FACULTY OF ENGINEERING
AND ARCHITECTURE

Physiology and Model-Based Methods for Precise Sensorineural Hearing-Loss Profiling and Personalized Auditory Models

Sarineh Keshishzadeh

Doctoral dissertation submitted to obtain the academic degree of
Doctor of Biomedical Engineering

Supervisors
Prof. Sarah Verhulst, PhD - Prof. Dick Botteldooren, PhD

Department of Information Technology
Faculty of Engineering and Architecture, Ghent University

July 2021

GHENT
UNIVERSITY

ISBN 978-94-6355-494-7
NUR 954
Wettelijk depot: D/2021/10.500/42

Members of the Examination Board

Chair

Prof. Filip De Turck, PhD, Ghent University

Other members entitled to vote

Prof. Torsten Dau, PhD, Technical University of Denmark, Denmark
Prof. Paul Devos, PhD, Ghent University
Prof. Ingeborg Dhooge, PhD, Ghent University
Prof. Jan Wouters, PhD, KU Leuven

Supervisors

Prof. Sarah Verhulst, PhD, Ghent University
Prof. Dick Botteldooren, PhD, Ghent University

Acknowledgement

The accomplishment of this dissertation would not have been possible without the unconditional support of individuals who through the years have encouraged me to follow my dreams and have helped me see that I can be more.

I would like to express my sincere gratitude to my PhD supervisor, Prof. Sarah Verhulst, for her invaluable support, endless encouragement and persistent guidance through the four years of my research at Hearing Technology group of Ghent University. I extend my thanks to my second supervisor, Prof. Dick Botteldooren, for his support, constructive comments and practical advices.

I am grateful to Dr. Viacheslav Vasilkov, who patiently answered my questions, anywhere and anytime. I express my thanks to Prof. Ingeborg Dhooge, Prof. Hannah Keppler, Tine Vande Maele, Nele De Poortere and Heleen Van Der Biest, who contributed to the part of the research conducted at Ghent University Hospital (UZ-Gent) and to Prof. Manfred Mauermann and Dr. Markus Garrett, who contributed to data collection in Oldenburg University.

It was a great privilege for me to work in the acoustics group of Ghent University. Definitely, I will never forget Luc's and Pieter's kind helps and advices, Tijmen's supportive attitude during my first days in Ghent, Karlo's technical supports, Deepak's Python instructions, Alejandro's psycho-acoustic discussions, Sarah D'Amico's italian cookies, scientific and friendly discussions with Fotis and Arthur's Dutch translations. I would also like to thank all current and former colleagues: Prof. Paul Devos, Prof. Timothy Van Renterghem, Prof. Bert De Coensel, Isabelle, Gemma, Kang, Ehsan, Emmanuel, Wout, Jorg, Marjoleen, Taiki and Meng for their help and support.

I gratefully acknowledge my former supervisor, Prof. Saeid Rashidi, who was the integral part of my pre-doctoral research.

My deepest appreciation belongs to my parents and brother- *Armen*, *Hermineh* and *Sarmen*- whose unconditional love and selfless encouragement supported me through these years, even being thousands of miles away from me. I am forever indebted to you for giving me the opportunities and experiences that have made me who I am.

Ghent, July 2021
Sarineh Keshishzadeh

Table of Contents

List of Abbreviations

A

ABR	Auditory Brainstem Response
ACC	Accuracy
AEP	Auditory Evoked Potential
AM	Amplitude Modulation
AN	Auditory Nerve
ANEP	Auditory Nerve Excitation Pattern
AudTH	Audiometric Threshold

B

BB	Broadband
BM	Basilar Membrane

C

CF	Characteristic Frequency
CGL	Cochlear-Gain-Loss
CMS	Common Mode Sense
CN	Cochlear Nucleus
CS	Cochlear Synaptopathy

D

DBEFR	Derived-Band Envelope Following Response
DPOAE	Distortion Product Otoacoustic Emission

DPTH	Distortion Product Otoacoustic Emission Threshold
DRL	Driven Right Leg

E

EEG	Electroencephalogram
EFR	Envelope Following Response
EHF	Extended High Frequency

F

FFT	Fast Fourier-Transform
FIR	Finite Impulse Response

H

HL	Hearing-Loss
HP	High-pass
HPD	Hearing Protection Device
HSR	High-Spontaneous-Rate

I

IC	Inferior Colliculus
IHC	Inner-Hair-Cell

K

kNN	k-Nearest Neighbour

L

LP	Low-pass
LSR	Low-Spontaneous-Rate

M

MF	Modulation Frequency
MSR	Medium-Spontaneous-Rate

N

NF	Noise-Floor
NH	Normal-Hearing
NHSR	Normal-Hearing Self-Reported
NIHL	Noise-Induced Hearing-Loss
NN	Neural Network

O

OAE	Otoacoustic Emission
OHC	Outer-Hair-Cell
oHI	Old Hearing-Impaired
oNH	Old Normal-Hearing

P

PLV	Phase-Locking Value
PTA	Pure-Tone Audiometry
PtN	Peak-To-Noise-floor

R

RAM	Rectnagularly Amplitude Modulated
RMS	Root-Mean-Square

S

SAM	Sinusoidal Amplitude Modulated
SPiN	Speech-in-Noise
SPiQ	Speech-in-Quiet
SNHL	Sensorineural-Hearing-Loss
SNR	Signal-to-Noise Ratio
SPL	Sound Pressure Level
SRT	Speech Reception Threshold
SD	Standard Deviation

T

TL	Transmission-Line
TTS	Temporary Threshold-Shift

W

WHO	World Health Organization

Y

yNH	Young Normal-Hearing

List of Publications

Publications in international conferences

- **Keshishzadeh S**, Vasilkov V, Verhulst S. Tonotopic sensitivity to supra-threshold hearing deficits of the envelope following response evoked by broadband stimuli. In23rd International Congress on Acoustics (ICA 2019) 2019 (pp. 6513-6518). Deutsche Gesellschaft für Akustik.

- **Keshishzadeh S**, Verhulst S. From derived-band envelope-following responses to individualized models of near-and supra-threshold hearing deficits. In-International Symposium on Auditory and Audiological Research (ISAAR 2019) 2019 (Vol. 7, pp. 13-20).

Publications in international journals

- **Keshishzadeh S**, Garrett M, Vasilkov V, Verhulst S. The derived-band envelope following response and its sensitivity to sensorineural hearing deficits, Hearing Research, 2020, doi: 10.1016/j.heares.2020.107979.

- **Keshishzadeh S**, Garrett M, Verhulst S. Towards Personalized Auditory Models: Predicting Individual Sensorineural Hearing-Loss Profiles From Recorded Human Auditory Physiology, Trends in Hearing, 2021, doi: 10.117 7/2331216520988406.

Published pre-prints

- Buran BN, McMillan G, **Keshishzadeh S**, Verhulst S, Bramhall N. Predicting synapse counts in living humans by combining computational models with auditory physiology, OSF Preprints, 2020, doi:10.31219/osf.io/f8bx6.

- Maele TV, **Keshishzadeh S**, De Poortere N, Dhooge I, Keppler H, Verhulst S. The variability in potential biomarkers for cochlear synaptopathy after

recreational noise exposure, BioRxiv, 2021, doi:10.1101/2021.01.17.427007.

Samenvatting

(Summary in Dutch)

In de afgelopen decennia is het aantal mensen met gehoorverlies wereldwijd gestegen. Recreatieve of beroepsmatige blootstelling aan lawaai, veroudering en gebruik van ototoxische geneesmiddelen zijn bekend als de belangrijkste oorzaken van gehoorverlies. De nadelige gevolgen van niet-gediagnosticeerd of onbehandeld gehoorverlies hebben tot een groot aantal problemen geleid, b.v. sociaal isolement en verminderde beroeps- of onderwijskansen. Eén van de redenen voor niet-gediagnosticeerde of onbehandelde gehoorstoornissen houdt verband met onze beperkte kennis over bepaalde gehoorstoornissen zoals cochleaire synaptopathie (CS). CS verlaagt de temporele coderingsgetrouwheid van supra-liminaal geluid als gevolg van een verminderd aantal gehoorzenuw (AN) vezels, die via de synaps aan de binnenste haarcellen gehecht zijn. CS heeft echter geen invloed op de algehele sensorische functie van het gehoor en daarom verwachten we dat patiënten met CS over moeilijkheden bij het spraakverstaan kunnen klagen, terwijl hun audiometrische drempels binnen het normale bereik liggen. Aan de andere kant is een directe kwantificering van neurale structuren alleen mogelijk in diermodellen en post-mortem bij mensen, daarom is de kracht van diagnostische hulpmiddelen voor CS bij levende mensen beperkt tot de kwaliteit van niet-invasieve en indirecte methoden zoals auditief geëvoceerde potentialen (AEP's). Op de hoofdhuid geregistreerde AEP's, b.v. de auditieve hersenstamrespons (ABR) en de omhullende-volgende respons (EFR), worden niet alleen beïnvloed door CS, maar ook door gehoorstoornissen zoals schade aan de buitenste haarcellen (OHC) Dit maakt het moeilijk om een differentiële diagnose van verschillende subtypes van perceptief gehoorverlies (SNHL), op te stellen. Dit proefschrift beoogt een modellering en experimentele benaderingen te combineren om zo frequentie-specifieke en geïndividualiseerde SNHL-profielen te ontwikkelen. In dit opzicht gebruiken we een numeriek model van de auditieve periferie, dat state-of-the-art kennis van dierfysiologie overbrugt naar menselijke toepassingen en dat een hulpmiddel biedt om te simuleren hoe OHC-verlies en CS, afzonderlijk en gecombineerd, de generatie van AEPs beïnvloedt. Op deze manier vormt de modelbenadering een aanvulling op de experimentele AEP-metingen en vergemakkelijkt het de ontwikkeling van sensitieve- en frequentie-specifieke methoden voor SNHL diagnose.

Op basis van dierstudies weten we dat de EFR als antwoord op een gemoduleerde stimulus een robuuste marker van CS biedt in de afwezigheid van OHC-verlies. De magnitude van deze EFR hangt echter af van de stimuluskenmerken

(bijv. omhullende, bandbreedte, ruis of toon, off-frequentie maskers) en de bijbehorende verspreiding basilair-membraan excitatie, en bemoeilijkt zo een frequentiespecifieke CS-diagnose. Ook is het geweten dat de EFR magnitude beïnvloed wordt door individuele factoren zoals geslacht of hoofdomvang. Om de behoefte aan gevoeligere diagnostische methoden aan te pakken, stellen we in dit proefschrift een afgeleide-band EFR's (DBEFR's) methode voor waarbij de EFR afgeleid wordt door de spectrale aftrekking van EFR's als reactie op amplitudegemoduleerde witte-ruis met verschillende bandbreedtes. Op basis van modelsimulaties, stelden we vast dat de [2-6] kHz-bandbreedte het meest dominante cochleaire frequentiegebied is voor de EFR-generatie. DBEFR's die uit die bandbreedte werden afgeleid, zijn dus frequentie-specifiek en bieden een relatieve maatstaf die de interindividuele variabiliteit veroorzaakt door bronnen die geen betrekking hebben op het gehoor, uitsluit. We vergeleken DBEFRs tussen jonge en oudere normaalhorende proefpersonen, evenals tussen oudere leeftijdsgroepen (normaalhorend en slechthorend) om het effect van leeftijd en gehoordrempel op DBEFR-magnitudes te bestuderen. In een tweede experiment werden DBEFR-magnitudes van jonge, normaalhorende personen die klagen over gehoorproblemen in lawaaierige omgevingen vergeleken met deze van de controlegroep. De resultaten toonden een significante impact van leeftijd en verhoogde gehoordrempels op DBEFR-magnitudes, terwijl zelf-gerapporteerde gehoorproblemen geen significant effect op de DBEFR hadden. Tegelijkertijd bood het simuleren van het respectieve effect van CS en OHC-schade op DBEFR's een hulpmiddel om de bijdrage van verschillende SNHL-bronnen tot de DBEFR-magnitude te ontwarren. De resultaten van de model simulaties toonden aan dat de DBEFR-magnitude een frequentie-specifieke diagnostische methode is voor CS in de afwezigheid van OHC-verlies. Ook zagen we dat OHC-schade een kleine invloed op de DBEFR magnitude kan hebben. De vermindering van de DBEFR-magnitude als gevolg van CS was echter groter dan de impact van OHC-schade op de magnitude.

Eerder hadden modelsimulaties al aangetoond dat de grootte van de EFR magnitude als antwoord op rechthoekig amplitudegemoduleerde (RAM) tonen een indicator van CS is, en slechts marginaal beïnvloed wordt door OHC-verlies. Daarom werden, met het doel op de ontwikkeling van gepersonaliseerde SNHL-profielen, RAM-EFR's gecombineerd met ABR-afgeleide markers om het individuele SNHL-patroon te bepalen dat de beste match biedt tussen de gemeten en gesimuleerde AEP's. In het kader van dit proefschrift werd onderzocht welke van de meest gebruikte technieken van OHC-schade kwantificering: audiogram, distortie-product otoakoestische emissie (DPOAE) drempels of DP-grams, de beste methode biedt om het cochleaire model te individualiseren. Ten tweede onderzochten we welke ABR/RAM-EFR-markers er het meest geschikt zijn om individuele CS-profielen van jonge normaalhorenden en oudere slechthorenden te voorspellen. Door middel van een voorwaarts-achterwaartse classificatietechniek, voorspelden we eerst CS-profielen van individuen op basis van een voorwaarste classificatie, waarna de prestatie van de classificatie geëvalueerd werd op basis van achterwaardse classificatie. We vonden dat de RAM-EFR-magnitude de beste prestatie leverde bij CS-voorspelling, zowel bij de audiogram en DPOAE-drempel gebaseerde indivi-

dualisering van het cochleaire model. De nauwkeurigheid van de CS-voorspelling was wel 15% hoger wanneer DPOAE-drempels werden gebruikt voor cochleair model individualisering. Hoewel de RAM-EFR-methode veelbelovende resultaten toonde bij SNHL-profilering van jonge en oudere luisteraars (met en zonder gehoorstoornis), veranderde de omvang ervan niet significant bij jonge normaalhorende volwassenen na recreatieve blootstelling aan lawaai. Bovendien werd er geen verband waargenomen tussen spraakverstaan drempels en RAM-EFR-magnitudes van dezelfde personen. Het is mogelijk dat de blootstellingsdosis voor recreatief geluid in het huidige onderzoek niet hoog genoeg was om CS te veroorzaken, aangezien de werkelijke geluidsdosissen niet objectief werden geregistreerd tijdens de muziekfestivals. Het is ook mogelijk dat de RAM-EFR-magnitude niet gevoelig genoeg is om lawaai-geïnduceerde CS te detecteren bij normaalhorende jongvolwassenen.

Dit proefschrift suggereert dat DBEFR-magnitudes die uit de [2-6] kHz band afgeleid werden een frequentie-specifieke marker vormen voor CS in de afwezigheid van OHC-verlies. Op basis van deze resultaten verklaren we de verminderde DBEFR-magnitudes van oudere normaalhorende luisteraars door leeftijdsgebonden CS, dat een gecompromitteerd vermogen om veranderingen in de omhullende van geluid te coderen met zich meebrengt. Bovendien kan de ontwikkelde methode voor de individualisering van SNHL-modellen op basis van RAM-EFR- en DPOAE-drempels gegeneraliseerd worden naar andere experimentele proefopstellingen en cohorten, aangezien de validatie van de methode op basis van RAM-EFR's uit een nieuw cohorte veelbelovende resultaten opleverde. Deze individualisatie methode kan dus gebruikt worden als basis voor modelgebaseerde algoritmen voor spraakverbetering die rekening houden met de individuele mate van gehoorschade door CS- en OHC verlies. Deze modelgebaseerde methoden voor gehoorherstel kunnen de weg effenen voor de volgende generatie hoorapparaatalgoritmen die naar verwachting ten goede zullen komen aan personen van wie het spraakverstaanbaarheid niet voldoende hersteld wordt op basis van geluidsversterking alleen, en aan de groep normaalhorenden met een verminderd spraakverstaan die momenteel niet behandeld worden.

Summary

In the last decades, the number of people with hearing-loss has risen worldwide. Recreational or occupational noise-exposure, aging and use of ototoxic drugs are known to be the main causes of hearing-loss. The adverse consequences of undiagnosed or untreated hearing-loss have led to a multitude of problems, e.g. social isolation and reduced vocational or educational opportunities. One of the reasons for undiagnosed or untreated hearing loss is associated with our limited knowledge of certain hearing deficits such as cochlear synaptopathy (CS). CS degrades temporal envelope encoding fidelity of supra-threshold sound by compromising the auditory nerve (AN) fibers that synapse onto inner-hair-cells, but it does not affect the overall sensory function of the cochlea. Therefore, patients who complain about difficulties in understanding speech, but who have audiometric thresholds within the normal range, may suffer from CS. On the other hand, a direct assessment of neural structures is only possible in animal models and post-mortem human studies, and hence the power of CS diagnostic tools in live humans is limited to the quality of non-invasive and indirect methods, such as auditory evoked potentials (AEPs). Scalp-recorded AEPs, e.g. auditory brainstem responses (ABRs) or envelope following responses (EFRs), are not merely affected by CS, but co-existing hearing deficits such as outer-hair-cell (OHC) loss can also impact them, which renders it difficult to establish a differential diagnosis of two important sensorineural hearing-loss (SNHL) sub-types, i.e. OHC-loss and CS. This dissertation aims to combine modelling and experimental approaches to develop frequency-specific individualized SNHL profiles. In this regard, we use a computational model of the auditory periphery that bridges state-of-the-art knowledge of animal physiology to human applications and that offers a tool to simulate how OHC-loss and CS, in isolation and combination, affect the source generators of AEP-derived metrics. In this way, the modelling approach complements experimental AEP measurement methods and facilitates the development of novel frequency- and deficit-specific diagnostic tools for SNHL.

Based on animal studies, the EFR strength to a modulated stimulus provides a robust marker of CS in the absence of OHC-loss. However, its amplitude depends on the stimulus characteristics (e.g. envelope, bandwidth, noise or tone, off-frequency maskers) and the associated spread of basilar-membrane excitation confounds a frequency-specific CS diagnosis. Moreover, the EFR magnitude is influenced by subject-specific factors such as gender or head-size. To address the need for more frequency-specific diagnostic metrics, we propose derived-band EFRs (DBEFRs), which are constructed by spectral subtraction of EFRs in re-

sponse to amplitude-modulated white-noise carriers of different bandwidths. Benefiting from model simulations, we identified that the [2-6] kHz frequency region is the most dominant cochlear frequency region for EFR generation. Thus, DBEFRs derived from that stimulus bandwidth are frequency-specific and provide a relative metric that rules out inter-subject variability caused by hearing-unrelated sources. We compared extracted DBEFRs between young and older normal-hearing subjects, as well as between elderly age-matched groups (older normal-hearing and hearing-impaired) to study the effect of age and hearing-threshold on DBEFR magnitudes, respectively. In another experiment, DBEFR magnitudes of young normal-hearing listeners who complain about hearing difficulties in noisy environments were compared to those of a control-group. The results showed a significant impact of age and elevated hearing thresholds on DBEFR magnitudes, whereas self-reported hearing difficulties did not cause a significant difference. At the same time, simulating the respective effect of OHC-damage and CS on DBEFR provided a tool to disentangle the contribution of different SNHL sources to DBEFR magnitudes. Model simulations showed that the DBEFR magnitude offers a frequency-specific diagnostic tool for CS in absence of OHC-loss, but that co-occurring OHC deficits can also affect the response. However, the reduction of DBEFR magnitude due to CS was greater than the impact of OHC-loss.

Previously, model simulations have shown that the magnitude of the EFR to rectangularly amplitude-modulated (RAM) pure-tones is an indicator of CS, and that it is only marginally affected by OHC-loss. Hence, with the aim of developing personalized SNHL profiles, we recorded and simulated RAM-EFRs by combining them with ABR-derived metrics to determine the individual SNHL pattern that provides the best match between recorded and simulated AEPs. Within the framework of this dissertation, it was investigated which of the following commonly used techniques of OHC-damage quantification: audiogram, distortion-product otoacoustic emission (DPOAE) thresholds or DP-grams, provided the best method to individualize the cochlear-model of young normal-hearing and older normal-hearing/hearing-impaired listeners. We also investigated which ABR/RAM-EFR markers, or combinations thereof, are best suited to predict individual CS profiles. Proposing a *forward-backward* classification technique, we were able to predict individualized CS profiles of the study participants using a *forward* classification and then evaluate the prediction performance using the *backward* classification approach. The results showed that the RAM-EFR magnitude yielded the best performance in CS prediction, both for audiogram and DPOAE-threshold-based individualization of the cochlea. However, the accuracy of the CS prediction was 15% higher when DPOAE-thresholds were used for cochlear model individualization. Although the RAM-EFR metric showed promising results in SNHL profiling of young and older listeners (with and without hearing impairment), its magnitude did not change significantly in young normal-hearing adults after recreational noise-exposure. Moreover, no relation was observed between speech reception thresholds and RAM-EFR magnitudes of the same listeners. It is possible that the recreational noise-exposure dose in the present study was not sufficiently high to cause CS, since actual noise doses were not assessed during the attendance

of experiment participants in the music festivals. Or there is the possibility that the RAM-EFR magnitude was not sensitive enough to detect noise-induced CS in normal-hearing young adults.

This thesis suggests that DBEFR magnitudes extracted from [2-6] kHz provide a frequency-specific marker of CS in the absence of OHC-loss. Hence, the degraded DBEFR magnitudes of the older normal-hearing listeners can be explained by age-induced CS that yields compromised supra-threshold temporal envelope encoding. Additionally, the developed method for the individualization of SNHL models based on RAM-EFR and DPOAE-thresholds is generalizable to other recording setups and cohorts, since validation of the method using recorded RAM-EFRs of a new cohort yielded promising results. Accordingly, the personalized SNHL profiles can be used as a basis for model-based speech enhancement algorithms, which take into account the individual degree of CS and OHC-damage. Model-based hearing restoration methods may pave the way for the next generation of hearing-aid algorithms which are expected to benefit individuals whose speech intelligibility is not adequately restored on the basis of gain prescription alone or who report degraded speech intelligibility without receiving treatment for these complaints.

1 Introduction

Hearing loss is the world's fourth leading cause of disability and it is estimated that 466 million people, i.e. 6.1% of the world population, are living with hearing loss [1]. Defined by World Health Organization (WHO), a normal-hearing person has hearing thresholds of 20 dB HL or better, in both ears. Thus, people who fall outside this range, have hearing loss. WHO identified several risk-factors that deteriorate the hearing ability of normal-hearing people: Over one billion young people aged between 12 and 35 years, are at risk for hearing loss caused by recreational exposure to loud sound. In addition, occupational noise-exposure, such as that from machinery and explosions, contributes to 22% of workplace-related health issues [2, 3]. Administration of some medications, such as Aminoglycocides for treatment of drug-resistant tuberculosis (10-50% of the cases) [4], or Cisplatin, as a therapeutic agent against cancer (75-100% of the cases) [5], were reported to cause ototoxic hearing damage [6]. Accumulated noise-exposure and side-effects of ototoxic drugs, may exacerbate the hearing problems of the elderly, given that aging itself, has degenerative effects on the auditory system [1]. Additionally, other factors such as genetic mutations, smoking, chronic disease and obesity may further contribute to the hearing impairment [7, 8]. Although implementing some preventive actions may reduce the adverse impact of the above-mentioned hearing-loss risk-factors, providing effective therapeutic interventions is of great importance. Leaving hearing loss untreated has a detrimental impact on the quality of life, since impaired hearing can lead to communication difficulties and gradual social isolation, reported by WHO.

However, providing a successful treatment entails powerful and objective di-

agnostic tools that can differentiate sources of hearing pathologies. Given that any portion of the auditory pathway can be targeted by various hearing-loss risk-factors and that co-occurring hearing deficits may confound the interpretation of diagnostic test results, it is crucial to develop metrics that precisely identify the underlying causes of the hearing impairment.

1.1 Peripheral hearing disorders

In general, peripheral hearing disorders are categorized into two groups: conductive and sensorineural. Conductive hearing-loss is caused by damaged outer- or middle-ear due to a reduced sound conduction to the inner ear. Cerumen impaction in the ear-canal, perforated eardrum, presence of fluid in the middle-ear or immobility of its bones are among the well-known contributors to conductive hearing-loss and are diagnosed via conducting otoscopy, tympanometry, air-bone conduction audiometry, Rinne's and Weber's tests [1, 9, 10]. On the other hand, sensorineural hearing-loss (SNHL) is the broad term that refers to the impaired organ of Corti or stria vascularis (sensory loss) and/or loss or dysfunction of spiral ganglion neurons (neural loss). The former is characterized by the elevated hearing thresholds, as well as reduced frequency-selectivity due to the widening of the auditory filters [11, 12], whereas the latter, which is called cochlear synaptopathy (CS), degrades the sound encoding [13]. Accordingly, CS hinders speech intelligibility in noise [14], but does not affect audiometric thresholds [13]. For this reason, CS is also called a *"hidden hearing loss"* [15]. In animal models, noise-induced SNHL can cause temporary threshold elevations due to reversible damage to cochlear outer-hair-cells (OHCs), which can lead to irreversible AN degeneration [16–19]. Apart from noise-induced SNHL, the use of ototoxic drugs increases lipid peroxidation or calcium influx and leads to apoptosis of hair cells, stria vascularis and AN fibers [20]. In addition to external contributors to SNHL, aging is assumed to be the major driver of SNHL in the elderly, since stiffening of the basilar membrane (BM), atrophy of the organ of Corti and AN in the basal end of the cochlea, atrophy of the stria vascularis and loss of neural population are natural causes of age-related hearing loss, known as presbycusis [21–23].

Contradictory to conductive peripheral hearing loss, establishing a robust differential diagnosis of sensory and neural sources of SNHL in humans is not straightforward. First of all, conventional hearing assessment techniques, e.g. audiometry and otoacoustic emissions (OAE), are not sensitive to CS [13]. Moreover, although CS has been demonstrated in various animal models [13, 16–19], its confirmation in humans requires post-mortem temporal bone analysis. Hence, developing non-invasive and indirect diagnostic methods, which are specific to CS, is a prerequisite for reaching a differential diagnosis of SNHL sub-types.

1.2 Objective diagnostic measures of sensorineural hearing deficits

Diagnosis of the OHC-loss component of SNHL, is routinely performed by measuring audiometric thresholds across audible frequencies or OAEs. Since, audiometric thresholds are behavioural measures of OHC-loss, they could be affected by ascending auditory processing stages. In contrast, OAEs are generated inside the cochlea and inform about OHCs health, without including ascending auditory pathways [24, 25]. However, their properties depend on the frequency-dependent location along the cochlear partition they are generated from and on the nonlinear characteristics of the OHC input-output function [26, 27]. A consequence of this compressive input-output function is that in response to two simultaneously presented pure-tones, the cochlea generates a third distortion product tone. The presence of the distortion-product otoacoustic emission (DPOAE) can be used as a localized predictor of the OHC-loss, near the frequency region of the f_1 and f_2 [28].

Different from OHC-loss, diagnosis of CS in live humans is not straightforward. Considering that CS targets the AN-fiber population, sound representation might degrade at the cochlear nucleus (CN) and auditory brainstem stages of the auditory pathway. Accordingly, elecrophysiological measurements, such as auditory evoked potentials (AEPs), that target different auditory processing stages, can be used in combination with psychoacoustic metrics to probe the auditory pathway from cochlea to cortex and to characterize the contribution of hearing deficits at ascending processing stages. The auditory brainstem response (ABR), a type of AEP evoked by transient clicks, reflects synchronous firing rate of the AN-fibers in the wave-I amplitude [29, 30]. In the context of AEPs, the envelope following response (EFR) is a phase-locked response to the envelope of amplitude modulated stimulus and reflects integrity of the temporal processing along the auditory pathway [31, 32]. In addition, CS can affect the neural-coding of speech and degrade speech intelligibility, particularly in noise [14].

AEP recordings have been studied in numerous animal models as a candidate diagnostic tool for CS and ABR wave-I amplitude [16, 33, 34] and EFR magnitude [31, 32] were shown to be affected by CS in mice and guinea pigs. These animal studies demonstrated that ABR wave-I amplitude, EFR strength and phase locking value (PLV) degrade as a consequence of noise-induced or age-related CS. Despite of these findings in animals, a direct translation of these measures for human diagnosis of CS yielded a mixed success [35–37].

Although human post-mortem studies have confirmed the existence of aging or noise-induced CS in humans [38–40], and findings of multiple human studies are consistent with animal models, e.g. decreased EFR strength due to recreational noise-exposure [41] or aging [42–44] and degraded ABR wave-I amplitude caused by lifetime [45–47] or occupational [48] noise-exposure, numerous human stud-

ies reported contradictory results. For example, no significant correlations were found between noise-exposure history and ABR wave-I amplitude or EFR strength or speech intelligibility in noise in [49–56]. Several factors may have contributed in generating such inconsistencies between animal and human studies. These discrepancies have been discussed in several studies, and are summarized as follows: (1) Partial synaptic repair might have occurred in some humans after exposure to noise, since a similar phenomenon was observed in guinea pigs [57, 58]. (2) Humans are less sensitive to noise-induced CS than animal models [18, 59], or higher noise levels are required to generate synaptic loss to human ears [37]. (3) AEPs are more variable in humans due to inter-individual differences in head-size, sex [60–62], cochlear duct length [62] and degree of awareness during the experiment (awake humans vs. anesthetized animals). In particular, variability of ABR wave-I amplitude in humans can be ascribed to the scalp electrodes, which are less robust than sub-dermal electrodes used in animals [35, 50]. (4) Adopted methods across different human studies for noise-exposure history estimation, as well as the considered time-intervals for those methods are not identical (e.g. [46,48,49,54]). (5) The lifelong noise-exposure dose is controlled precisely in control and target groups of animal models, however this is not possible in human studies where uncontrolled lifelong noise-exposures, specifically among control group individuals, might have confounded the interpretation of results [48]. (6) Degree of OHC deficits differs even among the control group individuals, and hence might have caused inconsistency across human study results, since OHC-loss decreases input to IHC and also affects the ABR wave-I amplitude [63,64] or EFR magnitude [32].

To rule out subject-specific factors in AEP-based SNHL diagnosis, several studies suggested the use of relative rather than absolute metrics: e.g., ABR wave-I amplitude growth as a function of stimulus intensity [16], the slope of EFR magnitude growth as a function of modulation depth [51, 65], ABR wave-I and V amplitude ratio [15,66,67] or latency difference [68–70]. Although it is assumed that employing relative metrics may reduce the ambiguity related to individual variability emanated from sources other than CS, the specific impact of OHC-loss and CS on AEPs, still remains unknown.

To overcome the above-mentioned disparities in human studies, and increase the diagnostic power of AEPs in the presence of mixed pathologies (among others OHC-loss and CS), multiple considerations need to be made. First of all, we need to identify the respective impact of each subcomponent of SNHL on AEPs and then develop stimuli that evoke AEPs that are specific to CS. Accordingly, derived AEP metrics need to isolate and quantify the CS aspect of SNHL, while remaining insensitive to OHC-loss, that may co-exist. At the same time it is desired that AEP markers provide a frequency-specific measure of CS in humans. In this regard, adopting auditory models which incorporate both sub-types of SNHL can be a powerful tool in the design of CS-sensitive diagnostic methods, as they can

simulate how different sub-types of SNHL interact and affect human AEP markers of hearing.

1.3 Individualized models of the auditory periphery

Over the past decades, different types of auditory models have been developed to study the functioning of normal and impaired auditory processing. Several models can simulate frequency-dependent SNHL [71–77], and can in this way offer a tool to investigate the impact of underlying hearing pathologies on AEP alterations [73, 78].

In this regard, the biophysically-inspired computational model of the auditory periphery, developed by Verhulst et al. [73], bridges knowledge from animal physiology to human applications. At the same time, it incorporates the middle-ear filtering, a nonlinear transmission-line (TL) model of the human cochlear mechanics [79, 80], a model of the IHC-AN complex [81], and a phenomenological description of the ventral CN and inferior colliculus (IC) neurons [82]. It provides a way to study the respective output of the peripheral processing stages in response to a given stimulus and applied SNHL profile, i.e. OHC-loss, CS, or co-existing OHC-loss and CS. Reducing the mechanical gain in a CF-dependent manner in the cochlear TL model yields wider cochlear filters with reduced gain and simulates the effect of OHC-damage on BM processing. To introduce CS, the population and types of AN fibers, that connect to each IHC, can be altered. Taking advantage of the developed model, the respective effect of CS and OHC-loss on EFRs and ABRs can be investigated [44, 64, 83]. Previously, model simulations suggested that OHC-loss reduces the ABR wave-I and V amplitudes. Simulating OHC-loss, also yields longer ABR wave-V latencies and steeper wave-V latency growth-functions, in line with several experimental studies (wave-V latency: [84], wave-V latency growth: [85, 86]). The EFR magnitude to sinusoidally amplitude-modulated pure-tones (SAM) increased due to OHC-loss, since OHC-damage reduces cochlear input to IHC-AN complex [63] and causes AN fibers to operate in a more sensitive amplitude-modulation-coding region. Therefore, enhanced SAM-EFR magnitudes are generated compared to an intact cochlea, in absence of CS. Introducing CS to the model, yielded degraded ABR wave-I amplitudes and SAM-EFR magnitudes, consistent with several studies (wave-I amplitude: [13, 16, 32], EFR: [31, 41, 42, 87]).

Model simulations provide a detailed insight into how acoustic stimuli of AEPs can be designed for differential diagnosis of SNHL. Benefiting from the modelling approach [73], Vasilkov et al. proposed an EFR stimulation paradigm in [44], that is sensitive to CS and maximally insensitive to possible co-existing OHC-loss. The optimal stimulus was a rectangularly amplitude-modulated (RAM) pure-tone with a 25% duty-cycle, that yielded stronger EFR magnitudes compared to the

conventional SAM stimulus. Using model simulations, the authors showed that the RAM stimulus yields stronger EFRs and more synchronized response of the AN fibers, by virtue of its sharply rising envelope shape. Simulating the effect of OHC-loss, marginally affected the RAM-EFR magnitude (compared to SAM-EFR), since the steeply rising envelop of the stimulus, together with prolonged inter-peak-intervals (25% duty-cycle), limited the impact of reduced cochlear amplification on the response. Introducing CS significantly degraded the magnitude of RAM-EFR, in absence of OHC-loss [44]. Taken together, model simulation combined with experimental validations in [44], support that the RAM stimulus can provide an enhanced sensitivity to the CS aspect of SNHL, in listeners with mixed SNHL sources.

Gaining advantage of such promising CS-sensitive metrics, together with our progressive knowledge of the functional aspects of different SNHL sub-types, stress the importance of developing personalized SNHL profiles using physiological measures, that target different auditory processing stages. These profiles can be incorporated within individualized models of the auditory periphery, which can be used in the development of personalised audio-signal processing for hearing-aids. Thus far, auditory profiling methods have been based on psychoacoustics (e.g. [76,88–90]), to yield a measure that cannot directly explain the exact location of the hearing deficit origin along the auditory pathway. With a view on hearing restoration, it is necessary to pursue an AEP-based quantification of the SNHL sub-types, as restoration will only be effective, when the functional consequences of each sub-type are considered.

1.4 Objectives

This doctoral research will tackle some of the shortcomings of present diagnostic tools for sensorineural hearing deficits in humans. Specifically, it aims to develop an AEP-based objective diagnostic tool, that provides a *differential diagnosis* of OHC-loss and CS sub-types of the SNHL, when these co-occur in the same subject. Using the state-of-the-art biophysically-inspired computational model of the auditory periphery, this project determines the frequency sensitivity of EFRs to different types of acoustic stimuli and develops techniques to yield more *frequency-specific* responses than those that presently exist. Additionally, the respective effect of OHC- and AN-damage on AEPs is studied in this thesis, since the adopted model simulates the OHC-loss and CS sources of the SNHL, in isolation and combination.

Benefiting from the model simulations and combining them with recorded human AEPs, this doctoral research identifies which of the thus far developed AEP-derived metrics (ABRs and EFRs) and physiological measurements are most suitable to build *individualized auditory periphery models* and develop *personalized*

SNHL profiles. In this regard, machine-learning techniques are proposed to tackle the issue of validating the predicted individualized CS profiles of human-subjects. What makes this approach unique is the *AEP-based CS profiling* of the human-listeners in presence of OHC-loss, that can be used in the future development of individualized hearing-aid algorithms to compensate for degraded sound-processing as a consequence of both OHC-loss and CS.

Since noise-exposure is one of the driving causes of CS and leads to a degraded speech perception, this thesis examines the effect of recreational noise-exposure on the AEP-based CS-markers and investigates how *speech intelligibility* relates to those measures.

1.5 Outline of the thesis

This dissertation is structured into three main parts and comprises five chapters, followed by the *Discussion* section:

The first part (Chapters 2 and 3) aims to develop frequency-specific EFR markers and determines their sensitivity to the OHC-loss and CS aspects of SNHL. Specifically, Chapter 2 introduces frequency-specific EFRs to five broadband stimuli with different low-cutoff frequencies and explores the tonotopy of those EFRs using model simulations. Model outcomes are compared to recorded EFRs to same stimuli. Chapter 3 builds on the findings in Chapter 2 and introduces derived-band EFRs (DBEFRs) by further limiting the frequency-content of EFRs developed in the previous chapter and then studies their sensitivity to CS, with or without co-existing OHC-loss, using model simulations. The model findings are compared with recorded EFRs from two independent experiments by taking into account the respective impact of age, elevated hearing thresholds and self-reported hearing difficulties of study participants in presence of normal audiograms.

The second part (Chapters 4 and 5) describes how simulated and recorded physiological measurements can be used to personalize auditory processing models to include individual AN- and OHC-damage patterns. Given that impaired functionality of these cochlear/neural elements affects AEP-derived metrics differently, the approach taken in Chapter 4 relies on the quality of recorded AEPs by combining them with a measure specific to OHC-damage, i.e. either audiogram or DPOAE. To this end, a machine-learning approach is developed to estimate the cochlear-gain-loss (CGL) parameters of the cochlear model using measured DPOAE thresholds, while audiogram-based CGL parameters are extracted following the method proposed in [64]. In this way, CGL parameters are hard-coded in the model for each individual, after which various AEP-derived metrics, e.g. EFR magnitudes, ABR amplitudes and latnecies, are simulated for different degrees of AN-damage. Then, a *forward-backward* classification technique is introduced to determine which AEP metric, or combination thereof, best predicts the sim-

ulated individualized CS profiles. The method is implemented for either of the cochlear model individualization methods and respective limitations are discussed by considering the obtained prediction errors. To further study cochlear model individualization methods, Chapter 5 takes DP-grams of the same listeners recorded in six different primary levels and estimates respective CGL parameters by implementing a similar machine-learning method introduced in Chapter 4. Then, it simulates individualized DP-grams, compares them with those obtained from the measurements and calculates the respective prediction errors. Considering the prediction errors of audiograms and DPOAE thresholds in Chapter 4, together with those of DP-grams measured in six levels, it determines the most suitable metric for cochlear model individualization, by taking into account limitations associated with the model.

The third part of this dissertation (Chapter 6) investigates the consequences of recreational noise-exposure on AEPs, audiometric thresholds and speech intelligibility. In particular, Chapter 6 studies the variability of those metrics in young normal-hearing listeners, after attending music events. RAM and SAM-EFR magnitudes, ABR wave-I and -V amplitudes, speech reception threshold (SRT) values in quiet and noise were recorded one day before (baseline session) and one, three and five days after attending festival. These metrics are compared across the sessions to identify probable effects of the exposure to loud sound on the potential biomarkers of CS.

The last chapter draws conclusion on the three above-mentioned topics and provides the future perspectives based on the major findings of this doctoral research.

In total, this thesis reports on three experiments. In the first experiment (Chapter 3), two groups were recruited: a young normal-hearing control group and a young normal-hearing group with self-reported hearing difficulties in the presence of background noise. AEPs of these groups were used (1) to determine the frequency sensitivity of EFRs to different broadband amplitude modulated stimuli in Chapter 2 and (2) to explore the effect of self-reported hearing difficulties on the recorded AEPs in Chapter 3 (first experiment). The second experiment (the second experiment of Chapter 3, first experiment of Chapter 4 and Chapter 5) includes data from young and older normal hearing, as well as older hearing-impaired subjects to investigate the effect of age and hearing impairment on the EFRs. Their data was used to test the derived-band EFR metric (Chapter 3) and to build individualized SNHL models (Chapters 4 and 5). Lastly, the third experiment considers young normal-hearing subjects, who attended multi-day music festivals. In this group, we explore the effect of loud sound exposure on the potential biomarkers of CS, by comparing their recordings before and after attending the music festivals (Chapter 6). Moreover, the first recording session of this cohort (measured one day before attending the music event), was used in Chapter 4 to validate the proposed

method for CS-profiling.

Lastly, It is worthwhile to mention that the framework of this thesis is partially structured based on the published papers, and hence brief overlaps can be noticed within the chapters.

2

Tonotopic Sensitivity to Supra-threshold Hearing Deficits of the Envelope Following Response Evoked by Broadband Stimuli

Sarineh Keshishzadeh, Viacheslav Vasilkov, Sarah Verhulst

Published in the Proceedings of the 23rd International Congress on Acoustics (ICA), 2019.[1]

This chapter explores the frequency specificity of envelope following responses (EFRs) in diagnosing supra-threshold hearing deficits. First, EFRs were recorded to sinusoidally amplitude-modulated band-pass filtered white noise carriers with different bandwidths in two participant groups: young normal-hearing control group and a group with self-reported hearing difficulties in noisy listening environments. Then, by adopting a biophysically-inspired model of the human auditory periphery, the origin of EFR variability to similar stimuli with different carrier bandwidths and modulation frequencies was studied. This study showed that despite the broadband cochlear excitation, the broadband EFR mostly reflects summed amplitude modulation coding strength in frequency channels above 2 kHz. Hence, frequency-specific information regarding the supra-threshold sound

[1] **SK**: running the experiment, model simulations, analysis, conceptualization, writing the original draft, review and editing, **VV**: conceptualization, review and editing, **SV**: supervision, conceptualization, review and editing, funding acquisition.

amplitude modulation coding (e.g. derived-band EFRs) can be extracted from EFRs to stimuli with frequency content of higher than 2 kHz.

2.1 Introduction

Threshold audiometry at frequencies between 250-8000Hz does not quantify the possible supra-threshold hearing deficits in listeners with normal hearing thresholds. One of the possible sources of supra-threshold hearing deficits is loss of auditory nerve (AN) fibers (i.e. cochlear synaptopathy) as a result of noise-exposure or aging. Cochlear synaptopathy (CS) leaves the outer-hair-cells (OHCs) intact and does not affect the audiometric thresholds [13]. Recently, auditory evoked potentials (AEPs) such as auditory brainstem responses (ABRs) and envelope following responses (EFRs) have been proposed to quantify these supra-threshold hearing deficits. EFRs reflect the phase-locking strength to the modulated sounds [31] and are sensitive to CS in animals [32]. Hence, EFR could be a more robust metric to measure the synchrony of AN fibers to the modulated stimulus [31]. However, EFRs are confounded by stimulus and subject-specific factors [60, 91], as well as the spread of the basilar membrane (BM) excitation [41]. To evaluate the diagnostic power and tonotopy of frequency-specific EFRs in humans, we investigated which cochlear frequency regions contribute to the scalp-recorded population response. We took an experimental approach in which we adopted, (i) white-noise carriers of various bandwidths with different low cut-off frequencies, (ii) stimuli with a high modulation frequency (MF; 120 Hz) to avoid cortical contributions to the response and (iii) a stimulus level of 70-dB-SPL to encompass the contribution of a range of low and high-threshold AN fibers to the EFR. We recorded EFRs in groups with normal audiograms, but with or without self-reported hearing difficulties in noisy environments. The former group might be representative of the target synaptopathy expected group without hearing thresholds deficits.

To further investigate the origin of the experimental EFRs differences to noise carriers of different bandwidths, we adopted a model-based approach [73]. The model offers a promising way to further explore the influence of different stimuli characteristics on the frequency sensitivity along the different processing stages of the auditory periphery (i.e. cochlea, AN and brainstem). Given that previous studies have shown that increasing the MF of a broadband noise, increases the behavioral AM detection thresholds [92] and decreases the EFR amplitudes [93], we focused on investigating the effect of different modulation frequencies in encoding the stimulus envelope in BM velocity and AN spike rates.

First, the experimental stimuli and the participated groups are described. After which, Sections 2.2 and 2.3 detail the implemented methods and show comparisons between experimental and simulated EFRs. Finally, the discussion states possible reasons for the lack of AM-coding in the experimental data at lower fre-

quencies and makes suggestions to improve the broadband EFRs frequency sensitivity based on the model simulations.

2.2 Methods

EFRs were recorded to five filtered white-noise carriers in [0.25-22], [0.5-22], [1-22], [2-22] and [4-22] kHz frequency bands and were 100% modulated with a 120 Hz pure-tone. The widest stimulus was presented at 70-dB-SPL and the other bandwidth stimuli had lower level to maintain a similar spectral level in all filtered stimuli. The stimulus epochs were 1.25-sec long and repeated 370 times. EFRs were recorded with a 64-channel Biosemi EEG system. Details of the experimental setup were described elsewhere [94]. Sixteen normal-hearing (NH) subjects aged 18-30 years (24.21±4.10 years, six females) and nine NH subjects with self-reported hearing difficulties in noisy environments (NHSR), aged 23-49 (33.78±8.57 years, three females) participated in the experiment. All participants had hearing thresholds below 25 dB-HL, except for one participant who had 30 dB-HL loss at 250 Hz. The best audiometric ear at 4 kHz was chosen for the monaural stimulation. During the recording, subjects were seated in a comfortable chair in an acoustically and electrically shielded sound booth watching silent movies with subtitles to stay awake. All participants were informed about the experimental procedure according to the ethical procedure at Ghent University and an informed consent was received.

2.2.1 Experimental data analysis

For each condition, the responses from the Cz-channel were filtered between 60 and 600 Hz with an 800th order Blackman window-based band-pass FIR filter in *MATLAB*, using the *filtfilt* function. Then, the signals were epoched in 1-s long blocks starting 0.25s after the stimulus onset to focus on the steady state part of the response. A baseline correction was applied by subtracting the average of each epoch, before averaging across trials. Thirty epochs with the highest peak-to-trough absolute values were removed to drop the noisiest epochs. In this way, an equal number of epochs were averaged for all conditions and subjects. We adopted a bootstrapping method to identify the EFR strength and signal-to-noise ratio in the frequency domain [95]. First the Fast Fourier-Transform (FFT) of each epoch was computed. Then, an averaged FFT over 340 epochs, randomly drawn with replacement out of total epochs was calculated. This procedure was repeated 200 times and resulted in a nearly Gaussian distribution of mean spectra of the responses. The absolute value of the mean of 200 averaged epochs, yielded the EFRs in the frequency domain. Secondly, the same procedure was followed with 1000 (of which half were phase-flipped) repeated averages to estimate the

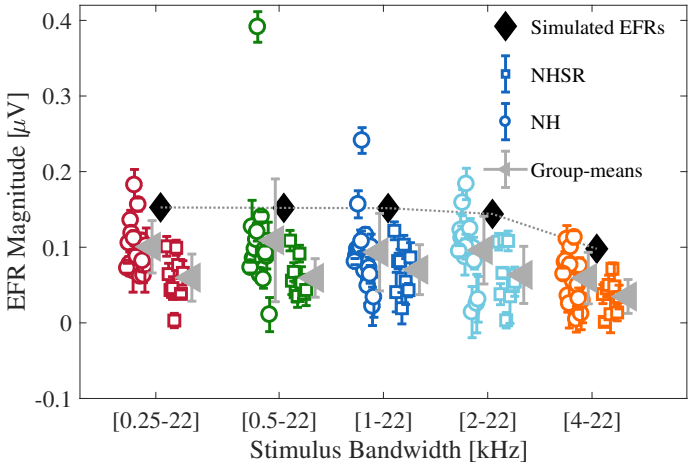


Figure 2.1: Recorded EFR magnitudes in NH (circles) and NHSR (squares) listeners. The corresponding group-means are shown with left-pointing triangles. Simulated EFRs for a normal-hearing model are shown by black diamonds.

noise-floor. Finally, the subtraction between the absolute values of the signal and noise-floor averages resulted in the stimulus-driven EFR spectrum. To calculate the EFR strength, we included the fundamental MF and any visible harmonics. Only two harmonics were visible in our data, we hence considered the magnitude of the fundamental frequency of the modulator (i.e. f_0=120 Hz) and two following harmonics:

$$\mathrm{EFR} = \sum_{i=0}^{2} |\mathrm{Magnitude}_{f_i}| \tag{2.1}$$

More detailed implementation steps in this regard can be found in Chapter 3, Section 3.3.

2.2.2 Modelling approach

We simulated EFRs using a biophysically-inspired model of the human auditory periphery [73]. The same stimuli as were adopted in the experiment, were presented to the model, but with a total duration of 600 ms. The last 400 ms of the simulated responses were used for the EFR calculation. Further analysis steps were identical to the experimental data analysis. Given that the stimulus carrier was noise, it was presented 100 times to the model with different seeds of the random generator. Before calculating the EFR, the simulated responses at the inferior colliculus (IC) stage of the model, were averaged over 100 stimulus presentations. The modelling approach, brings the benefit of allowing a frequency-specific inves-

tigation of the sources contributing to the EFRs to modulated broadband stimuli, at both the BM and AN processing levels. To this end, we explored the root-mean-square (RMS) and modulation strength of the response at each simulated characteristic frequency (CF) channel. We calculated the modulation strength at both BM and AN levels in the time domain by averaging the BM-velocity and AN spike-rates across all cycles for each epoch. For a MF of 120 Hz, each epoch encompasses 48 cycles, and taking into account the 100 repetitions of the stimulus representation, the half of the peak-to-peak amplitude of the 4800 averages were defined as the modulation strength of each CF channel. We followed the same procedure to calculate the modulation strength for the AN simulations to keep a fair comparison with the BM simulations.

2.3 Results

2.3.1 EFR magnitudes evoked by broadband stimuli

Figure 2.1, shows the EFR magnitudes for the different stimulus conditions and groups (i.e., NH and NHSR). Even though the data points for each condition show individual variabilities, the group-means were constant across conditions for stimuli of low cut-off frequencies up to 2 kHz. A decrement was observed for the last condition with the stimulus bandwidth of [4-22] kHz. The normal-hearing model-predicted EFRs (diamonds in Figure 2.1), revealed the same trend as the experimental group-means across conditions. The NHSR group-mean EFRs of the participants were lower than the EFRs of the NH participants in each condition, but the only significant EFR differences between the two groups were observed at [2-22] and [4-22] kHz conditions ($\mathrm{EFR}_{[2-22]}$: $\mathrm{t}_{(19)}$=3.36, p=0.003 and $\mathrm{EFR}_{[4-22]}$: $\mathrm{t}_{(19)}$=2.76, p=0.012). This finding is consistent with the idea that NHSR listeners may suffer from supra-threshold hearing deficits.

2.3.2 CF contributions to the EFR

To explore the origin of the frequency-dependent behavior of the EFRs, we studied the contribution of different CF channels for the different recording conditions to population responses at BM and AN levels of the model. Figure 2.2 shows the simulated excitation patterns and modulation responses of the BM-velocity (Figure 2.2a and 2.2b) and AN-responses (Figure 2.2c and 2.2d) to the experimental conditions. The excitation pattern shows increased velocity at CF channels corresponding to those contained in the stimulus frequency, while the BM modulation-response, shows a sloping response starting one octave below the excited CF channels, and a constant behavior at CF channels above the 4 kHz. A similar pattern can be observed at the AN level, where a sharp increase in spike rate is observed for frequencies contained in the stimulus in the excitation pattern. A sloping spike

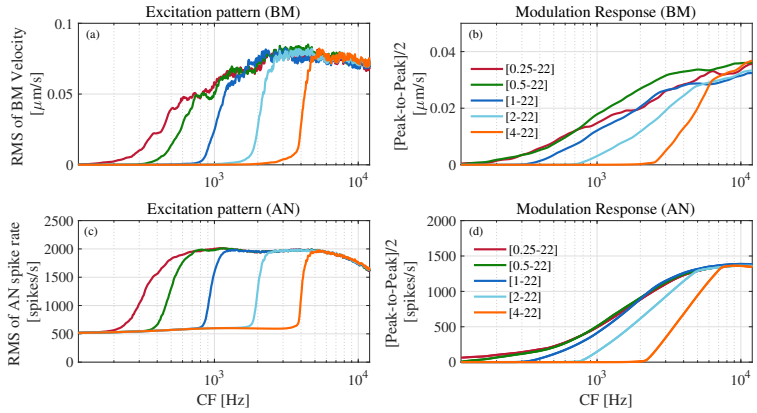


Figure 2.2: Simulated excitation patterns and modulation responses of the BM (a and b) and AN (c and d) processing levels of the model calculated for each CF channel to the experimental stimuli.

rate increment in the modulation-response was again observed when evaluating the modulated response. Since the EFR is a population response across the CF channels, the overlapping modulation response for the first three conditions, explains the equal group-means observed in the experimental data in Figure 2.1.

To investigate the absence of AM-coding at low CFs for 120 Hz modulator, we further evaluated the effect of MF and CF contributions to the modulated response in the broadest bandwidth condition, i.e., [0.25-22] kHz. First, a time-domain visualization from two CF channels is shown in Figure 2.3, for MFs of 20 and 120 Hz. CF channels of 500 and 4000 Hz are shown and are evaluated for their AM content at the BM-velocity (Figure 2.3a) and AN spike-rate (Figure 2.3b). Both responses to non-modulated and modulated broadband stimuli are shown and the response from CF=500 Hz reveals no noticeable difference between non-modulated and 120 Hz modulated noise at both BM and AN processing levels. However, at the CF=4000 Hz the effect of MF can be observed. A more direct comparison of the modulation strength at two CF channels was made in Figure 2.3c and d at BM and AN levels, where the 4800 averaged cycles at 120 Hz MF and 8×100=800 averaged cycles at 20 Hz MF are shown. These figures reveal the lack of modulation response at 120 Hz MF in low CF channel (500 Hz) at both BM and AN levels, which are in agreement with Figure 2.2b and d, where no modulation response can be seen at low CFs.

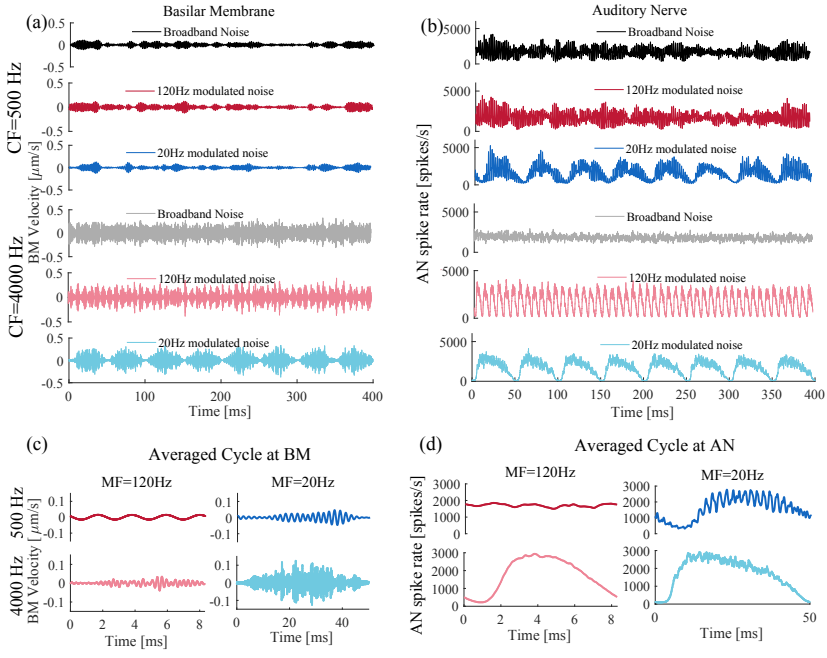


Figure 2.3: One hundred averaged simulated time-domain epochs in response to broadband stimuli modulated with different MFs in two CF channels at the (a) BM and (b) AN levels of the model. Panels (c) and (d) show the averaged cycles across each epoch at BM and AN levels.

2.4 Discussion

The simulation results in response to broadband stimuli confirmed that there is a reduced contribution from lower CF channels and higher modulation frequencies (120 Hz) at the AN level to the EFR. To study the impact of MF and different cochlea CF regions sensitivity to the EFR, Figure 2.4 compares the excitation pattern and modulation response of BM-velocity and AN-spike rates in response to a range of modulation frequencies between 20 and 120 Hz. Although there is an approximately equal excitation pattern in all MFs at both BM and AN levels (Figure 2.4a and 2.4c), the modulation responses show different behavior for different MFs. For a specific MF, the modulation response at BM and AN increases from the apical end to the base. Decreasing the MF, enhances the modulation response at all CFs, especially at the basal end and at the BM level (Figure 2.4b). Even though the AN level largely reflects the response from the cochlea, we observe that the sloping and increasing modulation response saturates for higher CFs to a fixed MF (Figure 2.4d). Therefore, in spite of an increased modulation response at BM level with decreasing MF, especially at higher CFs, the saturation mech-

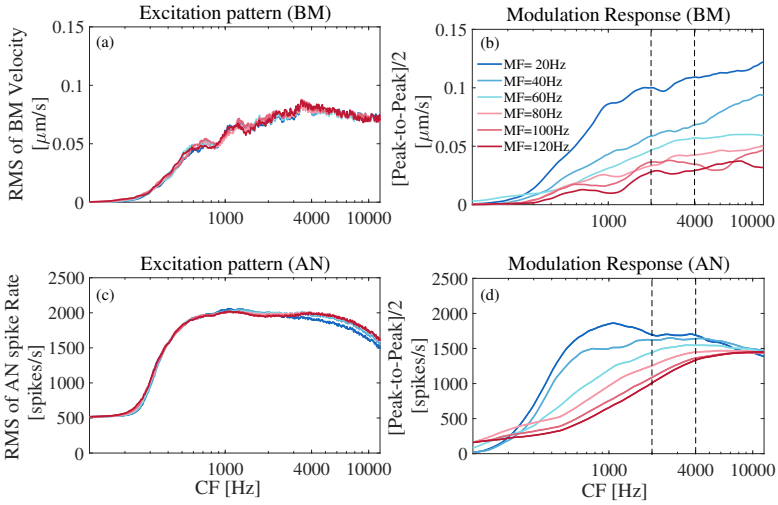


Figure 2.4: Excitation patterns and modulation responses at BM (a and b) and AN (c and d) levels for different MFs of the broadest bandwidth condition, i.e., [0.25-22] kHz.

anism at the AN level (due to saturating AN fibers) moves the EFR frequency sensitivity to the lower CFs. Considering a specific MF, at every CF channel associated with the stimulus frequency content, two side-bands are produced. Hence, one of the possible reasons for the enhanced modulated response of lower MFs at BM, could be the wider filters at higher CFs, where more side-bands of the noise-carrier fall inside the cochlear filters and increase the energy of the modulation response. In this context, increasing the MF moves the side-bands further away from the carrier component and decreases the modulation response within a single CF channel [96].

Based on the model simulations, we conclude that the equal magnitude EFRs for the first three experimental conditions is a consequence of small differences of the curve area of under MF=120 Hz AN modulation response and the saturation which occurs at about 4 kHz and higher CFs (Figure 2.4d). The decrement in the EFR for [4-22] kHz experimental condition is caused by the more significant difference between areas under the curves with the previous, broader conditions. The model study suggests that for broadband noise stimuli, modulation frequencies lower than 120 Hz can provide a better tonotopic-sensitive EFRs for lower CFs. However, there is a possibility of contribution from higher auditory processing levels to be involved in the response.

3

The Derived-Band Envelope Following Response and its Sensitivity to Sensorineural Hearing Deficits

Sarineh Keshishzadeh, Markus Garrett, Viacheslav Vasilkov, Sarah Verhulst

Published in *Hearing Research*, 2020. [1]
https://doi.org/10.1016/j.heares.2020.107979.

This chapter aims to improve frequency specificity of the envelope following response (EFR) by introducing a derived-band EFR (DBEFR) technique and investigate the effect of lifetime noise-exposure, age and outer-hair-cell (OHC) damage on DBEFR magnitudes. Additionally, a modelling approach is adopted to validate the frequency-specificity of the DBEFR and test how different aspects of sensorineural hearing-loss affect peripheral generators. The combined analysis of model simulations and experimental data proposes that (i) DBEFR extracted from the [2-6]-kHz frequency band is a sensitive and frequency-specific measure of synaptopathy in humans, (ii) older listeners consistently have reduced DBEFR magnitudes in comparison to young normal-hearing listeners, in correspondence to how age-induced synaptopathy affects EFRs and compromises temporal envelope encoding, (iii) to a lesser degree, OHC-damage affects the DBEFR magnitude

[1] **SK**: running the experiment (1), model simulations, analysis, conceptualization, writing the original draft, review and editing, **MG**: running the experiment (2) **VV**: conceptualization, review and editing, **SV**: supervision, conceptualization, review and editing, funding acquisition.

(based on model simulations). Hence the DBEFR metric should ideally be combined with a sensitive marker of OHC-damage to offer a differential diagnosis of synaptopathy in listeners with impaired audiograms.

3.1 Introduction

Struggling to understand speech in noisy environments is a prevalent complaint of the ageing population, even when they have normal audiometric thresholds. Although hearing thresholds are informative about the sensory function of the cochlea, they are insensitive to auditory-nerve (AN) fiber loss, which is the first sign of permanent hearing damage [13, 97] and related to supra-threshold hearing [87]. Recent animal studies have shown that ageing, ototoxicity and overexposure to noise can lead to an irreversible loss of AN synapses, i.e. cochlear synaptopathy (CS), and delayed degeneration of cochlear neurons, while leaving the cochlear sensory hair cells intact [13, 16–19, 57]. Even when the noise-exposure dose only causes a temporary threshold shift [13], noise-induced AN fibers degeneration can progress through the lifespan and yield an increased sensitivity of the ear to age-induced hearing dysfunction [98]. Additionally, reduced numbers of spiral ganglion cells in post-mortem histology of human temporal bones with preserved sensory cells, confirmed the existence of age-related CS in humans [38–40]. Thus, noise-exposure and ageing are important causes of CS, a deficit which compromises the temporal coding fidelity of supra-threshold sound as a result of a reduced number of afferent AN synapses innervating the inner-hair-cell (IHC) [41, 87].

Since the discovery of CS, several attempts have been made to associate changes in indirect and non-invasive measures of auditory function, such as scalp-recorded auditory evoked potentials (AEPs), to the histologically quantified degree of AN fibers loss in animals. For example, auditory brainstem responses (ABRs), evoked by transient stimuli and reflecting the synchronized onset responses of AN fibers [61] showed a decreased supra-threshold wave-I amplitude after synaptopathy due to noise-exposure [13, 17, 19], despite recovered normal distortion product otoacoustic emission (DPOAE) and ABR thresholds. The number of AN fibers can also be quantified using envelope following responses (EFRs), which capture how well AN fibers can phase-lock to the stimulus envelope [99]. The EFR can be extracted from scalp-electrodes in response to a sinusoidally amplitude modulated (SAM) pure-tone stimulus [87], and has been proposed as an AEP-based measure of CS [31, 32].

Despite the strong relation between AEP markers and CS in animal studies, the indirect nature of AEP recordings hinders a clear and direct interpretation of response strength in terms of CS. First of all, a mixture of sources contribute to scalp potentials, some of which are electrical activity induced by subject-specific

factors and unrelated to the sound-driven response (e.g. head size, age, sex, geometry of the generators and physiological noise level [35, 60, 87, 91]). Other sources relate to the sound-driven response but depend on outer-hair-cell (OHC) health [85] or cochlear tonotopy [61]. Lastly, the scalp-recorded AEP is strongly influenced by stimulus characteristics and the corresponding spread of basilar-membrane (BM) excitation, which can confound a frequency-specific diagnosis of CS [41, 73, 87, 100]. To address these issues, several studies have proposed differential/relative AEP-based metrics: the EFR amplitude slope as a function of modulation depth [41,51,87], ABR wave-V latency changes in different levels of background noise [101], or the combined use of noise-floor corrected EFRs with ABRs to segregate mixed hearing pathologies and normalize inter-individual variabilities [83]. Secondly, a number of techniques have been proposed to confine ABR generation to specific frequency bands: the use of simultaneous off-frequency masking paradigms, i.e. the derived-band ABR [61, 102], tone-burst ABRs [103] and notched noise paradigms [104]. Lastly, asynchrony of low-spontaneous rate (LSR) AN fibers to the transient stimulus [105] may limit the use of the ABR wave-I amplitude to capture all aspects of CS, as noise-induced CS might preferentially affect LSR AN fibers [16].

This study proposes the use of a relative derived-band EFR method (DBEFR), to confine the EFR to a specific frequency band. To construct DBEFRs, we changed the bandwidth of the stimulus on the low-frequency side rather than using off-frequency masking methods. In this way, a consecutive subtraction of responses to stimuli with various bandwidths will remove the contribution of the off-frequency channels caused by the stimulation paradigm and at the same time, it will yield a relative measure of supra-threshold sound encoding. We further hypothesize that the relative metric design of the DBEFR reduces the impact of subject-specific factors and increases its sensitivity to individual sensorineural hearing deficits. DBEFR magnitudes were extracted from individuals in four groups to study their applicability to diagnose sensorineural hearing deficits: (1) a young normal-hearing control group, (2) a group with self-reported hearing difficulties in noisy environments, (3) a group of older listeners with normal audiograms and (4) an age-matched group with sloping high-frequency audiograms. We assumed that the second group might be affected by CS due to noise overexposure or ageing and that the third group might be affected by age-induced CS, without co-occuring OHC damage. Aside from collecting DBEFRs, we assessed individual OHC function using audiometric and DPOAE thresholds. In line with animal studies of age-related and noise-induced synaptopathy, we expect that the DBEFR will be reduced in all but the control group.

Because a direct assessment of the individual degree of OHC- and AN-damage is presently experimentally impossible, we complemented our experimental work with a modelling approach to better understand the relationship between sen-

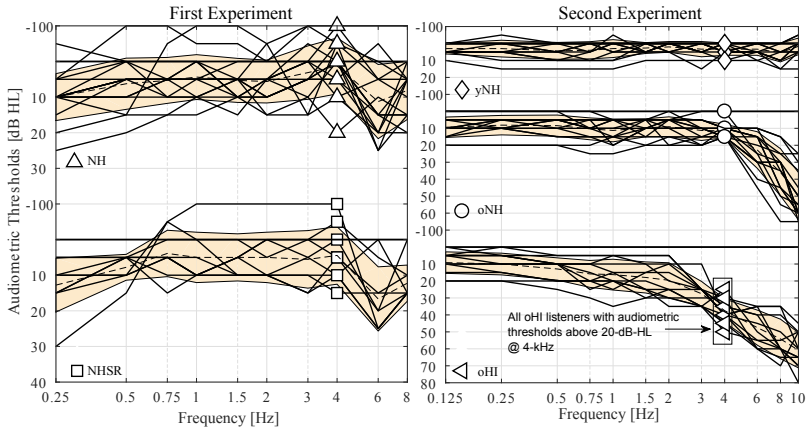


Figure 3.1: Measured audiograms in the first (left) and second (right) experiments. Markers indicate the audiometric threshold at 4 kHz. The dashed line is the averaged audiometric threshold at each group and the yellow shading corresponds to the standard deviation.

sorineural pathologies and their effect on the peripheral generators of the DBEFR. Models can study how AN fiber and sensory hair-cell damage impact the EFR generators to understand their respective roles for DBEFR generation [64,73,106]. We adopt a biophysically inspired model of the human auditory periphery calibrated for ABR and EFR simulation [73] and considered the simulations together with the data to interpret the implications of our findings for DBEFR-based hearing diagnostics.

3.2 Materials and methods

Two experiments were conducted at two recording locations. In the first experiment (Ghent University), normal-hearing (NH) and listeners with self-reported hearing difficulties (NHSR) participated. In the second experiment (Oldenburg University), a total of 43 participants were recruited in three groups: a young NH control group (yNH), an older NH group (oNH) and an older group with sloping high-frequency audiogram (oHI). Ethical approvals were obtained from Ghent and Oldenburg Universities and all participants were informed about the experimental procedures and signed an informed consent before the experiment.

3.2.1 Participants

Sixteen NH listeners with ages between 18 and 30 (NH: 24.21±4.10 years, five females) and 9 NH subjects with self-reported hearing difficulties (NHSR) with

ages between 23 to 49 (NHSR: 33.78±8.57 years, three females) participated in the first experiment. The NHSR participants were recruited using a flyer asking whether they had speech understanding difficulties in the presence of background noise, while not presently being treated for hearing disorders. Measurements were conducted in two sessions per subject, with a maximum sound exposure time of 90 minutes per session. The participants filled out a questionnaire, in which they were asked how often (yearly, monthly, weekly or daily) they had been playing a musical instrument in a band, attended festivals, concerts or discotheques and used noisy tools during their lifetime. Moreover, the total number of noise-exposed sessions, their duration and estimated noise loudness (a score between 1 to 5) were also assessed [107]. Audiograms were measured with an Interacoustics Clinical Computer Audiometer (AC5) at ten standard frequencies between 0.25 and 8 kHz.

The second experiment was conducted with three participant groups composed of: 15 young normal-hearing (yNH: 24.53±2.26 years, eight female), 16 old normal-hearing (oNH: 64.25±1.88 years, eight female) and 12 old hearing-impaired (oHI: 65.33±1.87 years, seven female) participants. All yNH participants had pure-tone thresholds below 20 dB-HL at all measured frequencies between 0.125 and 10 kHz (Auritec AT900, Hamburg, Germany audiometer). In both experiments, the audiometrically better ear was chosen for the experiment and stimuli were presented monaurally while participants were seated in a comfortable chair in an acoustically and electrically shielded sound booth, watching silent movies with subtitles to stay awake. Figure 3.1 shows audiograms of the subjects in all groups.

From here on, $\triangle$ stands for the NH group in the first experiment, $\square$ for NHSR group, $\Diamond$ for yNH in the second experiment, $\bigcirc$ for oNH and $\lhd$ for oHI group.

3.2.2 Distortion Product Otoacoustic Emissions (DPOAEs)

In the first experiment, DPOAEs were recorded to ten primary-level pairs, (L_1, L_2), at nine primary-frequency pairs: f_2 =[546, 780, 1002, 1476, 1998, 3012, 3996, 6006, 8003] and $f_1 = f_2/1.2$. L_2 ranged from 20 to 65 dB-SPL in 5 dB steps and L_1 = 0.4L_2 + 39 dB, according to the scissors paradigm [108]. The nine primary frequency pairs were chosen to have complete stimulus periods of the primaries in each pair. For each frequency and level pair, 45 repetitions were generated in *MATLAB 2016b* and an ER-10X extended-bandwidth Etymotic Research probe system was used to deliver the two pure tones via a loudspeaker/microphone probe inserted in the ear-canal using a silicone eartip. The response was recorded and digitized using a Fireface UCX external sound card (RME). The pure tones were calibrated separately using a B&K artificial ear and B&K sound level meter at each primary frequency, separately. The time-domain ear-canal recordings were converted to pressure using the microphone sensitivity (50 $\frac{\mathrm{mV}}{\mathrm{Pa}}$) and

pre-amplifier gain (40 dB). Then, I/O functions were calculated for the measured primary-frequency pairs by defining the L_{DP} as the averaged spectrum magnitude at the $2f_1$-f_2 cubic distortion frequency, multiplied by $\frac{2}{N\sqrt{2}}$, where N is the number of samples at each f_2 response. Finally, a linear function, i.e. $L_{DP} = aL_2 + b$, was fit to the bootstrapped data-points and the crossing point with L_{DP}=0 Pa was defined as the DPOAE threshold at the measured f_2 frequency. DPOAEs in the second experiment were acquired using a custom-made software [109], which implements a primary frequency sweep method at a fixed f_2/f_1 of 1.2 [110]. The primary frequencies were swept across an 1/3 octave range around the $f_2 = 4$ kHz geometric mean with a duration of 2s/octave. Primary levels were chosen according to the scissors paradigm [108]. DPOAE threshold at each frequency was calculated by fitting a linear function to the bootstrapped data-points and was extrapolated to cross L_{DP}=0 Pa. Additional details on the experimental procedure can be found in [64].

3.2.3 Envelope Following Responses (EFRs)

The EFR stimuli in the first experiment were five filtered white noise carriers, which were 100% modulated with a 120-Hz sinusoid. To generate them, the white noise was filtered between the following frequency regions: [0.25-22], [0.5-22], [1-22], [2-22] and [4-22] kHz, using a 1024th order FIR band-pass filter designed by the Blackman-window method. In each frequency band, a stimulus with a duration of 1.25 s was generated in *MATLAB 2016b*, windowed with a 1.25% cosine-tapered window and delivered monaurally over ER-2 earphones, connected to a Fireface UCX external sound card (RME) and a TDT-HB7 headphone driver. A uniformly-distributed random silence jitter was applied between consecutive epochs (200 ms$\pm$20 ms) of the 370 stimulus presentations. Stimuli with various bandwidths were calibrated to have the same spectral magnitude, i.e. the widest bandwidth stimulus was presented at 70-dB-SPL, while narrower bandwidth stimuli had lower sound pressure levels to preserve an equal spectral level in all conditions. The calibration was performed using a B&K sound-level-meter type 2606. Figure 3.2a illustrates the designed stimuli in the frequency domain. Scalp-recorded potentials were obtained with a 64-Channel Biosemi EEG recording system and a custom-built trigger-box using a sampling frequency of 16384 Hz. The electrodes were placed according to the 10-20 standard, using highly conductive gel (Signa gel). The Common Mode Sense (CMS) and Driven Right Leg (DRL) electrodes were placed on top of the head. Six external channels were used as well, i.e. two earlobe electrodes as reference and the remaining electrodes were placed on the forehead and cheeks to record electrical activity induced by horizontal and vertical eye movements. All channels were re-referenced to the average of the two earlobe electrodes.

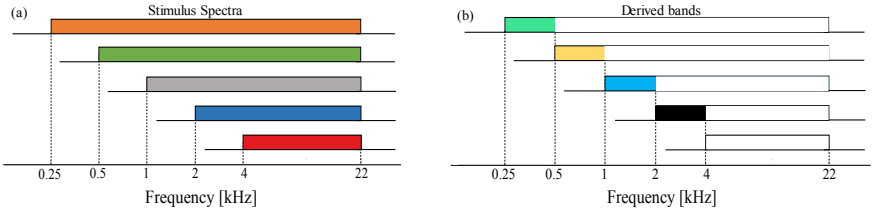


Figure 3.2: Spectra of the 120-Hz modulated stimuli and derived bands. (a) Designed stimulus spectra in different frequency bands and specified cut-off frequencies of the bandpass filter. (b) Derived bands from the EFRs recorded to the stimuli shown in (a) obtained by spectral subtraction.

In the second experiment, four EFR stimuli with white noise carriers were band-pass filtered using the same filter as in the first experiment in [0.3-16], [0.7-16], [2.8-16] and [5.6-16] kHz frequency regions. The precise lower cut-off frequencies employed in the band-pass filtering were $\frac{0.5}{\sqrt{2}}$, $0.5\sqrt{2}$, $\frac{4}{\sqrt{2}}$ and $4\sqrt{2}$ kHz, respectively. Stimuli were 95% modulated with a 120-Hz pure tone and presented at 70 dB-SPL using the same configuration as the first experiment. The stimuli had a duration of 400 ms, were 2.5% ramped with a tapered-cosine window and presented 1000 times using a uniformly distributed random inter-stimulus silence jitter of 100 ms$\pm$10 ms. The calibration was performed in the same way as for the first experiment, but using B&K sound level meter type 2610. A 64-channel Biosemi EEG system was adopted to record the responses using EEG caps with equidistant electrode spacing. The CMS and DRL electrodes were located on the fronto-central midline and on the tip of the nose of the participants, respectively.

3.3 EFR analysis

Acquired EFRs were first filtered with an 800th order Blackman window-based FIR filter between 60 and 600 Hz, using the *filtfilt* function of *MATLAB* to avoid time delays and phase shifts. Signals were broken into 1-s long epochs relative to the trigger onset, from 0.25 to 1.25s in the first and into 0.3-s long epochs, from 0.1 to 0.4 s in the second experiment. Baseline correction was applied before the epochs were averaged across trials. 30 and 100 epochs were rejected on the basis of the highest peak-to-trough values in the first and second experiment, respectively. Since the firing patterns of neurons are influenced by factors such as instantaneous external inputs, previous firing patterns and the general state of the system, the interpretation of the raw EFR spectrum resulting from the Fast Fourier Transform (FFT) of the averaged epochs is challenging. Synaptic delays and axon conduction limitations cause a $\frac{1}{f}$ behaviour in EEG (Chapter 10 in [111]) and it is crucial to suppress this noise-floor to analyse the stimulus-driven spectrum. The boot-

strapping approach proposed in [95] was employed to estimate the $\frac{1}{\mathrm{f}}$ noise-floor component. First, 340 epochs were drawn randomly with replacement, among the 340 epochs (900 epochs in the second experiment). Then, the FFT of these epochs were averaged. This procedure was repeated N_1=200 times (N_2=400 for the second experiment), resulting in a nearly Gaussian distribution of raw, averaged spectra. The average value of this distribution yielded the frequency domain representation of the EFRs. Afterwards, the same procedure with M_1=1000 repetitions (M_2=1200 for the second experiment) and phase-flipped (180°) odd epochs was followed to estimate the spectral noise-floor as a function of frequency. The idea behind this approach is that the time-locked response is suppressed if the averaging is repeated sufficiently across phase-inverted epochs. Finally, the averaged absolute values of the estimated noise-floors were subtracted from the averaged absolute values of the EFR spectra amplitudes ($\mathrm{EFR}_{\mathrm{raw}}$) to obtain the stimulus-driven EFR spectrum ($\mathrm{EFR}_{\mathrm{Spec}}$):

$$\mathrm{EFR}_{\mathrm{raw}}(f) = \frac{2}{n_p} \frac{\left|\sum_{i=1}^{N} \mathrm{FFT}(X_i)\right|}{N_p} \tag{3.1}$$

$$\mathrm{Noisefloor}(f) = \frac{2}{n_p} \frac{\left|\sum_{j=1}^{M} \mathrm{FFT}([-1]^j X_j)\right|}{M_p} \tag{3.2}$$

$$\mathrm{EFR}_{\mathrm{Spec}}(f) = \mathrm{EFR}_{\mathrm{raw}}(f) - \mathrm{Noisefloor}(f) \tag{3.3}$$

X represents the epochs vector, N the number of bootstrap repetitions, M the number of repetitions to estimate the noise-floor, p the experiment number (i.e. one or two) and n equals the number of FFT-points (n_1=16384 and n_2=8192). Figure 3.3 represents $\mathrm{EFR}_{\mathrm{raw}}$, Noisefloor and $\mathrm{EFR}_{\mathrm{Spec}}$ spectra of subject No. 8 from NH group in the first experiment. All $\mathrm{EFR}_{\mathrm{Spec}}$ peak values which were four standard deviations above the noise-floor ($\mathrm{EFR}_{\mathrm{SpecSD}}$) for frequencies corresponding to the modulation frequency (120 Hz) and its following two harmonics (240 and 360 Hz) were added to yield EFR magnitude of the corresponding condition.

$$\mathrm{EFR}_{\mathrm{PtN}} = \sum_{\mathrm{k}=0}^{2} \mathrm{EFR}_{\mathrm{SpecSD}}(f_k), \qquad f_k = 120 \times (k+1) \tag{3.4}$$

To construct DBEFRs, the calculated $\mathrm{EFR}_{\mathrm{PtN}}$ for each narrower-band condition was subtracted from the following wider-band condition using:

$$\mathrm{DBEFR}_{\mathrm{PtN}} = (\mathrm{EFR}_{\mathrm{PtN}})_{wide} - (\mathrm{EFR}_{\mathrm{PtN}})_{narrow} \tag{3.5}$$

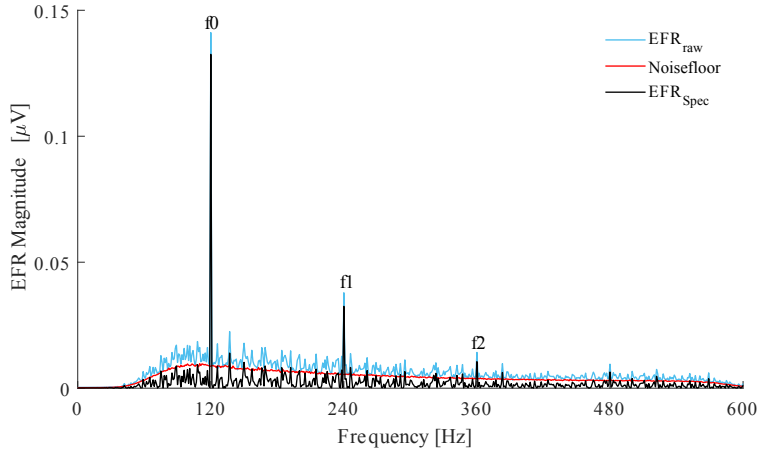
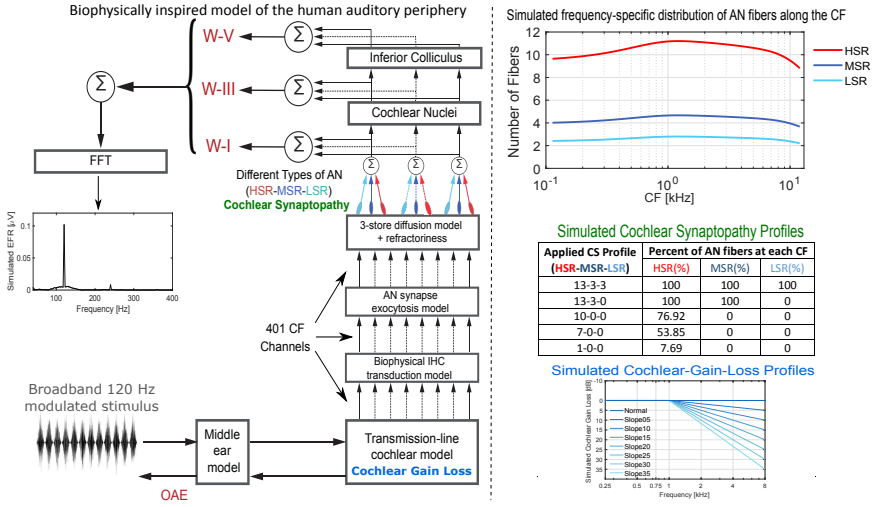


Figure 3.3: Spectrum of the $\mathrm{EFR_{raw}(f)}$ *(in blue),* $\mathrm{Noisefloor(f)}$ *(in red) and* $\mathrm{EFR_{Spec}(f)}$ *(in black) calculated for subject No. 8 from the first experiment. EFR spectra were evoked by the stimulus with the broadest bandwidth, i.e. [0.25-22] kHz. Peaks at the stimulus modulation frequency, and two harmonics (i.e.* $\mathrm{f_0 = 120Hz}$, $\mathrm{f_1 = 240Hz}$ *and* $\mathrm{f_2 = 360Hz}$*) are clearly visible above the noise-floor.*

Derived frequency bands from EFRs to the first experimental stimuli are shown schematically in Figure 3.2b.

3.4 Questionnaire analysis

The completed questionnaires from the participants in the first experiment were used to estimate the individual life-time noise-exposure dose. To this end, the collected individual data related to the frequency and duration of experienced noise-exposure were converted to a number of sessions per year multiplied by the duration and the personal estimated noise loudness scores, i.e. a number between 1 and 5. We followed the procedures as described in [107]. The scores were separately calculated for questionnaire categories: (i) playing musical instrument in a band, (ii) attending festivals, concerts and discotheques and (iii) using noisy tools. Outcomes were normalized across NH and NHSR groups participants by the highest reported dose, i.e. 30600, 18480 and 26000 hours in each category, respectively.



Applied CS Profile (HSR-MSR-LSR)	Percent of AN fibers at each CF		
	HSR(%)	MSR(%)	LSR(%)
13-3-3	100	100	100
13-3-0	100	100	0
10-0-0	76.92	0	0
7-0-0	53.85	0	0
1-0-0	7.69	0	0

Figure 3.4: Modeling approach. The block-diagram shows different levels of the auditory pathway modelled in the employed biophysical model of the auditory periphery [73]. The top-right graph indicates the simulated distribution of different types of AN fibers across CF. The table shows simulated CS profiles and the graph on the bottom right depicts simulated different degrees of cochlear-gain-loss. The corresponding simulated thresholds at 8 kHz are indicated by the legend.

3.5 Model simulations

A biophysical model of the human auditory periphery [73], schematically shown in Figure 3.4, was adopted to simulate the experimental conditions and to investigate the effect of different aspects of sensorineural hearing deficits on the $\mathrm{EFR}_{\mathrm{PtN}}$ and $\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitudes. The original implementation of the model is described in [73] and can be downloaded from *"https://github.com/HearingTechnology/ Verhulstetal2018Model"*. The parameters which determine the weights between the population AN, cochlear nucleus (CN) and inferior colliculus (IC) responses were adjusted along with the AN innervation patterns across characteristic frequency (CF) channels for the purpose of this study.

3.5.1 Auditory nerve-fiber distribution

The original model implementation introduced the same number of synapses between IHCs and AN fibers for all simulated CFs, whereas human and rhesus monkey innervation patterns show a bell-shaped pattern across CF. To make the model more realistic, the averaged synaptic counts of four control rhesus monkeys (seven ears) and nine frequencies [18] were mapped to corresponding fractional distances of the human cochlea using the monkey place-frequency map [112]. Fractional

distances from the base of cochlea, d_i, were calculated according to the measured frequency points ($\mathrm{f_{RM_i}}$):

$$\mathrm{f_{RM_i}[in\ Hz]} = 360(10^{2.1(1-d_i)} - 0.85), \qquad i = 1, 2, ..., 9 \tag{3.6}$$

The obtained d_is were substituted into the analogous Greenwood map equation for humans, yielding the corresponding frequency points ($\mathrm{f_{H_i}}$):

$$\mathrm{f_{H_i}[in\ Hz]} = 165.4(10^{2.1(1-d_i)} - 0.88), \qquad i = 1, 2, ..., 9 \tag{3.7}$$

To calibrate the model with the applied AN-pattern, a 70 dB-nHL click-train containing both stimulus polarities was presented at a rate of 11 Hz. To perform this calibration, simulated ABR wave amplitudes were matched to the experimental data on the basis of 55 averages. Specifically, the $\mathrm{M_1} = 4.6729 \times 10^{-14}$, $\mathrm{M_3} = 5.6885 \times 10^{-14}$ and $\mathrm{M_5} = 14.641 \times 10^{-14}$ parameters were adjusted on the basis of average NH ABR wave-I, III and V reference-data from [113], i.e. $\mathrm{w_I} = 0.15\mu\mathrm{V_p}$, $\mathrm{w_{III}} = 0.17\mu\mathrm{V_p}$ and $\mathrm{w_V} = 0.61\mu\mathrm{V_{pp}}$.

Using the synapse counts from rhesus monkey and the mapped frequency points for the human cochlea ($\mathrm{f_{H_i}}$), a *smoothing spline* curve was fit to estimate the number of synapses across all frequency channels in the model. Finally, to simulate different AN fiber types, i.e. high spontaneous-rate (HSR), medium spontaneous-rate (MSR) and LSR fibers, and their properties, the obtained population distribution was multiplied by the corresponding AN type proportion factor C, i.e. $\mathrm{C_{HSR}} = 0.60$, $\mathrm{C_{MSR}} = 0.25$ and $\mathrm{C_{LSR}} = 0.15$ (cat data in [114]), before responses were summed at each simulated CF and fed to the CN model. The simulated frequency-specific AN fibers distribution is shown on the top-right column of Figure 3.4.

3.5.2 Stimuli

The model stimuli were matched to the experimental conditions and had a duration of 600 and 400 ms for the first and second experiment, respectively. Twenty stimulus repetitions with different white noise iterations were applied to the model and simulations were averaged before the $\mathrm{EFR_{PtN}}$ was calculated using the same procedure as in Eq. 3.4. The amplitudes of the model stimuli were set based on the broadest condition, i.e. 0.25 to 22 kHz for the first experiment and 0.3 to 16 kHz for the second experiment to yield an input of 70 dB-SPL. The narrower band stimuli were calibrated relative to the broadest condition, such that they had the same spectral level as the broadband condition, but with a different SPL.

3.5.3 Simulating sensorineural hearing loss

In the employed cochlear transmission-line model, pole-trajectory of the BM admittance can be determined as a function of stimulus intensity to model the compressive BM-response growth, and in a CF-dependent manner to account for human cochlear frequency tuning [115]. Accordingly, OHC-damage was simulated by changing the CF-dependent mechanical gain of the cochlea by moving poles of the BM admittance function to yield a filter gain reduction corresponding to a desired dB-HL-loss, which also yielded wider cochlear filters. The inset in Figure 3.4 shows the simulated cochlear-gain-loss profiles. In the context of CS, different degrees of AN-damage were modelled by manipulating the number and types of AN fibers. The simulated CS profiles and their corresponding AN fiber types are shown in Figure 3.4. The table in Figure 3.4 shows the simulated CS profiles. Model-related procedures are further detailed in [64, 73].

3.6 Results

3.6.1 EFR and dependence on stimulus frequency

Figure 3.5 shows individual and group-mean $\mathrm{EFR}_{\mathrm{PtN}}$ magnitudes to different frequency bandwidths in the first (panel a) and second (panel b) experiments. Despite the within-group individual variability, experimental group-means revealed approximately constant $\mathrm{EFR}_{\mathrm{PtN}}$ magnitudes to stimuli with frequencies below 2 kHz and reduced magnitudes to frequencies above 2 kHz and 2.8 kHz in the first and second experiment, respectively. A paired-sample t-test with Bonferroni correction was applied to compare $\mathrm{EFR}_{\mathrm{PtN}}$ magnitudes to stimuli with different frequency bandwidths in each group. In the first experiment, a single significant difference was observed between the $\mathrm{EFR}_{[2-22]}$ and $\mathrm{EFR}_{[4-22]}$ conditions in NH group ($t_{(11)}$=7.02, p<0.000; specificed by # in Figure 3.5a), which disappeared for the NHSR group (t(8)=3.13, p=0.014). In the second experiment, a paired-sample t-test with Bonferroni correction gave a significant difference between $\mathrm{EFR}_{[2.8-16]}$ and $\mathrm{EFR}_{[5.6-16]}$ in yNH ($t_{(12)}$=7.86, p<0.000; specified by $+$ in Figure 3.5b) and oNH groups ($t_{(12)}$=6.21, p<0.000; specificed by $++$ in Figure 3.5b), but not in the oHI group ($t_{(9)}$=2.03, p=0.072). Simulated NH-EFRs are shown in hexagons in Figure 3.5 and corroborate experimental findings by showing a minor contribution of stimulus frequencies below 2 kHz on the EFR generation.

3.6.2 Derived-Band Envelope Following Responses (DBEFRs)

$\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitudes calculated using Eq. 3.5 are shown in Figure 3.6 for the first (panel a) and second (panel b) experiment. A paired-sample t-test with Bon-

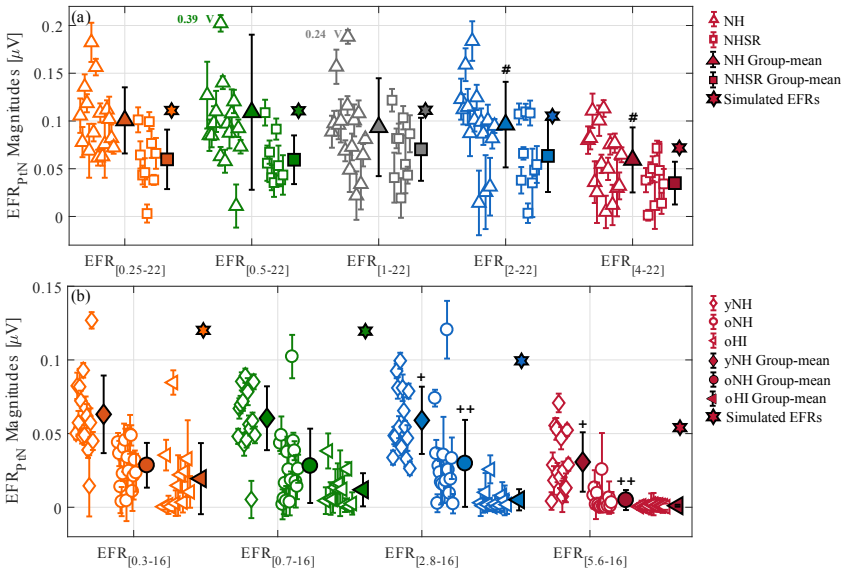


Figure 3.5: $\mathrm{EFR_{PtN}}$ *magnitudes to 120-Hz modulated stimuli with different white-noise-carrier bandwidths in the (a) first and (b) second experiment. Individual data-points are depicted with open symbols. Respective standard deviations were obtained using a bootstrapping procedure [95]. Filled symbols reflect the group-means and their corresponding standard deviations. Simulated EFRs from a NH model were added in filled hexagons. Significant effects of considered frequency-band on* $\mathrm{EFR_{PtN}}$ *magnitudes are specified by: (#) in the NH-group (first experiment), (+) in the yNH-group and (++) in the oNH-group (second experiment). To enhance the visualization of differences, panel (a) was plotted on narrower y-axis range, therefore the real values of lowered* $\mathrm{EFR_{PtN}}$ *magnitudes are specified next to the corresponding data-points.*

ferroni correction comparing the $\mathrm{DBEFR_{PtN}}$ magnitudes in each group revealed only a significant difference between the [1-2] and [2-4] kHz conditions in the NH group ($t_{(11)}$=-6.05, p<0.000; specified by # in Figure 3.6a). In the second experiment, paired-sample t-test showed significant difference between [0.3-0.7] and [2.8-5.6]-kHz conditions only in yNH group ($t_{(12)}$=-7.00, p<0.000; specificed by + in Figure 3.6b). In support of our experimental findings, simulated NH-DBEFR magnitudes in both experiments (shown by hexagons in Figure 3.6a and b) were equal for derived-bands below 2-kHz and increased for $\mathrm{DBEFR_{[2-4]}}$ (in the first experiment) and $\mathrm{DBEFR_{[2.8-5.6]}}$ (in the second experiment). In line with $\mathrm{EFR_{PtN}}$ findings in Section 3.6.1, experimental and simulated $\mathrm{DBEFR_{PtN}}$ magnitudes in both experiments showed an increased contribution of the [2-6] kHz derived frequency band to the EFR generation. Unlike the model simulations, negative DBEFR magnitudes were observed in a few experimental cases. How-

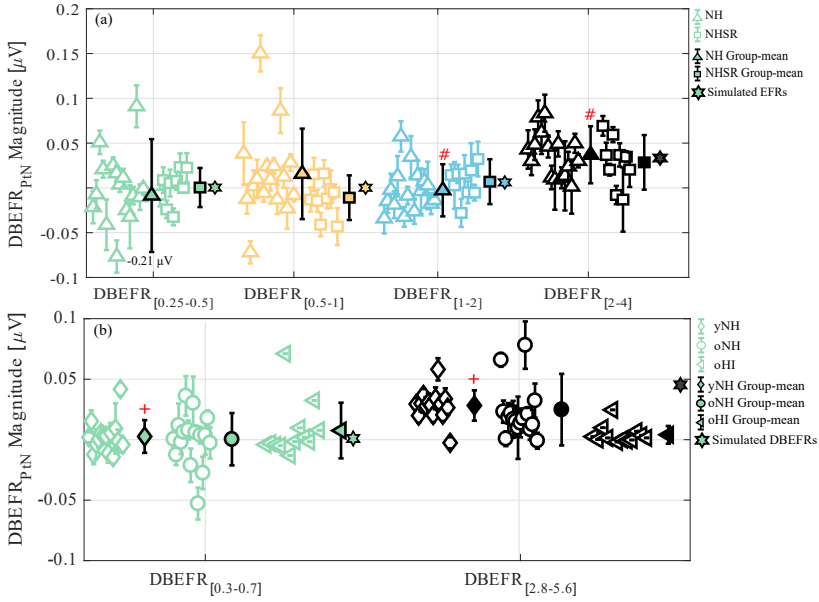


Figure 3.6: $\mathrm{DBEFR_{PtN}}$ *magnitudes derived using Eq. 3.5 for 120 Hz modulated stimuli with different white-noise-carrier bandwidths in the (a) first and (b) second experiment.* $\mathrm{DBEFR_{PtN}}$ *for each frequency band was obtained from a wider and narrower width stimulus. Standard deviations were calculated using a bootstrapping procedure. Simulated DBEFRs are shown with stars, which were calculated by averaging the responses from 20 stimulus iterations. Group means and respective standard deviations are depicted using filled symbols. Significant effects of considered frequency-band on NH-group in the first experiment and yNH-group in the second experiment are specified by (#) and (+), respectively. To enhance the visualization of differences, figures were plotted on narrower y-axis range, therefore the real values of lowered* $\mathrm{DBEFR_{PtN}}$ *magnitudes are specified next to the corresponding data-points.*

ever, the corresponding standard errors of the negative cases in $\mathrm{DBEFR}_{[2-4]}$ and $\mathrm{DBEFR}_{[2.8-5.6]}$ conditions cross zero, and hence can be interpreted by the measurement variability.

3.6.3 Possible origins of individual EFR differences

Previous studies have shown a dependency of the scalp-recorded AEP magnitude to head size, sex and age [60, 83, 91]. Hence, the spread of data-points within different recorded test-groups and spectral bandwidths could be explained by subject-specific factors unrelated to hearing or hearing-related factors associated with the main factors for grouping: (i) self-reported hearing difficulties in noisy environments in the first experiment, (ii) age and (iii) elevated hearing thresholds in the

second experiment.

Pooling together the NH and NHSR $\mathrm{EFR}_{\mathrm{PtN}}$ magnitudes, a regression analysis was conducted to investigate the effect of age, 4 kHz threshold, head-size and DPTH_{3000} on the $\mathrm{EFR}_{[2-22]}$ (Figure 3.7, left column) and $\mathrm{DBEFR}_{[2-4]}$ magnitude (Figure 3.8, left column). None of the regressions showed a relation between tested variables, suggesting that other factors than those reported were responsible for the individual variability among listeners. The regression analysis on $\mathrm{EFR}_{\mathrm{PtN}}$ and $\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitudes combined from all experimental groups in the second experiment (Figure 3.7 and 3.8, right column) showed a meaningful correlation of age, threshold, head-size and DPTH_{4000} with the $\mathrm{EFR}_{[2.8-16]}$ magnitude. However, extracting the $\mathrm{DBEFR}_{[2.8-5.6]}$, reduced the correlation with age and 4-kHz threshold and suppressed any meaningful correlation with head-size and DPTH_{4000}. Moreover, excluding the oHI group from the correlation analysis, led to a reduced and insignificant correlation coefficient (R=-0.377, p=0.090) between 4-kHz threshold and $\mathrm{DBEFR}_{[2.8-5.6]}$. These results suggest that the proposed DBEFR metric is not affected by head-size. Moreover, individual variabilities between the yNH and oNH groups in the second experiment might be related to degraded temporal envelope-coding as a consequence of CS [41], given the insignificant correlations of DBEFRs with the 4-kHz threshold, DPTH_{4000} and head-size.

3.6.4 $\mathrm{EFR}_{\mathrm{PtN}}$ and $\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitude variability across tested groups

To investigate the separability of the recruited groups by means of their DBEFR magnitudes, we analysed the group-mean differences in each experiment. In the first experiment, an independent two-sample t-test comparison between the means of stimulated frequency bandwidths in the NH and NHSR group (Figure 3.5a), showed a significant difference only between the [2-22] and [4-22]-kHz conditions ($\mathrm{EFR}_{[2-22]}$: $\mathrm{t}_{(19)}$=3.36, p=0.003 and $\mathrm{EFR}_{[4-22]}$: $\mathrm{t}_{(19)}$=2.76, p=0.012). However, significant mean-differences disappeared between similar conditions in the NH and NHSR groups after extracting DBEFR magnitudes in Figure 3.6a ($\mathrm{DBEFR}_{[2-4]}$: $\mathrm{t}_{(19)}$=1.15, p=0.265). The insignificant difference across groups and insignificant correlation coefficients of $\mathrm{DBEFR}_{[2-4]}$ with subject-specific factors observed in Figure 3.8, might partly be explained by the different amounts of experienced lifetime noise-exposure reported in the questionnaires and might point to various degrees of noise-induced CS. Calculated noise scores in Figure 3.9 revealed a weak correlation with $\mathrm{DBEFR}_{[2-4]}$ magnitudes (R=0.13, p$<$0.000). However, certain cases appeared to be inconsistent with our noise-induced synaptopathy hypothesis, i.e. (i) high noise scores in the NH group, e.g. subject No. 12 and (ii) low noise scores in the NHSR group, e.g. subject No. 1. We suggest

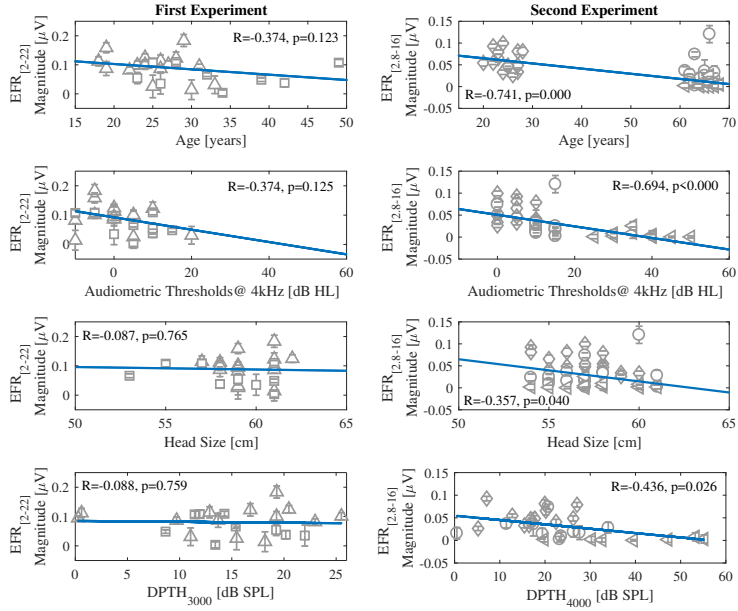


Figure 3.7: Correlation analysis of $\mathrm{EFR}_{[2-22]}$ *(*$\mathrm{EFR}_{[2.8-16]}$*) with age, audiometric threshold at 4 kHz, head-size and* DPTH_{3000} *(*DPTH_{4000}*) in the first (left) and second (right) experiments. Correlation between EFR magnitudes and all factors but age were reported using the Pearson's correlation coefficient. The Spearman's correlation coefficient was calculated to study the effect of age in the second experiment.*

that the insignificant group-mean differences could be explained by (i) subject-dependent unreliable discriminating factor between NH and NHSR group [116], (ii) variability in answering lifetime noise-exposure dose in questionnares [48, 49], (iii) an insufficient number of samples or (iv) a limited sensitivity of the $\mathrm{DBEFR}_{\mathrm{PtN}}$ metric to noise-induced CS.

In the second experiment, an independent two-sample t-test was applied to investigate the effect of age between the yNH and oNH groups, and elevated high-frequency thresholds between the oNH and oHI groups. This comparison showed a significant effect of age on all frequency bandwidths and a significant effect of hearing threshold on all frequency bands except for the [5.6-16] kHz band ($\mathrm{t}_{(21)}$ = -1.81, p = 0.084). The same comparison for the DBEFR magnitudes revealed a significant effect of age and hearing threshold only in the [2.8-5.6]-kHz derived band condition ($\mathrm{t}_{(24)}$ = 3.03, p=0.004 and $\mathrm{t}_{(21)}$ = -4.33, p = 0.002, respectively), consistent with the correlation presented in Figure 3.8. Detailed *t* and *p* values of independent two-sample t-tests, evaluating the effect of age and hearing thresholds

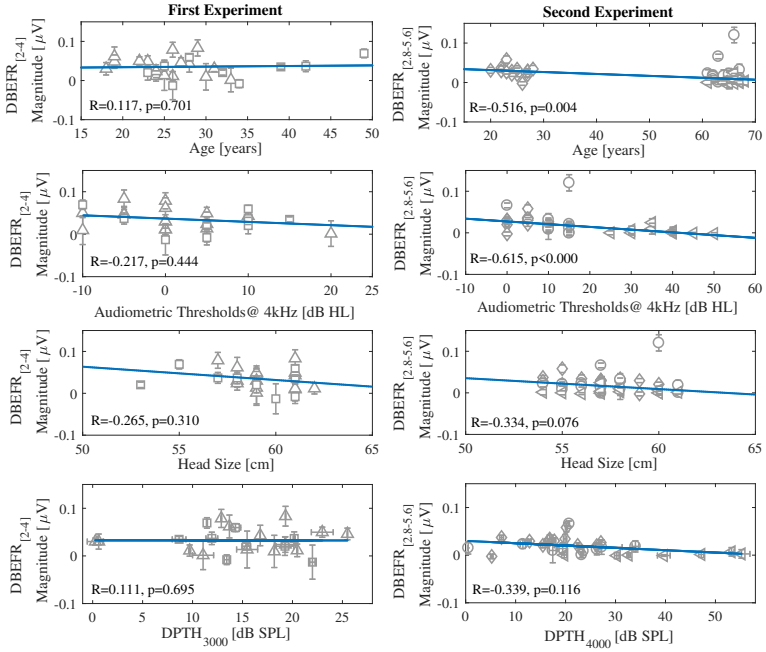


Figure 3.8: Correlation analysis of $\mathrm{DBEFR}_{[2-4]}$ *(*$\mathrm{DBEFR}_{[2.8-5.6]}$*) with age, audiometric threshold at 4 kHz, head-size and* DPTH_{3000} *(*DPTH_{4000}*) in the first (left) and second (right) experiments. Correlation between DBEFR magnitudes and all factors but age were reported using the Pearson's correlation coefficient. The Spearman's correlation coefficient was calculated to study the effect of age in the second experiment.*

on EFR and DBEFR magnitudes, are listed in Table 3.1.

Our group-mean results combined with the correlation analysis in Section 3.6.3 suggests that the DBEFR metric removes inter-subject variability unrelated to hearing between yNH and oNH groups, but leaves individual magnitude differences within a group meaningful, given the often non-overlapping standard deviations. Consequently, the significant group-mean difference between yNH and oNH groups might reflect individual degrees of sensorineural hearing loss. To investigate the diagnostic sensitivity, it is of course necessary to understand the respective role of OHC deficits and CS on DBEFR magnitudes. Given that oHI listeners may suffer from both OHC deficits and CS, it is important to study the impact of OHC-damage and CS, both independently and concomitantly.

Table 3.1: The results of a two-tailed t *test show the effect of age and hearing threshold on EFR and DBEFR magnitudes in the second experiment.*

Metric	Frequency Bandwidth [kHz]	Age Effect yNH vs. oNH	Threshold Effect oNH vs. oHI
EFR	[0.3-16]	$t_{(24)}$=5.812 p<0.000	$t_{(21)}$=-3.020 p=0.006
	[0.7-16]	$t_{(24)}$=6.632 p<0.000	$t_{(21)}$=-2.175 p=0.041
	[2.8-16]	$t_{(24)}$=5.836 p<0.000	$t_{(21)}$=-4.498 p<0.000
	[5.6-16]	$t_{(24)}$=4.734 p<0.000	$t_{(21)}$=-1.811 p=0.084
DBEFR	[0.3-0.7]	$t_{(24)}$=-1.35 p=0.191	$t_{(21)}$=-0.61 p=0.541
	[2.8-5.6]	$t_{(24)}$=3.03 p=0.004	$t_{(21)}$=-4.33 p=0.002

3.6.5 The EFR relationship to different aspects of sensorineural hearing-loss

Since OHC-damage and CS might both affect the EFR magnitude [42,83], we employed a computational model of the auditory periphery to simulate how different degrees of CS affected the $\mathrm{EFR}_{\mathrm{PtN}}$ magnitude, both in presence and absence of high-frequency sloping OHC-loss above 1 kHz (simulated high-frequency sloping audiograms in Figure 3.4). The most sensitive regions of the cochlea responding to a 120-Hz modulated broadband noise were identified to lie between the CFs of 2 and 6 kHz [117]. As a result, we only considered two EFR conditions of each experiment, namely $\mathrm{EFR}_{[2-22]}$ and $\mathrm{EFR}_{[4-22]}$ in the first experiment (Figure 3.10a) and $\mathrm{EFR}_{[2.8-16]}$ and $\mathrm{EFR}_{[5.6-16]}$ in the second experiment (Figure 3.10b). Model simulations showed that CS, when no other hearing deficits co-occur, reduces the EFR and DBEFR magnitudes. Applying sloping high-frequency OHC-damage increased the DBEFR magnitudes in both experiments (Figure 3.10c and d). According to the simulations, the NH DBEFR magnitude reduced by 46% as a consequence of removing 47% of the AN fibers (i.e., the 10-0-0 CS profile defined in Figure 3.4), while the Slope20 OHC-damage (defined in Figure 3.4) increased the NH DBEFR magnitude by 27%. Hence, the effect of OHC-damage on the DBEFR magnitude is smaller than that of CS alone, however it is not negligible. Therefore, the experimental range of individual EFR and DBEFR magnitudes can be explained by different degrees of variation simulated by CS and OHC-damage.

Our simulations explained the experimental differences between yNH and oNH groups on the basis of age-induced CS, not OHC-damage induced differences.

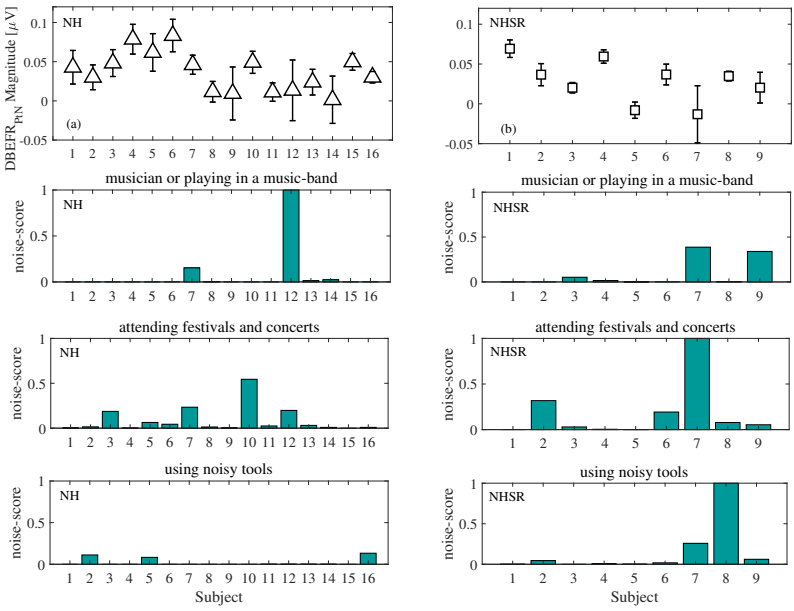


Figure 3.9: Bar-plots of noise scores acquired from questionnaires of NH and NHSR groups, classified in three categories, i.e. experience noise as a consequence of (i) playing a musical instrument in a band, (ii) attending festivals or concerts and (iii) using noisy tools. Results are shown normalised, where the score of 1 corresponds to 30600, 18480 and 26000 hours of accumulated noise dose in the considered categories, respectively.

Furthermore, the simulations suggest that oNH and oHI listeners might both suffer from CS. Results are less clear for the NHSR group where there is a strong overlap with the NH group. However, the noise-scores from the questionnaires in Figure 3.9, could ascribe some of the spread in DBEFR magnitudes within the NH and NHSR groups to noise-induced CS, and to a lesser degree to OHC-damage given that all had normal hearing thresholds.

It is worthwhile to note that EFR magnitudes in both experiments (Figure 3.10a and b), decreased as a result of CS alone and increased by applying high-frequency OHC-damage with a severity of less than 20 dB-HL at 8 kHz. However, higher degrees of OHC-damage reduced the EFR magnitudes. We explain this non-monotonic behaviour on the basis of the AN fiber discharge rate-level curve, where increased simulated $\mathrm{EFR}_{\mathrm{PtN}}$ magnitudes (Figure 3.10 c and d) and amplitude-modulated (AM) responses (Figure 3.11b) to supra-threshold stimuli (70 dB-SPL) caused by OHC-damage, might stem from the extended dynamic range of the AN fibers for less effective AN-driving levels (Figure 3c in [87]). Given that experimental and simulated stimuli were calibrated to have equal spectral magnitudes

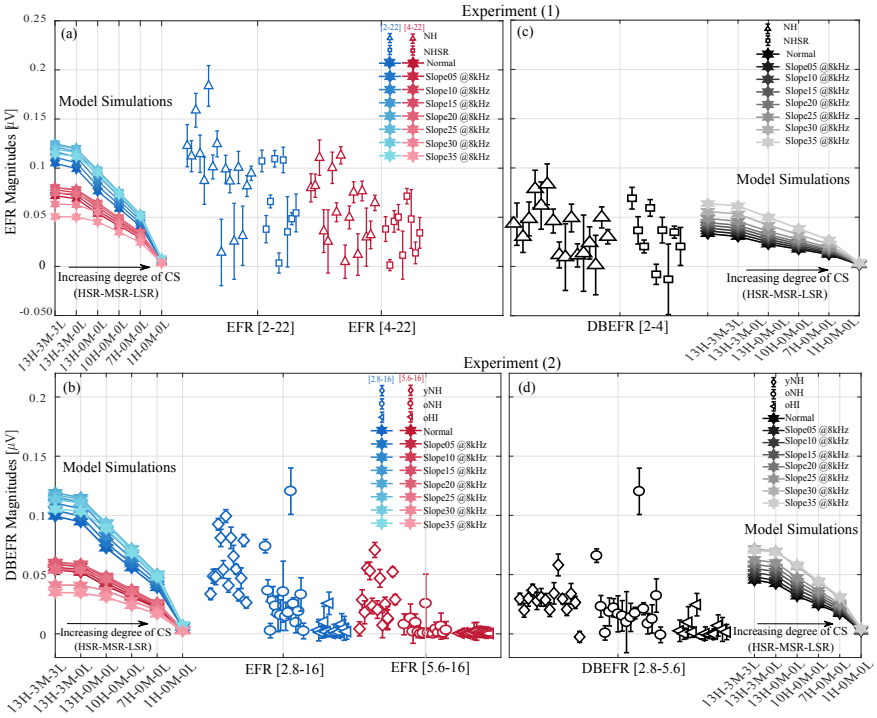


Figure 3.10: Experimental $\mathrm{EFR}_{\mathrm{PtN}}$ *and* $\mathrm{DBEFR}_{\mathrm{PtN}}$ *magnitudes (colored open symbols): (a)* $\mathrm{EFR}_{\mathrm{PtN}}$ *to [2-22] and [4-22] kHz, (b)* $\mathrm{EFR}_{\mathrm{PtN}}$ *to [2.8-16] and [5.6-16] kHz and (c)* $\mathrm{DBEFR}_{\mathrm{PtN}}$ *at [2-4] kHz and (d)* $\mathrm{DBEFR}_{\mathrm{PtN}}$ *at [2.8-5.6] kHz. Simulated* $\mathrm{EFR}_{\mathrm{PtN}}$ *(a,b) and* $\mathrm{DBEFR}_{\mathrm{PtN}}$ *(c,d) magnitudes are shown in each panel using filled hexagons and degrees of CS as indicated on the x-axis and CF-dependent patterns of OHC-damage as given by the legend.*

for all stimulus bandwidths, the narrowest bandwidth stimulus was presented at a lower overall sound level than the 70 dB-SPL broadband stimulus. Thus, applying more severe OHC-loss, lowered the AN discharge rate and envelope synchrony strength (Figure 5 in [73]) and decreased the EFR magnitudes (Figure 7 in [73]). However, DBEFR magnitudes increased monotonically for all simulated degrees of OHC-damage (Figure 3.10c and d).

3.7 Discussion

3.7.1 Tonotopic sensitivity of the EFR generators

Despite the individual variability within groups, experimental group-mean $\mathrm{EFR}_{\mathrm{PtN}}$ magnitudes to broadband stimuli with different bandwidths (Figure 3.5a), were

equal at frequencies below 4 kHz and reduced in response to [4-22] kHz condition. In the second experiment (Figure 3.5b), the EFRs remained equal at frequencies below 5.6 kHz and degraded when the [5.6-16] kHz band was added. Consequently, equal $\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitudes were obtained for frequencies below 2 kHz. Individual variability was best observed for the $\mathrm{DBEFR}_{\mathrm{PtN}}$ extracted from the [2-4] kHz (first experiment, Figure 3.6a) and [2.8-5.6] kHz (second experiment, Figure 3.6b) frequency bands. Simulated EFRs to the experimental stimuli shown with hexagons in Figure 3.5 and 3.6, confirmed observed experimental $\mathrm{EFR}_{\mathrm{PtN}}$ and $\mathrm{DBEFR}_{\mathrm{PtN}}$ frequency-dependent behaviour. In addition, the model can be used to study which CF regions along the cochlea contributed strongly to the population EFR response. To this end, we calculated the AM (Figure 3.11a) and derived-band AM (DBAM) responses at each CF (Figure 3.11b) as follows:

$$\begin{gathered}\mathrm{AM}_{\mathrm{AN}}(N_{CF}) = \frac{1}{n}\sum_{i=0}^{2}\left[2\left|\mathrm{FFT}(AN_{N_{CF}})\right|\right]_{f_i}, \\ N_{CF} = 1, 2, ..., 401,\, f_i = 120 \times (i+1)\end{gathered} \tag{3.8}$$

$$\mathrm{DBAM}_{\mathrm{AN}}(N_{CF}) = \left|\mathrm{AM}_{\mathrm{AN}}(wide) - \mathrm{AM}_{\mathrm{AN}}(narrow)\right| \tag{3.9}$$

$AN_{N_{CF}}$ is the AN-response at N_{CF} channel and $n = n_1$ as was defined in Eq. 3.1. These simulations corroborate the experimentally-observed minor contribution of low-frequency CF channels to the EFR generation.

In a previous modelling study [117], we investigated the tonotopic sensitivity of $\mathrm{EFR}_{\mathrm{PtN}}$ to broadband stimuli and ascribed the poor low-frequency AM-coding to a combination of the chosen modulation frequency (120 Hz) and the narrower bandwidth of apical cochlear filters compared to the higher CF filters [118]. Model simulations in response to the spectrally broadest condition, i.e. [0.25-22] kHz, modulated with a range of lower modulation frequencies than 120 Hz, showed that the saturation properties of AN fibers limited the modulation response at all modulation frequencies at higher CFs, despite an enhanced modulated response at the BM. This resulted in a degraded response at CFs above 4 kHz and shifted the frequency sensitivity of AM-coding to the lower CFs at low modulation frequencies. Since the brain response to modulation frequencies below 70 Hz may contain cortical as well as brainstem contribution (Chapter 10 of [93,113]), employing low modulation frequencies might render EFR-based CS diagnosis insensitive, even though an improved frequency-sensitivity can be obtained from the apical regions using these lower modulation frequencies. Therefore, the employed experimental modulation frequency, i.e. 120-Hz in combination with a broadband carrier, might be able to establish a frequency-specific CS diagnosis at frequencies above 2 kHz. In this context, the proposed DBEFR method showed a notable contribution of the [2-4] kHz CF region to the EFR generation by showing a significantly

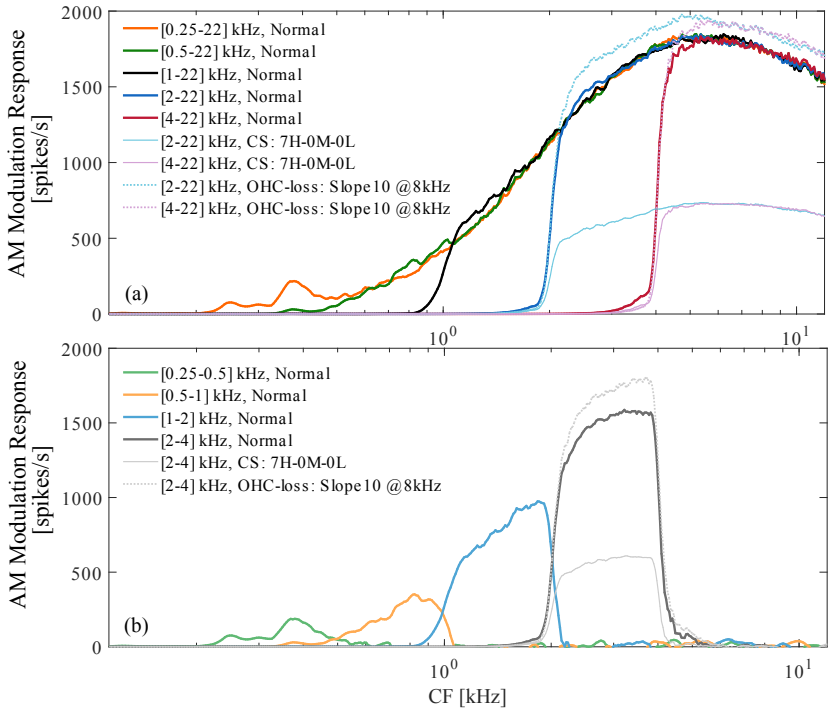


Figure 3.11: Modulated responses calculated at each CF using Eq. 3.8 and 3.9 to different experimental conditions for normal listeners and different sensorineural hearing losses at the AN processing level of the model, (a) broadband and (b) derived-band. In both panels, dotted lines show AM-responses to sloping 10 dB-HL OHC-loss at 8 kHz and lighter colors indicate AM responses to certain degree of CS.

stronger $\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitude compared to lower derived-band conditions in the NH group.

3.7.2 Diagnostic Applications

The measured DBEFR magnitudes are individually separable and above the noise-floor even for HI listeners, whose group-mean was significantly above the noise-floor. In addition, the DBEFR offers a frequency-specific metric to assess supra-threshold temporal coding of the population of AN fibers and brainstem neurons in the [2-6] kHz region. Despite these promising results, the diagnostic sensitivity of DBEFRs also has limitations. The proposed DBEFR magnitude is sensitive to CS alone, when no other coexisting hearing deficits occur and is hence applicable for use in ageing listeners with normal audiograms and those with self-reported hearing difficulties or prone to noise-exposure. According to the model simu-

lations, OHC-damage affected DBEFR magnitudes to a lesser degree than CS, which makes the metric also suitable for CS diagnosis of listeners with audiometric thresholds lower than 20 dB-HL (Figure 3.10). The metric hence needs to be complemented with another supra-threshold metric, specifically sensitive to OHC-damage to allow a separation of the CS and OHC aspect of sensorineural hearing damage from the recorded DBEFRs of listeners with impaired audiograms.

Lastly, the employed high modulation frequency, i.e. 120 Hz, suppresses cortical contributions to the $\mathrm{EFR}_{\mathrm{PtN}}$ magnitudes, but also degrades AM-coding from lower CFs and thereby limits the tonotopic sensitivity of the $\mathrm{EFR}_{\mathrm{PtN}}$ to frequencies above 2 kHz. Consequently, apical-end supra-threshold hearing deficits would not be reflected in the proposed $\mathrm{DBEFR}_{\mathrm{PtN}}$ metric even for stimuli which contain frequencies below 2 kHz. These results are consistent with the source generators of derived-band ABRs (DBABR), which reduce in amplitude for frequency bands below 2 kHz [61]. This predominant basal origin of the ABR also confines the potential of ABR/DBABR-based CS diagnosis to basal cochlear regions (e.g. wave-I amplitude).

3.8 Conclusion

We proposed the use of a relative $\mathrm{DBEFR}_{\mathrm{PtN}}$ metric to render the $\mathrm{EFR}_{\mathrm{PtN}}$ frequency-specific and rule out subject-specific factors unrelated to hearing to apply it in the study of identifying the origins of sensorineural hearing deficits and clarifying their role in supra-threshold temporal envelope-encoding. $\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitudes from two experiments were analysed and compared to model simulations to conclude that the frequency-sensitivity of $\mathrm{DBEFR}_{\mathrm{PtN}}$ magnitudes to broadband stimuli is limited to the [2-6] kHz bandwidth. Secondly, we showed that the DBEFR metric eliminates inter-subject variability caused by hearing-unrelated sources. Model simulations (Figure 3.10) explained the significant difference between yNH and oNH listeners on the basis of CS, which could result from age-induced CS as identified from human post-mortem studies [38–40]. Supported by model predictions (Figure 3.10d), the significant difference between age-matched oNH and oHI groups was explained by OHC-damage and coexisting CS as a consequence of ageing. Accordingly, profound OHC-damage may confound DBEFR-based clinical applications of CS diagnosis. Despite this limitation in the differential diagnosis of CS and OHC deficits on the basis of the DBEFR magnitude, the proposed metric can be used to diagnose CS in a frequency-specific manner in listeners with thresholds below 20 dB-HL. Moreover, it provides an objective marker of supra-threshold temporal envelope coding, which can be used to study its role in sound perception studies. Lastly, our results clearly demonstrate that older listeners with or without impaired audiograms suffer from degraded temporal envelope-coding at frequencies above 2 kHz.

4

Towards personalized auditory models: predicting individual sensorineural-hearing-loss profiles from recorded human auditory physiology

Sarineh Keshishzadeh, Markus Garrett, Sarah Verhulst

Published in *Trends in Hearing*, 2021. [1]

This chapter investigates how simulated and recorded auditory evoked potentials (AEPs) can be used to derive individual auditory-nerve or outer-hair-cell (OHC) damage patterns and personalize auditory processing models. First, the cochlear-model parameters are individualized using common methods of frequency-specific OHC-damage quantification, after which AEPs are simulated for different degrees of AN-damage. Using a classification technique, the recorded AEP metric that best predicts the simulated individualized CS profiles is determined. The proposed method is cross-validated using the dataset at hand, but also the trained classifier is applied to recorded AEPs from a new cohort to illustrate the generalisability of the method.

[1] **SK**: model simulations, analysis, conceptualization, methodology, writing the original draft, review and editing, **MG**: running the experiment, analysis, **SV**: supervision, conceptualization, review and editing, funding acquisition.

4.1 Introduction

Auditory Evoked Potentials (AEPs) are widely adopted as markers of sensorineural hearing-loss (SNHL) in clinical and research settings. In research animals, auditory brainstem response (ABR) or envelope-following response (EFR) amplitudes can be used to quantify auditory-nerve (AN) fiber damage, i.e. cochlear synaptopathy (CS), [13, 16, 31, 34]. However, applying the same AEP markers for CS diagnosis in humans has yielded mixed success, since AEP amplitudes can be affected by (i) other coexisting SNHL aspects such as outer-hair-cell (OHC) damage [42,61,64,85,94,119,120] and (ii) subject-specific factors such as age, gender, and head-size [60, 91, 121]. Moreover, the sensitivity of AEPs to different degrees of OHC-loss and CS is unclear, and a direct quantification of AN fiber damage through histopathology is impossible in live humans [87]. These problems hinder the study of the specific impact of OHC-damage and CS on recorded AEPs, and render an AEP-based quantification of AN fiber damage difficult in listeners with mixed hearing pathologies. However, this last step is crucial when developing personalized models of auditory processing for use within numerical closed-loop hearing restoration systems.

Even though several auditory models incorporate sources of SNHL (e.g. [71–77,122]), methods to individualize the AN-damage pattern on the basis of recorded AEP metrics are non-existent. Here, we investigate the potential of common AEP markers to individualize the frequency-specific AN-damage profile of personalized auditory models with or without other co-occurring aspects of SNHL. Specifically, we present a combined experimental-modelling method in which (i) individual cochlear-gain-loss (CGL) parameters are extracted from either the audiogram or distortion-product otoacoustic emissions (DPOAEs), and (ii) a feature set of recorded AEP metrics is compared to simulated AEP metrics to derive periphery models with different CS profiles. Using a classifier that was trained on simulated AEPs for different SNHL profiles, we selected the individual AN profile, that best explained the recorded AEP features from a test subject. We tested this method on 35 participants, which were separated into groups of young normal-hearing (yNH), older normal-hearing (oNH) and older hearing-impaired (oHI) listeners [123]. Validation of our method to predict individual AN-damage profile from recorded AEPs was performed on data from a new cohort.

Before we describe the classification method in detail, we summarize which AEP markers are promising to include. Among the hitherto proposed AEP-derived metrics of AN-damage, the ABR wave-I is known to degrade as a consequence of CS in subjects with intact sensory hair cells [13, 32], however this metric is highly variable in humans [35, 46] when the contribution of between-subject variability sources such as head-size or tissue resistance are not considered [50]. Even though we can assume that any hearing deficit reflecting on the ABR wave-I would

travel through the auditory pathway to reflect on the ABR wave-V as well, homeostatic gain changes between AN fibers and inferior colliculus (IC) may affect the wave-V amplitude [15,124–126], and hence its diagnostic power for CS diagnosis. Another AEP marker, the EFR amplitude, which reflects the strength of a phase-locked AEP response to an amplitude-modulated (AM) stimulus, was shown to degrade as a consequence of CS in mice histological studies [31, 32] and as a consequence of age in human listeners [44, 127]. EFRs offer a more robust measure of the AN-fiber population than the ABR wave-I, when recorded in the same animals [31, 32, 35]. However, similar to the ABR wave-V, EFR generators have latencies associated with IC processing [93], thus differences in central auditory processing may reflect on the EFR magnitude to mask individual synaptopathy differences [124, 125, 128, 129]. To address these issues, relative EFR and ABR metrics were proposed in several studies to cancel out subject-specific factors and isolate the CS component of SNHL in listeners with coexisting OHC-loss: ABR wave-I amplitude growth as a function of stimulus intensity [16], ABR wave-I and V latency difference [68–70], the wave-V and I amplitude ratio [15, 66, 67], EFR amplitude slope as a function of modulation depth [51, 65], the derived-band EFR [94], or the combined use of the ABR wave-V and EFR [83]. While these relative metrics are promising, it is not known how OHC-loss and CS differentially impact AEPs. Recent modelling approaches have shown promise to design EFR stimuli which are maximally sensitive to CS in the presence of OHC damage [44], but conclusive histopathological evidence is to date not available. To make use of the listed metrics to build personalized hearing profiles for a broad population with various SNHL etiologies, two requirements need to be fulfilled. We need to (i) use AEP markers that are maximally sensitive to the CS aspect of SNHL and (ii) combine them with a sensitive marker of OHC deficits to individualize the OHC-damage and CS aspects of SNHL. We thus considered various AEP markers (a total of 13) encompassing spectral magnitudes, time-domain peaks, latencies and relative metrics, and combinations thereof, to identify which markers best predict the simulated individualized CS profiles and can be used for reliable auditory profiling.

4.2 Experimental design

ABR, EFR and OHC-damage markers were derived from recordings of two experimental setups in different locations. These recordings were used for development and validation of the proposed method.

4.2.1 Participants

The dataset that was used to develop the auditory profiling method included recordings from a total of 43 subjects. They were recruited into three groups: 15 young normal-hearing (yNH: 24.53±2.26 years, 8 female), 16 older normal-hearing (oNH: 64.25±1.88 years, 8 female) and 12 older hearing-impaired (oHI: 65.33±1.87 years, 7 female) groups. Two oNH subjects were omitted from our study due to non-identifiable ABR waveforms. The hearing thresholds of the participants were assessed at 12 standard audiometric frequencies between 0.125 and 10 kHz (Auritec AT900, Hamburg, Germany audiometer). AEP stimuli were presented monaurally to the ear with the best 4 kHz threshold. Audiometric thresholds were below 20 dB-HL at all measured frequencies in the yNH group and below 25 dB-HL for frequencies up to 4 kHz in the oNH group. The oHI listeners had sloping high-frequency audiograms with 4-kHz thresholds above 25 dB-HL (Figure 4.1a). The AEP recordings were conducted in an electrically and acoustically shielded booth, while subjects were sitting in a comfortable chair and watching silent movies.

The second experiment, which was used to validate our method on a new cohort, had 19 yNH subjects, aged between 18 and 25 years (21.6±2.27 years, 12 female). Volunteers with a history of hearing pathology or ear surgery were excluded based on a recruitment questionnaire. Audiograms were measured in a double-wall sound-attenuating booth, using an Interacoustics Equinox Interacoustics audiometer. All participants had audiometric thresholds below 25 dB-HL within the measured frequency range, i.e. [0.125-10] kHz, and the best ear was determined on the basis of their audiogram and tympanogram. The experiment protocol included AEP measurements with a maximum duration of 3 hours and we only considered one AEP metric for validation purposes in the present study. AEP recordings were conducted in a quiet room while subjects were seated in a comfortable chair and watching muted movies. To minimize the noise intrusion level, both ears were covered with earmuffs and all electrical devices other than the measurement equipment (Intelligent Hearing Systems) were turned off and unplugged.

Participants of both experiments were informed about the experimental procedure according to the ethical guidelines at Oldenburg University (first experiment) or Ghent University Hospital (UZ-Gent, second experiment) and were paid for their participation. A written informed consent was obtained from all participants.

4.2.2 Distortion Product Otoacoustic Emission (DPOAE)

In the first experiment, DPOAEs were acquired and analyzed using a custom-made *MATLAB* software [109]. Stimuli were delivered through ER-2 earphones coupled to the ER-10B+ microphone system (Etymotic Research) using a pri-

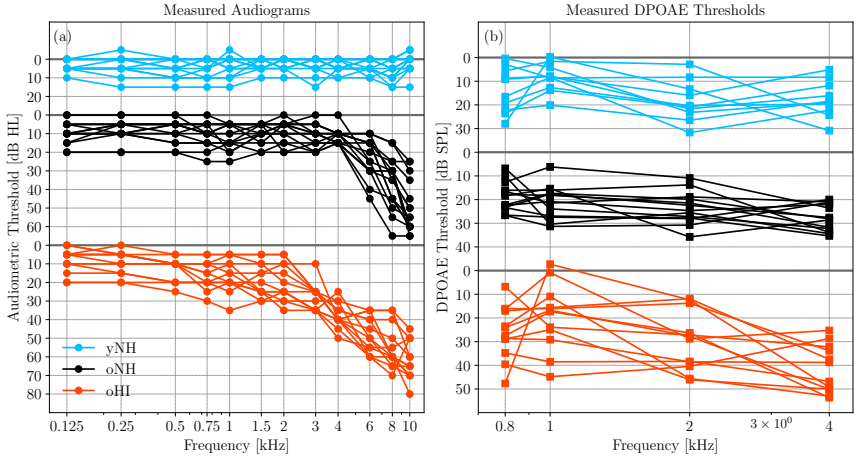


Figure 4.1: (a) Audiograms and (b) DPOAE thresholds (DPTHs) of the participants in the first experiment.

mary frequency sweeping procedure at a fixed f_2/f_1 ratio of 1.2. The implemented DPOAE paradigm, continuously swept the primary frequencies with a rate of 2s/octave within a 1/3 octave range around the geometric mean of $f_2 \in \{0.8, 1, 2, 4\}$ kHz [110]. The L_2 primary levels ranged between 30-60 dB-SPL for the yNH and oNH groups, using a 6-dB step. The level range was different for the oHI group: 30-72 dB-SPL. L_1 levels were determined according to the scissors paradigm [108]. For a given f_2 primary frequency, the DP-component ($\mathrm{L_{DP}}$) growth function was plotted as a function of L_2 and a cubic curve was fit to the $\mathrm{L_{DP}}$ data-points using a bootstrapping procedure to include the standard deviation of the individual $\mathrm{L_{DP}}$ data-points in the fit [64]. The L_2 level at which the cubic curve crossed -25 dB-SPL was determined for each bootstrap average to yield the DPOAE threshold (DPTH) and its standard deviation at a given f_2 [130]. Derived experimental DPTHs of the yNH, oNH and oHI groups are shown in Fig 4.1b. DPOAEs were not available for the subjects of the validation experiment.

4.2.3 EEG measurements

ABR and EFR stimuli were generated in *MATLAB* and were digitized with a sampling rate of 48 kHz for the first dataset. Afterwards, they were delivered monaurally through a Fireface UCX external sound-card (RME) and a TDT-HB7 headphone driver connected to a shielded ER-2 earphone. The electroencephalogram (EEG) signals were recorded with a sampling frequency of 16384 Hz via a 64-channel Biosemi EEG system using an equidistantly-spaced electrode cap. All active electrodes were placed in the cap using highly conductive gel. The

common-mode-sense (CMS) and driven-right-leg (DRL) electrodes were attached to the fronto-central midline and the tip of the nose, respectively. A comprehensive explanation of the experimental configuration can be found in [42].

AEPs of the validation experiment were recorded using the SmartEP continuous acquisition module (SEPCAM) of the Universal Smart Box (Intelligent Hearing System, Miami, FL, United States). EFR stimuli were generated in MATLAB using a sampling rate of 20-kHz and stored in a "*.wav*" format. These files were loaded in SEPCAM and converted to the "*.stm*", SEPCAM compatible format. AEP stimuli were presented monaurally through a shielded ER-2 earphone (Etymotic Research) and AEPs were recorded at a sampling frequency of 10 kHz via Ambu Neuroline 720 snap electrodes connected to vertex, nasion and both earlobes. The electrodes were placed after a skin preparation procedure using NuPrep gel. The skin-electrode impedance was kept below 3 kΩ during the recordings.

4.2.3.1 EFR stimuli

We recorded EFRs in response to a 400-ms-long stimuli consisting of a 4-kHz pure-tone carrier and a 120-Hz rectangular-wave modulator with 25% duty cycle (i.e. the RAM25 in [44]). The stimulus waveform is visualized in the inset of Figure 4.2b and we considered a modulation depth of 95%. Stimuli were presented 1000 times (500 times in either positive or negative polarity) and had a root-mean-square (RMS) of 68.18 dB-SPL. The calibration of the stimulus was performed to have the same peak-to-peak amplitude as a 70-dB-SPL sinusoidal amplitude modulated (SAM) 4-kHz pure-tone. The Cz channel recording was re-referenced to the average of the ear-lobe electrodes and 400-ms epochs were extracted relative to the stimulus onset. The mean-amplitude of each epoch was subtracted to correct for the baseline-drift. See [44] for further details on the frequency-domain bootstrapping and noise-floor estimation method. The noise-floor corrected spectral magnitudes (M_{f_k}) at the modulation frequency $\mathrm{f}_1 = 120$ Hz and four harmonics, i.e. f_2 to f_5, were summed up to yield the EFR.

$$\text{RAM-EFR} = \sum_{k=1}^{5} M_{f_k}, \quad f_k = 120 \times k \tag{4.1}$$

Figure 4.2a depicts a typical NH RAM-EFR spectrum and corresponding noise-floor. The arrows show the derived peak-to-noise-floor magnitudes at the modulation frequency and following harmonics. The energy of EFR peak is reduced for the oHI subject, shown in the panel (b).

The RAM-stimulus in the second experiment (i.e. the validation database) was a 100% rectangularly amplitude-modulated 4-kHz pure-tone with a modulation frequency of 110 Hz (25% duty cycle). The 500-ms stimulus was presented 1000 times with alternating polarity (500 each) and had a 70 dB-SPL level. The

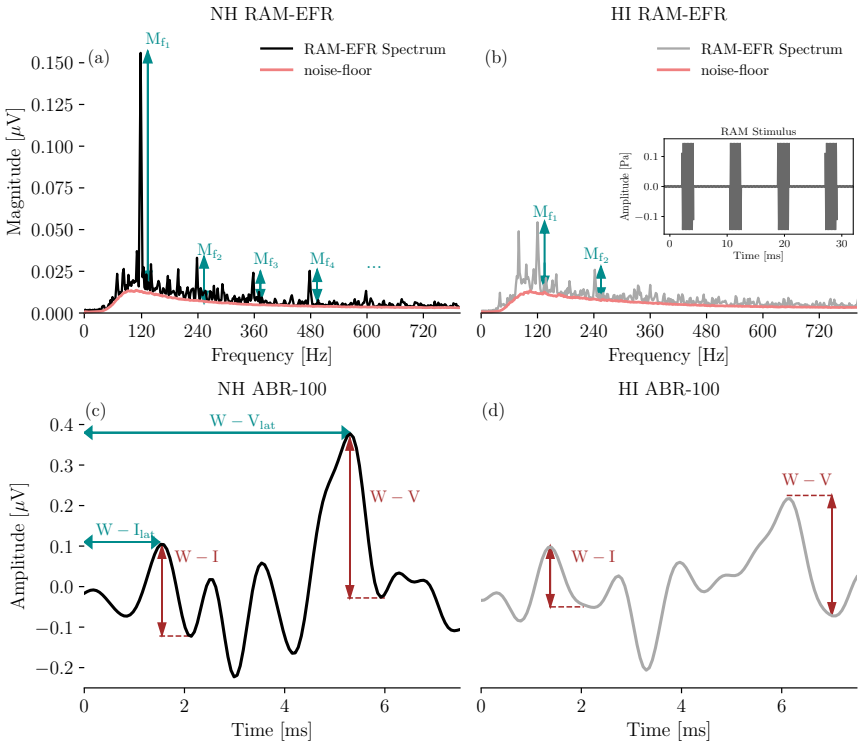


Figure 4.2: Comparison of exemplary NH and HI RAM-EFRs and ABRs. (a) RAM-EFR of a yNH subject (yNH15) and the corresponding noise-floor (NF). Arrows specified by $\mathrm{M_f}$*, show the peak-to-noisefloor magnitudes at the modulation frequency, i.e., 120 Hz, and the following harmonics. (b) RAM-EFR of an oHI subject (oHI) and the corresponding NF. (c) ABR of a yNH subject (yNH15). Arrows show the extracted wave-I and V amplitudes and latencies. (d) ABR of an oHI subject (oHI12).*

acquired AEPs were initially saved in "*.EEG.F*" format on SEPCAM and were afterwards converted to "*.mat*" format using the custom-made "*sepcam2mat*" *MATLAB* function for offline analysis. EFRs recorded from the vertex electrode were re-referenced to the ipsilateral earlobe electrode and were filtered between 30 and 1500 Hz using an 800^{th} order Blackman-window based finite-impulse-response (FIR) filter. Epoching was applied to the steady state part of the response, i.e. 100 to 500 ms of the response relative to the stimulus onset. The baseline drift was corrected by subtracting the mean of each epoch, afterwards 200 epochs with the highest peak-to-trough values were rejected. The amplitudes of the remained epochs did not exceed 100 μV. A frequency-domain bootstrapping approach was adopted to estimate the noise-floor and to remove it from the averaged trials using the method proposed in [95]. To this end, we calculated the Fast Fourier-Transform

(FFT) of 800 epochs to generate 400 mean spectra by randomly sampling the 800 epochs with replacement (keeping an equal number of polarities in the draw). Averaging the resampled spectra yielded the *i*-th mean-EFR spectrum ($\mathrm{EFR}_{\mathrm{raw_i}}$):

$$\mathrm{EFR}_{\mathrm{raw}_i} = \frac{2}{n} \left| X_i \right|, \quad i = 1, ..., 400 \tag{4.2}$$

where, X_i stands for the *i*-th averaged resampled spectra and n is the number of FFT points (n=10000). To calculate the spectral noise-floor, we repeated the resampling procedure 1500 times, but used phase-flipped odd epochs:

$$\mathrm{NF}_j = \frac{2}{n} \left| Y_j \right|, \quad j = 1, ..., 1500 \tag{4.3}$$

In Eq. 4.3, Y_j is the *j-th* averaged resampled spectra with phase-flipped odd epochs. Lastly, we subtracted the NF mean ($\overline{\mathrm{NF}}$), from each of the 400 bootstrapped mean-EFRs ($\mathrm{EFR}_{\mathrm{raw}_i}$) to derive 400 NF-corrected EFR spectra:

$$\mathrm{EFR}_{\mathrm{Spec}_i} = \mathrm{EFR}_{\mathrm{raw}_i} - \overline{\mathrm{NF}}, \quad i = 1, 2, ..., 400 \tag{4.4}$$

The peaks of the $\mathrm{EFR}_{\mathrm{raw}_i}$ at the modulation frequency of stimulus ($\mathrm{f_1}$=110 Hz), and the next three harmonics were identified if they were above the $\overline{\mathrm{NF}}$. We defined the RAM-$\mathrm{EFR_i}$ by summing the magnitudes of the identified peaks for each $\mathrm{EFR}_{\mathrm{Spec}_i}$. The RAM-EFR metric mean and variability was defined by the mean and standard deviation of RAM-$\mathrm{EFR_i}$ over 400 samples.

4.2.3.2 Auditory Brainstem Responses

ABRs were recorded to 80-μs-long alternating polarity clicks presented at 70 and 100 dB-peSPL. Stimuli were presented through the setup explained in [131] and repeated 3000 times with a rate of 10 Hz using a uniformly distributed random inter-stimulus interval of 100 ms$\pm$10 ms. Cz-channel recordings were re-referenced to the contra-lateral earlobe electrode and filtered between [100-1500] Hz. 25 ms-long epochs, i.e. -5 to 20-ms relative to the stimulus onset, were extracted and corresponding mean values were subtracted to perform a baseline correction. Then, each positive polarity epoch was averaged with the following negative epoch and 100 paired-averages with the highest peak-to-trough values were rejected. The remaining pair-averaged epochs had amplitudes below 25 μV. To include ABR variability in our analysis and to estimate the ABR noise-floor, we applied the bootstrapping approach of [95], in the time domain. 2000 and 4500 epochs were drawn for the signal and noise-floor estimation, respectively. Half of the noise-floor-estimation epochs (i.e. 2250 pair-averaged drawn epochs with replacement) were multiplied by -1, before final averaging. Finally, the estimated noise-floor mean was subtracted from the 2000 averaged epochs to yield mean

Table 4.1: Extracted AEP-metrics definitions and corresponding standard deviations. In the last column, σ represents the standard deviation. σ_{boot} is the standard deviation of the bootstrapped metric.

Metric	Symbol	Definition	Measure of Variability
rectangular-wave amplitude-modulated EFR	RAM-EFR	Eq.4.1	$\sigma_{\mathrm{boot(RAM-EFR)}}$
ABR-70 wave-I amplitude	w-I_{70}	w-$I_{70(\mathrm{peak})}$−w-$I_{70(\mathrm{trough-after})}$	$\sigma_{\mathrm{boot(peak-to-trough)}}$
ABR-100 wave-I amplitude	w-I_{100}	w-$I_{100(\mathrm{peak})}$−w-$I_{100(\mathrm{trough-after})}$	
ABR-70 wave-V amplitude	w-V_{70}	w-$V_{70(\mathrm{peak})}$−w-$V_{70(\mathrm{trough-after})}$	
ABR-100 wave-V amplitude	w-V_{100}	w-$V_{100(\mathrm{peak})}$−w-$V_{100(\mathrm{trough-after})}$	
ABR-70 wave-I latency	w-I_{lat70}	w-$I_{70(\mathrm{peak})}$latency	$\sigma_{\mathrm{boot(latency)}}$
ABR-100 wave-I latency	w-I_{lat100}	w-$I_{100(\mathrm{peak})}$latency	
ABR-70 wave-V latency	w-V_{lat70}	w-$V_{70(\mathrm{peak})}$latency	
ABR-100 wave-V latency	w-V_{lat100}	w-$V_{100(\mathrm{peak})}$latency	
ABR wave-I amplitude growth	w-I-growth	$\frac{\mathrm{w-I_{100}-w-I_{70}}}{100-70}$	$\frac{1}{\mathrm{N}}\sqrt{\sigma^2_{\mathrm{boot(w-I_{100})}}+\sigma^2_{\mathrm{boot(w-I_{70})}}}$
ABR wave-V amplitude growth	w-V-growth	$\frac{\mathrm{w-V_{100}-w-V_{70}}}{100-70}$	$\frac{1}{\mathrm{N}}\sqrt{\sigma^2_{\mathrm{boot(w-V_{100})}}+\sigma^2_{\mathrm{boot(w-V_{70})}}}$
ABR wave-I latency growth	w-I_{lat}-growth	$\frac{\lvert\mathrm{w-I_{lat100}-w-I_{lat70}}\rvert}{100-70}$	$\frac{1}{\mathrm{N}}\sqrt{\sigma^2_{\mathrm{boot(w-I_{lat100})}}+\sigma^2_{\mathrm{boot(w-I_{lat70})}}}$
ABR wave-V latency growth	w-V_{lat}-growth	$\frac{\lvert\mathrm{w-V_{lat100}-w-V_{lat70}}\rvert}{100-70}$	$\frac{1}{\mathrm{N}}\sqrt{\sigma^2_{\mathrm{boot(w-V_{lat100})}}+\sigma^2_{\mathrm{boot(w-V_{lat70})}}}$

noie-floor-corrected ABR waveforms. ABR wave-I and -V peak and trough amplitudes and corresponding latencies were determined by visual inspection from the mean ABR waveform and were confirmed by an audiologist. Figure 4.2 (panels c and d) compares ABR waveforms of a yNH and oHI subject from the cohort and indicates the identified ABR peaks and latencies. To extract peak latencies and amplitudes from the bootstrapped data, wave maxima and minima were detected in 1, 1.8, 0.5 and 1.5 ms intervals around the wave-I_{70}, wave-V_{70}, wave-I_{100}, wave-V_{100} peaks and troughs identified from the mean ABR waveform. The interval ranges were determined based on visual inspection. ABR wave-I and V latencies were shifted by 1.16 ms to compensate for the delay introduced by the sound-delivery system.

We used a total of 13 ABR and EFR markers in the development phase and one EFR marker in the validation phase. Table 4.1 details the definition of each metric and lists the corresponding abbreviations used in this paper. The last column defines the variability metric associated with each marker, which were obtained from the earlier described bootstrapping procedure. To determine the measurement variability of ABR growth-slopes, we applied error propagation to account for the standard deviations of two different measures from the same listener, e.g., ABR-70 and ABR-100. In this case, the bootstrapped metrics were drawn from the 95% confidence interval of a normal distribution characterized by the mean of the metric and its bootstrapped standard deviation. The bootstrapping technique described in this section, provided a tool to estimate the variability of AEP-derived metrics and to incorporate them in the proposed classification approach. Obtained standard deviations from bootstrapping can be used to measure the CS-profiling prediction robustness of the study participants.

4.3 Individualized auditory periphery model

To simulate individualized SNHL profiles that would match the histopathology of the study participants, we used a computational model of the auditory periphery [73, 132]. In the first step, we personalized the cochlear model parameters on the basis of OHC markers of SNHL (audiogram or DPTH). Afterwards, we simulated AEPs for different degrees of CS and compared the simulations to the recordings to develop and test our auditory profiling method. Figure 4.3 schematizes the auditory model individualization.

4.3.1 Cochlear Model Individualization

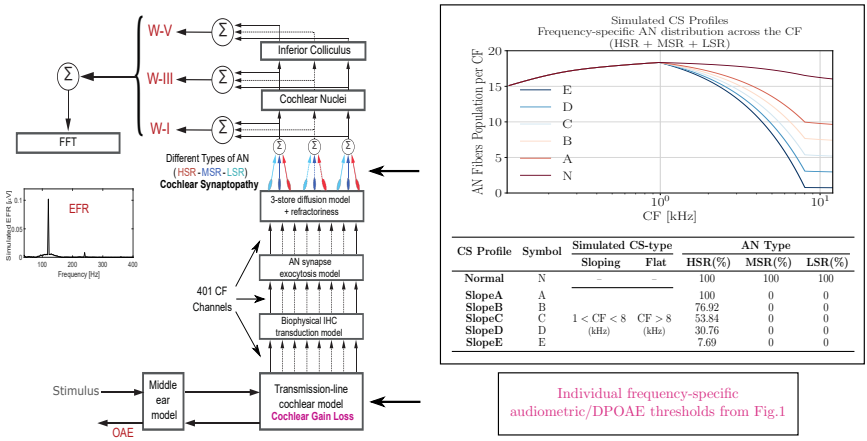
Measured audiograms and DPTHs were used independently to determine the individual CGL parameters (in dB-HL) of the cochlear transmission-line (TL) model, shown in pink in Figure 4.3. In our approach, CGL determines the double-pole of the cochlear admittance through the gain and tuning of the cochlear filters [79]. We thus model the consequence of OHC-damage or presbycusis without specifically accounting for damage of the stereocilia or sensory cells. From here on, mAudTH and sAudTH refer to measured and simulated audiometric thresholds, respectively. Likewise, mDPTH and sDPTH stand for measured and simulated DPOAE thresholds.

4.3.1.1 Audiogram-based cochlear filter pole-setting

Here, we translated the frequency-specific audiometric dB-HL (Figure 4.1a) into cochlear filter gain loss. These values were translated into double-pole values of the cochlear admittance function across CF (see [64]). Specifically, at a CF corresponding to a measured audiometric frequency ($\mathrm{CF} = f_{aud}$), the power spectrum of the NH basilar membrane (BM) impulse response, $\mathrm{H_{NH}}(f_{aud})$, served as reference before the gain loss was applied. Among a range of cochlear filter pole-values in [0.036,0.302], the pole-value, $\alpha^*_{\mathrm{A}}(f_{aud})$, that causes a relative gain-loss equal to $\mathrm{mAudTH}(f_{aud})$, was assigned. Thereby, the CGL at $\mathrm{CF} = f_{aud}$ is given by:

$$\mathrm{CGL}(f_{aud}) = \mathrm{H_{NH}}(f_{aud}) - \mathrm{H}_{\alpha^*_A}(f_{aud}) \tag{4.5}$$

where $\mathrm{H}_{\alpha^*_{\mathrm{A}}}(f_{aud})$ equals the power spectrum of the BM impulse response at $\mathrm{CF} = f_{aud}$ with a pole value of α^*_{A} that causes a CGL equal to $\mathrm{mAudTH}(f_{aud})$. This procedure was repeated for all CF channels corresponding to measured audiometric frequencies and individualized cochlear filter pole-functions were obtained by interpolating the pole-values across CF [64]. We employed the predicted pole functions to simulate individual audiograms and to evaluate the prediction error. To this end, individualized AN excitation patterns (ANEP) were simulated in response to 500-ms pure-tones presented at audiogram frequencies (f_{aud}) using 62



CS Profile	Symbol	Simulated CS-type		AN Type		
		Sloping	Flat	HSR(%)	MSR(%)	LSR(%)
Normal	N	-	-	100	100	100
SlopeA	A			100	0	0
SlopeB	B			76.92	0	0
SlopeC	C	$1 < CF < 8$ (kHz)	$CF > 8$ (kHz)	53.84	0	0
SlopeD	D			30.76	0	0
SlopeE	E			7.69	0	0

Figure 4.3: Auditory model individualization. The block-diagram on the left depicts the different stages of the employed auditory periphery model [73]. Experimentally measured audiometric thresholds were inserted to the transmission-line cochlear model to adjust BM admittance function poles. The box on top-right corner, shows the non-uniform AN population distribution across the CF for simulated different degrees of CS profiles. The profile without CS is shown in dark brown (indicated with N) and higher degrees of CS are shown according to the color-map.

intensity levels (L) between -5 and 55 dB-SPL. We defined ANEP as the RMS of the AN firing rate at each CF $\in f_{aud}$ and determined on-CF peaks of the presented level series as $\mathrm{ANEP}_{(f_{aud},L)}$. We simulated NH ANEPs using NH pole-function at the threshold of audibility in a frequency-specific manner ($\mathrm{L_{NH}}(f_{aud})$), i.e. the zero-phon curve of the equal-loudness-contour (ISO 226:1987). From this reference NH curve, we calculated the simulated audiometric thresholds (sAudTH) of the experiment participants as follows:

$$\mathrm{L_{min}}(f_{aud}) = \underset{L\in[-5,55]}{\arg\min}\,[|\mathrm{ANEP}_{(f_{aud},\mathrm{L_{NH}}(f_{aud}))} - \mathrm{ANEP}_{(f_{aud},L)}|] \tag{4.6}$$

$$\mathrm{sAudTH}(f_{aud}) = \mathrm{L_{min}}(f_{aud}) - \mathrm{L_{NH}}(f_{aud}) \tag{4.7}$$

Figure 4.4a shows grand-averaged mAudTHs and sAUDTHs across the yNH, oNH and oHI groups. Additionally, Figure 4.4c, compares the sAudTH (dashed lines) and mAudTH (solid lines) of an example yNH and oHI subject. Note that simulating CGLs greater than 35 dB-HL is impossible in our cochlear model, which has a maximal applicable cochlear mechanical gain of 35 dB. In the last step, we estimated the absolute prediction error as follows:

$$\mathrm{err_{audio}}(f_{audio}) = |\mathrm{mAudTH}(f_{aud}) - \mathrm{sAudTH}(f_{aud})| \tag{4.8}$$

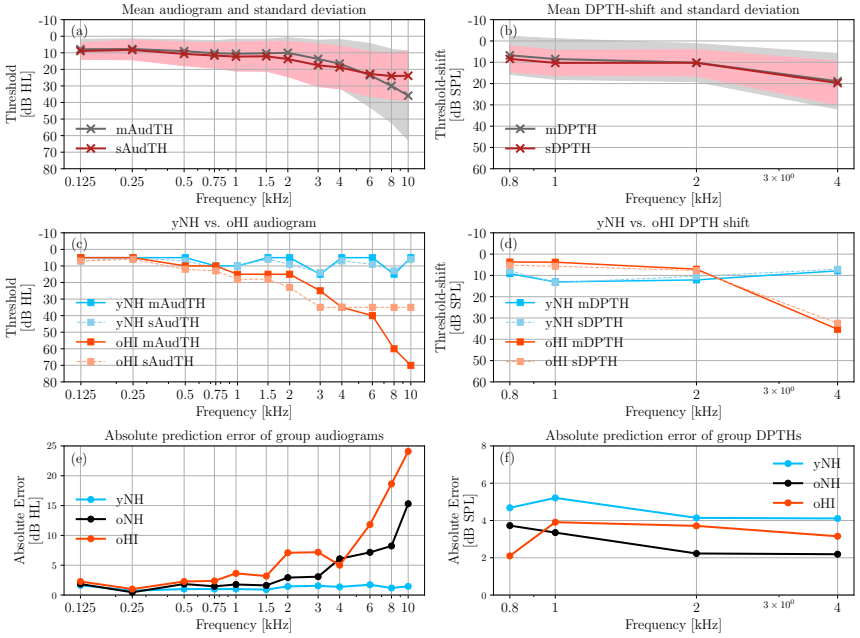


Figure 4.4: A comparison between the measured and simulated AudTHs and DPTHs. The average (solid) and standard deviation (shaded area) of the measured (grey) and simulated (red) AudTHs and DPTHs are shown in panel (a) and (b), respectively. A comparison between sAudTH and mAudTH of a yNH and oHI listener is shown in panel (c). Panel (d) compares the sDPTH (dotted) and mDPTH (solid) of the same yNH and oHI listeners. Frequency-specific group-averaged absolute prediction errors of AudTH and DPTH are shown in panels (e) and (f), respectively (yNH: blue, oNH: black, oHI: orange).

Figure 4.4e compares the mean absolute errors on a group-level basis. The elevated error of the oHI group at high frequencies is due to the model limitation in simulating gain-losses greater than 35 dB-HL.

4.3.1.2 DPTH-based cochlear filter pole-setting

Implementing DPTH-based cochlear model individualization was complicated by the fewer DPTHs we had available, i.e. four frequencies, compared to 12 AudTHs. Hence, a simple interpolation to determine poles between the measured frequencies, yielded large prediction errors. Additionally, the longitudinal filter coupling and associated gain propagation along the cochlear partition complicated matters. To tackle these issues, we trained a machine-learning algorithm to map DPTHs via cochlear travelling waves to corresponding cochlear filter pole-functions across CF. Once trained, we need only a few measured DPTHs to make a relatively accurate prediction of individual pole values. Figure 4.5 illustrates the complete

procedure.

First, we constructed the training data (Figure 4.5a) using 26 sets of random cochlear filter pole-functions. Each set contained 1001 CF-dependent poles and random pole-values lay between 0.036 and 0.302, covering the pole-values associated with both NH and HI profiles. Additionally, three reference pole-functions were included as part of the training: $\mathrm{NH_{poles}}$ (NH poles), $\mathrm{flat_{min}}$ with across-CF poles of 0.036 (maximally intact cochlea) and $\mathrm{flat_{max}}$, with across-CF poles of 0.302 (35 dB-HL across CF). We employed the generated pole-functions and simulated DP amplitudes ($\mathrm{sL_{DP}}$: the magnitude of $2f_1-f_2$) to train the mapping function. The considered f_2 primary frequencies, i.e. 0.8, 1, 2 and 4 kHz ($f_1 = f_2/1.2$) corresponded to the recordings we had available and L_2 levels (-10 to 70 dB-SPL, with a step of 5-dB). We simulated DPOAE input-output (I/O) functions at each f_2 frequency and determined the sDPTH as the L_2 level at which the $\mathrm{sL_{DP}}$ growth function crossed the $\mathrm{L_{DP}}$ of -10 dB-SPL. We chose a -10 dB-SPL threshold for our simulations, given that the conventional experimental -25 dB-SPL crossing point yielded inconclusive sDPTH, in particular for pole values associated with greater CGLs. sDPTH values for 26 sets of pole-functions at four primary frequencies were fed into the neural network after normalization ($\mathrm{sDPTH_{norm}}$, Figure 4.5b) to train it to map frequency-specific $\mathrm{sDPTH_{norm}}$ values (input) to CF-dependent pole-functions (output).

The architecture of the designed neural network is shown in Figure 4.5b, and consists of an input-layer of four neurons, two hidden-layers of 150 neurons and an output layer of 1001 neurons. A standard *sigmoid* activation function (i.e., between 0 and 1), was applied to the hidden layers. A customized *sigmoid* activation function (between 0.036 to 0.302), was employed in the output layer to yield the desired range of the cochlear model pole-functions. An ADAM optimizer with a learning rate of 0.001 was applied to minimize the mean-square-error (MSE) of the learning algorithm. The method was developed in Python using Keras library and Tensorflow back-end.

The trained neural network was employed to predict individualized pole-functions given DPTHs of the experimental cohort (Figure 4.5c). Prior to the prediction, mDPTHs needed to be pre-processed to determine a suitable experimental range of DPTHs for the mapping. Among the 41 subjects, six subjects (yNH: three, oNH: two and oHI: one) without complete mDPTH values at all measured frequencies were dropped. In each of the three recruited groups, the 99% confidence interval around the frequency-specific group-means were specified and mDPTH values that either exceeded or fell below of those intervals were set to extremum values. Then, mDPTHs were mapped to the range of the sDPTH associated with reference $\mathrm{flat_{min}}$ ($\mathrm{sDPTH_{flat_{min}}}$) and $\mathrm{flat_{max}}$ ($\mathrm{sDPTH_{flat_{max}}}$) pole-functions. Afterwards, mapped mDPTHs ($\mathrm{mDPTH_{map}}$) were normalized ($\mathrm{mDPTH_{norm}}$) and were given to the trained neural network to predict personalized pole-functions. To

assess the prediction error, the predicted pole-functions ($\mathrm{Poles}_{\mathrm{pred}}$ in Figure 4.5c), were used to simulate individualized sDPTHs that were compared to the individual mDPTHs at f_2 primary frequencies. mDPTHs and sDPTHs were referenced to the simulated DPTHs of a model with $\mathrm{NH}_{\mathrm{poles}}$ as follows:

$$\mathrm{sDPTH}_{\mathrm{ref}}(f_2) = \mathrm{sDPTH}(f_2) - \mathrm{sDPTH}_{\mathrm{NH}}(f_2) \tag{4.9}$$

$$\mathrm{mDPTH}_{\mathrm{ref}}(f_2) = \mathrm{mDPTH}_{\mathrm{map}}(f_2) - \mathrm{sDPTH}_{\mathrm{NH}}(f_2) \tag{4.10}$$

$\mathrm{sDPTH}_{\mathrm{NH}}(f_2)$ refers to the frequency-specific sDPTH values simulated using the model with $\mathrm{NH}_{\mathrm{poles}}$. Obtained $\mathrm{sDPTH}_{\mathrm{ref}}$ and $\mathrm{mDPTH}_{\mathrm{ref}}$ from Eq. 4.9 and 4.10 were mapped back to the experimental range according to Eq. 4.11 and 4.12, and corresponding grand-averages and standard deviations are shown in Figure 4.4b. More specifically, Figure 4.4d compares measured and simulated DPTH-shifts for a yNH and oHI subject.

$$\mathrm{sDPTH}_{\mathrm{shift}}(f_2) = \mathrm{sDPTH}_{\mathrm{ref}}(f_2)\frac{max[\mathrm{mDPTH}(f_2)] - min[\mathrm{mDPTH}(f_2)]}{\mathrm{sDPTH}_{\mathrm{flat}_{\mathrm{max}}}(f_2) - \mathrm{sDPTH}_{\mathrm{flat}_{\mathrm{min}}}(f_2)} \tag{4.11}$$

$$\mathrm{mDPTH}_{\mathrm{shift}}(f_2) = \mathrm{mDPTH}_{\mathrm{ref}}(f_2)\frac{max[\mathrm{mDPTH}(f_2))] - min[\mathrm{mDPTH}(f_2)]}{\mathrm{sDPTH}_{\mathrm{flat}_{\mathrm{max}}}(f_2) - \mathrm{sDPTH}_{\mathrm{flat}_{\mathrm{min}}}(f_2)} \tag{4.12}$$

Lastly, the prediction error was calculated as in Eq. 4.13. The absolute mean error for each group is shown in Figure 4.4f.

$$\mathrm{err}_{\mathrm{dpth}}(f_2) = |\mathrm{mDPTH}_{\mathrm{shift}}(f_2) - \mathrm{sDPTH}_{\mathrm{shift}}(f_2)| \tag{4.13}$$

The developed machine-learning approach can be used to personalize cochlear model parameters based on an objective measure of OHC-damage (DPTH) and predict individual CS profiles. CS-profiling can be compared for either the DPTH or AudTH-based cochlear model individualization method, and when no DPTHs are available the standard audiogram-based method can be adopted.

4.3.2 Simulating cochlear synaptopathy profiles

We employed the AudTH- and DPTH-based individualized CGL models to simulate EFRs and ABRs for different CS profiles. To introduce CS, the simulated normal-hearing AN fiber population, (the *N* CS profile in Figure 4.3) was reduced in a CF-specific manner. Five additional CS profiles were simulated by proportionally lowering the number of different AN types, starting from low- and medium-spontaneous-rate (LSR and MSR) fibers in profile *A* to the most severe AN-loss in

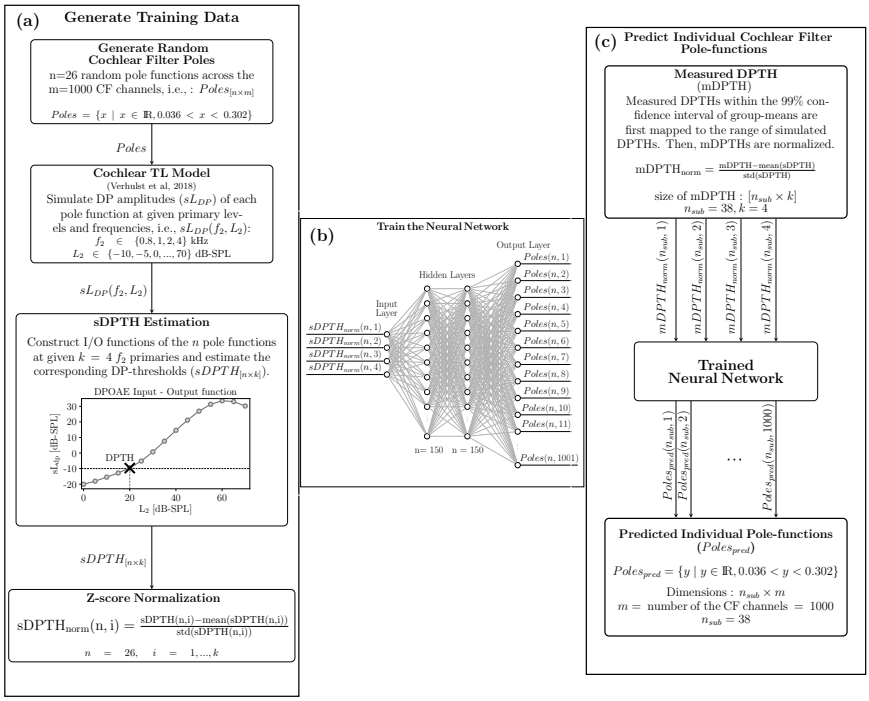


Figure 4.5: Neural network-based cochlear model individualization using measured and simulated DPTHs. (a) Random cochlear filter poles are generated and corresponding DPTHs are simulated using TL model (sDPTH). (b) The normalized sDPTH ($\mathrm{sDPTH_{norm}}$*) at four frequencies are introduced to the neural-network as input. The random pole values generated in (a) are served as training target for* $\mathrm{sDPTH_{norm}}$*. (c) Measured DPTHs (mDPTHs) are fed into the trained neural network after pre-processing and individualized cochlear filter pole-functions are predicted.*

E that only kept 7.69% of the high-spontaneous-rate (HSR) fiber population. The table in Figure 4.3 details the AN-fiber numbers and types considered for each of the six simulated CS profiles. IHC-related dysfunctions were not considered in this study, given that low degrees of CS do not cause IHC-damage [13, 16, 31]. However, removing all AN fibers from an IHC in the model would functionally correspond to IHC-damage. The CF dependence of the AN population was considered in two steps: (1) Following the CF-dependent AN distribution observed in rhesus monkey [18, 94], we applied a non-uniform NH AN-fiber population. (2) CF-specific AN-damage profiles were simulated. The former was achieved by mapping the counted CF-dependent AN fibers population in the rhesus monkey [18] to the human cochlea, using a distribution of $\mathrm{N_{HSR}} = 68\%$, $\mathrm{N_{MSR}} = 16\%$ and $\mathrm{N_{LSR}} = 16\%$ at each CF [114]. Then, sloping high-frequency AN-fiber loss was

applied across CF with the assumption that CS starts from the higher frequencies first [133]. We ran EFR/ABR simulations for different AN-fiber-damage profiles, which were characterized by a sloping loss of between 1 and 8-kHz. Above 8 kHz, we applied a frequency-independent loss.

For every subject we simulated AEPs for each CS profile, after which we personalized the cochlear models using either the AudTH-or DPTH-based method. The stimuli adopted for these simulations were identical to those adopted experimentally, but were digitized using a sampling rate of 100 kHz, rather than 48 kHz. Simulated instantaneous firing rates from the AN, cochlear nucleus (CN) and IC model stages, namely ABR wave-I, III and V, respectively, were added up to simulate EFRs (Figure 4.3). RAM-EFR magnitudes were calculated using Eq. 4.1.

To simulate ABRs, 80-μs clicks were presented to the model with a continuous sequence of 50 repetitions of alternating polarities (100 in total) and a rate of 10 Hz. Sequential stimulus presentation was adopted to account for the adaptation properties of AN-fibers. Individual ABR wave-I and V latencies and amplitudes were extracted by averaging the peak-to-trough values of the response to the last, i.e. 50^{th}, positive and negative clicks. The simulated ABR wave-I and V latencies were respectively shifted by one and three ms to match latencies of recorded ABRs. These values were determined to match the measured yNH group-mean ABR wave-I and V latencies (at 100 dB-peSPL) with the grand-average individualized ABR simulations across the yNH group. Given that simulated ABR latencies were not impacted by CS, the applied latency shifts will not confound the CS prediction.

4.4 Individual synaptopathy profile predictions

In previous sections, cochlear model parameters of the subjects were determined using either AudTH- or DPTH-based methods and 13 personalized AEP-derived metrics were simulated for six CS profiles of each experiment participant. Here, we develop a classification approach, forward-backward classification, to predict the simulated CS profile that best matches recorded individual AEP metrics and determine the AEP metric that gives the most accurate segregation of simulated individualized CS profiles. This step was implemented separately for either of the cochlear individualization methods. After excluding eight subjects from the cohort (six without complete DPTHs and two with undetectable ABRs), we developed our individual SNHL-profiling method on data from 35 subjects (yNH: 12, oNH: 12 and oHI: 11).

Before classification, we first normalized the 13 AEP metrics (Table 4.1) derived from measured (M) and simulated six CS profiles per individual (S). The normalized S and M were calculated using Eq. 4.14 and 4.15.

$$\mathrm{S_{norm}} = \frac{S - \bar{S}}{\sigma_S} \tag{4.14}$$

S is the matrix of simulated AEP metrics and contains 210 rows (35 subjects with six CS profiles) and 13 columns, the number of derived AEP metrics. $\bar{S}$ and σ_S refer to the mean and standard deviation of S, respectively.

$$\mathrm{M_{norm}} = \frac{M - \bar{S}}{\sigma_S} \tag{4.15}$$

In Eq. 4.15, M refers to the matrix of measured AEP metrics with a dimension of 35×13. We created 8191 feature-sets using all possible combinations of 13 metrics ($\sum_{i=1}^{13} \binom{13}{i} = 8191$). Metrics combination was performed separately for $\mathrm{M_{norm}}$ and $\mathrm{S_{norm}}$. The number of metrics in each feature-set varied between one and 13. From here on, **F** refers to the constructed 8191 feature-sets of AEP-derived metrics and $\mathrm{F_i}$ with $i \in \{1, ..., 13\}$ indicates a subset of **F** that has $\binom{13}{i}$ feature-sets and each feature-set contains a combination of i metrics. In the following paragraphs we explain the classification approach for an exemplary feature-set, f_e, selected from **F**. The train and test datasets required for classification were constructed by choosing f_e of all participants from $\mathrm{S_{norm}}$ and $\mathrm{M_{norm}}$ and we called them $\mathrm{S_{train}}$ and $\mathrm{M_{test}}$, respectively. The proposed *forward-backward* classification method, comprised of two identical k-nearest-neighbor (kNN: k=1, Euclidean distance) classifiers. Classifier(1) in *forward* classification was trained by $\mathrm{S_{train}}$ in six classes with known class labels from the model simulations ($\mathrm{L_S}$), i.e. the six simulated CS profiles previously described in Figure 4.3. Then, individual CS profiles were predicted by testing the trained classifier with the $\mathrm{M_{test}}$. Figure. 4.6a visualizes the different steps in *forward* classification. In this step, the evaluation of classification performance is unfeasible, since the actual CS degree of experiment participants are unknown. To address this issue, we interchanged the train-test datasets of the *forward* classification and implemented a second classification approach, called *backward* classification to assess the performance of the classifier(1) based on a second classifier (Figure 4.6b). In this regard, we took the output of *forward* classification, i.e. the predicted CS degrees of experiment participants ($\mathrm{L_M}$ in Figure 4.6), and corresponding measured AEP metrics ($\mathrm{M_{test}}$) to train the classifier(2) of Figure 4.6b. Afterwards, $\mathrm{S_{train}}$, with known CS labels ($\mathrm{L_S}$) from the simulated individualized CS profiles, was used to test the trained classifier(2). The vector of predicted CS labels by classifier(2) ($\mathrm{L_{S_{Pred}}}$) was compared to $\mathrm{L_S}$ and corresponding prediction accuracy (acc) was calculated as follows:

$$\mathrm{acc} = \frac{\sum_{q=1}^{n} [\mathrm{L_S}(q) == \mathrm{L_{S_{Pred}}}(q)]}{n} \tag{4.16}$$

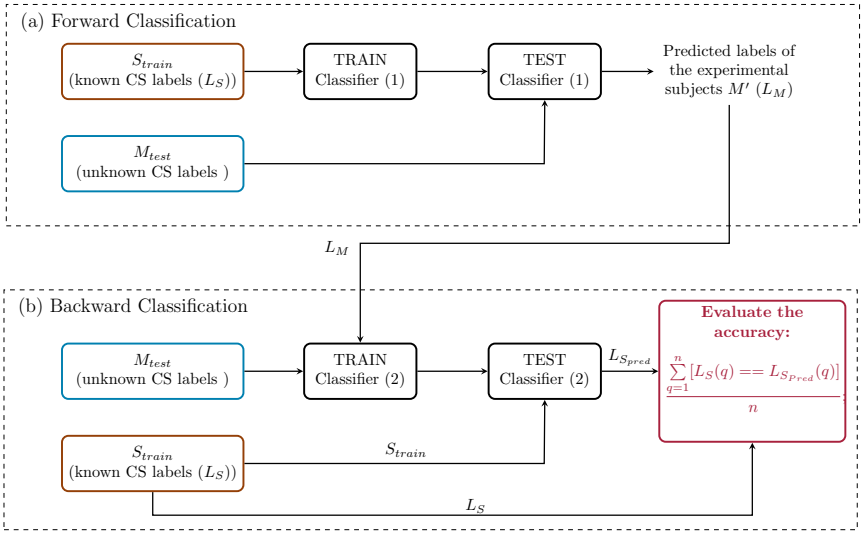


Figure 4.6: The forward-backward classification method. (a) Forward classification: Classifier(1) is trained with individualized simulated AEP-derived metrics ($\mathrm{S_{train}}$) for six CS profiles ($\mathrm{L_S}$) and is tested with measured AEP-derived metrics ($\mathrm{M_{test}}$). The predicted labels ($\mathrm{L_M}$) for the study participants are entered to block (b). The backward-classification in (b) trains classifier(2) using measured AEP-derived metrics, i.e., $\mathrm{M_{test}}$, and labels predicted by the forward classification i.e., $\mathrm{L_M}$. Classifier(2) is tested by $\mathrm{S_{train}}$ and corresponding labels ($\mathrm{L_S}$) are used to assess the classifier performance.

where n is equal to 210 (35 subjects with six CS profiles). Thus, the *backward* classification offers the possibility to calculate the accuracy of predicted CS profiles of study participants based on model simulations. We then repeated the *forward-backward* classification over all possible combinations of the derived metrics, i.e. 8191 feature-sets, and calculated the prediction accuracy of each feature-set according to Eq. 4.16. In this respect, the *backward* classification method gives the insight that to which degree classifier(1) was accurate in predicting CS degrees of experimental participants. Our classification approach makes use of combined simulated and recorded data to predict CS-profiles and can test the accuracy of these methods, even though a direct and actual validation of the CS histopathology still remains hidden due to experimental difficulties.

4.5 Results

We applied *forward-backward* classification for each of the cochlear model individualization methods (AudTH and DPTH) and calculated the prediction accuracy of all feature-sets in **F**. For each cochlear profiling method, first, we determined

the feature-set in each $\mathrm{F_i}$ ($i \in \{1, ..., 13\}$), that had the highest classification accuracy. $\mathrm{F_i}$ consisted of feature-sets with *i* AEP-derived metrics. Then, the prediction variability was estimated using *forward-backward* classification by including the standard deviations of selected feature-sets. Lastly, we report individually predicted CS profiles belonging to those feature-sets.

4.5.1 Combination of AEP-derived metrics

To determine the best combination of metrics for CS profiling, the *forward-backward* classification was performed on the mean AEP-derived metrics of experiment participants and corresponding classification accuracy was reported as $\mathrm{acc_{mean}}$. Thus, we calculated $\mathrm{acc_{mean}}$ values of the predictions for 8191 feature-sets in **F** and determined the feature-set that yielded the highest $\mathrm{acc_{mean}}$ among all feature-sets in $\mathrm{F_i}$, with i combined metrics ($i \in \{1, ..., 13\}$). Accordingly, 13 feature-sets were selected among 8191 in **F**. Table 4.2 and 4.3 list those feature-sets and corresponding $\mathrm{acc_{mean}}$ values for AudTH and DPTH-based methods, respectively. The RAM-EFR metric yielded the highest $\mathrm{acc_{mean}}$ values for both cochlear model individualization methods. The obtained 83.81% $\mathrm{acc_{mean}}$ of DPTH-based method, was higher than that of the AudTH-based method (68.57%), suggesting that methods which assess OHC-damage more directly (i.e. DPTH vs. AudTH) yield a better classification accuracy in predicting simulated individualized CS profiles.

4.5.2 Prediction variability

The impact of subject-specific factors and measurement noise reflect on inter- and within-subject variability of the AEP recordings and can have an impact on the accuracy of the classification method. To measure this effect, the *forward-backward* classification was repeated, this time by extracting metrics from the bootstrapped average trials, rather than from the mean of trials. This resulted in distributions for each specific metric and each subject, with standard deviations as given by the last column of Table 4.1. Then, 100 samples were randomly drawn from the distribution of each metric. Thus, for every feature-set in Table 4.2 and 4.3, the corresponding metrics samples were combined to yield 100 variations of each feature-set. Afterwards, the CS profile prediction was repeated 100 times with each feature-set for each subject, and prediction accuracy was assessed in every repetition. Lastly, the standard deviation of the calculated accuracies ($\mathrm{acc_{SD}}$) was determined over the 100 repetitions of each feature-set and listed in the last column of Table 4.2 and 4.3.

For the best predictor metric (RAM-EFR), $\mathrm{acc_{SD}}$ values of 2.95% and 2.66% were obtained for the AudTH- and DPTH-based methods, respectively. The lowest $\mathrm{acc_{SD}}$ was obtained when combining the RAM-EFR with the w-$\mathrm{V_{lat100}}$ metric in both cochlear model individualization methods (AudTH: 1.73% and DPTH:

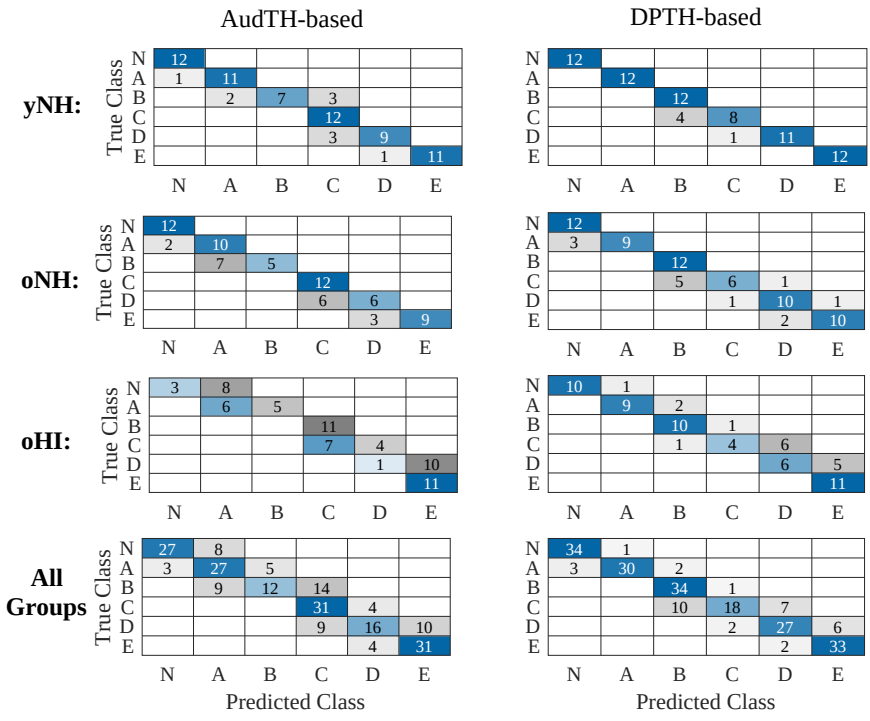
Table 4.2: Combination of metrics with the highest mean accuracy ($\mathrm{acc_{mean}}$) values in each $\mathrm{F_i}$, with i combined metrics. The standard deviation of obtained accuracies are shown in $\mathrm{acc_{sd}}$ column. The reported results are based on AudTH-based cochlear model individualization method.

Involved Metrics	Involved Subjects	Best Combination of Metrics	acc (%)	
			$\mathrm{acc_{mean}}$	$\mathrm{acc_{sd}}$
1	35	RAM-EFR	68.57	2.95
2	35	RAM-EFR, w-$\mathrm{V_{lat100}}$	64.76	1.73
3	35	RAM-EFR, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$	53.33	7.86
4	35	RAM-EFR, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$	51.90	9.28
5	35	RAM-EFR, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{70}}$	52.86	8.69
6	35	RAM-EFR, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{70}}$, w-V-growth	51.43	6.97
7	35	RAM-EFR, w-$\mathrm{V_{lat}}$-growth, w-V-growth, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{V_{100}}$	45.24	6.79
8	35	RAM-EFR, w-$\mathrm{V_{lat}}$-growth, w-V-growth, w-$\mathrm{V_{lat100}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$,w-$\mathrm{V_{100}}$	45.24	6.59
9	35	RAM-EFR, w-V-growth, w-I-growth, w-$\mathrm{I_{lat100}}$, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$	36.19	7.11
10	35	RAM-EFR, w-V-growth, w-I-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{V_{lat100}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$	32.86	6.67
11	35	RAM-EFR, w-V-growth, w-I-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{I_{lat}}$-growth, w-$\mathrm{V_{lat100}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$	27.62	6.49
12	35	RAM-EFR, w-V-growth, w-I-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{I_{lat}}$-growth, w-$\mathrm{V_{lat100}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$, w-$\mathrm{I_{lat70}}$	18.10	6.65
13	35	RAM-EFR, w-V-growth, w-I-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{I_{lat}}$-growth, w-$\mathrm{V_{lat100}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$, w-$\mathrm{I_{lat70}}$, w-$\mathrm{I_{lat100}}$	17.14	6.75

Table 4.3: Combination of metrics with the highest mean accuracy ($\mathrm{acc_{mean}}$) values in each $\mathrm{F_i}$, with i combined metrics. The standard deviation of obtained accuracies are shown in $\mathrm{acc_{sd}}$ column. The reported results are based on DPTH-based cochlear model individualization method.

Involved Metrics	Involved Subjects	Best Combination of Metrics	acc (%)	
			$\mathrm{acc_{mean}}$	$\mathrm{acc_{sd}}$
1	35	RAM-EFR	83.81	2.66
2	35	RAM-EFR, w-$\mathrm{V_{lat100}}$	58.57	1.34
3	35	RAM-EFR, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$	54.29	8.34
4	35	RAM-EFR, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$	61.90	8.22
5	35	RAM-EFR, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{70}}$	58.10	8.90
6	35	RAM-EFR, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{70}}$, w-I-growth	48.10	7.40
7	35	RAM-EFR, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{V_{100}}$, w-$\mathrm{V_{70}}$, w-V-growth, w-$\mathrm{I_{70}}$	40.95	6.96
8	35	RAM-EFR, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{V_{100}}$, w-$\mathrm{V_{70}}$, w-I-growth, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{100}}$	35.71	7.12
9	35	RAM-EFR, w-I-growth, w-V-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{I_{lat70}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$	34.29	7.30
10	35	RAM-EFR, w-I-growth, w-V-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{I_{lat70}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$	29.52	6.53
11	35	RAM-EFR, w-I-growth, w-V-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{V_{lat70}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{V_{100}}$	17.14	6.13
12	35	RAM-EFR, w-I-growth, w-V-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{I_{lat}}$-growth, w-$\mathrm{I_{lat70}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{V_{lat100}}$, w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$	16.67	2.63
13	35	RAM-EFR, w-V-growth, w-I-growth, w-$\mathrm{V_{lat}}$-growth, w-$\mathrm{I_{lat}}$-growth, w-$\mathrm{V_{lat100}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$, w-$\mathrm{I_{100}}$, w-$\mathrm{V_{100}}$, w-$\mathrm{I_{lat70}}$, w-$\mathrm{I_{lat100}}$	16.67	2.93

1.34%). However, the respective $\mathrm{acc_{mean}}$ values were considerably lower than those of the RAM-EFR by itself, particularly in DPTH-based method. To assess the performance of the RAM-EFR based CS profile prediction in sub-groups, we show confusion tables in Figure 4.7 for AudTH- and DPTH-based cochlear model individualization methods. The diagonals of each table reflect how often the classifier assigned a CS profile ($\mathrm{L_{S_{Pred}}}$: predicted class) that matched with that of in simulated individualized CS profiles ($\mathrm{L_S}$: true class). Off-diagonal values show the number of instances that $\mathrm{L_{S_{Pred}}}$ and $\mathrm{L_S}$ were not identical. Detailed prediction accuracy values of each sub-group are summarized in the tables in Figure 4.7. The highest and lowest prediction accuracy values relate to the yNH and oHI group, respectively for both AudTH- and DPTH-based methods. Comparing the cochlear model individualization methods, it is seen that the DPTH-based approach outperforms the AudTH-based method on both group- and sub-group levels.



Groups	$\sum \frac{Diagonal}{Total} \times 100$
yNH	86.11%
oNH	75.00%
oHI	42.42%
All	68.57%

Groups	$\sum \frac{Diagonal}{Total} \times 100$
yNH	93.06%
oNH	81.94%
oHI	69.44%
All	83.81%

Figure 4.7: Confusion tables at subgroup and group-levels for both AudTH and DPTH-based cochlear model individualization methods. The tables summarize the accuracy of classifier(2) in Figure 4.6b for subgroups, as well as all groups together.

4.5.3 Cochlear synaptopathy profile prediction based on individualized classifiers

Table 4.4 lists the predicted individual CS profiles from the RAM-EFR metric (best prediction accuracy) for both AudTH- and DPTH-based cochlear individualization methods. The reported profiles are the output of the *forward* classification step, i.e. L_M shown in Figure 4.6. Considering the AudTH and DPTH columns of Table 4.4, lower degrees of AN-damage were predicted for the yNH group than for the oNH and oHI groups. Additionally, the range of predicted CS profiles in the yNH group shows that yNH listeners might also suffer from different degrees of CS. The oHI

group, which was assumed to suffer from mixed OHC-damage and CS pathologies, were predicted to have the highest degree of CS among the cohort.

Thus far, the reported individualized CS profiles for RAM-EFR were predicted by training a single classifier with simulated individualized CS profiles of the whole experimental cohort. This has drawbacks for individual profiling in a clinical context, because it would be ideal if the profiling could be performed using only recordings from the tested individual. Hence, to establish more accurate predictions of the individual CS degrees, we took one step further and designed individualized classifiers, which were trained and tested with the RAM-EFR metric of the same listener. If $\mathrm{RAM_s}$ stands for the six simulated CS profiles of a nominal subject and $\mathrm{RAM_m}$ for the measured RAM-EFR metric, we first normalized $\mathrm{RAM_s}$ and $\mathrm{RAM_m}$ values by the $\overline{\mathrm{RAM_s}}$ and $\sigma_{\mathrm{RAM_s}}$ (mean and standard deviation of $\mathrm{RAM_s}$). Then we trained and tested the classifier, with the same characteristics as classifier(1) and (2), using normalized $\mathrm{RAM_s}$ and $\mathrm{RAM_m}$ values, respectively. This procedure was repeated for all listeners in the cohort and for both AudTH and DPTH-based cochlear model pole-setting methods. The predicted individualized CS profiles were listed in Table 4.4 (columns: $\mathrm{AudTH_{ind}}$ and $\mathrm{DPTH_{ind}}$). Considering either of the AudTH- or DPTH-based methods, designing individualized classifiers revealed only minor differences in the predicted CS profiles of individual listeners compared to those predicted by a single classifier trained with simulated individualized RAM-EFRs. However, the CS profiles reported in $\mathrm{AudTH_{ind}}$ and $\mathrm{DPTH_{ind}}$ columns might be more reliable than the group-based predictions, since the former were predicted by individualized classifiers that were trained on the basis of personalized cochlear simulations.

To provide a demonstration of the implemented method, and to show to which extent the model simulations imitate the experimental measurements, we compare simulated and measured AEPs of a yNH subject in Figure 4.8. Panel (a) depicts simulated RAM-EFR spectra for the different considered CS profiles. Based on the experimental RAM-EFR (panel (d)) and *forward* classification, we predicted that this subject had a *"N"* CS profile, i.e., no AN-damage. The CGL parameters of the individualized model were adjusted based on DPTHs of the same yNH listener. Panels (b) and (c), depict the simulated personalized ABR waveforms for the predicted *"N"* CS degree. Experimental ABR waveforms to 70 and 100 dB-peSPL clicks are shown in panels (e) and (f), respectively. Details regarding the value of extracted metrics from the measurements and simulations are provided in a table at the bottom of the Figure 4.8. Even though our classifier did not consider ABR metrics, the applied personalized OHC and AN profiles predicted w-$\mathrm{I_{lat100}}$, w-$\mathrm{I_{70}}$, w-$\mathrm{V_{70}}$ and w-$\mathrm{I_{100}}$ markers that fell within the standard deviation of the corresponding recorded values. The remaining simulated ABR metrics, i.e. w-$\mathrm{I_{lat70}}$, w-$\mathrm{V_{lat70}}$, w-$\mathrm{V_{lat100}}$ and w-$\mathrm{V_{100}}$, only minimally deviated the range of respective measurements, showing that our method accurately predicts AEP features to

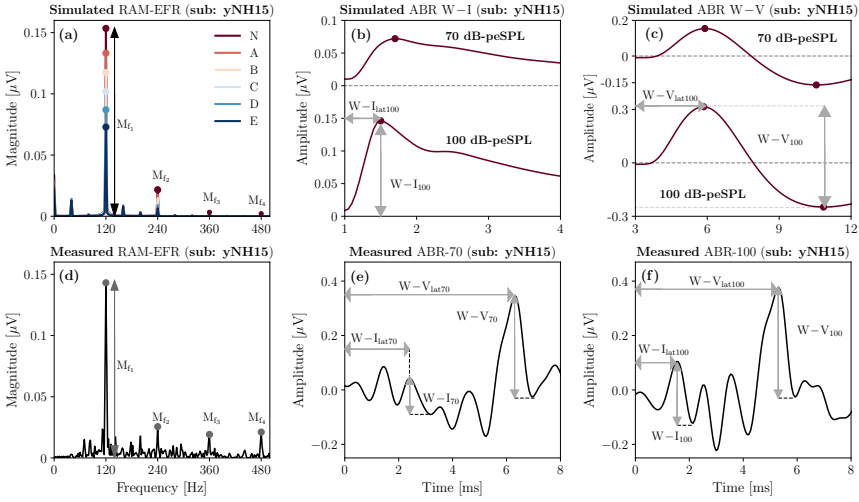
Table 4.4: Predicted individuals CS profiles obtained from AudTH and DPTH based cochlear individualization methods, based on RAM-EFR metric. Columns $\mathrm{AudTH_{ind}}$ *and* $\mathrm{DPTH_{ind}}$*, list the predicted CS profiles by designing individualized classifiers based on RAM-EFR metric.*

Group	**No.**	**AudTH**	$\mathrm{AudTH_{ind}}$	**DPTH**	$\mathrm{DPTH_{ind}}$
yNH	1	C	B	B	B
	2	A	A	A	B
	5	N	N	N	N
	7	N	N	N	N
	8	N	N	N	N
	9	N	N	N	N
	10	N	N	N	A
	11	A	B	B	B
	12	N	N	N	A
	13	A	A	A	A
	14	N	N	N	N
	15	N	N	N	N
oNH	1	D	D	C	D
	3	E	E	E	E
	4	D	E	D	D
	6	D	D	D	D
	7	C	D	D	D
	8	E	E	E	E
	9	N	A	N	A
	10	B	B	B	B
	11	C	D	D	D
	12	N	N	N	N
	13	E	E	E	E
	14	C	D	C	C
oHI	1	E	E	E	E
	2	E	D	E	D
	3	E	E	E	E
	4	E	E	E	E
	5	E	D	E	E
	7	E	D	E	E
	8	E	E	E	E
	9	E	E	E	E
	10	E	E	E	E
	11	E	E	E	E
	12	E	E	E	E

stimuli which were not included in the classifier.

4.5.4 Method validation

To validate the proposed method and its generalizability to other cohorts and other measurement equipment, we applied the developed classifier in *backward* classification step to RAM-EFRs recorded in a second experiment. Figure 4.9 schematizes the implementation of the validation method. Considering the different experimental setup and recording location of the second experiment, the measured RAM-EFRs of both experiments were scaled between zero and one, prior to classification. Given that only yNH listeners participated in the second experiment,



Metric Type	RAM − EFR [μV]	w − I_{lat70} [ms]	w − V_{lat70} [ms]	w − I_{lat100} [ms]	w − V_{lat100} [ms]	w − I_{70} [μV]	w − V_{70} [μV]	w − I_{100} [μV]	w − V_{100} [μV]
Measured value	0.1838±0.007	2.3800±0.241	6.347±0.128	1.5255±0.111	5.3097±0.123	0.1378±0.135	0.3747±0.158	0.2259±0.128	0.4036±0.132
Simulated individualized value	0.1896	1.7500	5.9500	1.5500	5.7500	0.0727	0.3127	0.1469	0.5895

Figure 4.8: A comparison between simulated and measured AEPs for a yNH subject (yNH15). This subject was predicted to have a normal (N) CS profile, i.e. without CS. (a) Simulated RAM-EFR spectra for six CS profiles. The sum of the drawn arrows yields the RAM-EFR magnitude metric. (b) Simulated ABR wave-I to 70 and 100 dB-peSPL clicks. Waveforms were shifted by 1 ms to match the experimental data. (c) Simulated ABR wave-V to 70 and 100 dB-peSPL clicks. Waveforms were shifted by 3ms to match the experimental data. The specified arrows in (b) and (c) indicate the extracted metrics. (d) Measured RAM-EFR of the same listener (yNH15). Shown arrows, indicate the peak-to-noisefloor values. Akin to (a), the measured RAM-EFR metric was calculated by summing the of arrow amplitudes. (e) Measured ABR waveform to 70 dB-peSPL clicks. (f) Measured ABR waveform to 100 dB-peSPL clicks. Arrows in (e) and (f), determine the extracted metrics. The shown simulated waveforms were predicted based on the DPTH-based cochlear individualization method. The table shows the exact value of EFR and ABR metrics derived from recordings and predicted CS-profile, "N"*, of the same listener.*

we employed the smallest RAM-EFR magnitude recorded from oHI listeners (as part of another study) recorded with the same setup as the second experiment to scale the RAM-EFRs. The scaled RAM-EFRs of the first experiment were used to train the classifier(1) in Figure 4.6 and afterwards, the trained classifier was tested with the scaled RAM-EFRs of the second experiment. The predicted CS profiles are illustrated as a bar-plot in Figure 4.9. 84.21% of the 19 yNH participants of the second experiment were classified as *N*, i.e. without CS, and the rest were predicted to have mild CS. These predictions show that a classifier designed on our cohort can be applied to other cohorts to predict individual CS degrees based on

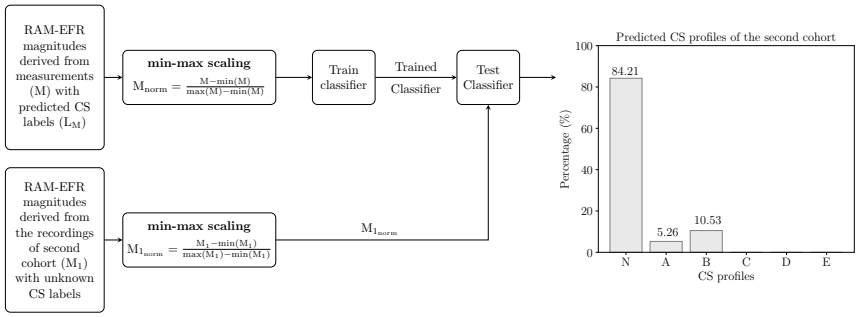


Figure 4.9: Implementation of the validation method. Measured RAM-EFRs (M) with predicted labels in Figure 4.6 (L_M) are scaled between zero and one to train a kNN classifier. The trained classifier is tested with scaled RAM-EFRs recorded from the second cohort comprised of yNH listeners. The bar-plot shows the predicted CS profiles for the second cohort listeners. The CS profiles labels in the bar-plot are similar to those defined in Figure 4.3.

the RAM-EFR. In line with expectations, the classifier predicted that most yNH subjects were synaptopathy free.

4.6 Discussion

By combining experimental ABR and EFR measurements with a modelling approach, we were able to develop a classifier that can assign one out of six CS profiles to listeners with mixed SNHL pathologies. The classifier considered 8191 feature-sets, of which our *forward-backward* classification method identified that the RAM-EFR metric yielded the best performance in both AudTH- and DPTH-based cochlear individualization methods. We tested both a group and individually based method and showed that our method can generalize to other cohorts and measurement setups. Taken together, we have high hopes that this method can find its way to clinical hearing diagnostics, since a single AEP metric is required to yield a CS-profile prediction, given the audiogram or at least four DPTHs.

4.6.1 Implications for RAM-EFR-based synaptopathy profiling prediction

On the one hand, predicting the CS-degree from AEP metrics is controversial in listeners with coexisting OHC deficits and on the other, validation of the predicted CS profiles with temporal bone histology is impossible in humans. Without these means, models of the human auditory periphery an AEP generators can provide a tool to bridge this experimental gap. The similarity between predicted AEP degradations for a known CS profile and experimental AEP degradations can be used

to predict the CS profile of individuals. In a previous study, we tested the potential of the derived-band EFR as a CS predictor in NH listeners using a fuzzy c-means clustering method, and validated our CS predictions using an another AEP-derived metric (wave-V amplitude growth-slope) recorded from the same listener. We evaluated the method based on the percent of subjects that were predicted and validated to have the same CS profile, i.e. 61% [134]. However, the performance of this method is easily impacted by the characteristics of the adopted predictor and validation metrics, e.g. different generator sources, degree of sensitivity to subtypes of SNHL and tonotopic susceptibility.

The interdisciplinary approach we took in this study tackled this validation issue by proposing a *forward-backward* classification approach and applying the trained classifier to AEPs from a new cohort to test its generalizability. Moreover, we were able to determine the most accurate AEP-derived metric for CS-degree prediction, given a range of 13 possible AEP-derived metrics. Among the considered AEP-based metrics and combinations thereof, we found that the RAM-EFR magnitude showed the best performance in segregating simulated individual CS profiles. At the same time, RAM-EFR metric was involved in all feature-sets that yielded the highest $\mathrm{acc}_{\mathrm{mean}}$ among feature-sets that had equal number of combined metrics (Table 4.2 and 4.3). This finding is consistent with the outcome of [32] and [44], showing that EFRs to SAM or RAM are sensitive to CS. Moreover, the combined modelling and experimental study of [44] showed that the adopted RAM-EFR marker (RAM with a 25% duty-cycle), is minimally impacted by OHC-damage. The sharp envelope combined with the long silence intervals between stimulus peaks generates more synchronized AN-fiber responses compared to conventional SAM stimulus to yield a stronger EFR with extended dynamic range across subjects. Lastly, the RAM-EFR is a more sensitive marker of CS than ABR [32]. Taken together, our results indicate that the RAM-EFR magnitude is an appropriate AEP-based metric to predict individual CS-degree of listeners in the presence of OHC-loss.

4.6.2 The effect of cochlear model individualization method on predicting cochlear synaptopathy profiles

In this study, we determined the CGL model parameters using either measured audiometric or DPOAE thresholds, and assessed the classifier performance of each method in the *backward* classification step. Comparing the resulting $\mathrm{acc}_{\mathrm{mean}}$ values for each cochlear individualization method can inform which of the two methods yielded the most accurate AEP simulations for a given CS profile. The $\mathrm{acc}_{\mathrm{mean}}$ values of RAM-EFR metric showed that setting cochlear filter pole-functions on the basis of measured DPTHs outperforms the AudTH method for all experimental groups (Figure 4.7, Tables 4.2 and 4.3). This outcome is consistent with literature

studies showing that OAEs are a more sensitive measure of noise-induced cochlear dysfunction in humans [135–138]. Moreover, OAEs are not influenced by IHC/AN damage [139], whereas behaviourally measured audiometric thresholds, particularly extended high-frequency thresholds, could be affected by extreme neural degeneration [37, 140, 141]. Consequently, given the varied susceptibility of AudTHs and DPOAEs to different aspects of SNHL, it was expected that we would obtain non-identical predictions of CS profiles for a nominal subject (Table 4.4). Comparing the AudTH and DPTH columns in Table 4.4, we found a mismatch between individually predicted CS profiles for 14.28% of subjects (yNH: two, oNH: three). The mismatch degree increased to 20% (yNH: three, oNH: two and oHI: two) when the individual CS profiles were predicted using personalized classifiers ($\mathrm{AudTH_{ind}}$ and $\mathrm{DPTH_{ind}}$ columns).

It is noteworthy that the DPTH-based cochlear individualization was implemented using DPTHs from only four frequencies (0.8 to 4 kHz), whereas the AudTH-based method considered audiometric thresholds measured at 12 frequencies (0.125 to 10 kHz). This difference may have resulted in less accurate CGL model parameters for the DPTH-based method, despite a better performance of *forward-backward* classification. In future implementations of this method, we intend to incorporate more frequencies in the DPTH measurements, especially at higher frequency regions. Employing DPgrams instead of DPTHs is another option, as these require a shorter measurement time. In both cases, we suggest to include lower stimulus levels as well, given that noise-induced OHC deficits can be identified earlier at lower stimulus levels [37].

4.6.3 Method limitations

The proposed method for AEP-based CS-profiling, relies on the interactive use of recordings and model simulations. Hence, shortcomings in either aspect could have caused performance limitations of the method. The following sections summarize a number of these limitations:

4.6.3.1 Experimental limitations

(1) ABRs in humans are recorded using vertex electrodes placed on the scalp, which yields smaller and more variable wave-I amplitudes than when they are recorded in animals using subdermal electrodes. The measured ABR w-$\mathrm{I_{70}}$ amplitude in our measurement produced a mean standard deviation of 0.198 μV across the cohort, which is fairly large with respect to the mean amplitude of 0.146 μV (yNH: 0.1964$\pm$0.1436 μV, oNH: 0.1304$\pm$0.203 μV, oHI: 0.1071$\pm$0.243 μV). Compared to w-$\mathrm{I_{70}}$, w-$\mathrm{I_{100}}$ amplitudes showed less variability, i.e. 0.2503$\pm$0.2056 μV. Variability of the w-$\mathrm{I_{100}}$ was considerably lower only for yNH group (0.350$\pm$ 0.143 μV). Per subgroup, variability increased for older groups (oNH: 0.205$\pm$

0.247 μV, oHI: 0.180$\pm$0.235 μV). Given these variabilities, adding the w-I_{100} metric to the second feature-set (RAM-EFR, w-$V_{\mathrm{lat}100}$), suddenly increased the $\mathrm{acc_{SD}}$ (Tables 4.2 and 4.3). (2) Although adopting relative ABR metrics, such as growth-functions might factor out individual differences, the standard deviation of the derived relative metric is influenced by the propagated error of the absolute metric. (3) ABRs to clicks presented at 100 dB-peSPL should yield higher wave-I and V amplitudes, than when the stimulus was presented at 80 dB-peSPL. Nevertheless, the opposite was observed in a few subjects.

4.6.3.2 Model limitations

(1) The adopted computational model of the auditory periphery allows for OHC deficit simulation on a CF-dependent basis, but not for CGLs above 35 dB, since the maximum possible BM filter gain is 35 dB in the model [73]. This constraint led to elevated absolute prediction errors for high-frequency audiometric thresholds in the oNH and oHI (above 4 kHz) groups (Figure 4.4e). The increased absolute errors were mainly observed for the audiometric threshold predictions, since DPTHs were only measured for frequencies up to 4 kHz. Thus, the individualized hard-coded OHC-loss component for the oHI group might lead to similar and less accurate CS profile prediction for oHI participants with audiometric losses greater than 35 dB-HL.
(2) In the adopted method, we hard-coded the CGL using the individual hearing thresholds and related the remaining AEP alterations to CS. An alternative way would be to run the model iteratively and simultaneously optimize both CGL and CS profile parameters on the basis of the experimental data to obtain the best OHC-loss and CS profiles. However, we did not further explore this route due to the high computational cost of running the adopted TL cochlear model in an iterative optimization procedure.

4.7 Conclusion

In this study, we proposed an integrated modelling and experimental approach to build personalized auditory models and predict the AN-damage profile of listeners with mixed SNHL profiles. To develop individualized cochlear models, we implemented two different methods on the basis of measured AudTHs and DPTHs. Next, we developed a classification-based approach to predict individual CS profiles and determined which AEP metric (or combinations thereof) yielded the highest prediction accuracy. Afterwards, we evaluated the implemented CGL and CS-profile individualization methods on the development dataset, as well as on a new cohort. Our study suggests that a DPTH-based cochlear model individualization approach combined with a RAM-EFR recording predicts individual CS

profiles most accurately among the 8191 possible combinations of 13 AEP markers. Additionally, we tested the applicability of the proposed method by applying the trained classifier to the recorded RAM-EFRs of a new cohort of yNH listeners. The classifier predicted that these listeners mostly had mild forms of CS, which supports that our method is generalizable to other recording setups and cohorts. Training the classifier again on larger cohorts may further increase the generalizability of the method. We hope that this method, or variations thereof, can be used in a clinical diagnostic context, as the number of needed AEP recordings to yield an individual CS-profile is small (i.e. 10-15 minutes). Individualized models of SNHL are an important step for the development of hearing aid algorithms that compensate for both the OHC- and AN-damage aspects of SNHL.

5

A comparison among DPOAE-based cochlear model individualization methods

Sarineh Keshishzadeh, Markus Garrett, Sarah Verhulst, Manfred Mauermann

The content of this chapter is based on the recent results and has not been published.[1]

This chapter adopts the machine-learning approach presented in Chapter 4 to examine the applicability of the distortion product otoacoustic emissions (DPOAEs) for developing individualized cochlear models. Here, we test whether individualized cochlear models based on DPOAE measurements (DPOAE thresholds and DP-grams) can predict the measured DPOAEs and audiograms of young and older normal-hearing, as well as hearing-impaired listeners. The outcome of this study shows that cochlear models individualized based on DP-grams measured at low stimulus levels or DPOAE thresholds, yield the smallest prediction errors. These promising results support the use of the developed DPOAE-based model individualization method in the future hearing-aid fitting algorithms.

[1] **SK**: model simulations, analysis, conceptualization, methodology, writing the original draft, review and editing, **MG**: running the experiment, **SV**: supervision, conceptualization, review and editing, funding acquisition, **MM**: analysis.

5.1 Introduction

Individualized models of hearing-impaired auditory processing have been widely adopted in the design of hearing-aid algorithms. The hitherto developed hearing-aid fitting procedures are largely based on audiometric thresholds or psychoacoustic metrics (e.g. speech intelligibility and loudness perception) [90, 142–145], and do not account for the cochlear synaptopathy (CS) aspect of the sensorineural hearing-loss (SNHL). However, determining individual CS parameters is controversial, as diagnostic metrics of CS are presently based on auditory evoked potentials (AEPs). These AEP measures do not necessarily yield a frequency-specific quantification of CS and might be affected by both outer-hair-cell (OHC) and auditory nerve (AN) damage [13, 32, 42, 44, 94]. Hence, developing individualized cochlear models, which precisely simulate how OHC-damage affects cochlear-gain-loss (CGL) is a key requirement for making progress towards AEP-based CS quantification. Specifically, incorporating such individualized cochlear models within the biophysically-inspired auditory model framework that account for both OHC-loss and AN-damage aspects of SNHL, will enable us to simulate how either of the aspects affects AEP markers and conversely will enable to use recorded AEPs to derive frequency-specific personalized SNHL profiles. These profiles can be incorporated within the personalized auditory models that can be used as a basis to develop personalized SNHL profiles and use them as a basis for hearing restoration algorithms.

Individual CGL parameters can be derived from audiograms measured within the standard audiometric frequencies or at extended high-frequency regions, as well as distortion product otoacoustic emissions (DPOAEs). While the audiogram reflects a behavioural response, DPOAE yields an objective metric that is a byproduct of cochlear amplification and its amplitude is informative of OHC-damage [146, 147]. In Chapter 4, the applicability of audiograms and DPOAE thresholds (DPTHs) as candidate metrics for developing individualized SNHL profiles were assessed based on a classification approach. The corresponding DPTH-based individualized cochlear models yielded a better accuracy in predicting CS profiles of a cohort consisted of normal-hearing and hearing-impaired young and older listeners [43]. However, an accurate estimation of the DPTH at a given frequency, requires DPOAE amplitudes measured over a range of primary levels, which is time-consuming. In addition, achieving an appropriate curve-fit that leads to a robust estimation of the DPTHs is challenging.

In this chapter, the use of the DP-gram, as an alternative metric for DPTH is evaluated in developing individualized cochlear models. DP-gram is characterized by the amplitude of DPOAEs measured across a range of primary frequencies of fixed level and provides a frequency-dependent objective measure of OHC-damage within a more reasonable acquisition time than DPTHs. Therefore, this study aims

to determine which of the individualized cochlear models (based on either DPTHs or DP-grams), best predicts the corresponding metrics from the measurements, and also estimates the behavioural audiograms of the study participants with an acceptable error. In addition, we explore whether the primary level, at which DP-grams are measured, plays a role in developing accurate individualized cochlear models.

5.2 Experiment participants

The experimental cohort adopted in this study consists of 12 young normal-hearing (yNH: 25.08±1.93), 12 older normal-hearing (oNH: 64.58 ±1.88) and 11 older hearing-impaired (oHI: 65.27±1.95) listeners. Audiograms of the participants were measured at 12 standard audiometric frequencies, between 0.125 and 10 kHz, and the ear with the lower threshold at 4 kHz was chosen for DPOAE measurements. The recruitment criterion of the experiment participants was based on audiometric thresholds. Audiometric threshold of the yNH group did not exceed 20 dB-HL at all measured frequencies. The oNH group had audiometric thresholds below 25 dB-HL at frequencies below 4 kHz and oHI listeners were suffering from sloping high-frequency hearing-loss with 4-kHz thresholds greater than 25 dB-HL. Audiograms of the three experimental groups are shown in the first column of Figure 5.1 (a: yNH, b: oNH, c: oHI). Further details regarding the experimental design can be found in Section 4.2.1 of Chapter 4.

5.3 DPOAE threshold and DP-gram extraction

To record DPOAEs, two pure-tones with f_1 and f_2 primary frequencies (using a fixed ratio of f_2/f_1=1.2) were simultaneously presented to the ear through ER-2 earphones coupled to the ER-10B+ microphone system (Etymotic Research). During the stimulation, the primary f_2 frequency was swept continuously [110] with a rate of 2s/octave over a 1/3 octave range around the determined center frequencies, i.e. 0.8, 1, 2 and 4 kHz [64]. The L_2 primary level varied between 30 and 60 (for yNH and oNH listeners) or 72 (for oHI listeners) dB-SPL, with a 6-dB step, and L_1 primary level was determined based on the *scissors* paradigm, $L_1 = 0.4L_2 + 39$ [108]. For a given primary frequency and level, the distortion product (DP) amplitude was calculated at $2f_1 - f_2$ frequency (L_{DP}). To extract DPTH at a given f_2, a cubic function was fit to L_{DP} values measured at different L_2 levels (i.e. DPOAE input-output (I/O) function at f_2), and the L_2 level at which the curve reaches to L_{DP} of -25 dB SPL, was estimated as the DPTH at f_2 [64, 130]. Figure 5.2 shows an exemplar DPOAE I/O function measured at $f_2 = 0.8$ kHz and illustrates the implemented method for extracting the

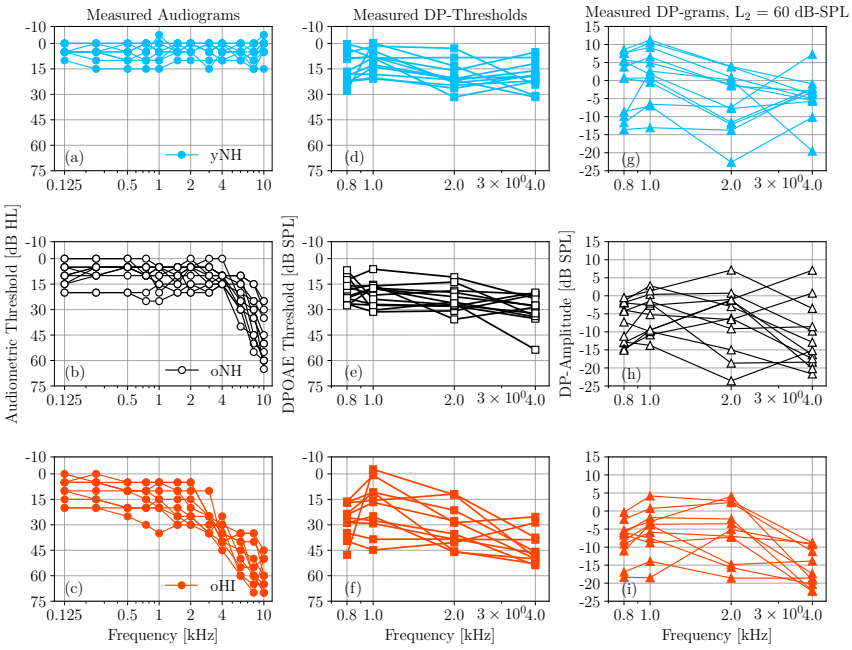


Figure 5.1: Measured audiograms (first column: (a) yNH, (b) oNH, (c) oHI), DPOAE thresholds (DPTHs: second column, (d) yNH, (e) oNH, (f) oHI) and DP-grams (third column: (g) yNH, (h) oNH, (i) oHI) of the experiment participants.

corresponding DPOAE threshold. This procedure was repeated for each f_2, and corresponding DPTHs were plotted as a function of f_2 (the second column of Figure 5.1).

DP-grams are L_{DP}s as a function of frequency (f_2) for a given L_2. For each subject DP-grams were measured at six primary L_2 levels. Measured DPgrams at $L_2 = 60$ dB-SPL are shown in the third column of Figure 5.1. The DPOAE recording paradigms and analysis were performed using a custom-made *MATLAB* program [109].

5.4 Estimating individualized cochlear-gain-loss parameters

Cochlear model individualization was implemented by predicting CF-dependent CGL parameters, based on recorded DPOAEs of the study participants. To simulate CGL caused by OHC-damage, poles of the BM admittance function of the transmission-line (TL) model [73, 79] were estimated using recorded DPTHs and

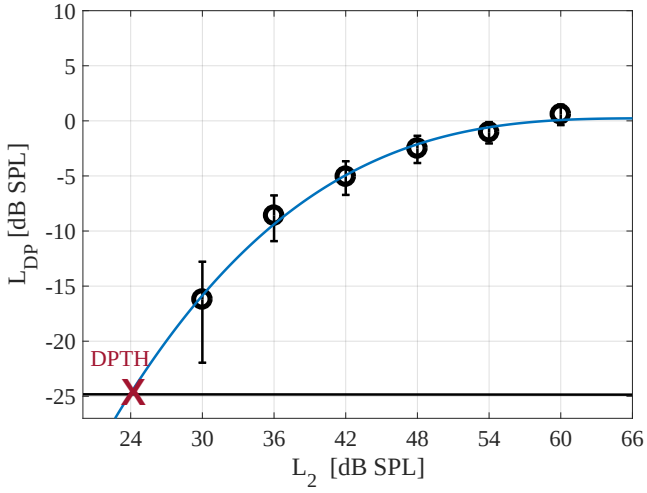


Figure 5.2: Measured DPOAE input-output function of a yNH subject at $f_2 = 0.8kHz$. The cubic curve-fit is shown in blue and the intersection with $L_{DP} = -25$ dB SPL is specified with red cross, which indicates the DP-threshold (DPTH) at 0.8 kHz.

DP-grams of six primary levels, independently. Each pole-function comprises of 1001 CF-channels and corresponding pole values (α^* in [73]) range between 0.036 and 0.32, which determine the gain and width of each simulated cochlear filter at low stimulation level. At each CF, α^* follows a stimulus level-dependent trajectory that accounts for the BM compression and wider cochlear filters with level increment [79]. Therefore, considering a constant stimulation level, by increasing α^* value without changing the level-dependent pole-trajectory function, the sensitivity of the cochlear model will reduce [64, 73]. Pole-function with constant across-CF pole-values of 0.036 or 0.32, accounts for an intact ($flat_{min}$) or completely damaged cochlea ($flat_{max}$: 35 dB-HL), respectively. To set CF-dependent pole-values, a machine-learning approach similar to that described in Section 4.3.1.2 of Chapter 4, was implemented. To construct the training data for the neural network (NN), 26 random pole-functions with pole-values accounting for both NH and HI cochleae (between 0.036 and 0.32) were generated. For each pole-function, two pure-tones with the same characteristics as the experimental stimuli were presented to the model, after which DP amplitude at $2f_1 - f_2$ frequency was calculated (sL_{DP}) to simulate DP-grams for all 26 pole-functions. This process was repeated for all primary frequencies and levels adopted in the experiment protocol. Accordingly, six DP-grams, each including sL_{DP} values at four f_2 primary frequencies, were simulated for a given random pole-function. To simulate DPTHs, DPOAE I/O functions were constructed for f_2 frequencies included in the recordings, and sL_{DP} values corresponding to L_2 levels between

-10 and 70 dB-SPL (with a 5-dB step) were calculated. Corresponding DPTHs of each simulated I/O growth-function was determined as the L_2 level at which the interpolated growth-function crossed the sL_{DP} of -10 dB-SPL.

Simulated DP-grams and DPTHs (sDP-gram and sDPTH) for the 26 random pole-functions were used to train seven different NNs with identical architectures that invert the model: one for the DPTH-based method and six for DP-grams of different levels (one for each level). The NNs were made up of one input-layer of four neurons, two hidden-layers of 150 neurons (150 for each layer) and an output-layer of 1001 neurons. While a standard Sigmoid activation function was applied to the hidden-layers, in the output layer a customized Sigmoid activation function was used to yield pole-values within the acceptable range of the cochlear model, i.e. between 0.036 and 0.32. The NNs were trained independently to learn how to map the given simulated DP-gram or DPTHs in the input layer to CF-dependent pole-values in the output. In this way, trained NNs can be adopted to predict the individualized pole-functions of the experimental cohort, given their measured DP-grams or DPTHs (mDP-gram and mDPTH) recorded at four primary frequencies.

However, it was required to apply a pre-processing step to the recorded DPOAEs prior to making the predictions of the corresponding pole-function. sDP-grams for a $flat_{min}$ profile in this study, overestimated the amplitudes of the recorded DP-grams of the yNH listeners. In addition, a -10 dB-SPL threshold was chosen for the simulations, since applying the -25 dB-SPL threshold (recorded I/O functions), yielded inconclusive sDPTHs, in particular for pole-values associated with greater CGLs. To bypass these limitations, measured DPOAE metrics were mapped to the range of the lower and higher extremum of the corresponding simulated DPOAE metrics. Later on, respective prediction errors were calculated based on their difference relative to the simulated profile without OHC dysfunction. For this purpose, first the outliers of the mDP-grams and mDPTHs were identified following the explanation provided in Section 4.3.1.2 of Chapter 4, after which the recordings were mapped to the range of associated sDP-grams and sDPTHs of the $flat_{min}$ and $flat_{max}$ profiles. Lastly, mapped mDP-grams and mDPTHs were normalized with the mean and standard deviation of the training data (Z-scores) and were fed into the trained NNs. The predicted individualized pole-functions in the output of the NN were used to simulate the individualized DPOAE metrics, sDPTH or sDP-gram, and evaluate the prediction errors.

5.5 Results

To determine which of the implemented cochlear individualization methods yielded more accurate predictions of the individual OHC damage, in the following sections mDP-grams/mDPTHs are compared with sDP-grams/sDPTHs. Moreover, it is tested that DPOAE-based cochlear model individualization methods how well

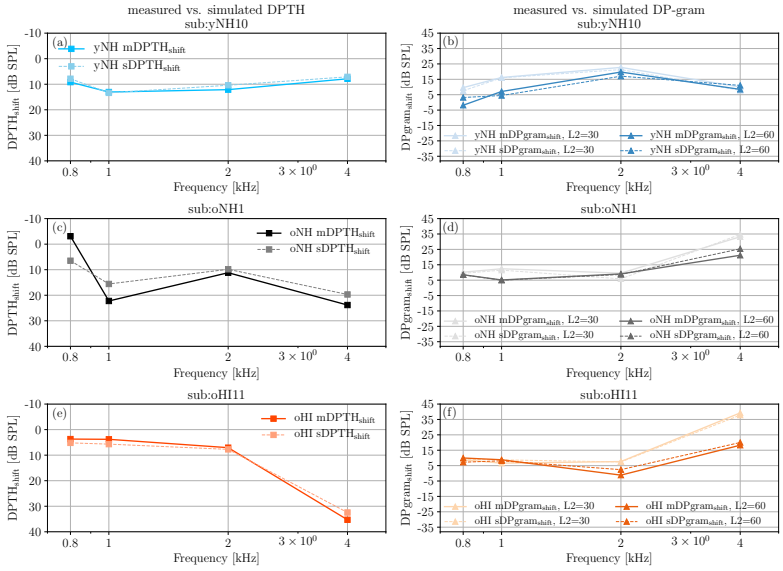


Figure 5.3: Comparison between measured (solid lines) and simulated (dashed lines) DPTHs/DP-grams. $sDPTH_{shift}$ and $mDPTH_{shift}$ values are compared across the measured frequencies for an exemplar of (a) yNH, (c) oNH and (e) oHI subject. $sDPgram_{shift}$ and $mDPgram_{shift}$ values in response to $L_2 = 30$ (light) and 60 (dark) dB SPL are shown for the same exemplar subjects: (b) yNH, (d) oNH and (f) oHI groups.

can predict the measured behavioural audiogram of the study participants.

5.5.1 DPTH and DP-gram prediction

To make a comparison between recorded and simulated DPOAEs, the type-1 error was defined as the absolute difference between mDP-grams/mDPTHs and sDP-grams/sDPTHs, which were simulated using individualized predicted pole-functions. In the previous step, recorded DPOAE metrics were mapped to a comparable range with the model simulations. In addition, model simulations overestimated DPOAE metrics. Therefore, before estimating prediction errors, a few more steps were required to be considered. To keep the formulation of these steps simple, we will refer to sDP for simulated DPOAE metrics, mDP for measured DPOAE metrics and $\mathrm{mDP_{map}}$ for measured DPOAE metrics which were mapped to the range of $flat_{min}$ and $flat_{max}$ profiles. In the first step, the sDP and mDP values were subtracted from the corresponding DPOAE metrics simulated for a normal-hearing profile ($\mathrm{sDP_{NH}}$).

$$\mathrm{mDP_{ref}} = (-1)^n(\mathrm{sDP_{NH}} - \mathrm{mDP_{map}}) \tag{5.1}$$

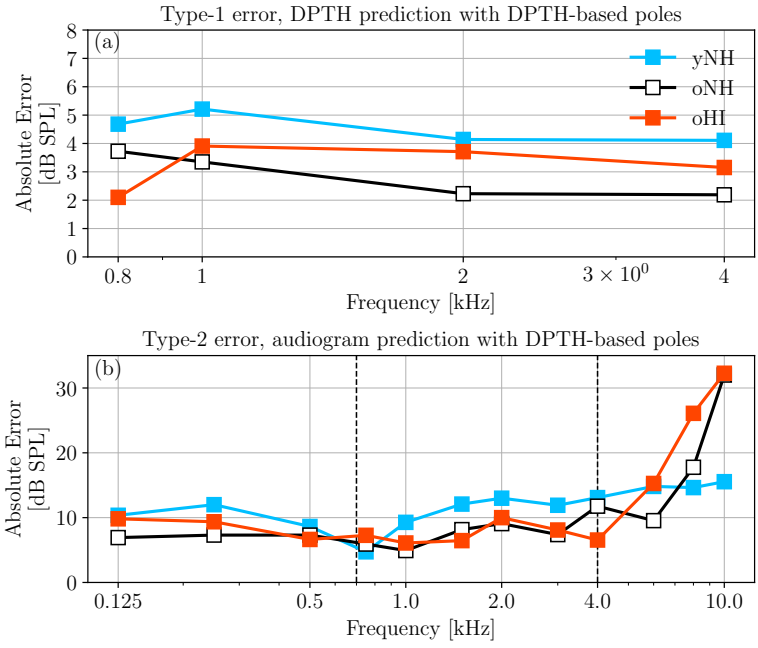


Figure 5.4: Averaged (a) type-1 and (b) type-2 errors of the DPTH-based cochlear model individualization across the yNH, oNH and oHI listeners. Two vertical dashed lines in panel (b), specify the frequency region of the audiogram which matches with the primary frequencies of the recorded DPTHs.

$$\mathrm{sDP}_{\mathrm{ref}} = (-1)^n(\mathrm{sDP}_{\mathrm{NH}} - \mathrm{sDP}) \tag{5.2}$$

In the above equations, if sDP and mDP stand for DPOAE thresholds, then n equals to one, otherwise $n = 0$. The referenced mDP and sDP ($\mathrm{mDP}_{\mathrm{ref}}$ and $\mathrm{sDP}_{\mathrm{ref}}$) amounts were mapped back to the range of measured DPOAE metrics as follows:

$$\mathrm{sDP}_{\mathrm{shift}}(f_2) = \mathrm{sDP}_{\mathrm{ref}}(f_2)\frac{max[\mathrm{mDP}(f_2)] - min[\mathrm{mDP}(f_2)]}{\mathrm{sDP}_{flat_{max}}(f_2) - \mathrm{sDP}_{flat_{min}}(f_2)} \tag{5.3}$$

$$\mathrm{mDP}_{\mathrm{shift}}(f_2) = \mathrm{mDP}_{\mathrm{ref}}(f_2)\frac{max[\mathrm{mDP}(f_2))] - min[\mathrm{mDP}(f_2)]}{\mathrm{sDP}_{flat_{max}}(f_2) - \mathrm{sDP}_{flat_{min}}(f_2)} \tag{5.4}$$

Figure 5.3 makes a comparison between the calculated $\mathrm{mDP}_{\mathrm{shift}}$ and $\mathrm{sDP}_{\mathrm{shift}}$. In the first column (panels a, c and e), measured (solid line) and predicted (dashed

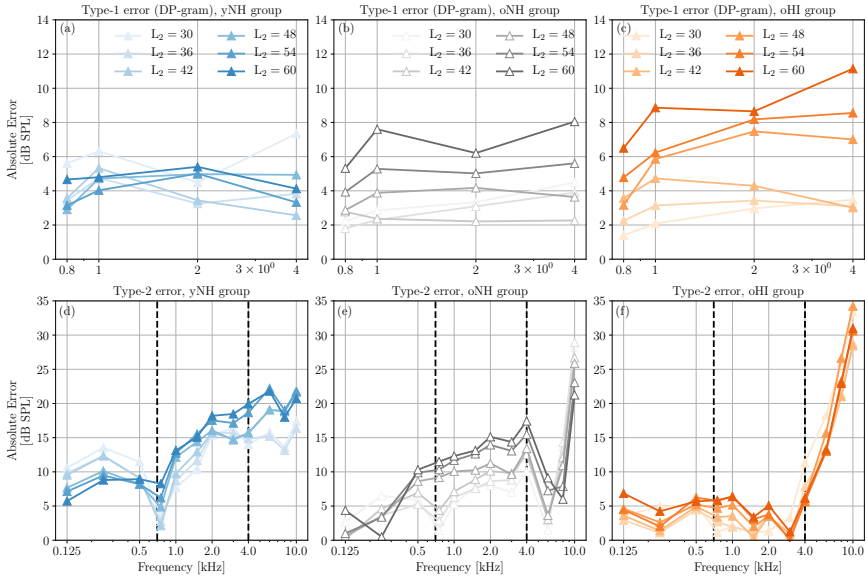


Figure 5.5: Averaged type-1 and type-2 errors of (a,d) yNH, (b,e) oNH and (c,f) oHI listeners for DP-gram based individualized cochlear models. In each panel, prediction errors based on DP-grams measured at L_2 levels between 30 (light colors) and 60 (dark colors) dB SPL are presented. Two vertical dashed lines in panel d, e and f, specify the frequency region of the audiogram which matches with the primary frequencies of the recorded DP-grams.

line) DPTHs of one subject from each of the yNH, oNH and oHI groups are shown. A similar comparison was made between $\mathrm{mDPgram_{shift}}$ and $\mathrm{sDPgram_{shift}}$ of different primary levels (30 and 60 dB SPL) in the second column of Figure 5.3 (panels b, d and f).

To quantify the difference between simulated and measured DPOAE metrics, the type-1 error was defined as the absolute frequency-specific difference between $\mathrm{sDP_{shift}}$ and $\mathrm{mDP_{shift}}$.

$$\mathrm{err1}(f_2) = |\mathrm{mDP_{shift}}(f_2) - \mathrm{sDP_{shift}}(f_2)| \tag{5.5}$$

Figure 5.4a compares the DPTH prediction mean errors (type-1 error) across the frequency and experiment groups. Type-1 errors of DP-gram prediction for different stimulation levels are shown in the first row of Figure 5.5 and each panel represents the respective averaged type-1 error of each group (a: yNH, b: oNH and c: oHI). When comparing the type-1 errors of the DPTHs with those of DP-grams (Figure 5.4a and Figure 5.5(a-c)), we conclude that cochlear models individualized

by DP-grams measured with low primary levels and DPTHs performed equally well. However, the prediction error increased when higher stimulation levels were used in the DP-gram measurements. To further investigate this, type-1 errors of each method were averaged over the four frequency points to provide an overview of how type-1 error alters across the adopted methods and stimulation levels. Boxplots of the type-1 error, averaged across the measured frequencies, are shown in Figure 5.6a. The individual variation of the type-1 error across the methods is indicated with thin lines. Pooling the type-1 errors of all groups together, a one-way ANOVA test confirmed that the differences between type-1 errors of the DP-grams measured in response to increasing primary levels are significant ($F_{(5,204)}$=9.95, p<0.000). Moreover, paired-sample t-test with Bonferroni correction showed that the only significant difference between the type-1 error of the DPTH and DP-grams was observed for $L_2 = 60$ dB SPL ($t_{(34)}$=-3.71, p<0.000). This confirms that cochlear model individualized based on DP-grams measured at lower stimulus levels performs equally well than the DPTH-based method.

5.5.2 Audiogram prediction

In the previous section, err1 was used to evaluate how well trained NNs with sDPTH or sDP-gram predict the mDPTH or mDP-gram. Apart from DPOAE measurement predictions, it is also of great interest to assess whether NNs trained using DPOAE metrics, can predict the measured audiograms. This would offer an additional and independent evaluation technique to assess the efficiency of the proposed DPOAE-based cochlear model individualization method. For this purpose, the audiometric thresholds were simulated by determining the stimulation intensity that was required to minimize the energy difference between the AN excitation-pattern (ANEP) of a considered participant and the ANEP of a NH-profile (L_{NH}, Figure 5.7). In this regard, 500-ms pure-tones were presented to the model with the same frequencies as of the measured audiograms (f_{aud}) and respective AN spike-rates were simulated for a range of stimulus levels, $L \in [-5, 55]$. The ANEP to the stimulus frequency was calculated by determining the root-mean-square (RMS) of each CF channel. This procedure was implemented for the predicted individualized pole-functions based on either of the DPTH or DP-gram methods. In addition, the NH ANEP at the threshold of audibility (based on equal-loudness-contour: ISO 226:1987) was simulated using the NH pole-function, which served as a reference to calculate the simulated audiometric thresholds ($\mathrm{sAudTH_{DP}}$) as follows:

$$\mathrm{L_{min}}(f_{aud}) = \underset{L\in[-5,55]}{\arg\min} \left[\left| \mathrm{ANEP_{DP}}_{(f_{aud}, L_{NH}(f_{aud}))} - \mathrm{ANEP_{DP}}_{(f_{aud}, L)} \right| \right] \quad (5.6)$$

$$\mathrm{sAudTH_{DP}}(f_{aud}) = \mathrm{L_{min}}(f_{aud}) - \mathrm{L_{NH}}(f_{aud}) \quad (5.7)$$

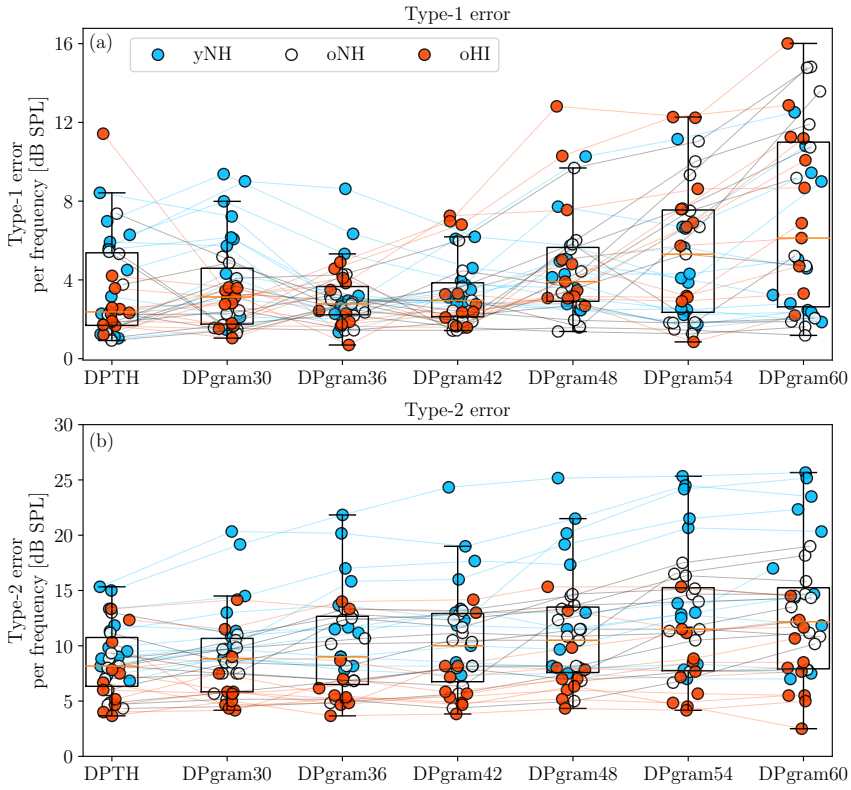


Figure 5.6: Averaged (a) type-1 and (b) type-2 errors across the frequencies within the [0.8-4] kHz. Corresponding errors of individuals are indicated by circles (yNH: blue, oNH: white, oHI: orange) and respective variations across different cochlear individualization methods are marked with thin lines.

In Eq. 5.6, $\mathrm{ANEP}_{\mathrm{DP}}$ refers to a simulated ANEP using a DPOAE-based (either DPTH or DP-gram) individualized pole-function. Figure 5.7 depicts the audiogram prediction process of two subjects at 8 kHz, using their individualized cochlear models based on DP-grams measured at $L_2 = 36$ dB SPL: yNH in blue and oNH in orange. Each circle in Figure 5.7, represents the peak of the AN spike-rate RMS at CF=8 kHz, in response to 8-kHz pure-tones of different levels. The lightest circle of each color indicates the lowest stimulus level. When considering the yNH subject (blue), the hearing threshold at 8 kHz was determined by the difference between the stimulus level of the closest ANEP peak-value to the threshold of audibility at 8 kHz (the black thick line), which is 12 dB SPL (according to equal-loudness-contour). Thus, an audiometric threshold of 5-dB-HL was estimated for the considered yNH subject at 8 kHz. Following the same

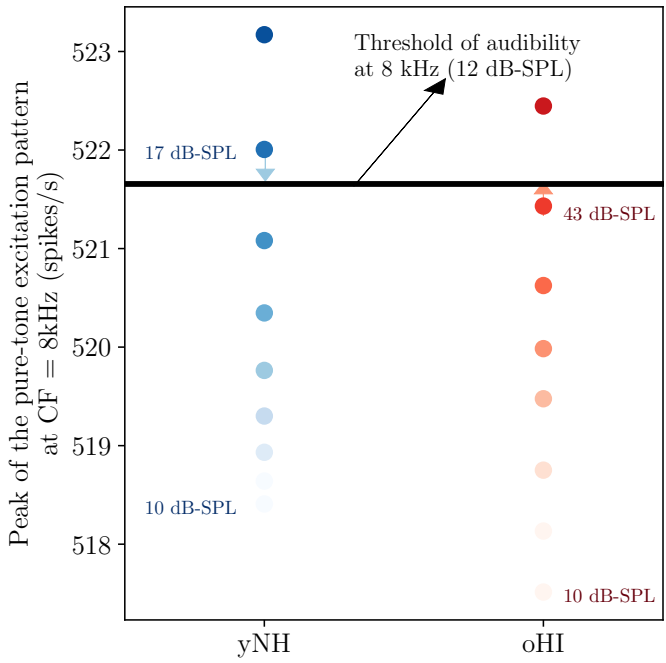


Figure 5.7: AN excitation pattern (ANEP) peaks (circles) at CF=8 kHz for an exemplar of yNH (blue) and oHI (orange) subject. The darker circles are associated with higher stimulation levels and vice-versa. The thick black line is the indicator of threshold of audibility at CF=8 kHz, i.e. 12 dB SPL.

method for the oHI subject (orange), an audiometric threshold of 31-dB-HL was predicted. Following this procedure we predicted individual audiograms measured at 12 frequencies based on DPOAE metrics of four frequencies.

To quantify the difference between measured and predicted audiograms, the type-2 error was defined as the absolute difference between $\mathrm{sAudTH_{DP}}$ and measured audiometric thresholds $\mathrm{mAudTH_{DP}}$ of the study participants.

$$\mathrm{err2}(f_{aud}) = |\mathrm{sAudTH_{DP}}(f_{aud}) - \mathrm{mAudTH_{DP}}(f_{aud})| \tag{5.8}$$

The group-mean of type-2 errors corresponding to DPTH and DP-gram based cochlear individualization methods are shown in Figure 5.4b and the second row of Figure 5.5 (panels c, d and e), respectively. Elevated high-frequency type-2 errors were observed, especially for oNH and oHI groups in both DPTH and DP-gram based individualized methods. This high-frequency error increment relates to the model limitation in simulating CGLs higher than 35 dB HL, since the maximal amount of the cochlear mechanical filter gain is 35 dB [73]. Consistent with the

type-1 error, we observed lower type-2 errors when DP-grams were recorded to lower primary levels. Given that DPOAEs were measured at four frequency points (between 0.8 and 4 kHz), only audiometric thresholds between 0.75 and 4 kHz frequency region were considered for further analysis of audiogram prediction errors. This frequency region is specified by vertical dashed lines in Figure 5.4b and the second row of Figure 5.5. In this context, Figure 5.6b illustrates the average individual type-2 errors across the six frequencies within the range of 0.75 and 4 kHz. The alteration of the averaged type-2 error for each individual across different methods are shown with thin lines. All subjects pooled together, a one-way ANOVA test showed a significant effect of the stimulation level on type-2 error of the DP-gram ($F_{(5,204)}$=3.14, p=0.009). A paired-sample t-test with Bonferroni correction showed that the type-2 error of DPTH was significantly different from that of the DP-gram measured at all levels, except at $L_2 = 30$ dB SPL ($t_{(34)}$=-1.43, p=0.16). This finding is consistent with that of the type-1 error, and indicates that individualized cochlear models based on either DP-grams measured at low stimulus levels or DPTHs yield the lowest type-1 errors. Moreover, they predict individual audiograms with a smaller error (type-2) than DP-grams of higher stimulus levels.

Lastly, to make an overall comparison between type-1 and type-2 errors, Figure 5.8, compares the type-1 and type-2 errors of the different implemented methods, as averages across frequency (between 0.75 and 4 kHz) and subject group. The figure shows that the average type-2 errors are always higher ($\approx$ 5 dB) than the type-1 error, and that an accurate prediction of DPOAE measurements (smaller type-1 errors), results in a more precise estimation of the individual audiogram (smaller type-2 errors).

5.6 Discussion

In this chapter, DPOAE measurements, as an objective diagnostic tool for OHC-damage, were evaluated to build individualized cochlear models. These models can then be used as pre-processors for hearing-aid or machine-hearing applications. In this regard, DPTHs and DP-grams measured at six primary levels were adopted to estimate individualized poles of the BM admittance function using a machine-learning approach. Personalized CGL parameters, i.e. CF-dependent pole-values of the BM admittance, were used to simulate individualized DPTHs/DP-grams. These simulations were compared to the recordings that were used to extract the CGL parameters (type1 error). In addition, DPTH and DP-gram based individualized models were adopted to predict audiograms of the study participants and their performance was compared (type-2 error). Considering these two error-types, we determined which of the adopted DPOAE-based metrics is more suitable for cochlear model individualization. Our analysis shows that DP-

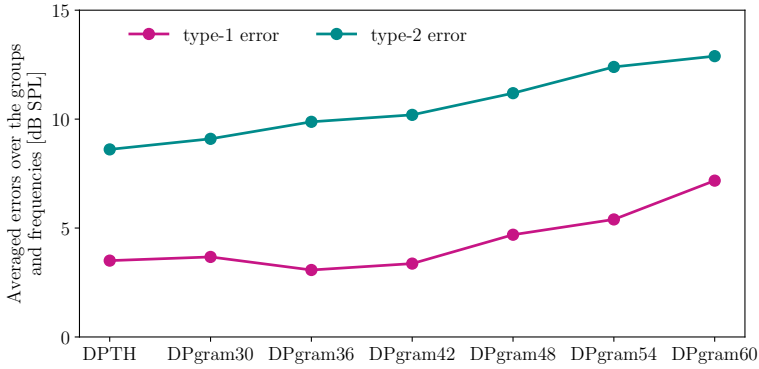


Figure 5.8: Comparison between the type-1 (purple) and type-2 (green) errors of different cochlear individualization methods, which are averaged across the frequency and study participants

grams measured at low primary levels, as well as the DPTHs, yielded the lowest type-1 and type-2 errors. To understand the reason, the following two aspects were considered: (1) At low stimulation levels, DPOAE amplitudes are maximally impacted by OHC amplification and corresponding losses. Near-threshold stimulation levels may hence provide more reliable DPOAE measurements [108]. This may partly explain the relatively high type-1 and type-2 errors we observed for some yNH subjects even at low stimulus levels, since their hearing thresholds were below 20 dB HL at all frequencies. At the same time, the smaller prediction errors of the HI listeners can be explained by their higher audiometric thresholds within the frequency range of [0.8-4] kHz, which are closer to the low stimulus levels. (2) The range of active pole values in the adopted cochlear TL model (i.e. the α_A^* in [73]), was constrained based on the low stimulus levels (below 35 dB SPL), since low stimulation intensities lead to sharper cochlear filters [73]. Consequently, using DP-grams of higher stimulation levels to set the CF-dependent α_A^* of the cochlear model, might have resulted in the higher prediction errors.

Regardless of the elevated high-frequency audiogram prediction errors caused by the maximal gain limitation of the cochlear model (i.e. 35 dB), in all implemented methods, type-2 errors were greater than the type-1 (Figure 5.8). The first reason for the observed difference can be related to the fact that cochlear models, which were individualized based on a DPOAE metric, i.e. an objective measure of OHC-loss, were used to predict audiogram, which is a behaviourally measured metric. Secondly, AN excitation patterns in response to pure-tones were used to predict individualized audiometric thresholds. AN excitation patterns are affected by the contribution of a large number of off-CF channels, and hence, compared to DPOAE-based metrics (i.e. more localized response, type-1 error), yielded less

accurate prediction of the audiometric thresholds.

The reader should be reminded that the compression slope of the cochlear model I/O function was not individualized, which together with above mentioned factors, could have contributed to generate additional prediction errors.

Taken together, the results of this chapter suggest that within the framework of this study, DP-grams recorded at low stimulation levels ($\mathrm{DPgram_{low}}$) (below 36 dB SPL) perform as well as DPTH-based methods for building individualized cochlear models. The NN trained with random CF-dependent pole-functions and corresponding simulated DPTH or $\mathrm{DPgram_{low}}$ metrics, predict the measured DPTH or $\mathrm{DPgram_{low}}$ with a reasonable error (below 5 dB/frequency). Additionally, the developed method can estimate individual audiograms (at frequencies involved in DPOAE measurements) with an average error of less than 10 dB/frequency. Considering the audiogram measurement resolution, i.e. 5 dB, the generated prediction error is acceptable and in this way, an additional independent test supports the efficiency of the proposed method. Despite these promising results, we suggest that conducting DPOAE measurements at more primary frequencies in the future, as well as taking into account the compression slope of the cochlear I/O function in the individualization process may lower the prediction errors. We hope that considering these two aspects may further improve the functionality of the personalized cochlear model predictions and can enhance the future individualized hearing aid algorithms.

6

Evaluation of the auditory evoked potentials and speech intelligibility after attending music festivals

Sarineh Keshishzadeh, Tine Vande Maele, Nele De Poortere, Ingeborg Dhooge, Hannah Keppler, Sarah Verhulst

Extracted from the pre-print entitled "The variability in potential biomarkers for cochlear synaptopathy after recreational noise-exposure". (https://doi.org/10.1101/2021.01.17.427007). [1]

This chapter reports and interprets the effect of recreational noise-exposure on the potential biomarkers of cochlear synaptopathy (CS), e.g. envelope following respones (EFRs), auditory brainstem responses (ABRs) and speech intelligibility. A cohort of 18 young normal-hearing attended music festival during the summer of 2019. The baseline auditory status of the experiment participants was assessed one day before, one, three and five days after attending the music festival. Changes in biomarkers from the first session to the follow-up sessions were non-significant, except for the speech audiometry, that showed a significant learning effect (performance improvement). This can indicate the absence of noise-exposure-driven CS

[1] **SK**: experiment design, analysis, conceptualization, writing the original draft, review and editing, **TVM**: experiment design, running the experiment, analysis, writing the original draft, **NDP**: analysis, **ID**: experiment design, supervision, conceptualization, review and editing, funding acquisition, **HK**: experiment design, supervision, conceptualization, review and editing, funding acquisition, **SV**: experiment design, supervision, conceptualization, review and editing, funding acquisition.

in the study cohort, or reflect that biomarkers were not sensitive enough to detect mild CS.

6.1 Introduction

Noise-exposure in young adults mainly takes place during leisure time activities such as watching movies, visiting nightclubs, attending music festivals or concerts, and listening to personal music-players. In this population, concerts and music festivals are the loudest activities [148–151]. Sound level measurements during such events confirm that noise levels can exceed 100 dBA [152–155]. Despite the potential risk of noise-induced hearing-loss (NIHL), the majority of young adults do not wear hearing protection devices (HPDs) when attending concerts (73.8%; [156]) or performances at music festivals (86%; [153]). After attending these music venues, 36-86% of young adults experience hearing related symptoms such as temporary hearing-loss, dullness or tinnitus [149, 151, 153].

The reduction of hearing sensitivity following noise-exposure has extensively been investigated in animal and human studies [37]. The incidence of NIHL depends on the noise characteristics: the intensity, frequency and duration of the exposure [157]. NIHL can cause permanent threshold shifts (PTS) due to mechanical damage of the cochlear hair cells or supporting cells and metabolic changes that cause cell degeneration and cell death [158, 159]. The reduction of hearing sensitivity after noise-exposure can also be reversible. Mechanical damage of outer-hair-cell (OHC) stereocilia, reversible changes in supporting cells, or a swelling of auditory nerve (AN) fiber terminals [158] can result in a temporary threshold shift (TTS), which recovers within minutes or hours [160,161] to weeks [162,163]. Research investigating the pathophysiology of TTS in rodents indicates that excessive glutamate release excitotoxicity is the source of AN fiber terminal swelling following noise-exposure [164, 165]. Recent studies have shown that TTS may not be as benign as previously thought, since it can be accompanied by permanent deficits at the synapse level where type-I AN fiber terminals connect to the inner-hair-cells (IHCs). Over the past years, this synapse loss has been named cochlear synaptopathy (CS) and has been linked to noise-exposure and aging in mice [13, 16, 34]. Damage to the IHC synapses due to noise-exposure, as seen in CS, can occur immediately after noise-exposure and can precede a slow degeneration of spiral ganglion cells [17]. CS could hence be a possible explanation for the pathophysiology of hidden hearing loss, where normal audiometric thresholds are accompanied by supra-threshold hearing deficits, such as decreased speech intelligibility in noise [87, 166]. Furthermore, it has been argued that this pathology may play a role in the origin of symptoms, such as tinnitus or hyperacusis [15, 97].

Animal and computational auditory modeling studies have attempted to define biomarkers for diagnosing CS non-invasively and have shown the potential of audi-

tory evoked potentials (AEPs). More specifically, the auditory brainstem response (ABR) and envelope following response (EFR) are promising tests, given that ABR- and EFR-based derived metrics relate directly to the number of histologically-verified synapse counts in animal models [13, 31, 32, 125]. Suprathreshold ABRs are generated by the synchronous firing of AN fibers and brainstem neurons in response to the rapid onset of a brief stimulus [167]. In rodents with CS, a reduced amplitude of ABR wave-I or reduced wave I/V amplitude ratio was observed, which captures the decreased integrity of the peripheral auditory nerve [13, 16, 17, 33]. The EFR is elicited by continuous, amplitude-modulated signals and the strength of the response to the stimulus envelope modulation frequency reflects temporal coding precision in the subcortical neurons [93]. Animals or models with CS showed reduced EFR-strength, which reflects a deficit in temporal coding fidelity of the auditory system [31, 32, 73, 83, 87], even when hearing sensitivity recovered or remained normal. Relating these AEP biomarkers to CS in living humans is far more difficult. Comorbidity with OHC-loss and inter-subject variability of AEP measures in humans, complicate the use of these methods. Moreover, CS diagnosis through direct synapse counts is not possible in living humans and only few human temporal bone studies have been conducted in relation to CS [38, 39]. Therefore, other methods are required to evaluate how AEP markers could relate to CS in living humans. Some studies were able to relate ABR and EFR strength to noise-exposure history [48,54,168,169], aging [42,83,170] or tinnitus [15, 171], but other studies could not confirm these findings [49, 51, 53, 172]. Longitudinal studies in humans, where AEP-measurements are used to evaluate auditory-nerve integrity before and after noise-exposure, are scarce. One study reported unchanged compound action potential amplitudes in electrocochleography, the analogue of ABR wave-I amplitude, in relation to recreational noise-exposure [52].

By recruiting a cohort of young normal-hearing adults, this study rules out age factors and focuses mainly on noise-induced CS, elicited by attending music events. The effect of recreational noise-exposure on potential biomarkers of CS and OHC-damage, as well as its probable consequences on speech intelligibility are investigated. By adopting a test battery that includes pure tone audiometry (PTA) at conventional and extended high frequencies (EHFs), speech audiometry tests in quiet (SPiQ) and in noise (SPiN) and AEP measurements, we hypothesize that speech intelligibility and AEP measurements may deteriorate due to noise-exposure, even if PTA is not affected or recovered shortly after noise-exposure. To this end, the variability of the above-mentioned metrics was analyzed one day before and in a period of five days after attending a music festival. Moreover, the relationship of speech intelligibility with AEP-derived metrics or PTA thresholds were assessed.

6.2 Methodology

PTA, ABR, EFR and speech intelligibility were used to evaluate short- and long-term consequences of recreational noise-exposure on the hearing and speech perception of a young normal-hearing cohort.

6.2.1 Participants

Young adults, aged 18 to 25 years, who planned to attend a music festival during the summer of 2019 were recruited for this study. Volunteers completed a recruiting questionnaire, which was used to exclude subjects with known hearing pathologies, history of ear surgery or tinnitus. Study participation consisted of four measurement sessions: one baseline measurement session (s1: session 1) between 1 to 3 days before the start of the festival and 3 follow-up sessions: 1 day, 3 days and 5 days after the end of the music festival. Further on, these follow-up sessions will respectively be referred to as session 2 (s2), session 3 (s3) and session 4 (s4). Subjects were asked not to expose themselves to noise other than that of the festival during the follow-up period and not to use party drugs. Participants were free to use HPDs. Twenty subjects participated in the first session, 8 males and 12 females (21.5 ± 2.24 years). Of those subjects, the best ear was chosen as the test ear, based on tympanometry and PTA. Of the 20 participants, two subjects with incomplete sessions were dropped. At every test session, subjects completed a test battery consisting of questionnaires, PTA, SPiQ and SPiN tests, and AEP measurements. The test protocol had a duration of maximum three hours per session (including breaks and information) and tests were conducted in the same order for every subject and at every session.

This study was approved by the ethical committee of the Ghent University Hospital (Belgium) and was performed following the statement of ethical principles of the Declaration of Helsinki. All subjects were informed about the testing procedures and signed an informed consent. Subjects received a financial compensation for their participation.

6.2.2 Pure-tone audiometry

PTA thresholds were measured while subjects were seated in a double-walled sound-attenuating measurement booth. At the baseline session, both ears were tested to select the test ear. Audiometric thresholds were measured at conventional octave frequencies 0.125, 0.25, 0.5, 1, 2, 4 and 8 kHz, half-octave frequencies of 3 and 6 kHz, and at EHFs of 10, 12.5, 14 and 16 kHz using the modified Hughson-Westlake procedure. An Equinox Interacoustics audiometer was used and stimuli were transmitted using Interacoustics TDH-39 headphones and Sennheiser HAD-200 headphones for conventional frequencies and EHF, respectively. Figure 6.1

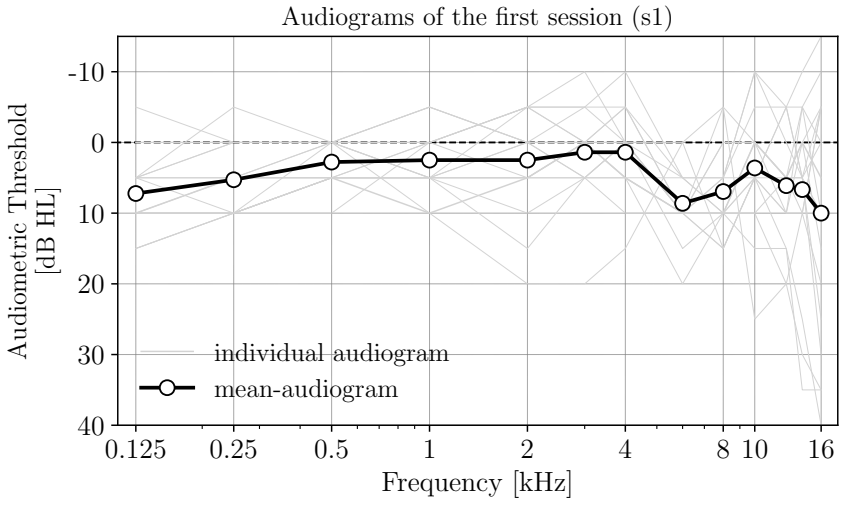


Figure 6.1: First session audiograms. Individual audiograms measured in the first session (s1: baseline measurement) are shown in thin grey lines. The thick black line represents the grand-averaged audiogram in the first session.

shows the individual audiograms of the experimental cohort and the respective group-mean, measured in session s1 (baseline measurement).

6.2.3 AEP measurements

AEP measurements were comprised of EFRs and ABRs. EFRs were evoked by two stimulus types. Both stimuli consisted of a 4-kHz amplitude-modulated tone and differed in their modulation waveform. Specifically, sinusoidal amplitude modulation (SAM) was applied to the first pure-tone and rectangular amplitude modulation (RAM) to the second (25% duty-cycle). The latter was adopted from [44], as this stimulus was found to yield strong EFRs, which are sensitive to detect CS in auditory model simulations. Both stimuli had a modulation frequency of 110 Hz, a modulation depth of 100% and a duration of 500 ms. Stimuli were presented 1000 times with a rate of 2 Hz at 70 dB SPL. ABRs were recorded to 5000 alternating polarity 80-μs clicks, presented at 80 dB peSPL.

At each session, AEP measurements were performed with the Universal Smart Box (Intelligent Hearing Systems) using SEPCAM software. Subjects were seated in a comfortable armchair in a quiet room and were watching a muted video during the recordings. The skin preparation was performed using NuPrep gel and Ambu Neuroline snap electrodes were placed on vertex, nasion and bilateral earlobes. The electrodes impedance never exceeded 3 kΩ. All stimuli were presented using

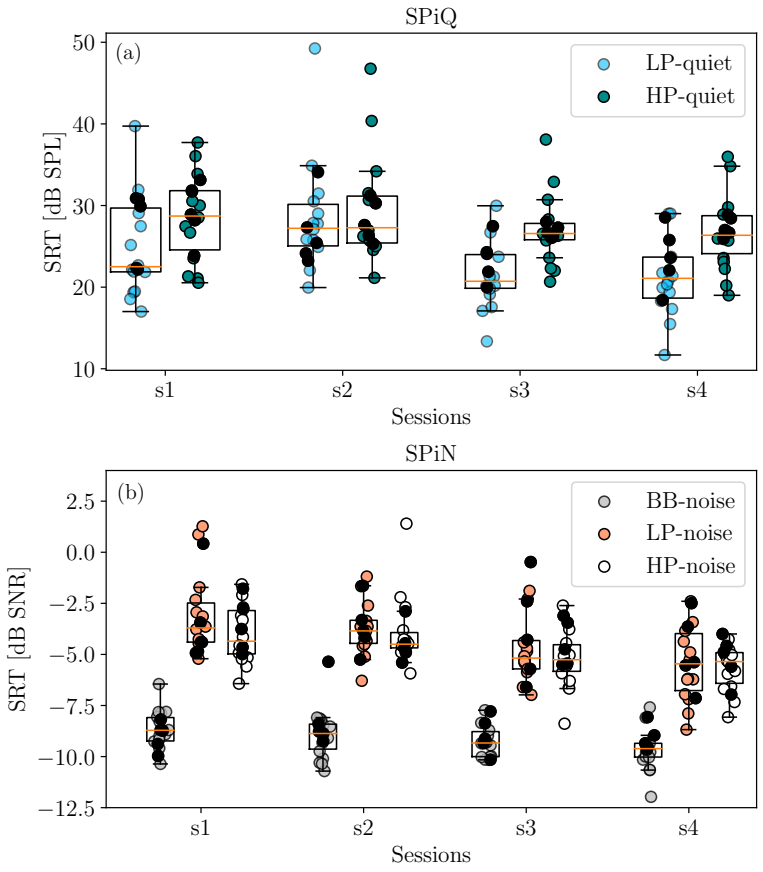


Figure 6.2: SRT values from different experiment conditions and sessions. (a) Low-pass (LP) and high-pass (HP) speech-in-quiet (SPiQ) SRT values in dB-SPL. Smaller SPLs reflect better performance. (b) Broadband (BB), LP and HP speech-in-noise (SPiN) SRT values in dB-SNR. Subjects that did not wear HPD during the festival, are specified by black circles. More negative SRTs reflect better performance.

ER-2 probes (Etymotic Research). Both ears were covered with earmuffs to minimize noise interaction and all inessential electrical devices were turned off during the measurements. Subjects were instructed to lean their head back and to relax as much as possible.

6.2.4 Speech intelligibility in quiet and in noise

At each session, SPiQ and SPiN tests were performed at the test ear using the Flemish Matrix sentence test [173] in Apex 3 software [174]. Sentences were

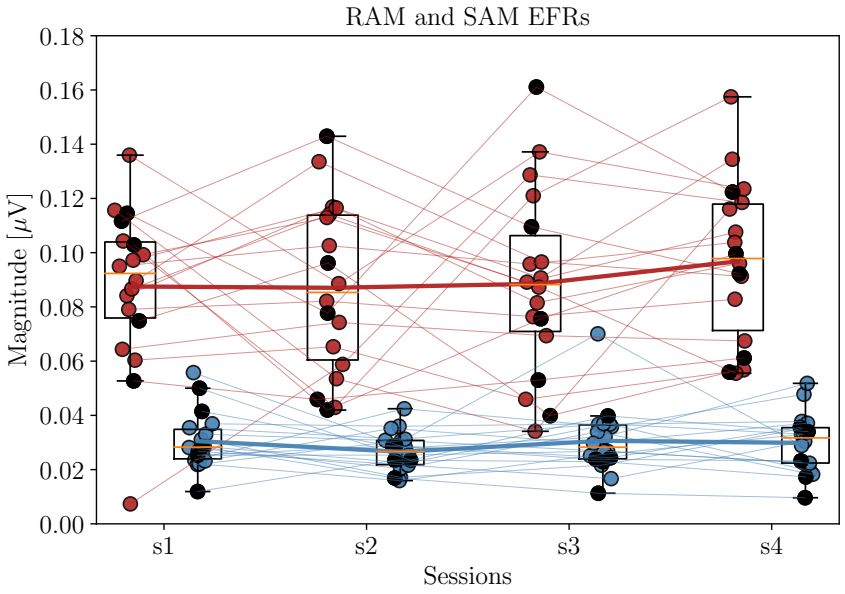


Figure 6.3: RAM- and SAM-EFR magnitudes derived from baseline (s1) and follow-up sessions (s2 to s4). Thin red and blue lines, illustrate variation of individual RAM- and SAM-EFRs across the sessions, respectively. Respective group-means are shown with thick lines (RAM: red, SAM: blue) and subjects that did not wear HPD during the festival, are specified by black circles.

presented in a relatively quiet room using a laptop connected to a Fireface UCX soundcard (RME) and HDA-300 (Sennheiser) headphones.

At each session, five randomly chosen experimental test-lists were presented in five conditions, presented in a random order. The five conditions consisted of two conditions in quiet, where speech was filtered with a zero-phase 256th-order FIR low-pass (LP) filter or with a zero-phase 1024th-order FIR high-pass (HP) filter with cutoff values of 1500 and 1650 Hz, respectively. Three conditions were presented in a speech-shaped noise with a fixed level of 70 dB SPL: a broadband condition (BB) where no filtering was applied, a LP and a HP condition, where speech, as well as, noise signals were filtered using the same filter cutoff values as for LP and HP conditions in quiet, respectively. Due to the filtering, the SPL levels of 0 dB-SNR LP and HP conditions were 70 and 53 dB SPL, respectively. During the first session, two training lists were presented in the BB-in-noise condition to minimize the known learning effect of this test [175]. For all test lists, subjects were asked to repeat the five-word sentences in a closed, forced-choice setting (10 options for each word were given). Speech was presented in an adaptive procedure using a staircase paradigm to determine the speech-reception threshold (SRT) with

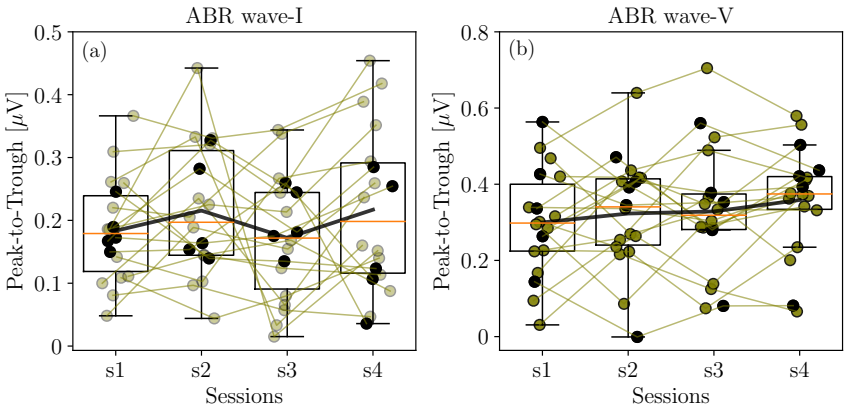


Figure 6.4: ABR wave-I (a) and -V (b) amplitudes derived from baseline (s1) and follow-up sessions (s2 to s4). Thin lines in each panel indicate individual ABR wave-I (a) and -V (b) variability across the sessions. Group-means are shown with thick lines and subjects that did not wear HPD during the festival, are specified by black circles.

a minimal step size of 0.1 dB. The mean signal level or mean SNR of the six last reversals was used as the SRT for the SPiQ and SPiN-tests, respectively. The SRT values of different conditions and sessions are shown in Figure 6.2.

6.3 AEP Processing

The raw AEP signals were stored in *"EEG.F"* format and converted to *"mat"* format using the custom-made *"sepcam2mat"* function. Data analysis was performed offline in *MATLAB*.

Recorded EFRs were bandpass filtered between 50 and 5000 Hz with an 800th order FIR filter using *"filtfilt" MATLAB* function to remove the filter delay. Then, 400-ms epochs were extracted from the 100 to 500-ms time-interval, relative to the stimulus onset and 200 epochs with the highest peak-to-trough values were rejected to remove the noisy epochs. The remaining 800 epochs were averaged and the corresponding spectrum magnitude was constructed using the Fast Fourier Transform (FFT). Additionally, a bootstrapping approach was adopted in the frequency domain to estimate the noise-floor and variability of the EFR. For a detailed explanation of the bootstrapping procedure see [43]. EFR strength was defined as the summation of the signal-to-noise spectral magnitude at fundamental frequency (110 Hz) and the following three harmonics, i.e. 220, 330 and 440 Hz, if they were above the noise-floor [44]. Extracted RAM- and SAM-EFR magnitudes from different sessions are shown in Figure 6.3.

Recorded ABR traces were bandpass filtered between 100 and 1500 Hz using

Table 6.1: Paired-sample t-test t- and p-values. Reported results are based on the comparison of RAM-EFR, SAM-EFR, ABR wave-I and -V amplitudes and averaged PTA thresholds over three frequency regions (extended high-frequency (EHF): 10-16 kHz, high-frequency (HF): 2-8 kHz and low-frequency (LF): 0.125-1 kHz) between sessions. t- and p-values relative to significant group-mean differences are shown in bold letters. No significant group-mean differences were observed in this regard.

Compared Sessions	RAM	SAM	wave-I	wave-V	Mean-Audiogram		
					EHF	LF	HF
s1-s2	$t_{(17)}$=0.05 p=0.98	$t_{(17)}$=1.34 p=0.20	$t_{(17)}$=-1.40 p=0.18	$t_{(17)}$=-0.70 p=0.48	$t_{(17)}$=-1.88 p=0.08	$t_{(17)}$=-1.54 p=0.14	$t_{(17)}$=-2.16 p=0.04
s1-s3	$t_{(17)}$=-0.10 p=0.92	$t_{(17)}$=-0.04 p=0.97	$t_{(17)}$=0.49 p=0.63	$t_{(17)}$=-0.85 p=0.41	$t_{(17)}$=0.65 p=0.52	$t_{(17)}$=-0.52 p=0.61	$t_{(17)}$=0 p=1
s1-s4	$t_{(17)}$=-1.20 p=0.28	$t_{(17)}$=0.23 p=0.82	$t_{(17)}$=-1.20 p=0.25	$t_{(17)}$=-1.79 p=0.09	$t_{(17)}$=0.21 p=0.84	$t_{(17)}$=-1.10 p=0.28	$t_{(17)}$=0.39 p=0.70

the same filter applied to EFRs. Twenty millisecond epochs were extracted relative to the stimulus onset and baseline correction was applied by subtracting the mean-value of each epoch. Two hundred epochs, equal number of each polarity, with the highest peak-to-trough values were rejected to remove the noisy epochs and the remaining 4800 epochs were averaged [94]. ABR peaks and troughs were identified manually and were confirmed by visual inspection of a second audiologist. The wave-I and V peak-to-trough values extracted from recordings of each session were considered for the analysis (Figure 6.4).

6.4 Results

Audiograms, speech reception thresholds, ABR wave-I and V amplitudes, RAM- and SAM-EFR magnitudes of 18 young normal-hearing subjects were extracted from four measurement sessions performed before (one session) and after attending the festival. To assess the effect of recreational noise-exposure on the extracted metrics and to investigate the respective variabilites within one, three and five days after the exposure, statistical Student t-test analysis was applied. We explored whether there are any statistically significant differences between metrics recorded in different sessions.

Applied paired-sample t-test analysis with Bonferroni correction compared RAM-EFR, SAM-EFR, ABR wave-I and -V amplitudes, audiogram and SRT metrics between each pair of the experimental sessions: s1-s2, s1-s3 and s1-s4. Considered session-pairs compare the short-term (s1-s2) and long-term (s1-s3 and s1-s4) effects of noise-exposure on each metric. Given that five of the 18 participants (27.77%) did not use HPDs during attending the festivals, we considered them separately in our statistical analysis, as well as mixed with the whole cohort. However, the population of non-HPD users was not large enough to draw any con-

Table 6.2: Paired-sample t-test t and p-values of subjects that did not use HPD during the music event. Reported results are based on the comparison of RAM-EFR, SAM-EFR, ABR wave-I and -V amplitudes and averaged PTA thresholds over three frequency regions (extended high-frequency (EHF): 10-16 kHz, high-frequency (HF): 2-8 kHz and low-frequency (LF): 0.125-1 kHz) between sessions. t- and p-values relative to significant group-mean differences are shown in bold letters. No significant group-mean differences were observed in this regard.

Compared Sessions	RAM	SAM	wave-I	wave-V	Mean-Audiogram		
					EHF	LF	HF
s1-s2	$t_{(4)}$=0.60 p=0.58	$t_{(4)}$=1.08 p=0.34	$t_{(4)}$=-0.62 p=0.57	$t_{(4)}$=0.35 p=0.74	$t_{(4)}$=-0.73 p=0.50	$t_{(4)}$=-1.55 p=0.19	$t_{(4)}$=-3.20 p=0.03
s1-s3	$t_{(4)}$=0.24 p=0.82	$t_{(4)}$=1.29 p=0.27	$t_{(4)}$=-0.37 p=0.73	$t_{(4)}$=0.48 p=0.66	$t_{(4)}$=0.99 p=0.40	$t_{(4)}$=-0.55 p=0.61	$t_{(4)}$=0.37 p=0.73
s1-s4	$t_{(4)}$=0.82 p=0.46	$t_{(4)}$=1.18 p=0.30	$t_{(4)}$=0.49 p=0.65	$t_{(4)}$=-0.41 p=0.70	$t_{(4)}$=-0.22 p=0.84	$t_{(4)}$=-2.06 p=0.11	$t_{(4)}$=-0.50 p=0.64

clusion. Besides the t-test analysis, Pearson correlation analysis was performed to evaluate the relationship between speech intelligibility and audiometric thresholds or AEP-derived metrics.

6.4.1 Variability of AEP-derived metrics and PTA thresholds

Paired-sample t-test analysis with Bonferroni correction, revealed no significant difference between RAM-EFR, SAM-EFR, ABR wave-I and -V amplitudes in non of the three above-mentioned paired-sessions (Table 6.1, Figure 6.3 and 6.4). Obtained results for ABR wave-I are in line with Grinn et al. findings in [52], who did not find ABR wave-I amplitude reductions in humans after attending a noisy event. Excluding outliers or HPD-users did not affect group-mean differences significantly. Table 6.1 details the *t*- and *p*-values of the paired-sample t-test, corresponding to each metric and experiment session. Results of similar analysis, but implemented only on non-HPD users are shown in Table 6.2. Evaluation of the audiogram variability between the sessions was performed on audiometric thresholds averaged over three frequency regions: (i) extended high-frequency (EHF: 10-16 kHz), (ii) high-frequency (HF: 2-8 kHz) and (iii) low-frequency (LF: 0.125-1 kHz). Akin to EFR and ABR metrics, insignificant differences were observed in similar frequency regions of paired-sessions (Table 6.1: considering all subjects, Table 6.2: non-HPD users). To make this comparison in a more frequency-specific manner, we applied the t-test between audiometric thresholds of each frequency, as well (Table 6.3). Panel (a) of Figure 6.5 illustrates group-mean audiograms of each session. Averaged EHF audiograms of individuals and corresponding variability over the four sessions are shown in panel (b) of Figure 6.5. However, the only significant difference was obtained at 10 kHz, between s1 and s2 sessions ($t_{(16)}$=-3.07, p=0.01).

Table 6.3: Paired-sample t-test results of PTA comparison between sessions. t- and p-values relative to significant group-mean differences are shown in bold letters.

Compared Sessions	Audiogram Frequencies [kHz]												
	0.125	0.25	0.5	1	2	3	4	6	8	10	12.5	14	16
s1-s2	$t_{(16)}$=-2.28 p=0.04	$t_{(16)}$=-1.70 p=0.11	$t_{(16)}$=-2.50 p=0.02	$t_{(16)}$=0 p=1	$t_{(16)}$=-0.22 p=0.83	$t_{(16)}$=-0.32 p=0.75	$t_{(16)}$=-0.52 p=0.61	$t_{(16)}$=-1.54 p=0.14	$t_{(16)}$=-1.94 p=0.07	$\boldsymbol{t_{(16)} = -3.10}$ $\boldsymbol{p = 0.01}$	$t_{(16)}$=-2.31 p=0.03	$t_{(16)}$=0.32 p=0.75	$t_{(16)}$=-0.82 p=0.42
s1-s3	$t_{(16)}$=0.16 p=0.88	$t_{(16)}$=-0.75 p=0.47	$t_{(16)}$=-0.22 p=0.82	$t_{(16)}$=1.32 p=0.21	$t_{(16)}$=0 p=1	$t_{(16)}$=-0.57 p=0.58	$t_{(16)}$=-0.76 p=0.45	$t_{(16)}$=-0.34 p=0.79	$t_{(16)}$=-0.27 p=0.79	$t_{(16)}$=-1.00 p=0.33	$t_{(16)}$=-0.61 p=0.55	$t_{(16)}$=1.38 p=0.19	$t_{(16)}$=1.05 p=0.31
s1-s4	$t_{(16)}$=0 p=1	$t_{(16)}$=0 p=1	$t_{(16)}$=0.82 p=0.42	$t_{(16)}$=0.70 p=0.50	$t_{(16)}$=-0.24 p=0.82	$t_{(16)}$=-0.24 p=0.82	$t_{(16)}$=-0.24 p=0.82	$t_{(16)}$=-1.51 p=0.15	$t_{(16)}$=-1.10 p=0.29	$t_{(16)}$=-0.98 p=0.34	$t_{(16)}$=-0.72 p=0.48	$t_{(16)}$=1.65 p=0.12	$t_{(16)}$=0.89 p=0.38

Table 6.4: Paired-sample t-test results of SRT comparison between sessions. t- and p-values relative to significant group-mean differences are shown in bold letters.

Compared Sessions	$SPiQ_{LP}$	$SPiQ_{HP}$	$SPiN_{BB}$	$SPiN_{LP}$	$SPiN_{HP}$
s1-s2	$t_{(17)}$=-1.65 p=0.12	$t_{(17)}$=-0.58 p=0.57	$t_{(17)}$=0.75 p=0.46	$t_{(17)}$=1.58 p=0.13	$t_{(17)}$=0.06 p=0.95
s1-s3	$\mathbf{t_{(17)} = 2.99}$ $\mathbf{p = 0.01}$	$\mathbf{t_{(17)} = 6.12}$ $\mathbf{p < 0.00}$	$t_{(17)}$=2.00 p=0.06	$\mathbf{t_{(17)} = 3.62}$ $\mathbf{p < 0.00}$	$\mathbf{t_{(17)} = 3.67}$ $\mathbf{p < 0.00}$
s1-s4	$t_{(17)}$=2.63 p=0.02	$\mathbf{t_{(17)}}$=2.75 $\mathbf{p = 0.01}$	$\mathbf{t_{(17)} = 3.49}$ $\mathbf{p < 0.00}$	$\mathbf{t_{(17)} = 7.82}$ $\mathbf{p < 0.00}$	$\mathbf{t_{(17)} = 7.04}$ $\mathbf{p < 0.00}$

Since the 10-kHz PTA thresholds of each non-HPD user increased by 5 dB-HL from session s1 to s2, leaving out HPD users, the corresponding t-values could not be determined. However, the other group-mean differences remained insignificant between the compared sessions.

6.4.2 Speech reception threshold variation across the sessions

Different from AEP-derived metrics and audiometric thresholds, more significant differences were observed between SRT values of respective conditions in different sessions (Figure 6.2). However, the short-term effect of the noise-exposure (between s1 and s2 sessions) was still absent in all speech audiometry conditions. Corresponding paired-sample t-test *t*- and *p*-values are reported in Table 6.4 and significant differences are specified in bold letters. On the contrary, group-mean differences were improved significantly (decreased SRT values) between s1 and s3, as well as between s1 and s4 sessions for all conditions, except for the BB in noise (s1-s3: $t_{(17)}$=2.00, p=0.06) and LP in quiet (s1-s4: $t_{(17)}$=2.63, p=0.02) conditions. Removing outliers did not affect group-mean differences (Table 6.5), whereas excluding HPD-users led to non-significant results in all conditions and paired sessions, except for the HP in quiet condition in s1-s3 paired session comparison (Table 6.6, $t_{(4)}$=4.11, p=0.01).

Our findings contrast those of the study of Grinn et al. [52], where the noise dose of an event was related to a reduction in SPiN intelligibility. Our study showed a significant improvement in SRT-values in the follow-up sessions, which could possibly be attributed to a learning effect. Figure 6.6 shows the declining trend (improvement) of different conditions' SRT-values across the sessions. Even though we used two training lists to overcome the learning effect, as suggested in the literature [175], significant improvements of SRT-values were observed across the sessions. As a consequence of this learning effect, possible deteriorations in speech intelligibility, shortly after noise-exposure could remain undetected.

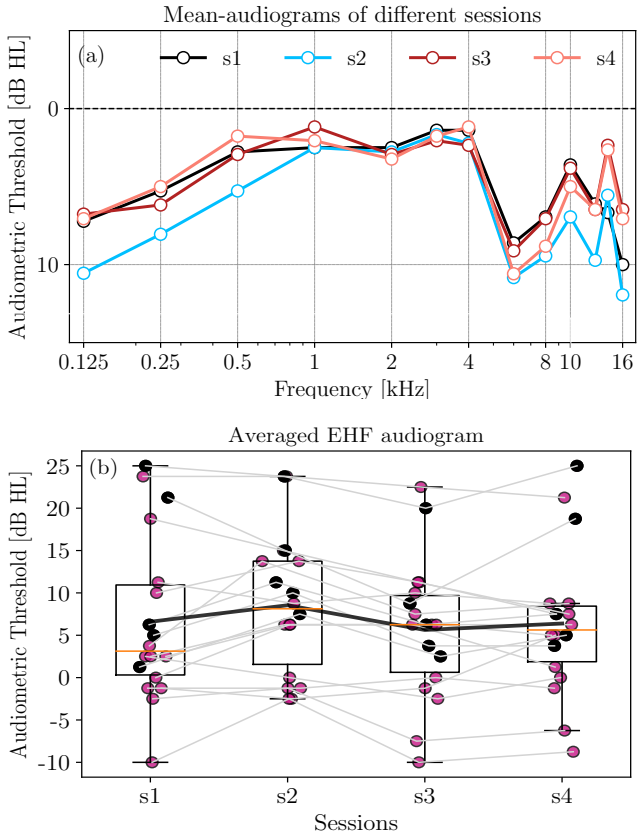


Figure 6.5: Variability of PTA thresholds across the sessions. (a) Comparison of grand-averaged PTA thresholds in different sessions. (b) Variability of averaged EHF audiograms across the sessions. Mean EHF audiomatric threshold of HPD- and non-HPD useres are shown in purple and black, respectively. Corresponding individual variabilties are specified by thin gray lines. The thick black line, represents the group-mean of averaged EHF audiometric thresholds.

6.4.3 Relationship between speech intelligibility and AEP-derived metrics

With a focus on the recreational noise-exposure effect on peripheral hearing and speech perception mechanisms, we explored if any relationship exists between speech intelligibility SRT values and AEP-derived metrics or audiograms. In this context, correlation analyses were performed (Pearson's *r*- and *p*- values), to determine the relation between SRT values of different conditions and EFR magnitudes, ABR markers and mean-audiometric thresholds of three frequency regions (Tables 6.7 and 6.8). Due to the learning effect observed in the speech intelligibil-

Table 6.5: Paired-sample t-test results of SRT comparison between sessions, after excluding outliers. t- and p-values relative to significant group-mean differences are shown in bold letters.

Compared Sessions	$SPiQ_{LP}$	$SPiQ_{HP}$	$SPiN_{BB}$	$SPiN_{LP}$	$SPiN_{HP}$
s1-s2	$t_{(16)}$=-1.27 p=0.22	$t_{(15)}$=0.54 p=0.59	$t_{(16)}$=1.98 p=0.09	$t_{(12)}$=-0.03 p=0.97	$t_{(14)}$=1.66 p=0.12
s1-s3	$t_{(16)}$=2.69 p=0.02	$\mathbf{t_{(12)} = 5.80}$ $\mathbf{p < 0.00}$	$t_{(17)}$=2.00 p=0.06	$\mathbf{t_{(13)} = 3.14}$ $\mathbf{p < 0.00}$	$\mathbf{t_{(16)} = 2.78}$ $\mathbf{p < 0.00}$
s1-s4	$t_{(17)}$=2.63 p=0.02	$\mathbf{t_{(16)} = 2.78}$ $\mathbf{p = 0.01}$	$\mathbf{t_{(14)} = 4.01}$ $\mathbf{p < 0.00}$	$\mathbf{t_{(14)} = 7.12}$ $\mathbf{p < 0.00}$	$\mathbf{t_{(17)} = 7.04}$ $\mathbf{p < 0.00}$

Table 6.6: Paired-sample t-test results of SRT comparison between sessions (non-HPD users). t- and p-values relative to significant group-mean differences are shown in bold letters.

Compared Sessions	$SPiQ_{LP}$	$SPiQ_{HP}$	$SPiN_{BB}$	$SPiN_{LP}$	$SPiN_{HP}$
s1-s2	$t_{(4)}$=0.14 p=0.89	$t_{(4)}$=0.41 p=0.70	$t_{(4)}$=-1.28 p=0.27	$t_{(4)}$=0.38 p=0.72	$t_{(4)}$=1.64 p=0.18
s1-s3	$t_{(4)}$=3.27 p=0.03	$\mathbf{t_{(4)} = 4.11}$ $\mathbf{p = 0.01}$	$t_{(4)}$=-0.09 p=0.93	$t_{(4)}$=0.83 p=0.45	$t_{(4)}$=1.53 p=0.20
s1-s4	$t_{(4)}$=1.05 p=0.35	$t_{(4)}$=0.98 p=0.38	$t_{(4)}$=0.41 p=0.70	$t_{(4)}$=2.37 p=0.08	$t_{(4)}$=3.67 p=0.02

ity test from the baseline session to the forth session (Figure 6.6), we only rely on the correlation analysis results of the baseline session (s1). Although, according to the paired-sample t-test analysis, SRT values did not change significantly between s1 and s2 sessions, the results of session s2 might still be affected by both the learning effect and the noise-exposure dose.

Correlation analysis in the baseline session, yielded two significant correlations. First, an unexpected *positive* correlation was observed between SAM-EFR and $SPiN_{HP}$, (r=0.61, p=0.01, n=18). After removing the outliers, this correlation became even stronger (Figure 6.7a: r=0.69, $p < 0.00$, n = 17). This finding is inconsistent with the *negative* correlation of SAM-EFR and $SPiN_{HP}$ reported in [123]. Given that in the same session, RAM-EFR, an EFR-based metric generated by almost same generators as in SAM-EFR, did not correlate to $SPiN_{HP}$ (Figure 6.7b), we suspected that the presence of variable minor degrees of OHC dysfunction among the subjects could have resulted in a positive significant correlation between SAM-EFR and $SPiN_{HP}$. Underlying this reasoning is the model simulations presented in [44], which showed that SAM-EFR is sensitive to OHC-damage, whereas the RAM-EFR is minimally affected by OHC deficits. The obtained positive correlation of SAM-EFR and $SPiN_{HP}$ in our study remained in the

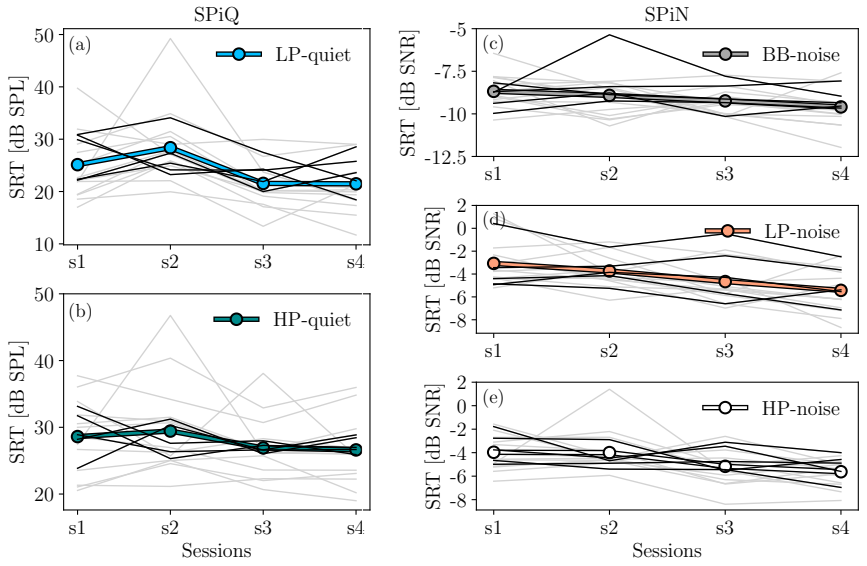


Figure 6.6: Individual variability of SRT values across the sessions in different conditions: (a) low-pass (LP) in quiet, (b) high-pass (HP) in quiet, (c) broadband (BB) in noise, (d) LP in noise and (e) HP in noise. Thin gray and black lines in each panel represent SRT-value variablity of HPD and non-HPD users across the sessions, respectively. Pooled average SRT values are shown with thick colored lines.

s2 follow-up session (r=0.68, $p < 0.00$, n=18), but disappeared after excluding the outliers (r=0.28, p=0.36, n=16). The second, positive correlation was found between the average audiometric thresholds of the HF region (2 to 8 kHz) and $SPiQ_{HP}$ in session s1 (r=0.67, $p < 0.00$, n=18; Table 6.8). However, this significant correlation vanished in the follow-up sessions. No significant correlations were observed in session s1 between RAM-EFR magnitudes, ABR wave-I or V amplitudes and either of the SRT-values in the different conditions.

6.5 Discussion

This study aimed to investigate the impact of recreational noise-exposure on potential biomarkers of CS in a group of 18 young normal-hearing subjects. In [176], we showed that lifetime noise-exposure history of our study cohort (calculated based on the questionnaires) does not correlate with either of the EFR, ABR, speech perception and audiogram metrics. Our inclusion criteria, specifically normal audiometric thresholds and the absence of hearing related symptoms, eliminated ears with OHC-dysfunction as much as possible. The assumption of a normal functioning cochlear amplifier facilitates interpretation of the other hearing-test results in

Table 6.7: Results of the correlation analysis. Pearson's r and p-values related to significant correlation between speech intelligibility SRT values and RAM-EFR, SAM-EFR, ABR wave-I and wave-V amplitudes are specified by bold letters.

Metric	Session	$SPiQ_{LP}$	$SPiQ_{HP}$	$SPiN_{BB}$	$SPiN_{LP}$	$SPiN_{HP}$
RAM	s1	r=0.25 p=0.31	r=0.21 p=0.40	r=0.16 p=0.52	r=0.31 p=0.20	r=0.34 p=0.17
	s2	r=-0.29 p=0.24	r=0.03 p=0.98	r=-0.04 p=0.86	r=0.23 p=0.36	r=-0.09 p=0.71
	s3	r=0.44 p=0.07	r=0.32 p=0.20	r=0.29 p=0.24	$\boldsymbol{r = 0.56}$ $\boldsymbol{p = 0.01}$	r=0.10 p=0.68
	s4	r=-0.04 p=0.88	r=-0.19 p=0.44	r=0.31 p=0.21	r=0.41 p=0.09	r=0.15 p=0.55
SAM	s1	r=0.15 p=0.54	r=0.23 p=0.36	r=0.11 p=0.67	r=0.40 p=0.10	$\boldsymbol{r = 0.61}$ $\boldsymbol{p < 0.00}$
	s2	r=0.45 p=0.06	r=0.36 p=0.14	r=0.15 p=0.54	r=-0.08 p=0.74	$\boldsymbol{r = 0.68}$ $\boldsymbol{p < 0.00}$
	s3	$\boldsymbol{r = 0.64}$ $\boldsymbol{p < 0.00}$	$\boldsymbol{r = 0.50}$ $\boldsymbol{p = 0.04}$	$\boldsymbol{r = 0.52}$ $\boldsymbol{p = 0.03}$	r=0.44 p=0.06	r=0.14 p=0.58
	s4	r=-0.17 p=0.50	r=-0.25 p=0.32	r=-0.04 p=0.86	r=0.07 p=0.77	r=-0.02 p=0.92
wave-I	s1	r=-0.02 p=0.93	r=0.02 p=0.95	r=-0.15 p=0.54	r=0.16 p=0.51	r=0.12 p=0.63
	s2	r=-0.15 p=0.54	r=-0.17 p=0.51	r=-0.08 p=0.74	r=-0.41 p=0.09	r=0.01 p=0.97
	s3	r=0.11 p=0.67	r=0.40 p=0.10	r=0.32 p=0.19	r=0.04 p=0.86	$\boldsymbol{r = 0.51}$ $\boldsymbol{p = 0.03}$
	s4	r=-0.08 p=0.73	r=0.27 p=0.27	r=0.13 p=0.60	r=0.08 p=0.74	r=0.03 p=0.90
wave-V	s1	r=0.38 p=0.12	r=-0.06 p=0.82	r=0.14 p=0.58	$\boldsymbol{r = 0.57}$ $\boldsymbol{p = 0.01}$	r=0.47 p=0.05
	s2	r=-0.04 p=0.87	r=-0.10 p=0.68	r=0.20 p=0.42	r=0.23 p=0.36	r=-0.17 p=0.49
	s3	$\boldsymbol{r = 0.53}$ $\boldsymbol{p = 0.02}$	r=0.27 p=0.28	r=0.30 p=0.22	$\boldsymbol{r = 0.53}$ $\boldsymbol{p = 0.02}$	r=-0.02 p=0.94
	s4	r=-0.01 p=0.95	r=0.19 p=0.44	r=0.16 p=0.51	r=0.24 p=0.34	r=-0.01 p=0.95

the context of hidden hearing-loss or CS. However, individual variabilities due to minor degrees of OHC-loss among the young normal-hearing listeners with audiometric thresholds within the normal range, should not be underestimated. Because thus far, the only available potential biomarkers of CS are based on indirect AEP measurements, and those are impacted by both CS and OHC-loss, and hence cannot be used for differential diagnosis of CS.

In Section 6.1, we hypothesized that even if audiometric thresholds remain unchanged or recover shortly after the recreational noise-exposure (TTS), metrics derived from electrophysiological measurements, i.e. EFRs and ABRs, and SRT

Table 6.8: Results of the correlation analysis. Pearson's r and p-values related to significant correlation between speech intelligibility SRT values and averaged PTA thresholds of three frequency regions are specified by bold letters.

Metric	Session	$SPiQ_{LP}$	$SPiQ_{HP}$	$SPiN_{BB}$	$SPiN_{LP}$	$SPiN_{HP}$
EHF	s1	r=-0.36 p=0.14	r=-0.05 p=0.83	r=0.15 p=0.56	r=-0.19 p=0.45	r=-0.04 p=0.87
	s2	r=-0.33 p=0.18	r=-0.31 p=0.21	r=0.02 p=0.93	r=0.03 p=0.92	r=-0.26 p=0.29
	s3	r=-0.18 p=0.49	r=-0.08 p=0.76	r=-0.15 p=0.57	r=-0.05 p=0.83	r=0.07 p=0.80
	s4	r=-0.01 p=0.96	r=-0.13 p=0.62	r=-0.17 p=0.50	r=0.07 p=0.79	r=0.32 p=0.20
HF	s1	r=-0.19 p=0.44	$\boldsymbol{r = 0.67}$ $\boldsymbol{p < 0.00}$	r=0.27 p=0.27	r=0.10 p=0.68	r=0.42 p=0.08
	s2	r=-0.24 p=0.34	r=0.21 p=0.39	$\boldsymbol{r = 0.55}$ $\boldsymbol{p = 0.02}$	r=0.38 p=0.12	r=-0.16 p=0.52
	s3	r=-0.20 p=0.43	r=0.27 p=0.30	r=-0.34 p=0.18	r=0.21 p=0.42	r=0.14 p=0.60
	s4	r=0.17 p=0.51	r=0.32 p=0.20	r=0.09 p=0.73	r=0.23 p=0.36	r=0.31 p=0.22
LF	s1	r=0.07 p=0.77	r=0.25 p=0.31	r=0.01 p=0.96	r=0.18 p=0.48	r=0.33 p=0.18
	s2	r=0.30 p=0.23	$\boldsymbol{r = 0.01}$ $\boldsymbol{p = 0.97}$	r=0.33 p=0.17	r=0.29 p=0.24	r=0.26 p=0.30
	s3	r=0.37 p=0.14	r=0.26 p=0.31	r=0.09 p=0.73	r=0.27 p=0.28	r=0.33 p=0.19
	s4	$\boldsymbol{r = 0.54}$ $\boldsymbol{p = 0.02}$	r=0.17 p=0.51	$\boldsymbol{r = 0.51}$ $\boldsymbol{p = 0.04}$	$\boldsymbol{r = 0.67}$ $\boldsymbol{p < 0.00}$	r=0.44 p=0.08

values might be affected. Consistent with our hypothesis in a group-level analysis, we did not observe any PTS at the tested frequencies (Figure 6.5). In the first follow-up session (s2: one day after attending festival), an average TTS of +3.33 dB-HL was observed at 10-kHz compared to the baseline session, which recovered within three days after attending the festival (according to the session s3 measurement). Based on findings in [52,160,162], the greatest auditory-threshold recovery occurs within two-to-four hours after the noise-exposure, and hence the possibility of a recovered TTS in the present study population should be considered.

In contrast with our hypothesis, RAM- and SAM-EFR magnitudes, as well as ABR wave-I and -V amplitudes group-means did not change between baseline session and sessions s2 to s4. Moreover, SRT group-means did not differ significantly one day after the noise-exposure at session s2, but improved when we compared those between baseline session and sessions s3 or s4. Gradually improved SRT values from session s1 to s4 were explained by the learning effect that participants had acquired after attending each experiment session. This fact also posed a limitation

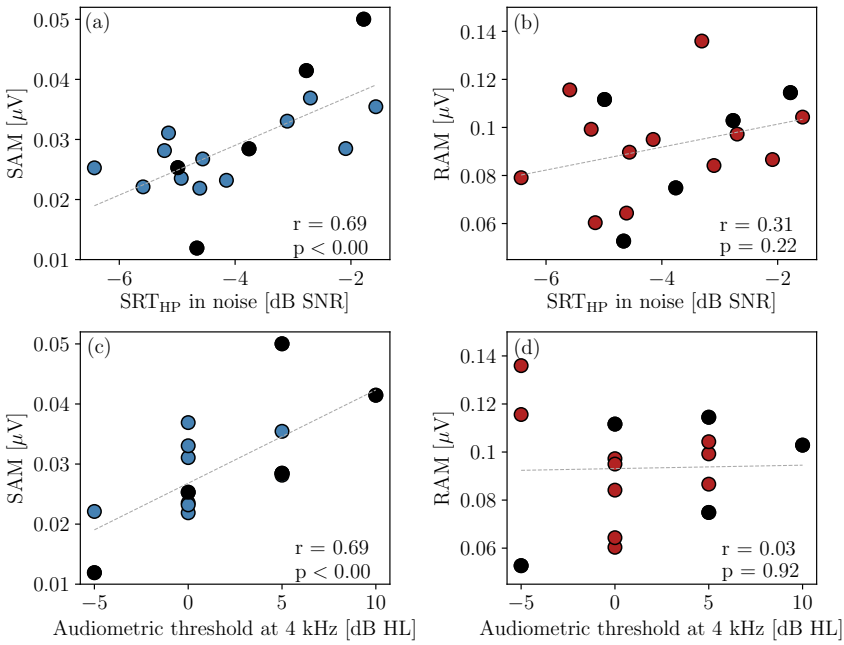


Figure 6.7: Correlation analysis: (a) correlation between SAM-EFR magnitude and $\mathrm{SPiN_{HP}}$*, (b) correlation between RAM-EFR magnitude and* $\mathrm{SPiN_{HP}}$*, (c) correlation between SAM-EFR magnitude and audiogram at 4 kHz, (d) correlation between RAM-EFR magnitude and audiogram at 4 kHz. Data-points corresponding to subjects that did not wear HPD during the festival, are specified by black circles*

to our correlation analysis, since we intended to investigate the effect of the noise dose of a music event on the relation between speech perception performance and EFR, ABR or audiogram. Although insignificant differences were found between SRT values of s1 and s2 sessions, the mixed effect of learning and noise-exposure may still confound the interpretation of the results. Thereby, despite of considering SRT values of all sessions in our correlation analysis (Table 6.7 and 6.8), we only relied on the Pearson's *r*- and *p*-values in the s1 session, where the results were not mixed-up with learning effect, and hence we could not draw any conclusion related to the effect of recreational noise-exposure on the relationship of SRT with other metrics.

In session s1, the observed significant correlation between SAM-EFR and $\mathrm{SPiN_{HP}}$ is inconsistent with a recent study of Garrett et al. [123], where SAM-EFR magnitudes negatively correlated with $\mathrm{SPiN_{HP}}$ SRT values. To further explore the source of this discrepancy, we considered the effect of OHC-loss on SAM-EFR metric, since a recent modelling and experimental study [44] showed

that SAM-EFR is affected by both, OHC-loss and CS. To this end, we compared the relationship of audiometric threshold at 4 kHz, the frequency which is in correspondence with the carrier frequency of SAM stimulus, with SAM-EFR (see Figure 6.7c) and $\mathrm{SPiN_{HP}}$. Correlation analysis revealed that the 4-kHz audiometric threshold significantly correlates with SAM-EFR magnitude (r=0.69, $\mathrm{p} < 0.00$, n=15) and $\mathrm{SPiN_{HP}}$ (r=0.51, p=0.04, n=17). These findings suggest that the strong correlation of SAM-EFR with $\mathrm{SPiN_{HP}}$ could be explained by the sensitivity of SAM-EFR to varying degrees of mild OHC-loss in the young normal-hearing cohort. In particular, obtained *positive* correlation coefficients indicate that increased audiometric thresholds yielded enhanced SAM-EFR magnitudes. This finding is consistent with model simulations in [44,83], where the authors showed that OHC-damage degrades cochlear input to IHC-AN complex and causes AN fibers to operate in a more sensitive amplitude modulation (AM)-coding region and generate enhanced SAM-EFR magnitudes compared to an intact cochlea, in absence of CS.

In this context, OHC-damage affected RAM-EFRs differently, since the sharp-rising stimulus envelope and short duty-cycle (25% duty cycle), limit the impact of reduced cochlear amplification on the response [44]. In line with [44], our RAM-EFR magnitudes in session s1, showed no significant correlation with audiometric thresholds at 4 kHz (See Figure 6.7d: r=0.03, p=0.92, n=15). However, the insignificant correlation of RAM-EFR with $\mathrm{SPiN_{HP}}$ (Figure 6.7b: r=0.31, p=0.22, n=17), is still in contrast with [123]. This discrepancy could arise from the target experimental group, since in [123], significant RAM-EFR and SRT correlations were observed when young and older normal-hearing groups were pooled together or when within the age controlled group (young normal-hearing), two subjects with high DPOAE thresholds were excluded.

6.6 Conclusion

We found no relation between the variability of potential AEP biomarkers of CS in normal-hearing young adults, after attending a music event. Baseline testing results and corresponding shifts from session s1 to session s2 did not differ between subjects. Thus, no strong conclusions can be made regarding the presence of noise-induced CS shortly after visiting a music festival in this cohort of young normal-hearing subjects. Possibly, the recreational noise-exposure dose in the present study was not sufficiently high to cause CS (76.2$\pm$7.82 dBA, ranged between 59.35 and 88.19 [176]). In addition, it is possible that employed biomarkers for CS are not sensitive or specific enough to detect or to monitor CS in normal-hearing young adults. The interpretation of speech intelligibility was complicated by a possible learning effect. Additionally, inter-subject variability due to head-size and sex [60] and intra-subject variability due to different degrees of awakeness [177] could have confounded the AEP-measurements.

Future studies may provide more insight into distinguishing between the absence of CS in the study cohort and a lack of specificity in the CS biomarkers. Hence, it is necessary to (1) further investigate inter-subject variability and test-retest reliability of the CS biomarkers in a cohort of normal-hearing subjects, (2) incorporate dosimetry and measure noise doses during the participation of subjects in noisy events, (3) consider logging of HPD-use during a noisy event and (4) run the hearing-tests immediately after the event. Finally, larger test groups will allow us to compare different possibly confounding factors, such as the use of HPD.

7

Overall Discussion and Conclusion

The doctoral research presented in this dissertation addressed the growing concern among the hearing research community regarding the availability of sensitive diagnostic tools that distinguish two sub-components of sensorineural hearing deficits, i.e. outer-hair-cell (OHC) loss and cochlear synaptopathy (CS), particularly when they co-exist. According to animal studies, CS targets auditory-nerve population and degrades supra-threshold sound encoding, while it leaves hearing thresholds unaffected. Additionally, CS precedes OHC-loss in the aging process [32, 34], and thus it is expected that a large population of people who have normal audiometric thresholds but complain about understanding speech in presence of noise, or those with impaired audiograms, may suffer from CS. With a perspective of rehabilitation through hearing aids, a gain prescription alone is not expected to enhance or restore the speech intelligibility of those people, which stresses the need for personalized audio-processing algorithms that compensate for different aspects of the hearing-loss. However, a deficit- and frequency- specific diagnostic tool of sensorineural hearing-loss (SNHL), which is the vital prerequisite for developing such individualized hearing restoration methods, was missing. Although a frequency-specific quantification of the OHC-loss aspect of SNHL is possible through distortion product otoacoustic emission (DPOAE) measurements or audiograms, CS quantification is not straightforward in live humans. The main reason that has thus far hindered the accomplishment of this task is the indirect nature of the available diagnostic tools for humans. While a direct estimation of auditory nerve (AN) damage (i.e. CS) in animal models is feasible through temporal bone histopathology and immunostaining, implementation of such invasive and direct

assessments is impossible in live humans. On the other hand, auditory evoked potentials (AEPs) as candidate non-invasive markers for SNHL diagnosis, are affected by both OHC-loss and CS, and hence their quality is compromised by the presence of mixed hearing pathologies. Therefore, while we ought to rely on these indirect and non-invasive measures for the SNHL diagnosis in humans, it is necessary to strengthen their discriminating power to isolate CS, while being maximally insensitive to possible co-existing OHC-damage.

Hitherto, animal studies have shown that the envelope following response (EFR) to a modulated stimulus is a robust marker of CS in the absence of OHC-loss, and respective changes due to CS are greater than that of the auditory brainstem responses (ABRs) [31, 32]. However, recent experimental studies on animal models (e.g. Parthasarathy et al. in [32]) and model simulations in [44], have shown that in addition to CS, co-occurring damage to OHCs can have an impact on EFR magnitudes. Furthermore, the spread of the basilar membrane (BM) excitation due to stimulation, confounds the frequency-specific diagnosis of the CS on the basis of EFRs [41, 73, 87, 100]. According to the experimental and modelling findings in Chapter 2 of this thesis, frequency sensitivity of the EFR to a broadband amplitude modulated stimulus is limited to frequencies above 2 kHz due to degraded amplitude-modulation (AM) coding at lower frequencies. In this regard, simulated EFR magnitudes fell in line with recorded EFRs and remained unchanged in response to broadband stimuli with lower cut-off frequencies below 2 kHz. Both simulations and recordings showed considerably reduced responses to stimuli with a bandwidth of [4-22] kHz. Model simulations exploring how different tonotopic regions of the cochlear partition contribute to the EFR generation, ascribed the lack of low-frequency AM-coding (below 2 kHz), to the relatively high modulation frequency of the stimulus, i.e. 120 Hz. Even though lowering the stimulus modulation frequency improved the AM-coding strength in both BM and AN stages of the model at low characteristic frequency (CF) channels, it is not recommended. Because the brain response to modulation frequencies below 70 Hz may involve generators from auditory processing levels beyond the brainstem, EFRs to lower modulation frequencies may be less sensitive to diagnose CS [117].

Having determined the sensitivity of EFRs to broadband stimuli, we aimed to develop a frequency-specific EFR-based metric, since the contributed off-CF channels to AM-coding can confound the frequency-specificity of the EFR as a diagnostic tool for CS (Chapter 2, Figure 2.2). With this purpose, Chapter 3 introduced the derived-band EFR (DBEFR), which is constructed by subsequent subtraction of EFRs to broadband stimuli with various bandwidths, akin a procedure earlier adopted for ABRs [61]. The DBEFR offers a frequency-specific metric in the [2-6] kHz frequency region (see Chapter 3 Figure 3.11) and minimizes possible effects of subject-specific factors, because it is based on a relative metric design. Experimentally, we found that DBEFR magnitudes extracted from

the [2-6] kHz frequency region, significantly decreased as a consequence of ageing (Experiment 2: younger normal-hearing (yNH) versus older normal-hearing (oNH)), elevated hearing thresholds (Experiment 2: oNH versus older hearing-impaired (oHI)), but not due to self-reported hearing difficulties (Experiment 1: NH versus self-reporteed hearing difficulties (NHSR)). To identify the sources of this variability, CS and OHC-loss effects on DBEFR magnitudes were simulated. Introducing CS to the model showed that DBEFR magnitudes extracted from the [2-6] kHz region reduce in absence of OHC-loss and applying OHC-loss alone increased the simulated DBEFR magnitudes. Even though CS had a much larger impact on the DBEFR magnitude than OHC-damage, the effect of the latter component was not negligible. The CS simulations were able to explain the significantly degraded DBEFR magnitudes of the oNH group. The suspected CS in those listeners could be related to either age-induced CS or accumulated lifelong noise-exposure, supported by human post-mortem studies [38–40, 133]. However, the significant difference between DBEFRs of oNH and oHI listeners were described based on both OHC-loss and CS, since model simulations showed that DBEFRs are affected by co-existing sources of SNHL. The DBEFR magnitude group-mean difference between NH and NHSR groups was insignificant. Moreover, no relationship was observed between the DBEFR magnitudes of NHSR group and respective noise-scores reported in individual questionnaires (Chapter 3, Figure 3.9), which might have stemmed from the variability in reporting lifetime noise-exposure dose [37, 48, 49].

Taken together, Chapter 3 showed that while DBEFR magnitudes extracted from [2-6] kHz can offer a frequency-specific diagnostic tool for CS in the presence of normal audiometric thresholds, they are not CS-specific when OHC-loss co-exists. A recent study by Vasilkov et al. [44], showed that the EFR to a rectangularly amplitude modulated (RAM) pure-tone, is a CS-specific metric, which is minimally affected by OHC-damage. The RAM-EFR, by virtue of the sharply rising stimulus envelope, yields a stronger EFR magnitude compared to that of the DBEFR. Nevertheless, the general sensitivity of AEPs to OHC deficits, even when this influence is very small (as in the RAM-EFR), may confound the accurate AEP-based CS diagnosis. Thus, with a view on precise AEP-based auditory profiling, it is necessary to combine AEPs with an another metric, which is explicitly sensitive to OHC-damage and at the same time frequency-specific. Thus far, audiograms and DPOAEs have been widely used for frequency-specific quantification of the OHC-damage in clinical applications and hence, when used together with AEP-based CS metrics they may be able to isolate the CS degree. To make this feasible, Chapter 4 took a model-based approach in combination with physiological measurements to build personalized SNHL (OHC-damage and CS) profiles. Specifically, the method was implemented following two fundamental steps: (1) perform a frequency-specific cochlear model individualization and (2) derive a

frequency-specific AN-pattern that in combination with (1), matches the recorded AEPs.

To develop individualized cochlear models, cochlear-gain-loss (CGL) parameters associated with OHC-damage were determined using either audiogram or DPOAE measurements. While the former was implemented by translating frequency-specific dB-HL into cochlear filter gain-loss, longitudinal cochlear filter coupling and associated gain propagation along the cochlear partition complicated DPOAE-based CGL parameters setting. To overcome this issue, a machine-learning approach was proposed that given the DPOAE thresholds/DP-grams recorded at certain frequencies, predicts corresponding CGL parameters across the CF channels. Adopting the developed machine-learning approach, the potential of DPOAE thresholds and DP-grams, as common objective metrics for diagnosis of OHC-damage in clinical applications, were assessed (Chapter 5). Following a thorough investigation, individualized cochlear models based on DP-grams measured at six primary levels were compared to those based on DPOAE thresholds. This evaluation assessed to which extent DPOAE threshold/DP-gram-based individualized cochleae were able to predict the measured (1) DPOAE thresholds/DP-gram and (2) audiogram. We showed that DPOAE-threshold-based cochlear model individualization performs equally well as the DP-gram method, when measured at low primary levels. Both approaches yielded the smallest error in predicting the respective DPOAE thresholds/DP-grams and audiogram, although audiogram prediction errors were overall higher. The observed smaller prediction errors for lower stimulation levels were consistent with the idea that DPOAEs measured at near-threshold primary levels are better indicators of OHC-damage [108]. On the other hand, the overall higher audiogram prediction errors were explained by considering the broadband characteristics of the AN excitation patterns used for predicting audiograms versus more localized responses measured by DPOAEs. Because DP-grams measured to low stimulation levels and DPOAE thresholds performed equally well, the latter was chosen for the developing individualized SNHL models, since it provides a metric that is calculated by involving both near- and supra-threshold stimulation levels. To compare DPOAE- and audiogram-based cochlear individualization methods, personalized cochlear models with DPOAE-thresholds (at four primary frequencies) and audiograms (at 12 frequencies) were adopted to build individualized SNHL profiles.

Having the personalized cochlear models, we simulated AN-damage patterns by altering the population and types of AN fibers, in the second step of the individualization procedure. Akin to the first step, this was implemented in a frequency-dependent manner, since the AN distribution [18,94] and the applied AN-fiber loss across the CF [133] were frequency-dependent.

To connect AEP measurements of study participants with simulated subject-specific AN-damage profiles, a range of AEP-derived metrics (e.g. RAM-EFR

strength, ABR amplitudes, latnecies and corresponding growth functions) were simulated. Afterwards, a *forward-backward* classification method was proposed to (i) determine which set of the AEP-derived metrics gives the highest accuracy in predicting individual CS profiles, (ii) cross-validate the predicted individualized CS profiles and (iii) specify which of the cochlear model individualization methods, either audiogram or DPOAE-thresholds, is more accurate to predict personalized SNHL profiles. In practice, the *forward-backward* classification method comprised two classifiers, of which the first was trained with individualized AEP-derived metrics simulated for different CS profiles, in order to determine which of the simulated CS profiles best matched the corresponding measured AEP-derived metrics (*forward* classification). Given that direct AN-synapse counts are not accessible for humans, the second classifier (*backward* classification) served as a validation tool. It was trained with the measured AEPs and CS-profile labels predicted by the first classifier in *forward* classification to predict the CS degree of simulated individualized AEPs. In this way, the proposed classification approach provided a tool to evaluate the accuracy of the adopted classifier in the *forward* classification, assess the efficiency of various sets of the AEP-derived metrics, and investigate the effect of different cochlear model individualization methods on auditory profiling.

Despite the small number of available DPOAE thresholds (i.e. four frequency points), setting the cochlear pole-functions based on DPOAE thresholds yielded higher CS-profile prediction accuracy than the audiogram-based method that used 12 frequency points. Focusing on the cochlear model individualization methods, this success could have stemmed from the notion that OAE is a more sensitive metric to deficits related to the cochlear-damage [135–138] and is not influenced by inner-hair-cell/AN damage [25], whereas the audiogram could have. Considering the accuracy of different AEP metrics in predicting individual CS profiles, the RAM-EFR outperformed other metrics, in both cochlear individualization methods. For a given individualized cochlear model, when low and medium spontaneous rate (MSR and LSR) AN fibers were removed (profile *A* in Figure 4.3), the magnitudes of simulated RAM-EFRs reduced by 15% compared to the normal-hearing profile (i.e. without CS). Besides the deletion of MSRs and LSRs, removing every 23% of the high spontaneous rate (HSR) AN fibers in profiles *B* to *E*, caused an additional 10% magnitude reduction of the simulated RAM-EFR magnitudes. In this regard, the proposed method yielded a 79.63% specificity in predicting simulated individualized CS profiles using a *backward* classifier. Consequently, the implemented method on yNH, oNH and oHI groups, suggested that RAM-EFR metrics combined with DPOAE-thresholds may provide a frequency and deficit-specific diagnostic tool to predict individualized SNHL profiles. Our method predicted that the yNH listeners of our study either had normal AN-fiber patterns across the CF or very mild degrees of CS, whereas higher degrees of CS

were predicted for oNH group participants. The latter can be ascribed to age-induced CS, which also confirms the findings in Chapter 3, where lower DBEFR magnitudes of the same oNH listeners were explained by their degraded supra-threshold envelope coding ability induced by age-related CS. In this regard, oHI listeners were predicted to suffer from severe degrees of CS. However, it is noteworthy that because of the limitation posed by the BM-filter gain in the model, simulating CGLs above 35 dB-HL was impossible. Hence, predicted CS profiles for individuals in the oHI group with audiometric losses greater than 35 dB HL, were less reliable.

By designing personalized classifiers, the implementation of the individualization process was further improved. Unlike the first implementation where we trained a single classifier for all individuals, in the improved version, subject-specific CS profiles were predicted using personalized classifiers, which were trained with simulated individualized RAM-EFRs. Even though the predicted CS profiles from either method did not differ considerably, it is expected that personalized classifiers might provide a better representation of the individualized CS profiles. Moreover, the proposed individualization technique was validated by applying the trained classifier (in the *backward* classification stage) to recorded RAM-EFRs of a new cohort of young normal-hearing listeners. This validation step used data that was not part of the training, and hence supports that the method could be generalizable to other recording setups and cohorts [43]. However, training the classifier with recordings of a larger and more diverse population, may provide more reliable prediction of individualized CS profiles. These promising results suggest that the RAM-EFR metric is sensitive to age-induced CS and that when combined with DPTH thresholds, can provide a reliable CS-specific measure in presence of OHC-loss. In addition, a recent study on the same dataset [123], showed that the RAM-EFR magnitude predicts the variation of speech-in-noise intelligibility across yNH, oNH and oHI participants of the experimental cohort. Furthermore, the sensitivity of the RAM-EFR magnitude to Kainic-acid (tested on budgerigar), i.e. ototoxic CS, has been confirmed in [123].

To further study the potential of the RAM-EFR as an AEP-based metric for CS diagnosis, its sensitivity to recreational noise-exposure was studied in Chapter 6. However, no significant difference was observed in RAM-EFR magnitudes of yNH listeners measured one day before attending a music festival and one, three and five days after. Besides the RAM-EFR, other measured metrics, i.e. the SAM-EFR magnitude, ABR wave-I and V amplitudes, standard and extended high frequency audiometric thresholds, neither changed significantly across the baseline and follow-up sessions. Albeit a 3.3 dB-HL threshold-shift at 10 kHz audiometric thresholds was observed from the baseline to the first follow-up session, this threshold-shift vanished in the third and forth follow-up sessions. These findings comply with the notion that humans are less sensitive to noise-induced

CS than animal models [18, 59]. Since, dosimetry was not performed during exposure to loud music, exact exposed noise levels during the festivals were not at hand. Hence, maybe the exposed noise actual levels (76.2±7.82 dBA [176]) were not high enough to cause synaptic losses [37]. Despite the use of two training lists in the speech intelligibility test conducted in the baseline session, significant learning effect was observed on speech reception thresholds (SRTs) in noise and quiet (i.e. lower SRT values) across the sessions [175]. It was of great interest to explore the effect of festival-related noise-exposure on the relationship between RAM-EFR strength and SRT values of the speech in quiet (SPiQ) and noise (SPiN) tests, but the observed learning effect hindered such type of investigation. Benefiting from the speech-test carried out in the baseline session, where the learning effect could not cause any confound on SRT values, the relation between SRTs and other measured metrics was studied. These analysis indicated a strong positive correlation between the high-pass SPiN condition and SAM-EFR magnitude, whereas no significant correlation was observed between RAM-EFR magnitude and the same speech intelligibility condition. These results are inconsistent with results of [123]. Further investigation revealed that the sensitivity of SAM-EFR metric to audiometric thresholds at 4 kHz, and at the same time the meaningful and positive correlation of audiometric threshold at 4 kHz and high-pass SPiN SRT values, have reflected on the significant correlation between SAM-EFR and high-pass SPiN. Different to the SAM-EFR, the RAM-EFR did not correlate to audiograms measured at 4 kHz, and confirmed that it is not affected by variable degrees of OHC-loss reflected on the audiometric thresholds within the normal range.

Limitations and future perspectives

The presented doctoral research in this thesis aimed to unravel the respective effect of SNHL sources on AEPs and develop objective diagnostic tools for CS, in presence of OHC-loss. Auditory profiling not only predicted individualized AN-damage patterns, but also for the first time provided individualized model of the auditory periphery based on AEP metrics. These personalized auditory periphery models can be used as a basis to develop hearing restoration algorithms after considering the impact of higher level auditory processing on AEPs, e.g. model-based speech enhancement methods that account for CS, as well as OHC-damage aspect of the SNHL. Nevertheless, there were some experimental and modelling limitations that constrained the efficiency of the proposed methods within the framework of this research. The following paragraphs address to those limitations and provide perspectives for the future.

- In the adopted computational model of the auditory periphery, the maximum

possible gain that can be applied to the BM filters to simulate CF-dependent CGL, is 35 dB. This limitation restricts the CGL simulations and makes it impossible to apply CGLs greater than 35 dB for listeners with audiometric thresholds above 35 dB HL. This issue can be addressed in the future by developing a BM filter-bank that provides gains higher than 35 dB, and at the same time remains stable [178]. In this regard, considering the effect of inner-hair-cell damage (100% loss of the AN fibers) may create additional gain-loss to simulate the larger threshold shifts than 35 dB HL. Incorporating the effect of IHC-damage in the model, will provide the extra benefit of simulating presbycusis, that affects the OHCs gain and receptor potential of IHCs. In this way, we will be able to compare the simulated sensory and neural presbycusis models with those proposed by Schuknecht et al. in [179].

- Using DPOAE thresholds measured at four primary frequencies, could have compromised the accuracy of the frequency-specific cochlear model individualization. Therefore, measuring them at more frequencies will provide more accurate individualized frequency-specific models.

- In the cochlear model, pole values of the BM admittance function change automatically as a function of stimulus level to simulate cochlear compression and level-dependent cochlear filter width. To develop the personalized cochlear model, we adjusted these pole values across the CF to simulate the effect of OHC-damage on cochlear filtering. Thus, by increasing pole values across the CF and leaving the level-dependent pole trajectory function untouched, the same applied stimulation level caused a reduced cochlear model output sensitivity, with a more linear I/O function. It is expected that considering the compression slope, as well as the BM filter pole-functions in the cochlear model individualization procedure, might improve the accuracy of the future individualized models of the auditory periphery.

- The high computational cost of running the adopted transmission-line (TL) cochlear model through the cochlear model individualization process, hindered simultaneous implementation of CGL and AN-damage individualization. However, executing cochlear model individualization process through the recently developed convolutional neural-network (CNN) based implementation of the TL cochlear model [180], will offer a faster procedure.

- The insignificant effect of the recreational noise-exposure on the potential biomarkers of CS observed in Chapter 6, could be related to the recruited small population of yNH listeners. This prevented a statistical meaningful comparison between measurements of hearing protection device (HPD) users and those who did not use HPD. Recruiting larger test groups will

provide the required statistical power to investigate the effect of such confounding factors.

- It is worthwhile to further investigate the RAM stimulus, since among the hitherto studied AEP-based metrics, the RAM-EFR metric has shown to be most affected by CS, and the least sensitive to OHC-loss. Hence, running model simulation for carriers at lower and higher frequencies can provide an insight into its sensitivity to SNHL sub-types at other frequencies and offer an informative tool for quantifying supra-threshold envelope coding at lower and higher frequencies than 4 kHz. In addition, conducting test-retest measurements is a necessary step to assess the reliability and robustness of the RAM-EFR metric, recorded to stimuli with different carrier frequencies. The results of such measurements combined with the validation using animal histopathology studies will play a significant role in future clinical applications.

References

[1] Lisa L Cunningham and Debara L Tucci. *Hearing loss in adults*. New England Journal of Medicine, 377(25):2465–2473, 2017.

[2] World Health Organization et al. *Hearing loss due to recreational exposure to loud sounds: a review*. World Health Organization, 2015.

[3] World Health Organization et al. *Addressing the rising prevalence of hearing loss*. World Health Organization, 2018.

[4] James A Seddon, Peter Godfrey-Faussett, Kayleen Jacobs, Adam Ebrahim, Anneke C Hesseling, and H Simon Schaaf. *Hearing loss in patients on treatment for drug-resistant tuberculosis*. European Respiratory Journal, 40(5):1277–1286, 2012.

[5] Mark J McKeage. *Comparative adverse effect profiles of platinum drugs*. Drug safety, 13(4):228–244, 1995.

[6] Debashree Mukherjea, Leonard P Rybak, Kelly E Sheehan, Tejbeer Kaur, Vickram Ramkumar, Sarvesh Jajoo, and Sandeep Sheth. *The design and screening of drugs to prevent acquired sensorineural hearing loss*. Expert opinion on drug discovery, 6(5):491–505, 2011.

[7] Jukka Starck, Esko Toppila, and Ilmari Pyykkö. *Smoking as a risk factor in sensory neural hearing loss among workers exposed to occupational noise*. Acta oto-laryngologica, 119(3):302–305, 1999.

[8] Karen J Cruickshanks, David M Nondahl, Dayna S Dalton, Mary E Fischer, Barbara EK Klein, Ronald Klein, F Javier Nieto, Carla R Schubert, and Ted S Tweed. *Smoking, central adiposity, and poor glycemic control increase risk of hearing impairment*. Journal of the American Geriatrics Society, 63(5):918–924, 2015.

[9] Jos J Eggermont. *Hearing loss: Causes, prevention, and treatment*. Academic Press, 2017.

[10] James A Stankiewicz and Harris J Mowry. *Clinical accuracy of tuning fork tests*. The Laryngoscope, 89(12):1956–1963, 1979.

[11] EW Johnson. *Tuning forks to audiometers and back again*. The Laryngoscope, 80(1):49–68, 1970.

[12] Brian R Glasberg and Brian CJ Moore. *Auditory filter shapes in subjects with unilateral and bilateral cochlear impairments*. The Journal of the Acoustical Society of America, 79(4):1020–1033, 1986.

[13] Sharon G Kujawa and M Charles Liberman. *Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss*. Journal of Neuroscience, 29(45):14077–14085, 2009.

[14] Enrique A Lopez-Poveda and Pablo Barrios. *Perception of stochastically undersampled sound waveforms: a model of auditory deafferentation*. Frontiers in neuroscience, 7:124, 2013.

[15] Roland Schaette and David McAlpine. *Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model*. Journal of Neuroscience, 31(38):13452–13457, 2011.

[16] Adam C Furman, Sharon G Kujawa, and M Charles Liberman. *Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates*. Journal of neurophysiology, 110(3):577–586, 2013.

[17] Harrison W Lin, Adam C Furman, Sharon G Kujawa, and M Charles Liberman. *Primary neural degeneration in the Guinea pig cochlea after reversible noise-induced threshold shift*. Journal of the Association for Research in Otolaryngology, 12(5):605–616, 2011.

[18] MD Valero, JA Burton, SN Hauser, TA Hackett, R Ramachandran, and MC Liberman. *Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta)*. Hearing research, 353:213–223, 2017.

[19] Edward Lobarinas, Christopher Spankovich, and Colleen G Le Prell. *Evidence of "hidden hearing loss" following noise exposures that produce robust TTS and ABR wave-I amplitude reductions*. Hearing research, 349:155–163, 2017.

[20] Kathleen CM Campbell, James Kalkanis, and F Robert Glatz. *Ototoxicity: mechanisms, protective agents, and monitoring*. Current Opinion in Otolaryngology & Head and Neck Surgery, 8(5):436–440, 2000.

[21] Harold F Schuknecht. *Presbycusis*. The Laryngoscope, 65(6):402–419, 1955.

[22] HAROLD F SCHUKNECHT. *Further observations on the pathology of presbycusis*. Archives of otolaryngology, 80(4):369–382, 1964.

[23] Harold F Schuknecht, Kozo Watanuki, Tadahiko Takahashi, A Aziz Belal Jr, Robert S Kimura, Diane Deleo Jones, and Carol Y Ota. *Atrophy of the stria vascularis, a common cause for hearing loss.* The Laryngoscope, 84(10):1777–1821, 1974.

[24] David T Kemp. *Stimulated acoustic emissions from within the human auditory system*. The Journal of the Acoustical Society of America, 64(5):1386–1391, 1978.

[25] P Trautwein, P Hofstetter, J Wang, R Salvi, and Al Nostrant. *Selective inner hair cell loss does not alter distortion product otoacoustic emissions*. Hearing research, 96(1-2):71–82, 1996.

[26] Christopher A Shera and John J Guinan. *Mechanisms of mammalian otoacoustic emission*. In Active processes and otoacoustic emissions in hearing, pages 305–342. Springer, 2008.

[27] Sarah Verhulst, James M Harte, and Torsten Dau. *Temporal suppression of the click-evoked otoacoustic emission level-curve*. The Journal of the Acoustical Society of America, 129(3):1452–1463, 2011.

[28] Patricia A Dorn, Dawn Konrad-Martin, Stephen T Neely, Douglas H Keefe, Emily Cyr, and Michael P Gorga. *Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears*. The Journal of the Acoustical Society of America, 110(6):3119–3131, 2001.

[29] ISAO Hashimoto, YOJI Ishiyama, TOMONOBU Yoshimoto, and SHIGERU Nemoto. *Brain-stem auditory-evoked potentials recorded directly from human brain-stem and thalamus.* Brain: a journal of neurology, 104(Pt 4):841–859, 1981.

[30] Aage R Møller and Peter J Jannetta. *Compound action potentials recorded intracranially from the auditory nerve in man*. Experimental neurology, 74(3):862–874, 1981.

[31] Luke A Shaheen, Michelle D Valero, and M Charles Liberman. *Towards a diagnosis of cochlear neuropathy with envelope following responses*. Journal of the Association for Research in Otolaryngology, 16(6):727–745, 2015.

[32] Aravindakshan Parthasarathy and Sharon G Kujawa. *Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing*. Journal of Neuroscience, 38(32):7108–7119, 2018.

[33] Sharon G Kujawa and M Charles Liberman. *Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss*. Hearing research, 330:191–199, 2015.

[34] Yevgeniya Sergeyenko, Kumud Lall, M Charles Liberman, and Sharon G Kujawa. *Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline*. Journal of Neuroscience, 33(34):13686–13694, 2013.

[35] Christopher J Plack, Agnès Léger, Garreth Prendergast, Karolina Kluk, Hannah Guest, and Kevin J Munro. *Toward a diagnostic test for hidden hearing loss*. Trends in hearing, 20, 2016.

[36] Hari M Bharadwaj, Alexandra R Mai, Jennifer M Simpson, Inyong Choi, Michael G Heinz, and Barbara G Shinn-Cunningham. *Non-invasive assays of cochlear synaptopathy–candidates and considerations*. Neuroscience, 407:53–66, 2019.

[37] Naomi Bramhall, Elizabeth Francis Beach, Bastian Epp, Colleen G Le Prell, Enrique A Lopez-Poveda, Christopher J Plack, Roland Schaette, Sarah Verhulst, and Barbara Canlon. *The search for noise-induced cochlear synaptopathy in humans: Mission impossible?* Hearing Research, 377:88–103, 2019.

[38] Chadi A Makary, Jennifer Shin, Sharon G Kujawa, M Charles Liberman, and Saumil N Merchant. *Age-related primary cochlear neuronal degeneration in human temporal bones*. Journal of the Association for Research in Otolaryngology, 12(6):711–717, 2011.

[39] Lucas M Viana, Jennifer T O'Malley, Barbara J Burgess, Dianne D Jones, Carlos ACP Oliveira, Felipe Santos, Saumil N Merchant, Leslie D Liberman, and M Charles Liberman. *Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue*. Hearing research, 327:78–88, 2015.

[40] PZ Wu, LD Liberman, K Bennett, V De Gruttola, JT O'Malley, and MC Liberman. *Primary neural degeneration in the human cochlea: evidence for hidden hearing loss in the aging ear*. Neuroscience, 407:8–20, 2019.

[41] Hari M Bharadwaj, Salwa Masud, Golbarg Mehraei, Sarah Verhulst, and Barbara G Shinn-Cunningham. *Individual differences reveal correlates of hidden hearing deficits*. Journal of Neuroscience, 35(5):2161–2172, 2015.

[42] Markus Garrett and Sarah Verhulst. *Applicability of subcortical EEG metrics of synaptopathy to older listeners with impaired audiograms*. Hearing research, 380:150–165, 2019.

[43] Sarineh Keshishzadeh, Markus Garrett, and Sarah Verhulst. *Towards Personalized Auditory Models: Predicting Individual Sensorineural Hearing-Loss Profiles From Recorded Human Auditory Physiology*. Trends in Hearing, 25, 2021.

[44] Viacheslav Vasilkov, Markus Garrett, Manfred Mauermann, and Sarah Verhulst. *Enhancing the sensitivity of the envelope-following response for cochlear synaptopathy screening in humans: The role of stimulus envelope*. Hearing Research, 400:108132, 2021.

[45] Joaquin T Valderrama, Elizabeth Francis Beach, Ingrid Yeend, Mridula Sharma, Bram Van Dun, and Harvey Dillon. *Effects of lifetime noise exposure on the middle-age human auditory brainstem response, tinnitus and speech-in-noise intelligibility*. Hearing research, 365:36–48, 2018.

[46] Greta C Stamper and Tiffany A Johnson. *Auditory function in normal-hearing, noise-exposed human ears*. Ear and hearing, 36(2):172, 2015.

[47] Greta C Stamper and Tiffany A Johnson. *Letter to the Editor: Examination of potential sex influences in Stamper, GC & Johnson, TA (2015). Auditory function in normal-hearing, noise-exposed human ears, Ear Hear, 36, 172-184*. Ear and hearing, 36(6):738, 2015.

[48] Naomi F Bramhall, Dawn Konrad-Martin, Garnett P McMillan, and Susan E Griest. *Auditory brainstem response altered in humans with noise exposure despite normal outer hair cell function*. Ear and hearing, 38(1), 2017.

[49] Garreth Prendergast, Hannah Guest, Kevin J Munro, Karolina Kluk, Agnès Léger, Deborah A Hall, Michael G Heinz, and Christopher J Plack. *Effects of noise exposure on young adults with normal audiograms I: Electrophysiology*. Hearing research, 344:68–81, 2017.

[50] Garreth Prendergast, Wenhe Tu, Hannah Guest, Rebecca E Millman, Karolina Kluk, Samuel Couth, Kevin J Munro, and Christopher J Plack. *Supra-threshold auditory brainstem response amplitudes in humans: Test-retest reliability, electrode montage and noise exposure*. Hearing research, 364:38–47, 2018.

[51] Hannah Guest, Kevin J Munro, Garreth Prendergast, Rebecca E Millman, and Christopher J Plack. *Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure*. Hearing research, 364:142–151, 2018.

[52] Sarah K Grinn, Kathryn B Wiseman, Jason A Baker, and Colleen G Le Prell. *Hidden hearing loss? No effect of common recreational noise exposure on cochlear nerve response amplitude in humans*. Frontiers in neuroscience, 11:465, 2017.

[53] Christopher Spankovich, Colleen G Le Prell, Edward Lobarinas, and Linda J Hood. *Noise history and auditory function in young adults with and without type 1 diabetes mellitus*. Ear and hearing, 38(6):724–735, 2017.

[54] John H Grose, Emily Buss, and Joseph W Hall III. *Loud music exposure and cochlear synaptopathy in young adults: Isolated auditory brainstem response effects but no perceptual consequences*. Trends in hearing, 21, 2017.

[55] Hannah Guest, Kevin J Munro, and Christopher J Plack. *Tinnitus with a normal audiogram: Role of high-frequency sensitivity and reanalysis of brainstem-response measures to avoid audiometric over-matching*. Hearing research, 356:116, 2017.

[56] Ingrid Yeend, Elizabeth Francis Beach, Mridula Sharma, and Harvey Dillon. *The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise*. Hearing Research, 353:224–236, 2017.

[57] Lijie Liu, Hui Wang, Lijuan Shi, Awad Almuklass, Tingting He, Steve Aiken, Manohar Bance, Shankai Yin, and Jian Wang. *Silent damage of noise on cochlear afferent innervation in guinea pigs and the impact on temporal processing*. PLoS One, 7(11):e49550, 2012.

[58] Lijuan Shi, Yin Chang, Xiaowei Li, Steven J Aiken, Lijie Liu, and Jian Wang. *Coding deficits in noise-induced hidden hearing loss may stem from incomplete repair of ribbon synapses in the cochlea*. Frontiers in neuroscience, 10:231, 2016.

[59] Robert A Dobie and Larry E Humes. *Commentary on the regulatory implications of noise-induced cochlear neuropathy*. International journal of audiology, 56(sup1):74–78, 2017.

[60] Curt Mitchell, David S Phillips, and Dennis R Trune. *Variables affecting the auditory brainstem response: audiogram, age, gender and head size*. Hearing research, 40(1-2):75–85, 1989.

[61] Metal Don and JJ Eggermont. *Analysis of the click-evoked brainstem potentials in man using high-pass noise masking*. The journal of the a stical society of America, 63(4):1084–1092, 1978.

[62] Hiroaki Sato, Isamu Sando, and Haruo Takahashi. *Sexual dimorphism and development of the human cochlea: computer 3-D measurement*. Acta otolaryngologica, 111(6):1037–1040, 1991.

[63] Peter Dallos, Jing Zheng, and Mary Ann Cheatham. *Prestin and the cochlear amplifier*. The Journal of physiology, 576(1):37–42, 2006.

[64] Sarah Verhulst, Anoop Jagadeesh, Manfred Mauermann, and Frauke Ernst. *Individual differences in auditory brainstem response wave characteristics: relations to different aspects of peripheral hearing loss*. Trends in hearing, 20, 2016.

[65] Hari M Bharadwaj and Barbara G Shinn-Cunningham. *Rapid acquisition of auditory subcortical steady state responses using multichannel recordings*. Clinical Neurophysiology, 125(9):1878–1888, 2014.

[66] Jianwen Wendy Gu, Barbara S Herrmann, Robert A Levine, and Jennifer R Melcher. *Brainstem auditory evoked potentials suggest a role for the ventral cochlear nucleus in tinnitus*. Journal of the Association for Research in Otolaryngology, 13(6):819–833, 2012.

[67] Ann E Hickox and M Charles Liberman. *Is noise-induced cochlear neuropathy key to the generation of hyperacusis or tinnitus?* Journal of neurophysiology, 111(3):552–564, 2014.

[68] Alfred C Coats and James L Martin. *Human auditory nerve action potentials and brain stem evoked responses: effects of audiogram shape and lesion location*. Archives of otolaryngology, 103(10):605–622, 1977.

[69] C Elberling and J Parbo. *Reference data for ABRs in retrocochlear diagnosis*. Scandinavian Audiology, 16(1):49–55, 1987.

[70] David Robert Watson. *The effects of cochlear hearing loss, age and sex on the auditory brainstem response*. Audiology, 35(5):246–258, 1996.

[71] Stephan D Ewert and Torsten Dau. *Characterizing frequency selectivity for envelope fluctuations*. The Journal of the Acoustical Society of America, 108(3):1181–1196, 2000.

[72] Michael G Heinz, Xuedong Zhang, Ian C Bruce, and Laurel H Carney. *Auditory nerve model for predicting performance limits of normal and impaired listeners*. Acoustics Research Letters Online, 2(3):91–96, 2001.

[73] Sarah Verhulst, Alessandro Altoè, and Viacheslav Vasilkov. *Computational modeling of the human auditory periphery: Auditory-nerve responses, evoked potentials and hearing loss*. Hearing research, 360:55–75, 2018.

[74] Muhammad SA Zilany and Ian C Bruce. *Modeling auditory-nerve responses for high sound pressure levels in the normal and impaired auditory periphery*. The Journal of the Acoustical Society of America, 120(3):1446–1466, 2006.

[75] Morten L Jepsen, Stephan D Ewert, and Torsten Dau. *A computational model of human auditory signal processing and perception*. The Journal of the Acoustical Society of America, 124(1):422–438, 2008.

[76] Morten L Jepsen and Torsten Dau. *Characterizing auditory processing and perception in individual listeners with sensorineural hearing loss*. The Journal of the Acoustical Society of America, 129(1):262–281, 2011.

[77] Thomas Rohdenburg, Volker Hohmann, and Birger Kollmeier. *Objective perceptual quality measures for the evaluation of noise reduction schemes*. In 9th international workshop on acoustic echo and noise control, pages 169–172, 2005.

[78] Manasa R Panda, Wendy Lecluyse, Christine M Tan, Tim Jurgens, and Ray Meddis. *Hearing dummies: Individualized computer models of hearing impairment*. International journal of audiology, 53(10):699–709, 2014.

[79] Sarah Verhulst, Torsten Dau, and Christopher A Shera. *Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission*. The Journal of the Acoustical Society of America, 132(6):3842–3848, 2012.

[80] Alessandro Altoè, Ville Pulkki, and Sarah Verhulst. *Transmission line cochlear models: improved accuracy and efficiency*. The Journal of the Acoustical Society of America, 136(4):EL302–EL308, 2014.

[81] Alessandro Altoè, Ville Pulkki, and Sarah Verhulst. *The effects of the activation of the inner-hair-cell basolateral K+ channels on auditory nerve responses*. Hearing research, 364:68–80, 2018.

[82] Paul C Nelson and Laurel H Carney. *A phenomenological model of peripheral and central neural responses to amplitude-modulated tones*. The Journal of the Acoustical Society of America, 116(4):2173–2186, 2004.

[83] Viacheslav Vasilkov and Sarah Verhulst. *Towards a differential diagnosis of cochlear synaptopathy and outer-hair-cell deficits in mixed sensorineural hearing loss pathologies*. medRxiv, 2019(preprint).

[84] James Jerger and Karen Johnson. *Interactions of age, gender, and sensorineural hearing loss on ABR latency*. Ear and hearing, 9(4):168–176, 1988.

[85] Michael P Gorga, Don W Worthington, Jan K Reiland, Kathryn A Beauchaine, and David E Goldgar. *Some comparisons between auditory brain stem response thresholds, latencies, and the pure-tone audiogram.* Ear and Hearing, 6(2):105–112, 1985.

[86] O Yamada, K Kodera, and T Yagi. *Cochlear processes affecting wave V latency of the auditory evoked brain stem response: A study of patients with sensory hearing loss*. Scandinavian audiology, 8(2):67–70, 1979.

[87] Hari M Bharadwaj, Sarah Verhulst, Luke Shaheen, M Charles Liberman, and Barbara G Shinn-Cunningham. *Cochlear neuropathy and the coding of supra-threshold sound*. Frontiers in systems neuroscience, 8:26, 2014.

[88] Christopher J Plack, Vit Drga, and Enrique A Lopez-Poveda. *Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss*. The Journal of the Acoustical Society of America, 115(4):1684–1695, 2004.

[89] Raul Sanchez Lopez, Federica Bianchi, Michal Fereczkowski, Sebastien Santurette, and Torsten Dau. *Data-driven approach for auditory profiling and characterization of individual hearing loss*. Trends in hearing, 22, 2018.

[90] Raul Sanchez-Lopez, Michal Fereczkowski, Tobias Neher, Sébastien Santurette, and Torsten Dau. *Robust data-driven auditory profiling towards precision audiology*. Trends in hearing, 24:2331216520973539, 2020.

[91] Dennis R Trune, Curt Mitchell, and David S Phillips. *The relative importance of head size, gender and age on the auditory brainstem response*. Hearing research, 32(2-3):165–174, 1988.

[92] Armin Kohlrausch, Ralf Fassel, and Torsten Dau. *The influence of carrier level and frequency on modulation and beat-detection thresholds for sinusoidal carriers*. The Journal of the Acoustical Society of America, 108(2):723–734, 2000.

[93] David W Purcell, Sasha M John, Bruce A Schneider, and Terence W Picton. *Human temporal auditory acuity as assessed by envelope following responses*. The Journal of the Acoustical Society of America, 116(6):3581–3593, 2004.

[94] Sarineh Keshishzadeh, Markus Garrett, Viacheslav Vasilkov, and Sarah Verhulst. *The derived-band envelope following response and its sensitivity to sensorineural hearing deficits*. Hearing Research, page 107979, 2020.

[95] Li Zhu, Hari Bharadwaj, Jing Xia, and Barbara Shinn-Cunningham. *A comparison of spectral magnitude and phase-locking value analyses of the frequency-following response to complex tones*. The Journal of the Acoustical Society of America, 134(1):384–395, 2013.

[96] PX Joris, CE Schreiner, and A Rees. *Neural processing of amplitude-modulated sounds*. Physiological reviews, 84(2):541–577, 2004.

[97] M Charles Liberman and Sharon G Kujawa. *Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms*. Hearing research, 349:138–147, 2017.

[98] Katharine A Fernandez, Penelope WC Jeffers, Kumud Lall, M Charles Liberman, and Sharon G Kujawa. *Aging after noise exposure: acceleration of cochlear synaptopathy in "recovered" ears*. Journal of Neuroscience, 35(19):7509–7520, 2015.

[99] Philip X Joris and Tom CT Yin. *Responses to amplitude-modulated tones in the auditory nerve of the cat*. The Journal of the Acoustical Society of America, 91(1):215–232, 1992.

[100] Gerard Encina-Llamas, James M Harte, Torsten Dau, Barbara Shinn-Cunningham, and Bastian Epp. *Investigating the Effect of Cochlear Synaptopathy on Envelope Following Responses Using a Model of the Auditory Nerve*. Journal of the Association for Research in Otolaryngology, pages 1–20, 2019.

[101] Golbarg Mehraei, Ann E Hickox, Hari M Bharadwaj, Hannah Goldberg, Sarah Verhulst, M Charles Liberman, and Barbara G Shinn-Cunningham. *Auditory brainstem response latency in noise as a marker of cochlear synaptopathy*. Journal of Neuroscience, 36(13):3755–3764, 2016.

[102] JJ Eggermont. *Analysis of compound action potential responses to tone bursts in the human and guinea pig cochlea*. The Journal of the Acoustical Society of America, 60(5):1132–1139, 1976.

[103] Daniel M Rasetshwane, Michael Argenyi, Stephen T Neely, Judy G Kopun, and Michael P Gorga. *Latency of tone-burst-evoked auditory brain stem responses and otoacoustic emissions: Level, frequency, and rise-time effects*. The Journal of the Acoustical Society of America, 133(5):2803–2817, 2013.

[104] Carolina Abdala and Richard C Folsom. *The development of frequency resolution in humans as revealed by the auditory brain-stem response recorded with notched-noise masking*. The Journal of the Acoustical Society of America, 98(2):921–930, 1995.

[105] Jerome Bourien, Yong Tang, Charlene Batrel, Antoine Huet, Marc Lenoir, Sabine Ladrech, Gilles Desmadryl, Regis Nouvian, Jean-Luc Puel, and Jing Wang. *Contribution of auditory nerve fibers to compound action potential of the auditory nerve*. Journal of neurophysiology, 112(5):1025–1039, 2014.

[106] Sarah Verhulst, Frauke Ernst, Markus Garrett, and Viacheslav Vasilkov. *Suprathreshold psychoacoustics and envelope-following response relations: Normal-hearing, synaptopathy and cochlear gain loss*. Acta Acustica united with Acustica, 104(5):800–803, 2018.

[107] Sofie Degeest, Paul Corthals, Bart Vinck, and Hannah Keppler. *Prevalence and characteristics of tinnitus after leisure noise exposure in young adults*. NOISE and HEALTH, 16(68):26–33, 2014.

[108] Peter Kummer, Thomas Janssen, and Wolfgang Arnold. *The level and growth behavior of the 2 f1- f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss*. The Journal of the Acoustical Society of America, 103(6):3431–3444, 1998.

[109] Manfred Mauermann. *Improving the usability of the distortion product otoacoustic emisssions (DPOAE)-sweep method: An alternative artifact rejection and noise-floor estimation*. In Proceedings of Meetings on Acoustics ICA2013, volume 19. ASA, 2013.

[110] Glenis R Long, Carrick L Talmadge, and Jungmee Lee. *Measuring distortion product otoacoustic emissions using continuously sweeping primaries*. The Journal of the Acoustical Society of America, 124(3):1613–1626, 2008.

[111] Gyorgy Buzsaki. *Rhythms of the Brain*, chapter Perceptions and Actions Are Brain-State Dependent. Oxford University Press, 2006.

[112] Donald D Greenwood. *A cochlear frequency-position function for several species—29 years later*. The Journal of the Acoustical Society of America, 87(6):2592–2605, 1990.

[113] Terence W Picton. *Human auditory evoked potentials*. Plural Publishing, 2010.

[114] M Charles Liberman. *Auditory-nerve response from cats raised in a low-noise chamber*. The Journal of the Acoustical Society of America, 63(2):442–455, 1978.

[115] Christopher A Shera, John J Guinan, and Andrew J Oxenham. *Otoacoustic estimation of cochlear tuning: validation in the chinchilla*. Journal of the Association for Research in Otolaryngology, 11(3):343–365, 2010.

[116] Steven S Coughlin. *Recall bias in epidemiologic studies*. Journal of clinical epidemiology, 43(1):87–91, 1990.

[117] Sarineh Keshishzadeh, Viacheslav Vasilkov, and Sarah Verhulst. *Tonotopic sensitivity to supra-threshold hearing deficits of the envelope following response evoked by broadband stimuli*. In 23rd International Congress on Acoustics (ICA 2019), pages 6513–6518, 2019.

[118] Brian CJ Moore and Brian R Glasberg. *Suggested formulae for calculating auditory-filter bandwidths and excitation patterns*. The journal of the acoustical society of America, 74(3):750–753, 1983.

[119] Anthony T Herdman and David R Stapells. *Auditory steady-state response thresholds of adults with sensorineural hearing impairments: Umbrales de las respuestas auditivas de estado estable en adultos con hipoacusia sensorineural*. International Journal of Audiology, 42(5):237–248, 2003.

[120] Guang-Di Chen, Chiemi Tanaka, and Donald Henderson. *Relation between outer hair cell loss and hearing loss in rats exposed to styrene*. Hearing research, 243(1-2):28–34, 2008.

[121] Ann E Hickox, Erik Larsen, Michael G Heinz, Leslie Shinobu, and Jonathon P Whitton. *Translational issues in cochlear synaptopathy*. Hearing research, 349:164–171, 2017.

[122] Stephan D Ewert, Steffen Kortlang, and Volker Hohmann. *A model-based hearing aid: Psychoacoustics, models and algorithms*. In Proceedings of Meetings on Acoustics ICA2013, volume 19, page 050187. Acoustical Society of America, 2013.

[123] Markus Garrett, Viacheslav Vasilkov, Manfred Mauermann, John L Wilson, Kenneth S Henry, and Sarah Verhulst. *Speech-in-noise intelligibility difficulties with age: the role of cochlear synaptopathy*. bioRxiv, 2020.

[124] Anna R Chambers, Jennifer Resnik, Yasheng Yuan, Jonathon P Whitton, Albert S Edge, M Charles Liberman, and Daniel B Polley. *Central gain restores auditory processing following near-complete cochlear denervation*. Neuron, 89(4):867–879, 2016.

[125] Dorit Mohrle, Kun Ni, Ksenya Varakina, Dan Bing, Sze Chim Lee, Ulrike Zimmermann, Marlies Knipper, and Lukas Ruttiger. *Loss of auditory sensitivity from inner hair cell synaptopathy can be centrally compensated in the young but not old brain*. Neurobiology of aging, 44:173–184, 2016.

[126] Kenneth S Henry and Kristina S Abrams. *Persistent auditory nerve damage following kainic acid excitotoxicity in the budgerigar (melopsittacus undulatus)*. Journal of the Association for Research in Otolaryngology, 19(4):435–449, 2018.

[127] Tine Goossens, Charlotte Vercammen, Jan Wouters, and Astrid van Wieringen. *Aging affects neural synchronization to speech-related acoustic modulations*. Frontiers in aging neuroscience, 8:133, 2016.

[128] Aravindakshan Parthasarathy, Edward L Bartlett, and Sharon G Kujawa. *Age-related changes in neural coding of envelope cues: peripheral declines and central compensation*. Neuroscience, 407:21–31, 2019.

[129] Aravindakshan Parthasarathy, Björn Herrmann, and Edward L Bartlett. *Aging alters envelope representations of speech-like sounds in the inferior colliculus*. Neurobiology of aging, 73:30–40, 2019.

[130] Paul Boege and Thomas Janssen. *Pure-tone threshold estimation from extrapolated distortion product otoacoustic emission I/O-functions in normal and cochlear hearing loss ears*. The Journal of the Acoustical Society of America, 111(4):1810–1818, 2002.

[131] Markus Garrett, Stefan Debener, and Sarah Verhulst. *Acquisition of subcortical auditory potentials with around-the-ear cEEGrid technology in normal and hearing impaired listeners.* Frontiers in neuroscience, 13:730, 2019.

[132] Alejandro Osses and Sarah Verhulst. *Calibration and reference simulations for the auditory periphery model of Verhulst et al. 2018 version 1.2*. arXiv preprint arXiv:1912.10026, 2019.

[133] Pei-zhe Wu, Jennifer T O'Malley, Victor de Gruttola, and M Charles Liberman. *Age-related hearing loss is dominated by damage to inner ear sensory cells, not the cellular battery that powers them*. Journal of Neuroscience, 40(33):6357–6366, 2020.

[134] Sarineh Keshishzadeh and Sarah Verhulst. *From derived-band envelope-following responses to individualized models of near-and supra-threshold hearing deficits*. In Proceedings of the International Symposium on Auditory and Audiological Research, volume 7, pages 13–20, 2019.

[135] Bo Engdahl, Ole Woxen, Atle R Arnesen, and Iain WS Mair. *Transient evoked otoacoustic emissions as screening for hearing losses at the school for military training*. Scandinavian audiology, 25(1):71–78, 1996.

[136] Wieslaw Konopka, Malgorzata Pawlaczyk-Luszczynska, Mariola Sliwinska-Kowalska, Antoni Grzanka, and Piotr Zalewski. *Effects of impulse noise on transiently evoked otoacoustic emission in soldiers Efectos del ruido impulsivo sobre las emisiones otoacústicas evocadas por transitorios en soldados*. International Journal of Audiology, 44(1):3–7, 2005.

[137] NS Seixas, B Goldman, L Sheppard, R Neitzel, S Norton, and SG Kujawa. *Prospective noise induced changes to hearing among construction industry apprentices*. Occupational and environmental medicine, 62(5):309–317, 2005.

[138] Lynne Marshall, Judi A Lapsley Miller, Laurie M Heller, Keith S Wolgemuth, Linda M Hughes, Shelley D Smith, and Richard D Kopke. *Detecting incipient inner-ear damage from impulse noise with otoacoustic emissions*. The Journal of the Acoustical Society of America, 125(2):995–1013, 2009.

[139] Patricia Trautwein. *Auditory neuropathy: diagnosis and case management*. In 4th ACFOS international conference, pages 8–10, 2002.

[140] Edward Lobarinas, Richard Salvi, and Dalian Ding. *Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas*. Hearing research, 302:113–120, 2013.

[141] M Charles Liberman, Michael J Epstein, Sandra S Cleveland, Haobing Wang, and Stéphane F Maison. *Toward a differential diagnosis of hidden hearing loss in humans*. PloS one, 11(9):e0162726, 2016.

[142] Harvey Dillon. *NAL-NL1: A new procedure for fitting non-linear hearing aids*. The Hearing Journal, 52(4):10–12, 1999.

[143] Denis Byrne, Harvey Dillon, Teresa Ching, Richard Katsch, and Gitte Keidser. *NAL-NL1 procedure for fitting nonlinear hearing aids: characteristics and comparisons with other procedures.* Journal of the American academy of audiology, 12(1), 2001.

[144] Leonard E Cornelisse, Richard C Seewald, and Donald G Jamieson. *The input/output formula: A theoretical approach to the fitting of personal amplification devices*. The Journal of the Acoustical Society of America, 97(3):1854–1864, 1995.

[145] Josephine E Marriage, Deborah A Vickers, Thomas Baer, Brian R Glasberg, and Brian CJ Moore. *Comparison of different hearing aid prescriptions for children*. Ear and hearing, 39(1):20–31, 2018.

[146] Sally A Gaskill and Ann M Brown. *The behavior of the acoustic distortion product, 2 f 1- f 2, from the human ear and its relation to auditory sensitivity*. The Journal of the Acoustical Society of America, 88(2):821–839, 1990.

[147] FP Harris. *Distortion-product otoacoustic emissions in humans with high frequency sensorineural hearing loss*. Journal of Speech, Language, and Hearing Research, 33(3):594–600, 1990.

[148] Sofie Degeest, Els Clays, Paul Corthals, and Hannah Keppler. *Epidemiology and risk factors for leisure noise-induced hearing damage in Flemish young adults*. Noise & health, 19(86):10, 2017.

[149] Hannah Keppler, Ingeborg Dhooge, and Bart Vinck. *Hearing in young adults. Part II: The effects of recreational noise exposure*. Noise & health, 17(78):245, 2015.

[150] Nicolae Petrescu. *Loud music listening*. McGill Journal of Medicine: MJM, 11(2):169, 2008.

[151] Pauline A Smith, Adrian Davis, Melanie Ferguson, Mark E Lutman, et al. *The prevalence and type of social noise exposure in young adults in England*. Noise and health, 2(6):41, 2000.

[152] M Jennifer Derebery, Andrew Vermiglio, Karen I Berliner, Marilee Potthoff, and Kirsten Holguin. *Facing the music: pre-and postconcert assessment of hearing in teenagers*. Otology & Neurotology, 33(7):1136–1141, 2012.

[153] V Mercier, D Luy, BW Hohmann, et al. *The sound exposure of the audience at a music festival*. Noise and Health, 5(19):51, 2003.

[154] Johanna Bengtsson Ryberg et al. *A national project to evaluate and reduce high sound pressure levels from music*. Noise and Health, 11(43):124, 2009.

[155] A Yassi, N Pollock, N Tran, and M Cheang. *Risks to hearing from a rock concert.* Canadian Family Physician, 39:1045, 1993.

[156] Sofie Degeest, Hannah Keppler, Paul Corthals, and Els Clays. *Epidemiology and risk factors for tinnitus after leisure noise exposure in Flemish young adults*. International journal of audiology, 56(2):121–129, 2017.

[157] William W Clark. *Recent studies of temporary threshold shift (TTS) and permanent threshold shift (PTS) in animals*. The Journal of the Acoustical Society of America, 90(1):155–163, 1991.

[158] Yong Wang, Keiko Hirose, and M Charles Liberman. *Dynamics of noise-induced cellular injury and repair in the mouse cochlea*. Journal of the Association for Research in Otolaryngology, 3(3):248–268, 2002.

[159] H Yamane, Y Nakai, M Takayama, H Iguchi, T Nakagawa, and A Kojima. *Appearance of free radicals in the guinea pig inner ear after noise-induced acoustic trauma*. European Archives of Oto-rhino-laryngology, 252(8):504–508, 1995.

[160] Edeltraut Emmerich, Frank Richter, Halgund Hagner, Frank Giessler, Sabine Gehrlein, and Hans-Georg Dieroff. *Effects of discotheque music on audiometric results and central acoustic evoked neuromagnetic responses.* The international tinnitus journal, 8(1):13–19, 2002.

[161] Bartosz Trzaskowski, W Wiktor Jedrzejczak, Edyta Pilka, Magdalena Cieslicka, and Henryk Skarzynski. *Otoacoustic Emissions before and after Listening to Music on a Personal Player*. Medical science monitor: international medical journal of experimental and clinical research, 20:1426, 2014.

[162] Colleen G Le Prell, Shawna Dell, Brittany Hensley, James W Hall III, Kathleen CM Campbell, Patrick J Antonelli, Glenn E Green, James M Miller, and Kenneth Guire. *Digital music exposure reliably induces temporary threshold shift (TTS) in normal hearing human subjects*. Ear and hearing, 33(6):e44, 2012.

[163] Allen F Ryan, Sharon G Kujawa, Tanisha Hammill, Colleen Le Prell, and Jonathan Kil. *Temporary and permanent noise-induced threshold shifts: a review of basic and clinical observations*. Otology & neurotology: official publication of the American Otological Society, American Neurotology Society [and] European Academy of Otology and Neurotology, 37(8), 2016.

[164] Nobuhiro Hakuba, Kenichiro Koga, Kiyofumi Gyo, Shin-ichi Usami, and Kohichi Tanaka. *Exacerbation of noise-induced hearing loss in mice lacking the glutamate transporter GLAST*. Journal of Neuroscience, 20(23):8750–8753, 2000.

[165] Jean-Luc Puel, Jérôme Ruel, Christine Gervais d'Aldin, and Rémy Pujol. *Excitotoxicity and repair of cochlear synapses after noise-trauma induced hearing loss*. Neuroreport, 9(9):2109–2114, 1998.

[166] SB Smith, J Krizman, C Liu, T White-Schwoch, T Nicol, and N Kraus. *Investigating peripheral sources of speech-in-noise variability in listeners with normal audiograms*. Hearing research, 371:66–74, 2019.

[167] James Wilbur Hall. *New handbook of auditory evoked responses*. Pearson, 2007.

[168] Brandon T Paul, Sajal Waheed, Ian C Bruce, and Larry E Roberts. *Subcortical amplitude modulation encoding deficits suggest evidence of cochlear synaptopathy in normal-hearing 18–19 year olds with higher lifetime noise exposure*. The Journal of the Acoustical Society of America, 142(5):EL434–EL440, 2017.

[169] Erika Skoe and Jennifer Tufts. *Evidence of noise-induced subclinical hearing loss using auditory brainstem responses and objective measures of noise exposure in humans*. Hearing research, 361:80–91, 2018.

[170] Dawn Konrad-Martin, Marilyn F Dille, Garnett McMillan, Susan Griest, Daniel McDermott, Stephen A Fausti, and Donald F Austin. *Age-related changes in the auditory brainstem response*. Journal of the American Academy of Audiology, 23(1):18–35, 2012.

[171] Naomi F Bramhall, Dawn Konrad-Martin, and Garnett P McMillan. *Tinnitus and auditory perception after a history of noise exposure: Relationship to auditory brainstem response measures*. Ear and hearing, 39(5):881, 2018.

[172] Angela NC Fulbright, Colleen G Le Prell, Scott K Griffiths, and Edward Lobarinas. *Leisure Noise and Hearing: Effects of Recreational Noise on Threshold and Suprathreshold Measures of Auditory Function*. In Seminars in hearing, volume 38, page 298. Thieme Medical Publishers, 2017.

[173] Heleen Luts, Sofie Jansen, Wouter Dreschler, and Jan Wouters. *Development and normative data for the Flemish/Dutch Matrix test*. 2014.

[174] Tom Francart, Astrid Van Wieringen, and Jan Wouters. *APEX 3: a multi-purpose test platform for auditory psychophysical experiments*. Journal of neuroscience methods, 172(2):283–293, 2008.

[175] Birger Kollmeier, Anna Warzybok, Sabine Hochmuth, Melanie A Zokoll, Verena Uslar, Thomas Brand, and Kirsten C Wagener. *The multilingual matrix test: Principles, applications, and comparison across languages: A review*. International Journal of Audiology, 54(sup2):3–16, 2015.

[176] Tine Vande Maele, Sarineh Keshishzadeh, Nele De Poortere, Ingeborg Dhooge, Hannah Keppler, and Sarah Verhulst. *The variability in potential biomarkers for cochlear synaptopathy after recreational noise exposure*. bioRxiv, 2021.

[177] J Pethe, K Begall, R Mühler, and JK Lottmann. *Amplitude Modulation Following Response (AMFR)-Eine Methode zur objektiven frequenzspezifischen audiologischen Diagnostik*. Laryngo-rhino-otologie, 75(01):23–28, 1996.

[178] Alessandro Altoe and Christopher A. Shera. *Nonlinear cochlear mechanics without direct vibration-amplification feedback*. Phys. Rev. Research, 2:013218, Feb 2020.

[179] Harold F Schuknecht and Mark R Gacek. *Cochlear pathology in presbycusis*. Annals of Otology, Rhinology & Laryngology, 102(1_suppl):1–16, 1993.

[180] Deepak Baby, Arthur Van Den Broucke, and Sarah Verhulst. *A convolutional neural-network model of human cochlear mechanics and filter tuning for real-time applications*. Nature Machine Intelligence, pages 1–10, 2021.