenberg B**, Avram A, Cardinaletti A, Del Bo L, Pascu A, Volpato F, Govaerts PJ. Development of the A§E 2009 Test Battery for Speech Pitch Perception Assessment in Cochlear Implant Users. **Cochlear Implants Int 2011**.

5. **Vaerenberg B**, Pascu A, Del Bo L, Schauwers K, De Ceulaer G, Daemers K, Coene M, Govaerts PJ. Clinical assessment of pitch perception. **Otol Neurotol 2011.**

6. Heeren W, Coene M, **Vaerenberg B**, Avram A, Cardinaletti A, Del Bo L, Pascu A, Volpato F, Govaerts PJ. Development of the A§E 2009 Test Battery for Speech Pitch Perception Assessment in Cochlear Implant Users. **Cochlear Implants Int. 2012.**

7. **Vaerenberg B**, Péan V, Lesbros G, De Ceulaer G, Schauwers K, Daemers K, Gnansia D, Govaerts PJ. Combined electric and acoustic hearing performance with Zebra® speech processor: speech reception, place and temporal coding evaluation. **Cochlear Implants Int. 2012.**

8. Schauwers K, Coene M, Heeren W, Del Bo L, Pascu A, **Vaerenberg B**, Govaerts PJ. Perception of Pitch Changes in Hearing-Impaired Adults with Aided and Unaided Hearing Loss. **Journal of Hearing Science 2012.**

9. **Vaerenberg B**, ten Bosch L, Kowalczyk W, Coene M, De Smet H, Govaerts PJ. Language-universal speech audiometry with automated scoring**. Proceedings of Interspeech 2013.**

10. Bermejo I, Diez FJ, Govaerts PJ, **Vaerenberg B**. A Probabilistic Graphical Model for Tuning Cochlear Implants. **Artificial Intelligence in Medicine 2013.**

11. **Vaerenberg B**, Heeren WFL, Govaerts PJ. Managed estimation of psychophysical thresholds. **Journal of Hearing Science 2013.**

12. **Vaerenberg B**, De Ceulaer G, Szlávik Z, Mancini P, Buechner A, Govaerts PJ. Setting and Reaching Targets with Computer-assisted Cochlear Implant Fitting. **Scientific World Journal 2014**.

13. **Vaerenberg B** et al. Cochlear implant programming: A global survey on the state of the art. **Scientific World Journal 2014.**

14. **Vaerenberg B**, Govaerts PJ, Stainsby T, Nopp P, Gault A, Gnansia D. A uniform Graphical Representation of Intensity Coding in Current Generation Cochlear Implant Systems. **Ear and Hearing 2014.**

15. Buechner A, **Vaerenberg B**, Gazibegovic D, Brendel M, De Ceulaer G, Govaerts PJ , Lenarz T. Evaluation of the “Fitting to Outcomes Expert” FOX with Established Cochlear Implant Users **Cochlear Implants Int. 2014 (in press).**

16. Battmer R-D, Borel S, Brendel M, Britz A, Büchner A, Cooper H, Fielden C, Gazibegovic D, Goetze R, Govaerts P, Lenarz T, Mosinier I, Muff J, Nunn T, **Vaerenberg B**, Vanat Z. Assessment of "fitting to outcomes expert" fox with new cochlear implant users in a multicentric study. **Cochlear Implants Int. 2014 (in review).**

POSTERS & PRESENTATIONS

1. **7 - 16/09/09** Advanced (doctoral level) individual training course in Psychoacoustics and Experimental Audiology. Dept d'Etudes Cognitives, Ecole Normale Supérieure, **Paris**, France
2. **7/10/09** The programming of implants: from craftsmanship to evidence based process control. Advanced Bionics Fox Phase II Study Investigator's Conference, **Paris**, France
3. **22/01/10** TEMA: Threshold Estimation by Managed Algorithm. DUAL-PRO Scientific Committee, **Nice**, France.
4. **11/03/10** Outcome matters. Advanced Bionics BERG, **Amsterdam**, The Netherlands
5. **12 - 13/03/10** Fox: Fitting to Outcome Expert. Advanced Bionics 9th European Investigators' Conference (EIC), **Amsterdam**, The Netherlands
6. **25/03/10** Programming cochlear implants. Neurelec fitting workshop, **Antwerp**, Belgium
7. **27/03/10** Fox & A§E 2009. CIrkel Symposium, **Oostende**, Belgium
8. **25/05/10** TEMA: Threshold Estimation by Managed Algorithm. Workshop 'Beyond Hearing', **Leiden**, The Netherlands
9. **30/06 - 3/07/10** How to assess pitch perception in daily clinical practice? 11th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, **Stockholm**, Sweden
10. **1/10/2010** Fox switch on trial: interim evaluation. Fox Investigator's meeting, **Cambridge**, England
11. **20/12/2010** Fox: example of an Intelligent Agent. Workshop 'Artificial Intelligence', **Antwerp**, Belgium
12. **19/03/11** Programming cochlear implants: the Eargroup approach. CIrkel Symposium, **Hasselt**, Belgium
13. **17/03/12** Otospeech: automated language-universal speech audiometry. CIrkel Symposium, **Gent**, Belgium

14. **03/05/12** Artificial Intelligence to assist the outcome-driven programming of cochlear implants. 12th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, **Baltimore**, USA
15. **02/06/12** Pitch perception and speech understanding in noise with MED-EL FS4 strategy. The 9th International Conference on Cholesteatoma and Ear Surgery , **Nagasaki**, Japan
16. **23-26/05/13** Otocube: desktop outcome assessment and programming of cochlear implants. European Sypmposium on Paediatric Cochlear Implantation 2013, **Istanbul**, Turkey
17. **25-29/08/13** Language-universal speech audiometry with automated scoring. Interspeech 2013, **Lyon**, France