Characterizing cochlear hearing impairment using advanced electrophysiological methods

Encina Llamas, Gerard

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

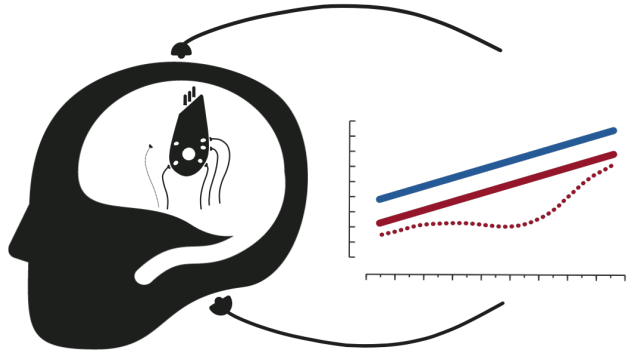
Technical University of Denmark

CONTRIBUTIONS TO
HEARING RESEARCH

Volume 30

Gerard Encina-Llamas

Characterizing cochlear hearing impairment using advanced electrophysiological methods



Hearing Systems
Department of Electrical Engineering

Characterizing cochlear hearing impairment using advanced electrophysiological methods

PhD thesis by
Gerard Encina-Llamas

Technical University of Denmark

2017

The defense was held on November 7th, 2017.

This PhD dissertation is the result of a research project carried out at the Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark (Kgs. Lyngby, Denmark). Part of the project was carried out at the Auditory Neuroscience Laboratory in the Center for Computational Neuroscience and Neural Technology (CompNet) at Boston University (BU) in Boston, MA (USA). Part of the project was done in collaboration with the group of Prof. Sharon G. Kujawa from the Massachusetts Eye and Ear Infirmary (MEEI), Harvard Medical School (HMS) in Boston, MA (USA).

The project was partly financed by the Oticon Centre of Excellence for Hearing and Speech Sciences (2/3) and by the Technical University of Denmark (1/3). The external stay in Boston, MA was further supported by a grant from the Erasmus Mundus program of the European Union (EU), Student Exchange Network in Auditory Cognitive Neuroscience (ACN).

Supervisors

Main supervisor

Assoc. Prof. Bastian Epp
Hearing Systems Group
Department of Electrical Engineering
Technical University of Denmark
Kgs. Lyngby, Denmark

Co-supervisors

Prof. Torsten Dau
Hearing Systems Group
Department of Electrical Engineering
Technical University of Denmark
Kgs. Lyngby, Denmark

James M. Harte, PhD
Interacoustics Research Unit (IRU)
Kgs. Lyngby, Denmark

"Руку платком обмотай и в венценосный шиповник,
В самую гущу его целлулоидных терний
Смело до хруста ее погрузи, –
Добудем розу без ножниц!
Но смотри, чтобы он не осыпался сразу – [...]"

"Wrap your hand with a handkerchief and boldly
Plunge it, till it cracks, in the depth of a royal wild rose,
Into its celluloid thorns. Let's get a rose without scissors.
Yet watch lest the petals fall at once [...]"

Путешествие в Армению *(Journey to Armenia)*, 1933.

Осип Мандельштам (*Osip E. Mandelstam*), Russian poet and essayist
(Warsaw, 1891 – Vtoraya Rechka transit camp, URSS, 1938)

Abstract

One of the impressive and not fully understood abilities of the healthy auditory system is the property of enabling successful communication in highly complex acoustical environments with high levels of background noise. In order to do so, the complex biological machinery in the auditory periphery must function with precision to transform the sound into neural activity that the brain can interpret. Dysfunction or loss of the cells that underlay such transformation leads to a disruption or degradation of the sound encoding. Restoration strategies through the use of prosthetic devices (i.e. hearing aids) have been proposed to compensate for those deficits, but these are not always entirely successful, and do not target the specific physiological deficit leading to the hearing loss or dysfunction. Thus, aided, hearing-impaired (HI) listeners commonly do not perform as proficiently as normal-hearing (NH) listeners, which can result in frustrating experiences. An accurate assessment of the precise deficits of the auditory system may allow for the development of more precise and individualized compensation strategies.

The gold-standard metric today to evaluate the peripheral auditory system is pure-tone threshold audiometry. This is still the case even with direct physiological evidence from animal models that a substantial loss or dysfunction of sensory inner hair cells (IHC) and/or auditory nerve (AN) fiber synapses is not strongly related to pure-tone threshold. On the other hand, loss or dysfunction of actuator outer hair cells (OHC) have been clearly related to the elevation of hearing thresholds. This might explain why some people with normal hearing sensitivity report difficulties in speech understanding and music perception in complex acoustical scenarios. In order to examine the potential effect of such "hidden" pathologies on supra-threshold perception, this thesis describes the investigation of supra-threshold processing by means of electrophysiological methods. More precisely, envelope following responses (EFR) recorded as a function of stimulus level (level-growth) were proposed as a method to estimate compression in the peripheral auditory system, and to investigate the potential effect of damage of AN fibers synapses.

Firstly, EFR level-growth functions using multiple carrier frequency stimuli were recorded in NH and mild-HI listeners to estimate peripheral compression. The results showed that EFRs provide similar estimates of auditory peripheral compression to those reported in the previous literature. However, the place-specificity of the compression estimates may be compromised due to

contributions from off-frequency (i.e. away from the characteristic place of the stimulus) neural populations to the EFR. Secondly, EFR level-growth functions were measured for strongly and shallowly modulated tones in NH threshold and mild-HI listeners to study the postulated presence of loss of AN fiber synapses (i.e. cochlear synaptopathy). The results revealed different patterns within the homogeneous group of young-NH threshold listeners and in the mild-HI listeners. Similar patterns were observed in noise-exposed mice in contrast to an unexposed control group. A well-established phenomenological computational model of the AN was used to investigate the potential relation between the recorded EFR level-growth functions in humans and mice and the postulated presence cochlear synaptopathy. The model simulations suggested that the contribution of off-frequency neural activity dominate the response of the AN to modulated tones at medium-to-high stimulus levels, and potentially also the recorded EFR.

In the light of the results described in this thesis it was suggested that the combination of non-invasive electrophysiology in humans and non-human mammals and computational modeling may be a promising approach to study and better characterize supra-threshold processing in the auditory system.

Resumé

En af de bemærkelsesværdige og ikke fuldt ud forståede egenskaber ved et velfungerende auditivt system er dets evne til at muliggøre kommunikation selv i særdeles komplekse akustiske miljøer med høj baggrundsstøj. Denne egenskab kræver at de komplekse biologiske mekanismer i den auditive periferi på nøjagtig vis omdanner lyd til neural aktivitet, som hjernen kan fortolke. Dysfunktion eller tab af de celler, der danner basis for denne omdannelse, fører til forstyrrelse eller forringelse i kodningen af det akustiske signal. Der eksisterer strategier til at rette op på disse problemer via hjælpemidler såsom høreapparater, men disse er ikke målrettet de specifikke fysiologiske problemer der fører til høretabet, og resultatet er ikke altid vellykket. Hørehæmmede personer med høreapparater hører derfor som regel ikke så godt som normalthørende personer, hvilket kan være en kilde til frustration. En mere præcis diagnose af de eksakte defekter ved en persons hørelse kan åbne muligheder for at udvikle mere nøjagtige og individuelle kompensationsstrategier.

I dag evalueres det perifære auditive system ved måling af høretærskler via rentone-audiometri. Dette til trods for at fysiologiske resultater fra dyremodeller har vist en mangel på sammenhæng mellem høretærskler og tab eller dysfunktion af indre hårceller og/eller de auditive nervefibres synapser. På den anden side er der påvist en klar sammenhæng mellem tab eller dysfunktion af ydre hårceller og forhøjede høretærskler. Dette kan muligvis forklare hvorfor nogle personer med normale høretærskler melder om vanskeligheder ved at forstå tale og nyde musik i komplekse akustiske scenarier. I denne afhandling beskrives det hvordan mulige virkninger af sådanne "skjulte" patologier er blevet undersøgt via elektrofysiologiske metoder ved brug af *supra-threshold* lyde (dvs. lydniveauer over høretærsklerne). Nærmere bestemt er det blevet foreslået at benytte *envelope following responses* (EFR) målt som en funktion af stimuluslydniveauet (*level-growth*) som en metode til at estimere kompression i det perifære auditive system samt til at undersøge de mulige følger af beskadigede nervefibre synapses.

Først blev EFR level-growth funktioner målt ved brug af multi-frekvensstimuli hos normalthørende og mildt hørehæmmede personer for at estimere perifær kompression. Resultaterne viste at EFRs giver estimater for auditiv perifær kompression, der svarer til resultater fra litteraturen; dog bliver det stedsspecifikke aspekt af kompressionsestimatet muligvis kompromitteret af bidrag fra *off-frekvens* neurale populationer (dvs. andre områder end i det karakteristiske stimulusområde i forhold til EFR'et). Dernæst blev EFR level-growth

funktioner målt med mere og mindre modulerede toner hos normalthørende og mildt hørehæmmede personer for at undersøge det mulige tab af auditive nervefibre og synapser (dvs. *cochlear synaptopathy*). Resultaterne viste forskellige mønstre internt i den homogene gruppe af unge personer med normale høretærskler og internt i gruppen af personer med mildt høretab. Tilsvarende mønstre blev observeret i hhv. mus der var udsat for støj og mus der ikke var udsat for støj. En anerkendt fænomenologisk computermodel der modellerer auditive nervefibre blev brugt til at undersøge potentielle relationer mellem de målte EFR level-growth funktioner i mennesker og mus og den mulige tilstedeværelse af synaptopathy. Modelsimuleringerne indikerede at bidragene fra off-frekvens neural aktivitet dominerer responsen på de modulerede toner ved medium-til-høje stimulus-niveauer fra hørenerven, og muligvis også fra de målte EFRs.

De resultater der redegøres for i denne afhandling belyser hvordan kombinationen af ikke-invasive elektrofysiologiske metoder brugt på mennesker og ikke-menneskelige pattedyr samt computermodellering er en lovende tilgang til at studere og mere nøjagtigt karakterisere supra-threshold processering i det auditive system. En sådan karakteristik vil kunne bidrage til en forståelse af hvordan det auditive system fungerer ved de lydniveauer der benyttes ved dagligdags kommunikation, og samtidig fremme udviklingen af nye og mere præcise redskaber til brug ved diagnosticering af høretab.

Resum

Una de les habilitats més increïbles i no plenament entesa del sistema auditiu humà és la seva facultat de permetre la comunicació en ambients acústics altament complexos amb alts nivells de soroll de fons. Per tal que això sigui possible, la complexa maquinaria biològica del sistema auditiu perifèric ha de funcionar amb exactitud pel que fa a la transformació del so en activitat neuronal que el cervell sigui capaç d'interpretar. La disfunció o pèrdua de les cèl·lules subjacents a aquesta transformació produeix la tergiversació o la degradació de la codificació del so. Encara que l'ús d'aparells protètics (és a dir, audiòfons) s'ha suggerit com a estratègia restaurativa per compensar aquestes pèrdues, tal estratègies no sempre esdevenen reeixides, i no reverteixen específicament el dèficit fisiològic que condueix a la disfunció o pèrdua auditiva. Així doncs, aquelles persones hipoacústiques (HA) que son usuàries d'audiòfons no son habitualment capaces de desenvolupar-se igual de competentment que aquelles altres normooients (NO), la qual cosa porta a experiències frustrants. Una avaluació més acurada dels dèficits concrets del sistema auditiu permetria de desenvolupament estratègies compensatòries més individualitzades i precises.

Ara per ara, la mesura considerada com a estàndard que s'usa per tal d'avaluar el sistema auditiu perifèric és l'audiometria tonal liminar. Això és així tot i que hi ha evidència fisiològica extreta a partir de models animals que mostra que una pèrdua o disfunció substancial de les cèl·lules ciliades internes (CCI) sensitives i/o de les sinapsis de les fibres del nervi auditiu (NA) no està estretament relacionada amb el llindar d'audició tonal. D'altra banda, la pèrdua o disfunció de les cèl·lules ciliades externes (CCE) transmissores de força, sí que s'ha relacionat clarament amb l'elevació del llindar auditiu. Això podria ser explicatiu de perquè alguna gent amb sensitivitat auditiva normal, es queixa d'experimentar dificultats pel que fa a la comprensió de la paraula i la percepció musical en escenaris acústics complexos. Per tal d'examinar els efectes potencials d'aquestes patologies "ocultes"relatives a la percepció supra-liminar, la present tesina descriu la investigació de tal percepció supra-liminar a través de mètodes electrofisiològics. Més acuradament, les respostes seguidores de l'envolupant (RSE) mesurades en funció del nivell d'estimulació foren proposades com a mètode per estimar la compressió del sistema auditiu perifèric, i per tal de investigar l'efecte potencial del dany en les sinapsis de les fibres del NA.

Primerament, les RSE en funció del nivell d'estimulació, emprant estímuls amb múltiples freqüències de la portadora, foren mesurades en persones NO i

en persones HA lleus, per tal d'estimar la compressió perifèrica. Els resultats mostraren que les RSEs proporcionen estimacions de la compressió auditiva perifèrica semblants a aquelles ja indicades en la literatura prèvia. No obstant, la especificitat del lloc d'estimació de la compressió pot veure's compromesa degut a les contribucions a la RSE dels grups neuronals corresponentsa les zones coclears de freqüències llunyanes (és a dir, aquelles freqüències remotes al lloc característic de l'estímul). En segon terme, les RSE en funció del nivell d'estimulació foren mesurades amb tons modulats fortament i suau en persones amb llindars de NO i en persones amb HA lleus, per tal d'estudiar la postulada pèrdua de les sinapsis de les fibres del NA (és a dir, sinaptopatia coclear). Els resultats revelaren diferents patrons dintre del grup homogeni de persones joves amb llindars de NO, i en el grup de persones de HA lleu. Un model computacional auditiu fenomenològic, plenament establert, fora emprat per tal d'investigar la relació en potència entre les mesures de les RSE en funcions del nivell d'estimulació en humans i en ratolins, i la postulada presència de sinaptopatia coclear. Les simulacions del model suggeriren que les contribucions de l'activitat neuronal de les freqüències llunyanes dominen la resposta del NA en ser excitat amb tons modulats a nivells mitjans alts; i potencialment també dominen les mesures de les RSE.

A la llum dels resultats descrits en la present tesina es suggerí que la combinació d'electrofisiologia no invasiva en humans i en mamífers no humans, juntament amb l'ús de models computacionals, podria esdevenir un enfoc prometedor per a l'estudi i la caracterització acurada del processament supraliminar del sistema auditiu.

Resumen

Una de las habilidades más increíbles y no enteramente comprendida del sistema auditivo humano es la facultad de permitir la comunicación en ambientes acústicos altamente complejos con altos niveles de ruido de fondo. Para que esto se produzca, la compleja maquinaria biológica del sistema auditivo periférico debe funcionar con exactitud en cuanto a la transformación del sonido en actividad neuronal que el cerebro sea capaz de interpretar. La disfunción o pérdida de las células subyacentes a dicha transformación produce la tergiversación o degradación de la codificación del sonido. Aunque el uso de aparatos protésicos (esto es, audiófonos) se ha sugerido como estrategia restaurativa para compensar tales pérdidas, estas estrategias no siempre son exitosas, y no revierten específicamente el déficit fisiológico que lleva a la disfunción o pérdida auditiva. Por lo tanto, aquellas personas con problemas auditivos, hipoacúsicos (HA), usuarias de audiófonos no suelen ser capaces de desarrollarse de forma igual de competente que aquellas otras con audición normal, normo-oyentes (NO), la cual cosa conduce a experiencias frustrantes. Una evaluación más detallada de los déficits concretos del sistema auditivo permitiría el desarrollo de estrategias compensatorias más individualizadas y precisas.

En la actualidad, la medición considerada como estándar usada para evaluar el sistema auditivo periférico es la audiometría tonal liminar. Esto es así incluso a pesar de la evidencia fisiológica adquirida a partir de modelos animales que muestra que una sustancial pérdida o disfunción de las células ciliadas internas (CCI) sensitivas y/o de las sinapsis de las fibras del nervio auditivo (NA), no está estrechamente vinculada con el umbral de audición tonal. Por otro lado, la pérdida o disfunción de las células ciliadas externas (CCE) transmisoras de fuerza, sí se ha relacionado claramente con la elevación de los umbrales auditivos. Esto podría explicar por qué cierta gente con sensibilidad auditiva normal, se queja de experimentar dificultades en cuanto a la comprensión de la palabra y la percepción musical en escenarios acústicos complejos. Con tal de examinar los efectos potenciales de aquellas patologías "ocultas"relativas a la percepción supra-liminar, la presente tesis doctoral describe el estudio de dicha percepción supra-liminar a través de métodos electrofisiológicos. Con más exactitud, respuestas seguidoras de la envolvente (RSE) medidas en función del nivel de estimulación fueron propuestas como método con el que estimar la compresión del sistema auditivo periférico, y para estudiar el efecto potencial del daño en las sinapsis de las fibras del NA.

Primero, las RSE en función del nivel de estimulación, utilizando estímulos

con múltiples frecuencias de la portadora, fueron medidas en personas NO y en personas de HA leves, para estimar la compresión periférica. Los resultados mostraron que las RSEs proporcionan estimaciones de la compresión auditiva periférica parecidas a aquellas previamente indicadas en la literatura. No obstante, la especificidad del sitio de estimación de la compresión podría verse comprometida debido a las contribuciones a la RSE de grupos de neuronas correspondientes con zonas cocleares de frecuencias lejanas (esto es, frecuencias remotas al lugar característico del estímulo). Segundo, las RSE en función del nivel de estimulación fueron medidas con tonos modulados fuerte y suavemente en personas con umbrales de NO y en personas de HA leves, para estudiar la postulada pérdida de las sinapsis de las fibras del NA (esto es, sinaptopatía coclear). Los resultados revelaron diferentes patrones dentro del grupo homogéneo de personas jóvenes con umbrales de NO, y en el grupo de personas de HA leve. Un modelo computacional auditivo fenomenológico, ampliamente reconocido, se utilizó para investigar la relación potencial entre las medidas de las RSE en función del nivel de estimulación en humanos y en ratones, y la postulada presencia de sinaptopatía coclear. Las simulaciones del modelo sugieren que las contribuciones de la actividad neuronal de frecuencias lejanas dominan la respuesta del NA al excitarse con tonos modulados a niveles medios altos; y potencialmente también dominan las mediciones de las RSE.

A la luz de los resultados descritos en esta tesis doctoral se sugirió que la combinación de electrofisiología no invasiva en humanos y en mamíferos no humanos, junto con el uso de modelos computacionales podría convertirse en un enfoque prometedor para el estudio y la esmerada caracterización del procesamiento supra-liminar del sistema auditivo.

Acknowledgments

Academic acknowledgment

To my supervisors at DTU

Digging in my notes, I could see that the first "official" meeting I had with my supervisors, James M. Harte, from Interacoustics Research Unit, and Torsten Dau and Bastian Epp, from the Hearing Systems group at DTU, was on March 6th, 2014. The PhD project officially started on February 14th, 2014, but I guess that the coincidence of ARO in late February and the fact that there was a smooth transition between my MSc thesis project, also at DTU, and the PhD project, postponed that meeting. I do not recall this meeting in particular, but I remember fresh and clearly the sensations that I experienced during those first meetings. Most of the time, only James, Bastian and I were present. It occurred in Bastian's office, room 118, first floor, building 352. I always sat at the end of the table. Bastian on my left, James on my right. Now the game could start! At the beginning Bastian would make a little joke, or explain an irrelevant anecdote regarding the construction of the new lab. Rarely would it be James who explained a curious recent personal situation. At some point someone, perhaps I, would say: "Ok. Let's start!". Then the real game started. They would start talking, vehemently, proposing this and that idea, considering this and that condition, hypothesizing, speculating, referring to this study or that other one that I had never heard of and that I wrote down in my notebook, making little confusing drawings about membranes moving up and down, nerve fibers spiking I don't know exactly what, talking about damage or disconnection or saying weird words like deafferentation of synapses and so on. I was there in the middle. Moving my head from left to right, like in a good clay ground tennis match. If you like football (or soccer, for my American dudes) and you have at some point watched a football match of Pep Guardiola's team when he was FC Barcelona's coach, it was exactly that. Bastian and James were Xavi and Iniesta.

I was the opponent, of course. They were passing the ball and I was running behind, trying to catch it, fighting a losing battle. At some point, they would stop the game and would look towards me and, normally Bastian, would say: "So, what do you think?". Damn it! They got me. "I think it's fine!", I would reply. I didn't know exactly what I was replying to, but I guessed it was fine. Then I would go to my desk, and I would continue reading papers to try to understand something.

Today is about three years and four month after that first meeting. The situation is different. Today, I am in the middle of the field, dressed with the blue and red characteristic colors of Barça's T-shirt, I bring the ball and I play with them almost as if I would be another member of the team. A junior one, but another member, after all. Today, I feel I participate and even sometimes lead the discussion. Before I was a spectator of a scientific discussion. Today I am *an actor* of the scientific discussion. This transition has occurred thanks to my supervisors. I've worked, of course. I dedicated effort and I invested time, but I would have never been able to achieve this on my own. And I am completely sure that without Bastian and James, who have always been accessible to me and have always gifted me with their wise advice, I wouldn't have come this far. Thank you for shaping my mind, transforming it from that of a somewhat normal human being to something that resembles that of a scientist. Thank you very much! Danke schön!

I also want to thank Torsten. Torsten's role has been a bit more away, in the distance, behaving like the first external reviewer but still in the house. This has been of great value because it has confronted me with reality and a certain skepticism regarding the ideas that we were proposing, without the need of going to a conference to discuss them. In short, if I couldn't convince Torsten that what we were doing made sense, that meant that it was not good enough. To have these argumentative crashes within the group with someone with the talent of Torsten is priceless. In addition, Torsten's support and feedback during the writing process at the very last stage of the PhD has been invaluable. And finally, thank you Torsten so much for leading this group, our group, with determination and passion, but also with comprehension and consideration. My project coincided in time with a situation were the department grew considerably and other advisers, like Bastian, had to take more responsibility. You allowed this to happen, which is not easy when coming from a situation in which everything relevant that occurred in the group essentially passed through you. Resistance

to this processes would have challenged my project in a very undesirable way, but it didn't, and I want to thank you also for that. Ultimately, thank you for trusting in me even when I was not entirely sure of myself, and for allowing me to take the journey of doing a PhD; and also for letting me continue a bit more after it.

To my colleagues at DTU

I really want to thank all my colleagues in the Hearing Systems group, and also in the Acoustic Technology group at DTU. I will start thanking Caroline, Eva, Jørgen, Tom, Nadia and more recently Henrik and Katrine for assisting me and all of us any time we need it, with a smile on their face. They make the machinery work and make our lives easier. I would like to thank the other senior researchers from our group, Jeremy, Sébastien, Jens-Bo, Tobias and specially Ewen, with whom I had nice and tough discussions about how the peripheral system should be assessed, in our eternal friendly fight between psychoacousticians and electrophysiologists. I also want to thank my PhD colleagues, especially the ones from office 111: Suyash, Christoph, François, Golbarg, Søren, Richard, Thomas, Raúl, Helia, Antoine and Josefine. I've loved those over-desk-panel crazy and often absurd discussions about anything that Suyash enthusiastically would propose to discuss about. That was great! I would like to thank Andreu in particular, first for allowing me to talk in Catalan in Denmark, and secondly for often kindly disagreeing with me about non scientific and lab related things. You know what I mean! And very specially I would like to thank Laurel Carney who stayed with us for six months in the spring of 2017, and who explained me that I was right in my interpretations of the model simulations and who helped us flip our brains in many aspects and undoubted assumptions related to auditory nerve processing, rate-level curves and the encoding of modulations. I am convinced that a great collaboration will occur due to Laurel's stay. Thank you so much.

To my advisers and colleagues in Boston

I would like to express my sincere gratitude to Barbara Shinn-Cunningham for allowing me to spend six month in the spring of 2016 in her lab at Boston University (BU). I really appreciate that she showed me another way of managing a group and doing science than the one I was used to at DTU, and that she very easily integrated me in her group. This last thing can be extended to all of

the Auditory Neuroscience Lab members, specially Robert, Darrin, Le, Scott, Hari, Axelle, Marissa and Lenny. I want to thank particularly Le for the good discussions about the implementation of the model and the interpretation of the simulated results. An also specially to Robert for the good time not only in the lab but mostly outside. I also want to thank my lovely frozen mates (due to the insane tradition of setting the AC to -50ºC in the US) Lia, Yuqi, Jasmine and Sophie, and also Lengshi who, even though he hardly ever talked to me during the six months I was sitting next to him, made me feel accompanied.

I would like to thank as well Steven Colburn from the Hearing Research Center at BU for permitting me to annoy Graham Voysey so that he could help me, and he helped me greatly, to learn Python and to use his implementation of the model of Sarah Verhulst. I also thank Sarah for her support on the use of the model she developed.

I would also like to thank Sharon G. Kujawa and Aravind Parthasarathy from the Massachusetts Eye and Ear Infirmary (MEEI) from Harvard Medical School for performing a similar experiment to the one I did in humans, but in mice. My supervisors know well that I had, as we say in Catalan, between my eyebrows the wish to contrast the electrophysiological recordings from humans with mice data. Sharon allowed this to happen and Aravind executed it. It was a great experience to be able to see how an animal physiology lab works, to be able to assist to all the talks in MEEI and see Liberman, Guinan, Delgutte, Shera and many others pass by. I specially thank Aravind for the close and warm collaboration, for wonderful discussions and for a fantastic dual-poster presentation at ARO in 2017. I hope this collaboration will continue in the near future.

Others

This work was founded by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) at the Technical University of Denmark (DTU). I want to acknowledge also the support provided by the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience which made my American dream financially possible. More generally, I would like to thank the three Danish hearing aids manufacturers, Oticon, GN Resound and Widex, for continuously funding the Hearing Systems group and giving all of us great conditions to develop our work. Thanks.

The head icon in the front page image was made by Freepik from Flaticon.

Personal acknowledgment

Til min danske familie (To my Danish family)

I arrived to Denmark on August 22nd, 2011, that is, a bit less than six years ago. I arrived alone, with one single suitcase, leaving behind my family, my friends, my dog, and everything that constituted my life. Initially it was only for two years, perhaps a bit more, for taking a MSc in Engineering Acoustics at DTU. After that I decided to continue with a PhD, the thesis of which you are reading right now. Just before starting the PhD project, I met Karen, who at the time of starting the PhD was my girlfriend, at the time of finishing the PhD was already my wife, and if the biological dice are well thrown and everything goes as it should, at the time of defending this thesis will be the mother of my daughter. With Karen came her family. Torsten, Birgit, Anne, Morten and the little Asger; who was very little when I started the project and is luckily becoming less and less little. All of them are Danish. I am not Danish. I am a Catalan living in Denmark. This, obviously, has brought complexity to the relationship with Karen due to the cultural difference and the linguistic challenge, but I am absolutely convinced that such complexity is beautiful and will provide a wider understanding of the world and its diversity to our children. We are accepting the differences and overcoming those challenges and we are moving forward. This, which has been an essential personal and emotional support during the whole PhD time, has been possible because Karen's family, which is now my family too, have accepted this situation and have welcomed me. They could have done the opposite. They could have faced it, fought it back, and perhaps they would have won. But they didn't. Their decision will definitely determine the lives of Karen and me. And their decision allowed this project to come to fruition, as I needed emotional stability to attempt to excel. It is for all that and for all the help and love they give me and Karen that I want to thank all of them from the bottom of my heart. Mange tak!

A la meva familia catalana (To my Catalan family)

I was born in 1986 in Barcelona. My parents, my grandparents, my sister and I have always lived in the same house in the Prosperitat neighborhood at the very periphery of the city of Barcelona. This part of the city started to grow around the second decade of the 20th century with immigrants coming mainly from

the southeastern part of Spain to Catalonia. My family was one of those. They sacrificed themselves for the good of the next generation. My great-grandfather, for instance, worked, among other places, in a foundry; and he died relatively young due to the smoke he breathed during years. My grandparent survived the Spanish Civil War and the bloody indiscriminate bombing of the city by the Italian fascist air force allied with the Francoist army. Specially my grandfather suffered from severe hunger. Some years after the war finished, luckily they could work. And they worked tremendously hard to move the family forward. They had a butchery. My mother left school at the age of 14 to work in the family's shop that my grandmother owned. They worked an infinite amount of hours, during very long days and during many years without any vacations, just with a too short rest on the Sundays. They earned some money, mainly because they didn't spent much. And when they spent something, they did it with extreme caution. My father lost his father at the age of 15. They have always said that he died of sadness. I do not know really what that means. But the thing is that my father had to start working to support my grandmother and he could just study in the evenings. He managed to get his high school diploma and later even started college. However, it was too much and he could not finish his university degree. All of them have worked so much with the ultimate goal that my sister and I had the best conditions to become whatever we wished. They couldn't become what they wished themselves. All this suffering and effort is also part of this thesis. I can very humbly but, with my head held high, say that I am the first one in my family that will hold a PhD. I am also the only one among my classmates at school that will hold a PhD. Perhaps I am not completely dumb. Maybe. But the true difference between the other kids that played in the late 90's in the streets of my neighborhood and I, is my family. The conditions that too many of those kids had at home were exasperating. In my first school, a public school, the mother of a girl, classmate of mine, was a prostitute that had to leave her child alone at home during the night while she was working. The father of another one had died of a heroin overdose while being in prison. As they were not economically sustainable, the mother and the kids had to live with their uncle, who abused the mother and beat up the children. The father of a third one was in prison for drug trafficking. Their conditions were not the most fortunate. My conditions where luckily different. But they were different because my parents and grandparents made them different. All that I am is due to them and what they have done. I think, very humbly, that with very little

resources, they have done a very good job. Thank you so much. Us estimo!

To Karen

When I wrote the acknowledgment of my MSc thesis, I reserved the last paragraph for my parents. I was placing them at the bottom of the acknowledgment text as a symbol of the foundations and the reference they are to me. Now, a bit less than four years after writing those lines, the person to whom the last paragraph must be dedicated to must change. Now, the person who conforms the pillar on which I lean and support myself to grow is Karen. The time during which I have been working on this PhD is the same period that we have dedicated to know each other and to decide that we want to build something together. We started being just a couple a bit before the PhD begun, and a few month before the end, we decided to get married, buy a house and we knew that we were waiting a child. A little girl, the name of whom we are not entirely sure yet. We have some ideas, but we don't know it yet. She must also be now at the bottom of this text. We do not yet know each other much. We just had a very short and awkward contact some days ago when it seemed that I could touch her head through Karen's belly. I am looking forward to meet you soon! And I am really looking forward to spend the rest of my life with both of you, and whomever else might come. It is true that we do not use to express it loud, and I always defend the silence arguing that we do not need to say it because we show it every day. But this could be a good place to write it for the posterity: Thank you Karen for having taken the decision of sharing the rest of your life with me. Jeg elsker dig!

Gerard Encina-Llamas, July 15th, 2017

Related publications

Journal papers

- Encina-Llamas, G., Dau, T., and Epp, B. (**under review**). "Estimates of peripheral compression using envelope following responses," JARO - J. Assoc. Res. Otolaryngol

- Encina-Llamas, G., Harte, J. M., Dau, T., Shinn-Cunningham, B. G., and Epp, B. (**under review**). "Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve," JARO - J. Assoc. Res. Otolaryngol

Conference papers

- Encina-Llamas, G., Harte, J. M., Epp, B., and Dau, T. (**2015**). "Evaluation of peripheral compression and auditory nerve fiber intensity coding using auditory steady-state responses," Proceedings of International Symposium on Auditory and Audiological Research (ISAAR): Individual hearing loss - Characterization, modelling, compensation strategies, (pp. 1–8).

Published abstracts

- Encina-Llamas, G., Epp, B., and Dau, T. (**2014**). "Comparison of peripheral compression estimates using auditory steady-state responses (ASSR) and distortion product otoacoustic emissions (DPOAE)," Association for Research in Otolaryngology (ARO), 37^{th} Mid-Winter Meeting, San Diego, CA (USA), February 2014.

- Encina-Llamas, G., Epp, B., and Dau, T. (**2014**). "Comparison of peripheral compression estimates using auditory steady-state responses (ASSR)

and distortion product otoacoustic emissions (DPOAE)," Dansk Teknisk Audiologisk Selskab (DTAS), Stouby, Denmark, September 2014.

- Encina-Llamas, G., Harte, J.M., Dau, T., and Epp, B. (**2015**). "Evaluation of cochlear processing and auditory nerve fiber intensity coding using auditory steady-state responses (ASSR)," XXIV Biennial Symposium of the International Evoked Response Audiometry Group (IERASG), Busan, South Korea, May 2015.

- Encina-Llamas, G., Harte, J.M., Dau, T., and Epp, B. (**2015**). "Evaluation of peripheral compression and auditory nerve fiber intensity coding using auditory steady-state responses (ASSR)," International Symposium on Auditory and Audiological Research (ISAAR), Nyborg, Denmark, August 2015.

- Encina-Llamas, G., Dau, T., Harte, J.M., and Epp, B. (**2016**). "Using auditory steady-state responses to evaluate auditory nerve integrity in normal-hearing and mild hearing-impaired listeners," Association for Research in Otolaryngology (ARO), 39^{th} Mid-Winter Meeting, San Diego, CA (USA), February 2016.

- Encina-Llamas, G., Dau, T., Harte, J.M., and Epp, B. (**2016**). "Using auditory steady-state responses to evaluate auditory nerve integrity in normal-hearing and mild hearing-impaired listeners," Frequency Following Response Workshop, Boston, MA (USA), May 2016.

- Encina-Llamas, G., Parthasarathy, A., Harte, J. M., Dau, T., Kujawa, S. G., Shinn-Cunningham, B., and Epp, B. (**2017**). "Hidden hearing loss with envelope following responses (EFRs): The off-frequency problem," Association for Research in Otolaryngology (ARO), 40^{th} Mid-Winter Meeting, Baltimore, ML (USA), February 2017.

- Parthasarathy, A., Encina-Llamas, G., Shinn-Cunningham, B., and Kujawa, S. G. (**2017**). "Temporal processing deficits due to noise-induced synaptopathy studied using envelope following responses," Association for Research in Otolaryngology (ARO), 40^{th} Mid-Winter Meeting, Baltimore, MD (USA), February 2017.

- Encina-Llamas, G., Parthasarathy, A., Harte, J. M., Dau, T., Kujawa, S. G., Shinn-Cunningham, B., and Epp, B. (**2017**). "Synaptopathy with enve-

lope following responses (EFRs): The off-frequency problem," Erasmus Mundus Symposium on Auditory Cognitive Neuroscience, Leipzig, Germany, April 2017.

- Encina-Llamas, G., Parthasarathy, A., Harte, J. M., Dau, T., Kujawa, S. G., Shinn-Cunningham, B., and Epp, B. (**2017**). "Synaptopathy with envelope following responses (EFRs): The off-frequency problem," XXV Biennial Symposium of the International Evoked Response Audiometry Group (IERASG), Warsaw, Poland, May 2017.

- Encina-Llamas, G., Parthasarathy, A., Harte, J. M., Dau, T., Kujawa, S. G., Shinn-Cunningham, B., and Epp, B. (**2017**). "Synaptopathy with envelope following responses (EFRs): The off-frequency problem," International Symposium on Auditory and Audiological Research (ISAAR), Nyborg, Denmark, August 2017.

- Encina-Llamas, G., Parthasarathy, A., Harte, J. M., Dau, T., Kujawa, S. G., Shinn-Cunningham, B., and Epp, B. (**2017**). "Synaptopathy with envelope following responses (EFRs): The off-frequency problem," Dansk Teknisk Audiologisk Selskab (DTAS), Stouby, Denmark, September 2017.

Datasets and Software

- Voysey, G., and Encina-Llamas, G. (**2016**). "Corti: Version 0.9," Modeling environment to explore cochlear synaptopathy with complex stimuli. July, 2016.

- Encina-Llamas, G., Dau, T., and Epp, B. (**2017**). "Estimates of peripheral compression using envelope following responses [Data set]," Zenodo. DOI 10.5281/zenodo.844834. August, 2017.

- Encina-Llamas, G., Harte, J. M., Dau, T., Shinn-Cunningham, B. G., and Epp, B. (**2017**). "Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve [Data set]," Zenodo. DOI 10.5281/zenodo.844850. August, 2017.

Contents

Abstract vii

Resumé på dansk ix

Resum en català xi

Resumen en castellano xiii

Acknowledgments xv

Related publications xxiii

Table of contents xxx

1 Introduction 1

2 Estimates of peripheral compression using envelope following responses 9

2.1 Introduction 10
2.2 Methods 13
2.2.1 Listeners 13
2.2.2 Apparatus 13
2.2.3 EFR recordings 14
2.2.4 DPOAE recordings 17
2.3 Results 18
2.3.1 EFR level-growth functions for normal-hearing and hearing-impaired listeners 18
2.3.2 DPOAE level-growth functions for NH listeners 24
2.4 Discussion 25
2.4.1 Compression estimates based on EFR level-growth functions 25
2.4.2 Compression estimates based on DPOAE level-growth functions in NH listeners 28

2.4.3 Comparison of compression estimates based on EFR and DPOAE level-growth functions 29
2.4.4 On- versus off-frequency contributions to compression estimates 30
2.5 Summary and Conclusion 31

3 Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve 37
3.1 Introduction 38
3.2 Methods 41
3.2.1 Listeners 41
3.2.2 Apparatus 41
3.2.3 EFR recordings and analysis 42
3.2.4 AN model 43
3.3 Results 45
3.3.1 EFR level-growth functions in human listeners 45
3.3.2 Simulating EFR level-growth functions in human listeners with and without hair-cell dysfunction 46
3.3.3 Simulating EFRs in NH threshold listeners and HI listeners with postulated synaptopathy 50
3.4 Discussion 52
3.4.1 EFR level-growth functions from strongly and shallowly modulated SAM tones 52
3.4.2 A model of the auditory nerve to investigate individual differences in EFR level-growth functions 53
3.5 Conclusions 58

4 Simulating envelope following responses using the model of Verhulst *et al.* (2015) 61
4.1 Introduction 62
4.2 Methods 64
4.2.1 Implementation of the model by Verhulst *et al.* (2015) . . . 64
4.3 Results 65
4.3.1 Simulated basilar-membrane (BM) and auditory nerve (AN) response 65
4.3.2 On- versus off-frequency contributions to simulated auditory nerve responses 68

4.4 Discussion 70
4.4.1 Over-represented simulated AN response at high frequencies 70
4.4.2 A suggestion for improving the AN response at high frequencies 72
4.5 Conclusion 77

5 Envelope following response in mice with and without cochlear synaptopathy investigated through a computational model of the auditory nerve 79
5.1 Introduction 80
5.2 Methods 82
5.2.1 Experimental animals 82
5.2.2 Physiological tests 82
5.2.3 AN model 83
5.3 Results 85
5.3.1 Validating noise-induced synaptopathy in mice 85
5.3.2 EFR level-growth functions in exposed versus non-exposed mice 85
5.3.3 Simulated EFR level-growth functions in non-exposed animals 88
5.3.4 Simulated EFR level-growth functions in exposed animals 90
5.4 Discussion 93
5.4.1 EFR level-growth functions in exposed and non-exposed animals 93
5.4.2 Simulated EFR level-growth functions 94
5.4.3 On the need of a species specific computational model to study synaptopathy 97
5.5 Conclusion 98

6 General discussion 101
6.1 Summary 101
6.2 Estimating peripheral compression using envelope following responses: revisited 102
6.3 On the challenge of investigating the potential presence of cochlear synaptopathy in human listeners 104

6.3.1 EFRs at supra-threshold stimulus levels: A general observation 110
6.3.2 Allying heuristic innovation with computational power: a suggestion for the generation of novel hypotheses 111
6.4 Perspectives 112

Bibliography **115**

Collection volumes **133**

1 Introduction

Motivation and background

The ability to enable communication, particularly in challenging acoustical environments with high levels of background noise, is a fascinating property of the healthy auditory system. To enable this communication, the auditory system has to convert the local air pressure vibrations that constitute a sound wave into electrical impulses that the brain cells, the neurons, can process. In the auditory periphery, the acoustic energy entering the ear canal is transformed into mechanical vibration of the tympanic membrane in the outer ear. This vibration is transferred through the auditory ossicles in the middle ear, which excite an incompressible watery fluid in the cochlea (perylimph) via the oval window. This generates a fluid pressure gradient within the cochlea in the inner ear, leading to motion of the basilar membrane (BM). The organ of Corti, a sensory organ located on top of the BM, contains one row of sensory cells, known as inner hair cells (IHC), and three rows of actuator cells, the outer hair cells (OHC). The vibrational motion of the BM leads to deflection of the stereocilia (the hair bundles) of the IHC. This, in turn, leads to a release of neurotransmitters in the synaptic cleft between the IHC and the auditory nerve (AN) synaptic receptor. The OHCs have electromotile properties, thought to induce a local active gain to the BM motion. The neurotransmitters trigger the generation of action potentials (spikes) in the AN. Thus, in this way, sound energy is converted into a spike train which is transmitted further along the auditory pathway.

Hearing impairment can, among other factors, be caused by the failure or dysfunction of any of those fragile cells and structures involved in the transduction of sound vibrations to neural spike trains. Occlusion or infection of the outer or middle ear (e.g. otitis media) causes a conductive hearing loss, as the transmission of the sound vibration to the inner ear is hindered. Loss or dysfunction of hair cells and/or AN fibers leads to a sensorineural hearing

loss, which refers to dysfunction in the cochlear processes. Conductive hearing losses are generally treated with pharmaceutical intervention or corrective surgeries (e.g. tympanoplasty). Sensorineural hearing losses are mainly treated with strategies that either intend to compensate for moderate cochlear dysfunction (e.g. amplification through hearing aids) or that mimic cochlear processing through the implantation of a cochlear implant in cases of a profound hearing impairment. For cases of moderate hearing loss when the use of a hearing aid is prescribed, the treatment strategy is mainly based on the restoration of audibility, that is, hearing thresholds.

In order to provide an optimal restoration of the system, each type of hearing impairment should be accurately diagnosed prior to treatment. The gold-standard metric to evaluate the status of the auditory system is pure-tone threshold audiometry. Hearing impairment is revealed as a permanent elevation of hearing threshold. Conductive hearing losses can be distinguished from sensorineural because the outer and middle-ear can be partially assessed through direct visual inspection, middle ear function can be assessed by tympanometry and can also be bypassed with the use of bone-conducted audiometry. However, when a sensorineural hearing loss is diagnosed, it remains unknown whether the dysfunction arises from OHC, IHC or AN neuronal damage, or a combination of them. Although it is not possible to directly assess the differential cellular damage in humans, it is considered ethically acceptable to do so in non-human animals. Studies in animal models have shown that the loss of OHCs was well correlated with poor hearing sensitivity (Ryan and Dallos, 1975; Stebbins et al., 1979), but hearing thresholds were not altered in the absence of significant amounts of IHCs (Lobarinas et al., 2013; Oxenham, 2016; Wang et al., 1997), AN fiber cell bodies known as spiral ganglion cells (SGC) (Dandy, 1934a,b; Neff, 1947; Schuknecht and Woellner, 1955; Wever and Neff, 1947) and/or AN fiber synapses (Kujawa and Liberman, 2009; Liberman and Kujawa, 2017); highlighting the shortcoming of currently used methods.

In audiometric practice, it has been widely reported that despite showing normal sensitivity to pure-tones, some human listeners complain about having difficulties in complex acoustical situations (Hind et al., 2011; Kumar et al., 2007; Saunders and Haggard, 1989; Tremblay et al., 2015). Typically, such listeners have been diagnosed as suffering from nonspecific pathologies such as obscure auditory dysfunction (Saunders and Haggard, 1989), King-Kopetzky syndrome (Zhao and Stephens, 2007), central auditory processing disorders (Jerger et al.,

1990), auditory disability with normal hearing (King and Stephens, 1992) or auditory neuropathy (Rance and Starr, 2015; Starr et al., 1996). It may be plausible though that a specific physiological insult or alteration not revealed in the audiogram could explain such hearing difficulties. It is clear that the audiogram provides valuable information regarding the status of the peripheral auditory system, but it is also broadly accepted that the assessment of hearing thresholds does not fully characterize the status of the peripheral auditory system. Furthermore, daily communication occurs at sound levels well above hearing threshold. Therefore, novel diagnostic methods that can evaluate the integrity of the auditory system at supra-threshold processing can help to diagnose hearing loss more accurately.

The aim of this thesis was to investigate the use of electrophysiological methods to evaluate hearing function at supra-threshold stimulus levels. Supra-threshold processing can be investigated using psychoacoustical methods, but these methods rely on numerous assumptions to relate the obtained results to specific physiological processes. Electrophysiological methods, on the other hand, assess neural activity objectively. Specifically, envelope following responses (EFR) as a function of stimulation levels (level-growth) were used. EFRs are currently used in clinical practice, where the name auditory steady-state responses (ASSR) is more commonly used, for assessing hearing thresholds objectively (e.g.; Dimitrijevic et al., 2002; Michel and Jørgensen, 2016; Ozdek et al., 2010; Picton et al., 2005). In this work, it was proposed to extend the use of EFRs towards supra-threshold levels to firstly estimate peripheral compression and secondly to investigate the potential effect of damage to the auditory nerve fiber synapses (i.e. cochlear synaptopathy, see Kujawa and Liberman, 2009).

Normal function of OHCs leads to a compressive growth of the BM velocity with increasing stimulation level (Ruggero et al., 1997). OHC dysfunction leads to a linearization of this BM velocity growth, or in other words, to a loss of compression. The reduction of gain by loosing the active mechanism associated with OHC electromotility reduces the overall system's sensitivity to low-intensity sounds. Therefore, hearing thresholds become elevated. Hearing aids aim to restore such compressive processing by applying amplification to the low-intensity sounds that arrive to the patient's ears. Therefore, an accurate estimate of frequency-dependent residual compression is of interest for audiologists to potentially improve hearing-aid fitting. Many studies have attempted to estimate cochlear compression using psychoacoustical methods (e.g.; Fer-

eczkowski et al., 2017a; Lopez-Poveda and Alves-Pinto, 2008; Nelson et al., 2001; Oxenham and Plack, 1997; Plack et al., 2004) or distortion product otoacoustic emissions (DPOAE) (e.g.; Dorn et al., 2001; Neely et al., 2003, 2009). These measures evaluate the system as a whole (psychoacoustics) or the mechanical processing of the cochlea (DPOAE), but do not reflect neural activity. EFR might be an alternative to these methods when one is interested in the neural response, as this is the ultimate form in which acoustic information is encoded. In this thesis, it was hypothesized that the BM compressive processing affects the processing of a sinusoidally amplitude modulated tone (SAM) presented to the cochlea. Furthermore, this processing is assumed to be reflected in the EFR recorded as a function of stimulation level. Compression estimates obtained using EFRs were compared to estimates obtained from DPOAE level-growth functions in the same listeners, as OHC dysfunction has shown to be reflected in a reduction of DPOAE amplitudes (Hofstetter et al., 1997; Trautwein et al., 1996).

EFR level-growth functions obtained using full and shallow depths of modulation were investigated as a potential biomarker for cochlear synaptopathy. Cochlear synaptopathy, defined as a loss of AN fiber synapses without hair-cell damage (i.e. normal hearing thresholds), was initially reported by Kujawa and Liberman (2009) and has been demonstrated in several non-human animals, such as mice (Furman et al., 2013; Kujawa and Liberman, 2009), guinea pigs (Lin et al., 2011; Liu et al., 2012), rats (Lobarinas et al., 2016) and rhesus macaques (Valero et al., 2017). In addition, Furman et al. (2013) proposed that cochlear synaptopathy is selective to AN fibers with medium- and low-spontaneous rate (SR), but does not affect high-SR fibers. Such a loss of AN fibers was also found in mice that were not exposed to noise, where AN fiber synapses seem to naturally degenerate with aging (Sergeyenko et al., 2013). Furthermore, degeneration seems to be accelerated with noise-exposure (Fernandez et al., 2015). The finding of cochlear synaptopathy has had a profound impact on the hearing research community, as it has challenged the classical view of sensorineural hearing loss in which the primary targets of noise exposure and aging are hair cells (Kujawa and Liberman, 2015). The results from animal models suggest that such primary targets might indeed be the AN fiber synapses, and that cochlear synaptopathy may be a precursor, leading to traditional overt hearing loss (Kujawa and Liberman, 2006; Liberman and Kujawa, 2017). It is plausible then that hearing impaired listeners might suffer from a loss of synapses at frequencies

showing hearing threshold within normal limits, in particular for the elderly listeners. In fact, human temporal bone studies in cadavers have suggested that such age-related synaptopathy might also be present in humans (Makary et al., 2011; Viana et al., 2015). However, the question of whether noise-induced synaptopathy in humans exists is to date unresolved (Bharadwaj et al., 2015; Grose et al., 2017; Le Prell et al., 2017; Lopez-Poveda et al., 2017; Mehraei et al., 2016; Prendergast et al., 2016a,b).

By considering the contributions of the different types of AN fibers to level encoding, a heuristically derived hypothesis was formulated in the present study on how synaptopathy may affect EFR level-growth functions. Specifically, it was investigated whether synaptopathy is reflected as a reduction of the EFR magnitude measured as a function of stimulation levels at different modulation depths. It was hypothesized that shallowly modulated SAM tones presented at high stimulus levels are mainly encoded by medium- and low-SR fibers because the firing rate of the high-SR fibers is saturated at such high levels. Thus, in the presence of synaptopathy, the magnitude of the EFR should be reduced when more shallowly modulated stimuli are used due to a degraded encoding of the envelope fluctuations. In order to link these findings to physiological recordings in non-human animal models, similar EFR level-growth functions were recorded in noise exposed and non-exposed mice where synaptopathy can be quantified. Finally, a phenomenological model of the AN (Zilany et al., 2009, 2014) was used to relate the findings and to further analyze the EFR level-growth functions obtained from individual listeners and from mice.

In order to relate the EFR to the AN activity, it is required to determine the relative contribution of specific groups of AN neurons to the EFR. Stimuli presented at medium-to-high intensities produce a broad excitation of the BM (Robles and Ruggero, 2001) and recruit a broad range of AN neurons (Müller and Robertson, 1991; Rose et al., 1971). This challenges the interpretation of evoked responses in terms of place-specific activity at the level of the cochlea. Even though SAM tones have a very narrow spectral energy distribution, they excite a large number of neurons when presented at high stimulus levels. Therefore the neuronal activity at frequencies away from the characteristic frequency of the stimulus (off-frequency) can significantly contribute to the total scalp-recorded EFR. The AN model was used to investigate the effect of such on- versus off-frequency contributions on the EFRs, and to systematically simulate the potential effect of a postulated loss of the different types of AN fiber synapses.

Overview of the thesis

Chapter 2 describes the potential use of EFR level-growth functions to estimate compression in the peripheral auditory system. EFR level-growth functions were recorded in normal-hearing (NH) and hearing-impaired (HI) listeners using a multi-frequency paradigm to stimulate four different regions of the BM simultaneously. The slope of the EFR level-growth function was used as an estimate of peripheral compression at each tested frequency. The compression estimates obtained from EFRs were compared to estimates of compression obtained using DPOAE level-growth functions. The frequency specificity of the compression estimates obtained using EFRs is discussed. This chapter is based on Encina-Llamas et al. (under review[a]).

Chapter 3 describes the use of EFR level-growth functions as a potential tool to investigate cochlear synaptopathy in human listeners. EFRs were recorded in NH threshold and mild HI listeners using a single SAM tone with either full or shallow modulation depth. The HI listeners were tested at an audiometric frequency showing no threshold elevation. It was hypothesized that, assuming that cochlear synaptopathy is a precursor of hair-cell damage, a listener with a mild threshold elevation at a higher frequency is likely to show AN fiber synapse loss at lower frequencies. The AN phenomenological model of Zilany et al. (2009, 2014) was used to simulate the different individual patterns obtained within the NH threshold listeners group and from the HI listeners. This computational model was also used to systematically investigate the on- versus off-frequency (i.e. near or away from the characteristic frequency of the stimulus respectively) contributions and the different types of AN fibers. This chapter is based on Encina-Llamas et al. (under review[b]).

Chapter 4 describes an analysis of the model of Verhulst et al. (2015) to simulate EFRs in a large range of stimulus levels. The different stages along the peripheral auditory pathway (i.e. BM, IHC and AN) were simulated in response to the same stimuli used in the study of *Chapter 3*. The model analysis focuses on the relative contributions of the on- versus off-frequency processing of the stimulus envelope. After finding an over-represented simulated activity of the AN at high characteristic frequencies (CF), a new fitting is suggested for one of the parameters in the AN section of the model. The original model versus the newly suggested implementation are analyzed and compared.

In *Chapter 5*, a similar study as the one presented in *Chapter 3* for human

listeners is described in noise exposed versus non-exposed mice [1]. EFR level-growth functions for fully and shallowly modulated tones were used to investigate the effects of noise-induced synaptopathy onto the EFRs. In contrast to humans, synapses can be directly counted in mice. EFR level-growth functions were recorded at a frequency where no loss of synapses were observed, and at a frequency where synaptopathy was maximal. An animal version of the AN model of Zilany et al. (2009, 2014) was used to simulate the EFR level-growth functions and to investigate the contributions of on- versus off-frequency processing and the different types of AN fibers.

Finally, *Chapter 6* provides a summary of the main findings. The effect of off-frequency contributions on electrophysiological evoked responses obtained using stimuli presented at medium-to-high intensities is further discussed, in relation to the estimates of peripheral compression using EFRs. The challenge of finding cochlear synaptopathy in humans is discussed in more detail. A general observation regarding the use of EFRs at high stimulation levels is included. Perspectives for future research are suggested.

Main contributions of this thesis

The following contributions are reported in this thesis for the first time:

- EFR level-growth functions were considered to estimate compression in the peripheral auditory system both in NH and HI listeners.
- The slopes derived from EFR level-growth functions were compared with the slopes derived from DPOAE level-growth functions in the same NH listeners. The slopes obtained from DPOAEs were higher (less compressive) than the slopes obtained from EFRs.
- EFR level-growth functions of high resolution (many data points) for fully and shallowly modulated tones were obtained for young NH threshold listeners.
- EFR level-growth functions for fully and shallowly modulated tones were recorded in mild HI listeners presenting the stimuli at an audiometric frequency at which the listener's hearing threshold is within the normal

[1] The mice data was recorded by Dr. Aravindakshan Parthasarathy at the Massachusetts Eye and Ear Infirmary, Harvard Medical School in Boston, MA (USA)

range. Different EFR level-growth functions were obtained for fully than for shallowly modulated tones.

- This thesis reports one of the few studies that has investigated the presence of cochlear synaptopathy in human listeners by combining either human electrophysiology and computational modeling (together with Paul et al., 2016) or human and animal electrophysiological work plus computational modeling (together with Mehraei, 2016).

- Cochlear synaptopathy was reflected in the EFR level-growth functions for fully and shallowly modulated tones recorded in noise exposed mice.

- The simulations from the computational model suggested that, considering a distribution of AN fibers types as in the cat (Liberman, 1978), EFRs were dominated by the activity of high-SR fibers.

- The simulations from the computational model suggested that at medium-to-high stimulation levels the envelope of the stimulus was better encoded by AN fibers tuned to frequencies away from the characteristic frequency of the stimulus (off-frequency), so that the off-frequency contributions dominated the total EFR.

- The computational modeling analysis challenged the established idea that cochlear synaptopathy is selective to medium- and low-SR fibers and that high-SR fibers are unaffected (Furman et al., 2013), as loss of medium- and low-SR fibers only did not affect the simulated EFR level-growth functions in the framework of the model.

2

Estimates of peripheral compression using envelope following responses[a]

Abstract

The ability of the peripheral auditory system to compress the input level range of incoming sounds is an important property that enables communication in complex acoustical environments. Hearing impairment typically leads to a decrease in sensitivity and a reduction in compression. While the sensitivity can be measured efficiently via audiometry in individual listeners, no one measure exists that provides fast and reliable compression estimates suitable for clinical investigation. In the present study, intensity level-growth functions of envelope following responses (EFR) evoked by multiple sinusoidally amplitude modulated tones were measured in normal-hearing (NH) and hearing-impaired (HI) listeners. For comparison, distortion-product otoacoustic emission (DPOAE) level-growth functions were measured in the same NH listeners. The median values of compression estimated from the EFRs in the NH listeners were consistent with previously reported compression estimates based on psychoacoustical measures and group-averaged DPOAE level-growth functions in human listeners as well as basilar-membrane compression values measured invasively in non-human mammals. The EFR level-growth functions for the HI listeners were less compressive than for the NH listeners, indicating that EFR might be sensitive to a reduction of compression. The compression estimated based on DPOAEs was lower than that estimated based on EFRs in the same NH listeners, suggesting different underlying processes. Overall, the results show that EFR level-growth functions might provide a robust and consistent measure of auditory

[a] This chapter is based on Encina-Llamas et al. (under review[a])

peripheral compression, although the place-specificity of the compression estimates may be compromised due to contributions from off-frequency neural populations to the EFR.

2.1 Introduction

An important characteristic of the healthy mammalian auditory system is the compressive transformation of the large dynamic range of input sound pressure levels to a smaller range of levels that can be processed by the sensory cells. Part of this compressive transformation is a consequence of the processing by the outer hair cells (OHC) in the inner ear, the cochlea. Although there is still some controversy about the precise mechanism underlying OHC function (Ashmore, 2008; Dallos, 2008; Dong and Olson, 2013; Guinan, 2013; Gummer et al., 2002), it is broadly accepted that OHC electro-motility provides a level-dependent gain to the movement of the basilar membrane (BM) in the healthy cochlea. This leads to a high sensitivity to low-level sound and a compressive input-output (I/O) characteristic (Robles and Ruggero, 2001). In addition to increasing sensitivity, OHC function has also been associated with high frequency selectivity and a normal loudness growth with sound pressure level (Plack, 2013).

Invasive physiological recordings in alive non-human mammals allow precise measures of place-specific BM velocity level-growth functions using pure-tone stimuli (e.g.; Nuttall and Dolan, 1996; Recio and Rich, 1998; Rhode and Robles, 1974; Rhode and Recio, 2000; Ruggero et al., 1997). For a pure tone, the envelope of the resulting traveling wave shows a maximum at one specific place. Level-growth functions measured at this place ("on-frequency") show a level-dependent amplification, resulting in a compressive growth. Basal and apical to this place ("off-frequency"), the level-growth functions show a passive and linear growth (e.g.; Robles and Ruggero, 2001). The combination of active and nonlinear on-frequency and passive, linear off-frequency level-growth functions leads to a level-dependent BM excitation pattern with sharp tuning at low levels and broader tuning towards higher levels. In the case of OHC dysfunction, on-frequency level-growth functions demonstrate reduced compression or are linearized at medium levels. This leads to a lower amplitude on-frequency and a less level-dependent BM excitation pattern.

Direct measurements of the BM velocity are not possible in humans with the existing non-invasive techniques. However, behavioral measurements using

forward-masking paradigms have been developed to estimate BM level-growth functions in humans (e.g., Nelson et al., 2001; Oxenham and Plack, 1997). These rely on the assumption that it is possible to investigate frequency-specific or, equivalently, place-specific processing at the level of the BM, even though behavioral measures always also involve retro-cochlear processes and decision making. Thus, behavioral estimates of cochlear compression may reflect additional retro-cochlear compressive processing. Furthermore, these behavioral measures are typically time consuming and thus not well suited for clinical applications.

An objective and non-invasive estimate of cochlear compression can be made in humans using otoacoustic emissions (OAE). OAEs are low-level sounds that are measured in the ear canal, but are produced in the cochlea as a side-effect of the non-linear electro-mechanical processing induced by the OHC's electro-motility. Distortion product otoacoustic emissions (DPOAE) represent one type of OAE. DPOAEs are the result of intermodulation distortion that arises when two simultaneous pure tones (referred to as primaries) at frequencies f_1 and f_2 are presented, with $f_1 < f_2$ (e.g.; Robinette and Glattke, 2002). Based on the fact that the generation of OAEs exclusively reflects cochlear processing, it has been proposed that cochlear compression could be obtained by measuring the cubic distortion $2f_1 - f_2$ of the DPOAE level-growth functions (e.g.; Neely et al., 2003, 2009). The shape of the DPOAE level-growth function was found to differ between normal-hearing (NH) and hearing-impaired (HI) listeners. Dorn et al. (2001) reported that group averaged DPOAE level-growth functions in NH listeners showed a strongly compressive growth as a function of the sound pressure level (SPL) of the primary f_2 in the range from 40 to 80 dB. At lower and higher levels, i.e. below 40 dB and above 80 dB, reduced compression was observed. The DPOAE level-growth functions for HI listeners showed a reduced range and amount of compression, together with a higher minimum stimulus level required to evoke a measurable DPOAE. Neely et al. (2003) suggested that the slope of the DPOAE level-growth functions could be used as an estimate of cochlear compression by fitting the DPOAE level-growth function with a two-slope function. However, while the average slope estimates matched well with the compression estimates from physiological and behavioral experiments, the individual values varied strongly across subjects, such that individual DPOAE level-growth functions do not seem to have sufficient predictive value for clinical applications. The variability of the individual DPOAE level growth functions

might be due to the complexity of the generation mechanisms of DPOAEs. It is now understood and widely accepted that DPOAEs represent the superposition of two main sources from within the cochlea: distortion and (coherent) reflection sources (Kalluri and Shera, 2001; Shera et al., 1999). Recently, techniques have been developed to separate the distortion source from the coherent reflection source in DPOAE recordings (e.g. Long et al., 2008), which were shown to reduce the variability in individually derived DPOAE level-growth functions (Mauermann and Kollmeier, 2004).

Envelope following responses (EFR) represent another objective measure to investigate auditory function. EFRs, also referred to as auditory steady-state responses (ASSR), are gross electroencephalographic (EEG) potentials elicited by populations of neurons that fire synchronously (phase-locked) to the envelope of an acoustic stimulus. The EFR amplitude varies depending on the modulation frequency, with a predominant peak at 40 Hz (the so-called 40-Hz potential) and a smaller peak around 80-100 Hz (the 80-Hz potential; Picton et al., 2003). It has been suggested that EFRs to 80-100 Hz modulations are mainly generated by brainstem midbrain sources while EFRs to 40 Hz modulations are thought to have more dominant sources at cortical levels (Herdman et al., 2002a).

Due to their narrow bandwidth, multiple sinusoidally amplitude modulated (SAM) tones have been used to record EFRs evoked at multiple cochlear regions simultaneously (Herdman et al., 2002b; John et al., 1998; Lins and Picton, 1995). While the carrier frequency of each SAM tone determines the cochlear region to be excited, different modulation frequencies produce different peaks in the recorded EFR spectrum (Picton, 2010), making it possible to separate the responses in the frequency domain. Lins and Picton (1995) demonstrated that EFRs can be recorded using four simultaneous SAM tones modulated between 80-100 Hz, a technique used in clinical systems to estimate hearing thresholds objectively (John and Picton, 2000). At supra-threshold levels, EFR level-growth functions show either a strictly monotonic growth (Kuwada et al., 1986; Picton et al., 1987) or a monotonic growth followed by a plateau for stimulus levels above about 60 dB SPL when using multiple simultaneous SAM tones modulated around 80 Hz (Picton et al., 2007).

In the present study, EFR level-growth functions evoked by 80-Hz SAM tones were measured in NH and HI listeners. Since EFRs evoked by a 80-Hz modulation are predominantly elicited at the level of the auditory brainstem, the compression estimates derived from the slope of the EFR level-growth functions

might reflect both cochlear and brainstem (i.e., retro-cochlea) compression and will therefore be referred to as the more general term "peripheral" compression. Assuming that the compressive cochlear processing will reduce the modulation depth of the stimulus envelope at the output of the cochlea, such reduction should also be reflected in the magnitude of the EFR. It was anticipated that the amount of peripheral compression will be higher in NH listeners compared to HI listeners due to the reduced cochlear compression in the HI listeners with sensorineural hearing loss. Since multi-frequency EFRs have successfully been used to estimate thresholds, it was assumed that also peripheral compression estimates can be obtained simultaneously at different cochlear regions. In order to disentangle the contribution of cochlear compression from potential retro-cochlear compressive sources, also the level-growth functions of the DPOAE generator component were measured in the same group of NH listeners.

2.2 Methods

2.2.1 Listeners

A total of 20 adult listeners (10 males and 10 females, 34.0 ± 15.9 years) participated in this study, separated into 13 normal-hearing (5 males and 8 females, 24 ± 3.2 years) and 7 hearing-impaired (5 males and 2 females, 56.2 ± 12.7 years) listeners. All normal-hearing (NH) listeners had thresholds below 15 dB HL at octave frequencies between 125 and 8000 Hz. All hearing-impaired (HI) listeners were selected to have normal-hearing (threshold ≤ 20 dB HL) below 4000 Hz and a mild hearing impairment at 4000 Hz and above, with audiometric thresholds between 20 and 45 dB HL.

2.2.2 Apparatus

The EFR and DPOAE recordings were performed in a dark, soundproof and electrically shielded booth, where the listeners were seated in a comfortable reclined armchair. The listeners were instructed to close their eyes and relax to avoid moving and were allowed to sleep. The recording and data analysis routines were implemented in MATLAB (The MathWorks, Inc., Natick, Massachusetts, USA). All acoustic stimuli were generated in MATLAB and presented using PLAYREC 2.1 (Humphrey, R., www.playrec.co.uk, 2008-2014) via a RME Fireface UCX soundcard (sampling rate $f_{s|sound} = 48$ kHz, 24 bits). The analogue

acoustic signal was attenuated by a headphone buffer (HB7, Tucker-Davis Technologies) with an attenuation of 6 dB (stimulus levels > 55 dB SPL) or 27 dB (stimulus levels ≤ 55 dB SPL). The attenuated signal was presented through a pair of ER-2 insert earphones (Etymotic Research Inc.) mounted on an ER-10B+ low-noise DPOAE microphone probe (Etymotic Research Inc.) with ER10-14 foam eartips.

EFRs were recorded using a Biosemi ActiveTwo system (sampling rate $f_{s|EFR}$ = 8192 Hz, 24 bits). Five active pin-type electrodes were used. Three electrodes were mounted at positions P10, P9 and Cz following the 10-20 system (American Clinical Neurophysiology Society, 2006). The remaining two electrodes (DRL and CMS) were placed at the center of the parieto-occipital coronal line (on either side of electrode POz). Conductive electrode gel was applied and the offset voltage was stabilized at < 20 mV for each electrode. The recorded EEG signals were down sampled by a factor of 2, and low-pass filtered with a bandwidth limit of $^1/_5{}^{th}$ of the final sampling frequency (about 820 Hz). The EEG data were stored to hard disk. The results shown in this study represent the Cz-P10 potential in response to right-ear stimulation, and the Cz-P9 potential in response to left-ear stimulation.

For the measurement of DPOAEs, the stimulus presentation apparatus was the same as for the EFR measurements, with the modification that the ER-10B+ ear probe was connected to the ER-10B+ pre-amplifier (with a gain of 20 dB). The microphone signal was bandpass filtered using a cascade of a high-pass filter (Rockland model 852, Butterworth 48 dB/octave, cut-off frequency 100 Hz) and a low-pass filter (cut-off frequency 9 kHz, same as HP-filter). The recorded signal was digitalized using the sound card at $f_{s|DPOAE}$ = 48 kHz with 24 bits resolution and stored to hard disk for post-processing.

2.2.3 EFR recordings

The EFR data were recorded in two sessions. In the first session (approx. two hours in duration), the EFR level-growth functions were recorded in the NH listeners using input levels in the range from 20 to 80 dB SPL, in steps of 5 dB. A multi-frequency stimulation paradigm was used as described below. The second recording session (approx. 45 minutes in duration) took place on a different day usually about one month later than the first session. Three input levels (35, 55 and 70 dB SPL) were recorded again in the same NH listeners to evaluate the repeatability of the results. In all NH listeners, the right ear was

Table 2.1: Duration of EFR stimuli for each input level used in the NH listeners.

Input level [dB SPL]	20	25	30	35	40	45	50	55	60	65	70	75	80
Duration [min]	12	12	11.2	10.13	8.53	8.53	7.73	7.2	7.2	6.67	5.6	5.6	5.6

stimulated. In the HI group, the multi-frequency recording was carried out in the level range from 30 to 80 dB SPL, in steps of 5 dB. Here, the recording ear was chosen depending on the individual listener's audiogram, such that the amount of sensitivity loss due to the hearing impairment was as similar as possible within the group. There was no second recording session to evaluate repeatability for this subject group.

The multi-frequency stimulus consisted of four SAM tones. The SAM tones had carrier frequencies of 498, 1000, 2005 and 4011 Hz (referred to as 500, 1000, 2000 and 4000 Hz throughout this work) modulated at 81, 87, 93 and 98 Hz, respectively. The modulation depth was set to $m = 85\%$. The four SAM tones were calibrated individually (B&K 4157 ear simulator) and added together, resulting in a final stimulus level that was 6 dB higher than that of each individual SAM tone. The stimuli were digitally generated as 1-s long epochs and continuously presented to the listener in a loop, where a trigger signal marked the beginning of a new epoch for later averaging. The total stimulus duration depended on the stimulus intensity to achieve a statistically significant EFR signal-to-noise ratio (SNR), based on a pilot study. Table 3.1 shows the stimuli duration used for each input level in the EFR recordings.

The recorded EEG data were filtered using a fourth-order Butterworth digital band-pass filter with cut-off frequencies of 60 and 400 Hz, applied in forward and backward direction to yield zero phase. All recorded epochs that exceeded a voltage amplitude of $\pm$ 80 µV in any of the channels were rejected. In order to increase the SNR, weighted averaging was applied (John et al., 2001). Sixteen 1-s long epochs were concatenated, to achieve a higher frequency resolution in the EFR spectrum analysis.

An F-test was used to identify statistically significant responses by comparing the spectral power at the modulation frequency (EFR frequency) to the noise power in the range of 3 Hz below and above the modulation frequency (Dobie and Wilson, 1996; Picton et al., 2003). The power ratio (*F-ratio*) was calculated as the power in the EFR frequency bin divided by the averaged power in 3 Hz below and above the modulation frequency (96 bins). The probability (p) of the EFR power being different from the noise power can be calculated as $1-F$,

with F representing the cumulative distribution function of the power ratio. The F-test was defined to be positive if $p \leq 0.01$, implying that the EFR frequency was statistically significant from the noise estimate.

To estimate the slopes of the EFR level-growth functions, a piecewise linear function with two segments was used. The model was fitted to each individual EFR level-growth data using a non-linear least squares fitting method described by:

$$f(L_s) = \begin{cases} s_1 \cdot (L_s - b_x) + b_y & \text{if } L_s < b_x \\ s_2 \cdot (L_s - b_x) + b_y & \text{if } L_s > b_x \end{cases} \tag{2.1}$$

where, L_s represents the stimulus input level, s_1 the lower slope, s_2 the upper slope and b_x and b_y represent the value on the abscissa and the ordinate at the breakpoint respectively. Motivated by the BM I/O function characteristics observed either in direct animal physiological recordings (e.g.; Ruggero et al., 1997) or human psychoacoustical estimates of cochlear compression (e.g., Nelson et al., 2001; Oxenham and Plack, 1997; Plack, 2013), the lower slope was forced to be larger than the upper slope (i.e. $s_1 > s_2$).

If a single-slope (first-order polynomial) was found to provide a better fit (based on an adjusted R^2 statistic) of a given EFR level-growth function than the two-slopes piecewise functions, this simpler model was used. Only statistically significant data points were included in the entire fitting procedure.

Repeatability of EFR responses at three stimulus levels was assessed as proposed by Bland and Altman (1986). The test-retest difference values were plotted against the mean response amplitude between two test runs. This method defines the test repeatability coefficient as twice the standard deviation of the differences. Repeatability can also be expressed as a percentage of the mean amplitude within a given frequency-level group (repeatability coefficient / mean of the group $\cdot$ 100), termed repeatability variability (e.g.; D'Haenens et al., 2008; Wilding et al., 2012).

In order to test whether the estimated EFR slopes at two different frequencies were statistically different, a two-sample permutation test for equality of the means was used (Ernst, 2004; Fisher, 1935). This was done to evaluate the hypothesis that the estimated EFR slopes for two given frequencies were a random partition of both frequency data added together, against the alternative hypothesis that the EFR slopes from one frequency were part of a population with a different mean than the other frequency. The test was performed using 1

million permutations.

2.2.4 DPOAE recordings

DPOAEs were only recorded in the NH group (excluding listener NH03, who could not participate). They were measured using a sweeping technique (Long et al., 2008) and the non-linear distortion source was unmixed using a time-windowing method (Kalluri and Shera, 2001; Knight and Kemp, 2001).

The sweeping primaries consisted of two upward sweeps of equal level with a duration of 10 seconds and a sweep rate of half an octave per second (Long et al., 2008). Primary f_2 started at 250 Hz and ended at 8000 Hz while keeping the frequency ratio fixed at $f_2/f_1 = 1.22$. The distortion source DPOAE level-growth functions of the $2f_1 - f2$ component were recorded at levels of the primaries ranging from 30 to 65 dB SPL, in 5 dB steps. The acoustic waveforms recorded in the ear canal were analyzed in overlapping, windowed time frames (Hanning window of 0.5 s, or 24000 samples, and a step size of 600 samples). Each windowed frame was analyzed using a least-squares-fit procedure which estimated the magnitude and phase of a sinusoid to the expected DPOAE component within that frame (Long et al., 2008).

The recorded sweeps were averaged to increase the SNR of the DPOAE. Prior to averaging, the SNR in each time frame was estimated to reject noisy frames. The noise was estimated by averaging two frames with inverting the phase of the second one by 2π rad to remove deterministic components. The SNR estimation was defined as the difference between the magnitudes of the DPOAE non-linear distortion component and the estimated background noise in the same frequency bins in the recording spectrum. Two stopping rules were defined to stop the recording. The recording was ended for each stimulus level either if the SNRs in all four frequency bins of interest (0.5, 1, 2 and 4 kHz) were above 10 dB or when 8 pairs of sweeps had been recorded.

A simple first-order polynomial model was fitted to the statistically significant data points with respect to the background noise estimates. From the slope (s) of the fitted polynomial, the compression estimate was calculated as its derivative, as proposed in Neely et al. (2003).

2.3 Results

The data reported in this chapter is publicly available here (Encina-Llamas et al., 2017a).

2.3.1 EFR level-growth functions for normal-hearing and hearing-impaired listeners

Normal-hearing listeners

Figure 2.1 shows EFR level-growth functions from one representative NH listener (NH01) for 500 Hz, 1000 Hz, 2000 Hz and 4000 Hz (panels A-D, respectively). The complete set of EFR level-growth functions for all NH individuals can be found in section 2.5 (Fig. 2.7). The recorded EFR magnitudes are shown as circles, represented in dB relative to 1μV. The EFR magnitudes obtained in the second recording session are indicated as squares. Filled symbols represent statistically significant responses (*F(2, 96); p* $\leq$ 0.01), whereas open symbols represent non-significant (*F(2, 96); p* $>$ 0.01) responses. The estimated EEG background noise is indicated by the light-gray shaded area. The fitted function is shown by the continuous dark-gray line. A linear reference with a slope of 1 dB/dB is included as the dotted line.

All EFR level-growth functions grew monotonically and compressively (with slopes $<$ 1) up to stimulus levels of about 50-65 dB SPL. EFR level-growth functions from the carrier frequencies 500, 1000 and 2000 Hz (panels A-C) showed a different trend than for the 4000 Hz (panel D) carrier. At 500, 1000 and 2000 Hz, the EFR magnitudes saturated or slightly decreased for stimulus levels above 50-65 dB SPL, leading to a breakpoint in the level-growth function. This is also reflected in Figure 2.2B (blue symbols) that shows boxplots with fitted breakpoint levels at the four carrier frequencies. The median values for the breakpoint levels varied between 50 to 65 dB SPL. In contrast, no breakpoint was observed at 4000 Hz. At this frequency, the level-growth functions grew monotonically with a single slope (see also table 2.2).

The median values of the EFR slopes were 0.24 dB/dB at 500 Hz, 0.31 dB/dB at 1000 Hz, 0.25 dB/dB at 2000 Hz and 0.21 dB/dB at 4000 Hz, as indicated in Figure 2.2A (blue symbols). A two-sample permutation test for equality of the means (Ernst, 2004; Fisher, 1935) revealed that the estimated EFR slopes in the NH listeners were not statistically different across frequency, except for the

Table 2.2: Fitted parameters to the EFR level-growth functions for all NH listeners at all frequencies. The values in the brackets indicate the lower and upper 95% confidence interval (CI).

	Frequency			
Listener	500 Hz	1000 Hz	2000 Hz	4000 Hz
NH01	s_1 = 0.49 (0.40, 0.57) s_2 = -0.02 (-0.47, 0.42) b_x = 65.0 (54.1, 75.9) b_y = -18.6 (-23.2, -14.1)	s_1 = 0.52 (0.40, 0.64) s_2 = -0.16 (-0.57, 0.25) b_x = 61.8 (53.4, 70.3) b_y = -16.0 (-19.8, -12.1)	s_1 = 0.37 (0.28, 0.46) s_2 = -0.01 (-0.13, 0.11) b_x = 53.8 (46.6, 60.8) b_y = -22.4 (-24.2, -20.6)	s = 0.33 (0.29, 0.37) b = -43.5 (-45.7, -41.3)
NH02	s_1 = 0.71 (-0.04, 1.46) s_2 = -0.01 (-021, 0.20) b_x = 41.7 (31.6, 51.9) b_y = -28.4 (-32.6, -24.3)	s_1 = 0.42 (0.17, 0.67) s_2 = 0.04 (-0.16, 0.24) b_x = 50.0 (33.9, 66.1) b_y = -24.8 (-28.8, -20.8)	s_1 = 0.38 (0.22, 0.54) s_2 = -0.11 (-0.43, 0.22) b_x = 57.0 (44.6, 69.3) b_y = -22.2 (-26.1, -18.4)	s = 0.17 (0.11, 0.23) b = -39.2 (-42.4, -36.0)
NH03	s = 0.24 (0.08, 0.40) b = -41.3 (-51.0, -31.7)	s_1 = 0.77 (0.12, 1.42) s_2 = 0.18 (0.11, 0.25) b_x = 43.0 (35.8, 50.1) b_y = -28.9 (-31.3, -26.4)	s = 0.32 (0.24, 0.40) b = -47.6 (-52.2, -42.9)	s = 0.31 (-0.10, 0.72) b = -47.7 (-71.3, -24.1)
NH04	s = 0.19 (0.12, 0.26) b = -41.1 (-45.3, -36.8)	s = 0.16 (0.07, 0.25) b = -39.9 (-45.7, -34.2)	s = 0.10 (0.00, 0.20) b = -39.3 (-45.7, -32.8)	s = 0.18 (-0.30, 0.67) b = -44.6 (-78.1, -11.0)
NH05	s_1 = 0.45 (0.11, 0.78) s_2 = 0.17 (0.05, 0.28) b_x = 55.0 (40.7, 69.3) b_y = -27.0 (-30.5, -23.6)	s_1 = 0.37 (0.19, 0.56) s_2 = 0.01 (-0.23, 0.25) b_x = 56.0 (42.7, 69.3) b_y = -23.2 (-27.1, -19.3)	s_1 = 0.29 (0.15, 0.42) s_2 = -0.28 (-0.60, 0.03) b_x = 53.8 (46.6, 60.8) b_y = -21.6 (-23.9, -19.3)	s = 0.24 (0.15, 0.32) b = -39.5 (-44.3, -34.7)
NH06	s = 0.20 (0.03, 0.37) b = -42.9 (-54.0, -31.9)	s_1 = 0.38 (0.30, 0.47) s_2 = -0.02 (-0.13, 0.09) b_x = 51.6 (45.2, 58.0) b_y = -25.4 (-27.2, -23.5)	s_1 = 0.31 (0.13, 0.49) s_2 = -0.48 (-1.31, 0.35) b_x = 64.9 (56.1, 73.7) b_y = -29.3 (-32.7, -25.8)	s = 0.16 (0.09, 0.24) b = -42.1 (-46.8, -37.4)
NH07	s = 0.13 (-0.39, 0.66) b = -36.3 (-68.3, -4.3)	s_1 = 0.21 (0.16, 0.25) s_2 = -0.24 (-0.47, -0.02) b_x = 69.4 (65.0, 73.7) b_y = -23.2 (-24.2, -22.2)	s_1 = 0.20 (0.09, 0.32) s_2 = -0.11 (-0.76, 0.56) b_x = 70.0 (49.7, 90.3) b_y = -23.8 (-26.9, -20.6)	s = 0.18 (0.13, 0.24) b = -39.6 (-42.5, -36.6)
NH08	s = 0.24 (0.01, 0.46) b = -41.3 (-55.3, -27.2)	s = 0.24 (0.19, 0.28) b = -41.1 (-43.6, -38.6)	s = 0.22 (0.14, 0.30) b = -39.2 (-43.7, -34.7)	s = 0.21 (0.15, 0.28) b = -41.8 (-45.6, -38.0)
NH09	s = 0.19 (0.06, 0.32) b = -39.3 (-47.7, -30.9)	s_1 = 0.86 (0.17, 1.55) s_2 = -0.13 (-0.44, 0.18) b_x = 57.7 (50.2, 65.2) b_y = -22.1 (-26.0, -18.2)	s_1 = 0.21 (0.14, 0.28) s_2 = -0.17 (-0.39, 0.03) b_x = 62.5 (55.0, 70.0) b_y = -25.4 (-26.9, -23.8)	s = 0.17 (0.06, 0.28) b = -36.7 (-43.5, -29.9)
NH10	s_1 = 0.61 (0.22, 1.01) s_2 = 0.18 (0.02, 0.34) b_x = 46.3 (34.2, 58.5) b_y = -26.2 (-31.3, -21.1)	s_1 = 0.31 (0.26, 0.37) s_2 = -0.08 (-0.25, 0.09) b_x = 63.7 (58.0, 69.4) b_y = -18.1 (-29.6, -16.7)	s = 0.20 (0.10, 0.29) b = -38.1 (-43.7, -32.5)	s = 0.26 (0.16, 0.36) b = -41.7 (-47.5, -35.9)
NH11	s = 0.19 (0.04, 0.34) b = -39.7 (-49.2, -30.2)	s = 0.15 (0.04, 0.26) b = -35.7 (-42.9, -28.6)	s = 0.04 (-0.07, 0.16) b = -30.5 (-37.8, -23.2)	s = 0.19 (0.05, 0.34) b = -40.0 (-49.2, -30.9)
NH12	s = 0.26 (0.23, 0.28) b = -46.4 (-48.0, -44.8)	s = 0.25 (0.14, 0.36) b = -43.8 (-49.9, -37.7)	s = 0.25 (0.14, 0.37) b = -43.6 (-50.7, -36.5)	s = 0.31 (0.25, 0.37) b = -47.5 (-51.1, -43.8)
NH13	s_1 = 0.28 (0.20, 0.35) s_2 = 0.00 (-0.25, 0.24) b_x = 63.6 (51.9, 75.4) b_y = -26.4 (-29.3, -23.5)	s = 0.24 (0.11, 0.38) b = -39.4 (-48.2, -30.7)	s_1 = 0.58 (0.38, 0.78) s_2 = 0.00 (-0.20, 0.20) b_x = 60.0 (54.4, 65.6) b_y = -24.6 (-27.3, -21.9)	s = 0.31 (0.19, 0.43) b = -45.8 (-53.8, -37.9)

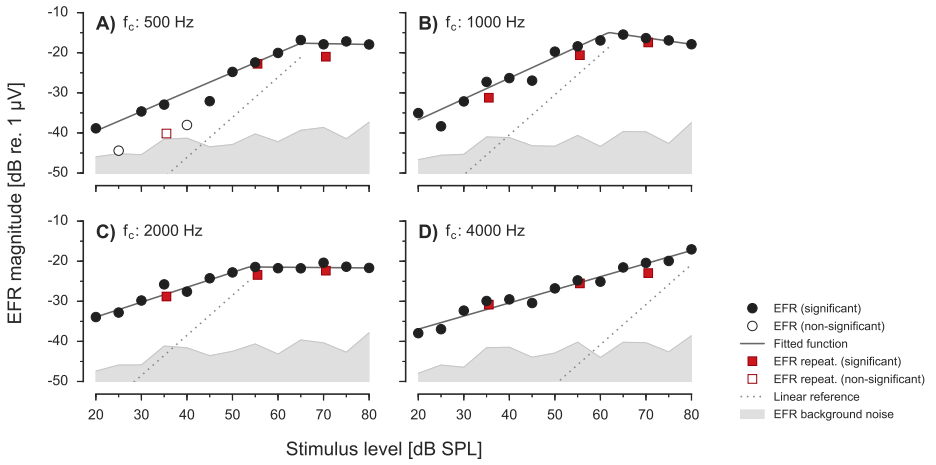


Figure 2.1: EFR level-growth function recorded in one representative NH listener (NH01) for the carrier frequencies of A) 500 Hz, B) 1000 Hz, C) 2000 Hz and D) 4000 Hz. EFR magnitudes are represented as filled symbols in case of a statistically significant response (positive F-test), and as open symbols in case of statistically non-significant (negative F-test) responses. Circles indicate the EFR magnitudes recorded in the first recording session and squares indicate EFR magnitudes recorded in the second recording session. EEG background noise estimates are shown as the gray shaded area. The best fitted curve is represented as a continuous dark-gray line. A linear reference with slope of 1 dB/dB is indicated by the dotted line.

conditions at 1000 Hz vs 4000 Hz ($p = 0.0272$).

Figure 2.3 shows the repeatability of the EFR amplitudes for all NH listeners and carrier frequencies of 500, 1000, 2000 and 4000 Hz (panels A-D) and stimulus levels of 35, 55 and 70 dB SPL. The repeatability coefficients of EFR amplitudes increase with stimulus level at 500 Hz and 4000 Hz, decrease with stimulus level at 1000 Hz, and hardly vary with stimulus level at 2000 Hz, thus showing no consistent trend across frequencies (see also table 2.3). The repeatability variability results obtained in this study were similar to those presented in previous studies, even though the EEG systems used and the stimuli were different (D'Haenens et al., 2008; Wilding et al., 2012).

Hearing-impaired listeners

Figure 2.4 shows the EFR level-growth functions from one representative HI listener (HI01) in the same representation as Figure 2.1. The complete set of EFR level-growth functions for all HI individuals can be found in Figure 2.8 in section 2.5. The EFR level-growth functions for 500, 1000 and 2000 Hz carrier frequencies (panels A-C) showed very similar trends as observed in the NH

Table 2.3: Repeatability values for all NH listeners at stimulation levels of 35, 55 and 70 dB SPL. EFR repeatability coefficient in nV. Mean in dB re 1 µV of the test-retest EFR amplitudes mean for each level group ± repeatability coefficient also converted to dB (in brackets). Repeatability variability in percentage and repeatability variability derived from D'Haenens et al. (2008).

	Frequency											
	500 Hz			1000 Hz			2000 Hz			4000 Hz		
	Level [dB SPL]			Level [dB SPL]			Level [dB SPL]			Level [dB SPL]		
	35	55	70	35	55	70	35	55	70	35	55	70
Repeatability coefficient [nV]		22.5	37.3	30.7	27.2	21.0	22.6	24.7	23.7	4.8	12.2	29.7
Group mean (± rep. coeff.) [dB re. 1µV]		-26.8 (-32.75, -23.36)	-25.4 (-35.63, -20.81)	-29.2 (-47.67, -23.67)	-24.3 (-29.38, -21.07)	-22.9 (-25.9, -20.65)	-30.2 (-41.65, -25.45)	-24.9 (-29.9, -21.8)	-25.0 (-29.7, -21.92)	-31.9 (-33.67, -30.39)	-29.2 (-32.9, -26.56)	-25.4 (-32.48, -21.6)
Repeatability variability [%]		49.4	69.3	88.1	44.5	29.3	73.2	43.5	42.0	18.7	34.9	55.5
Repeatability variability [%] (from D'Haenens et al., 2008)	144.3	55.0		53.5	40.5		57.7	41.6		72.2	44.8	

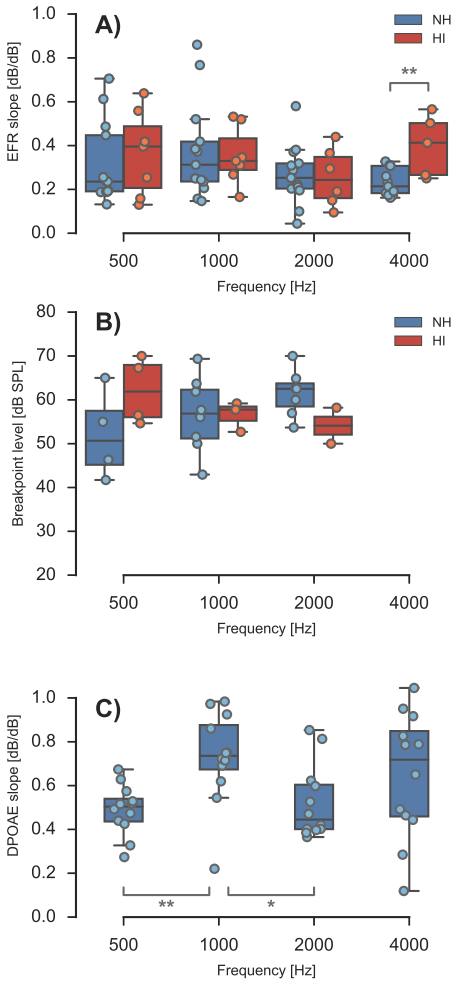


Figure 2.2: Fitted parameters to EFR level-growth functions in NH and HI listeners and to DPOAE level-growth functions to NH listeners. Panel A shows boxplots with the fitted slope at all four carrier frequencies in NH (blue) and HI (red) listeners. Panel B shows boxplots with the fitted breakpoint level at all four carrier frequencies in NH (blue) and HI (red) listeners when the two-slopes piecewise fit was used. Panel C shows boxplots with the fitted slope to the DPOAE level-growth functions in NH listeners. All boxplots are Tukey boxplots, where the bottom and top of the box are the first and third quartiles respectively, and the band inside the box is the second quartile (the median). Whiskers show 1.5 of the interquartile range (IQR) of the lower and upper quartile. The circles depicts the raw observations. Statistical significance is represented by the asterisks, where * corresponds to a $p\text{-}value \leq 0.05$ and ** corresponds to a $p\text{-}value \leq 0.01$.

listener (Fig. 2.1). At 4000 Hz (panel D), the EFR magnitudes for stimulus levels up to 60 dB SPL were not statistically different from the EEG background noise, whereas significant EFR magnitudes were obtained above 60 dB SPL, showing a compressive growth with level. This frequency is within the region of reduced sensitivity in this listeners' audiogram (red arrow in panel D). Overall, the EFR magnitudes recorded in some HI listeners showed a lower SNR than in the NH listeners, resulting in a larger number of statistically non-significant data points (table 2.4).

The slopes of the EFR level-growth functions at 500, 1000 and 2000 Hz (i.e. the audiometrically normal frequencies for all listeners) were not statistically different between the NH listeners and the HI listeners (See Fig. 2.2A). However,

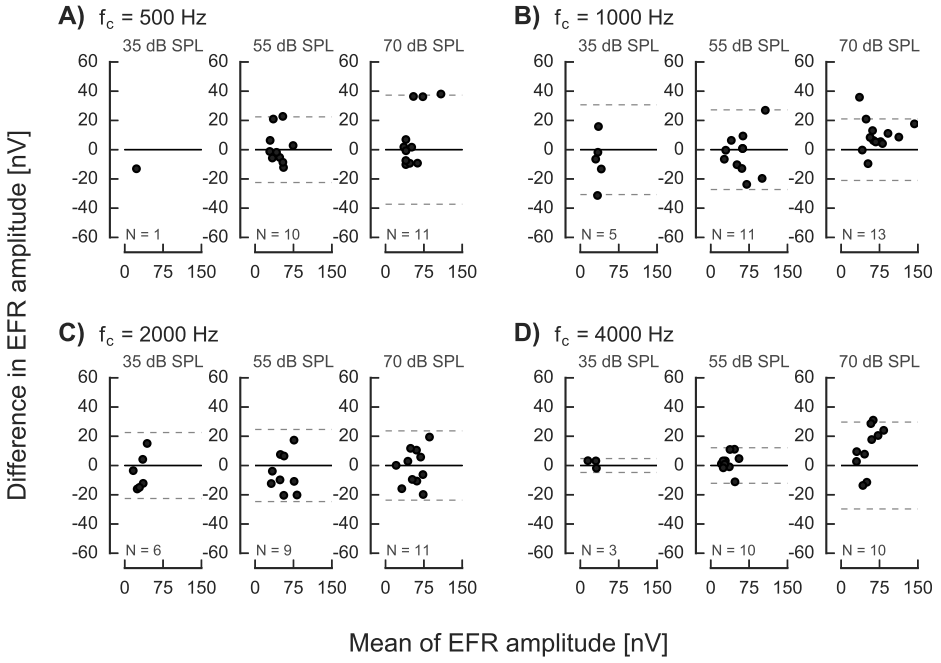


Figure 2.3: Repeatability of EFR amplitudes in NH listeners for carrier frequencies of 500, 1000, 2000 and 4000 Hz (panels A-D) and stimulus levels of 35, 55 and 70 dB SPL. For each plot, the data points show the mean of the EFR amplitude in nV on the x-axis vs. the test-retest difference EFR amplitudes on the y-axis. The upper and lower gray dashed lines show the ± calculated repeatability coefficient. N indicates the number of data points used on each plot (only statistically significant EFR responses were considered).

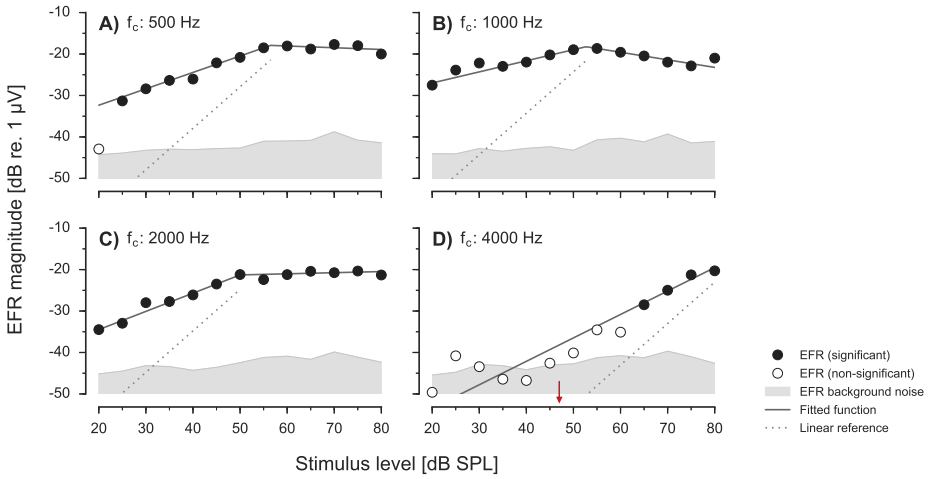


Figure 2.4: EFR level-growth function recorded in one representative HI listener (HI01) at the carrier frequencies of A) 500 Hz, B) 1000 Hz, C) 2000 Hz and D) 4000 Hz. Same as in Figure 2.1 but without the repeatability measurements. The small red arrow in panel D indicates the behavioral hearing threshold of the listener at 4000 Hz in dB SPL.

Table 2.4: Fitted parameters to the EFR level-growth functions for all HI listeners at all frequencies, including lower and upper 95% CI in brackets.

Listener	Frequency			
	500 Hz	1000 Hz	2000 Hz	4000 Hz
HI01	s_1 = 0.40 (0.32, 0.47) s_2 = -0.04 (-0.17, 0.09) b_x = 56.5 (50.9, 62.1) b_y = -17.9 (-19.7, -16.1)	s_1 = 0.27 (0.16, 0.37) s_2 = -0.18 (-0.32, -0.05) b_x = 52.7 (45.7, 59.6) b_y = -18.2 (-19.9, -16.6)	s_1 = 0.44 (0.34, 0.54) s_2 = 0.03 (-0.05, 0.11) b_x = 50.0 (44.2, 55.8) b_y = -21.3 (-22.8, -19.8)	s = 0.57 (0.18, 0.95) b = -64.8 (-93.0, -36.5)
HI02	s_1 = 0.64 (0.35, 0.93) s_2 = -0.52 (-1.83, 0.79) b_x = 70.0 (60.6, 79.4) b_y = -24.8 (-30.7, -18.9)	s = 0.35 (0.00, 0.69) b = -52.3 (-76.3, -28.3)	s = 0.30 (0.04, 0.55) b = -51.4 (-68.7, -34.1)	s = 0.25 (-1.96, 2.46) b = -48.8 (-211.5, 114.0)
HI03	s_1 = 0.42 (0.31, 0.52) s_2 = -0.72 (-0.96, -0.48) b_x = 67.3 (65.0, 69.6) b_y = -22.2 (-23.4, -20.9)	s_1 = 0.52 (0.04, 0.99) s_2 = -0.02 (-0.23, 0.20) b_x = 59.2 (48.9, 69.5) b_y = -22.7 (-25.3, -20.0)	s = 0.15 (0.04, 0.26) b = -35.2 (-42.1, -28.2)	s = 0.27 (0.01, 0.52) b = -48.2 (-65.7, -30.6)
HI04	s_1 = 0.56 (0.11, 1.01) s_2 = -0.14 (-0.38, 0.10) b_x = 54.7 (44.5, 64.8) b_y = -28.2 (-31.4, -25.0)	s_1 = 0.31 (0.26, 0.37) s_2 = -0.07 (-0.14, 0.01) b_x = 57.8 (54.0, 61.5) b_y = -23.7 (-24.5, -22.8)	s_1 = 0.41 (0.36, 0.47) s_2 = 0.04 (-0.03, 0.12) b_x = 55.6 (51.7, 59.6) b_y = -25.7 (-27.0, -24.4)	s = 0.41 (0.26, 0.57) b = -58.1 (-68.3, -47.9)
HI05	s = 0.13 (-0.83, 1.09) b = -34.6 (-99.8, 30.6)	s = 0.33 (-0.11, 0.77) b = -39.3 (-57.4, -21.3)		s = 0.50 (-3.32, 4.33) b = -61.9 (-349.2, 225.4)
HI06	s = 0.16 (-0.30, 0.61) b = -34.4 (-64.6, -4.2)	s = 0.53 (0.38, 0.68) b = -53.7 (-61.0, -46.5)	s = 0.19 (-0.02, 0.40) b = -35.3 (-46.7, -23.9)	
HI07	s = 0.26 (-0.31, 0.82) b = -45.8 (-85.0, -6.6)	s = 0.17 (0.06, 0.28) b = -38.7 (-45.1, -32.2)	s = 0.10 (-0.05, 0.24) b = -34.9 (-43.4, -26.4)	

the median HI listeners' slopes at 4000 Hz were significantly higher than the slopes in the NH listeners ($p = 0.0074$).

2.3.2 DPOAE level-growth functions for NH listeners

Figure 2.5 shows the estimate of compression based on the distortion source DPOAE level-growth functions recorded in one representative NH listener (NH05) for the four different carrier frequencies (panels A-D). The complete set of DPOAE level-growth functions for all NH individuals can be found in Figure 2.9 in section 2.5. The recorded DPOAE magnitudes as a function of the stimulus level are represented as circles. The filled symbols represent DPOAE responses with SNRs > 10 dB. The estimated background noise magnitudes are indicated by the light-gray shaded area. The fitted first-order polynomial function is represented by the continuous dark-gray line. A linear reference with a slope of 1 dB/dB is indicated by the dotted line.

The distortion source DPOAE level-growth functions showed a monotonic increase with level throughout the whole measured level range for all carrier frequencies (panels A-D). The DPOAE level-growth functions at 1000 and 4000

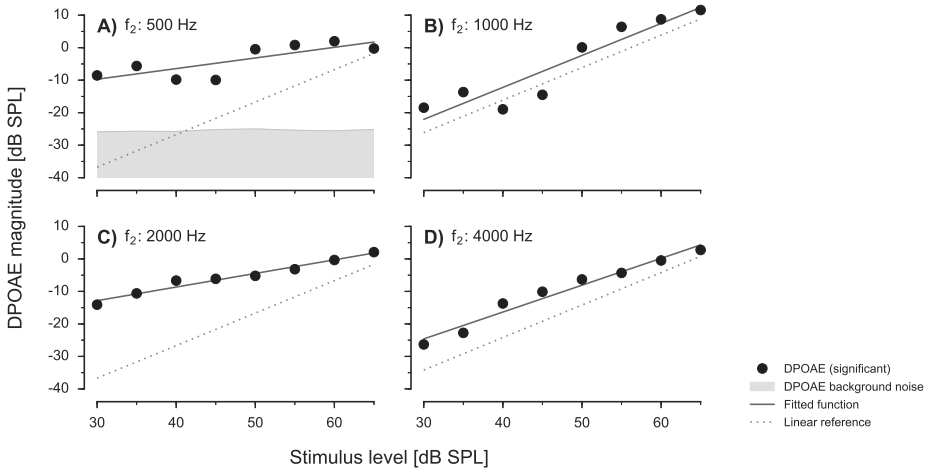


Figure 2.5: Distortion source DPOAE level-growth function recorded in one representative NH listener (NH05) at the carrier frequencies of A) 500 Hz, B) 1000 Hz, C) 2000 Hz and D) 4000 Hz. DPOAE responses with SNRs > 10 dB are represented as filled circles. Estimated background noise magnitudes are indicated by the light-gray shaded area. The fitted first-order polynomial model is represented as a continuous dark-gray line. A linear reference with slope of 1 dB/dB is indicated by the dotted line.

Hz (panels B and D) showed a weakly compressive growth, whereas the DPOAE level-growth functions at 500 and 2000 Hz (panels A and C) showed a more compressive growth for this listener. This difference in slope across carrier frequencies is also reflected in the group data summarized in Figure 2.2C. The median values of the DPOAE slopes in the NH listeners were 0.50 dB/dB at 500 Hz, 0.74 dB/dB at 1000 Hz, 0.45 dB/dB at 2000 Hz and 0.72 dB/dB at 4000 Hz. The estimated DPOAE slopes thus showed some variability across frequencies as well as a large variability at 4000 Hz. The DPOAE slopes at 500 Hz were statistically different from those at 1000 Hz ($p = 0.0025$). Likewise, the slopes at 1000 Hz were statistically different from those at 2000 Hz ($p = 0.015$). Table 2.5 shows the parameters of the first-order polynomial fitted to the DPOAE level-growth function for all tested NH individual listeners.

2.4 Discussion

2.4.1 Compression estimates based on EFR level-growth functions

The EFR level-growth function slopes for NH listeners varied between 0.2 and 0.35 dB/dB, which is consistent with cochlear compression estimates obtained

Table 2.5: Fitted parameters to the EFR level-growth functions for all NH listeners at all frequencies, including lower and upper 95% CI in brackets.

Listener	Frequency			
	500 Hz	1000 Hz	2000 Hz	4000 Hz
NH01	s = 0.27 (0.11, 0.44) b = -18.0 (-26.1, -10.0)	s = 0.69 (0.38, 1.00) b = -39.8 (-54.9, -24.6)	s = 0.62 (0.23, 1.02) b = -47.7 (-67.0, -28.4)	s = 0.95 (0.80, 1.10) b = -57.8 (-65.0, -50.7)
NH02	s = 0.49 (-0.09, 1.07) b = -19.2 (-47.7, 9.2)	s = 0.72 (0.40, 1.05) b = -43.7 (-59.7, -27.8)	s = 0.85 (0.40, 1.31) b = -52.4 (-74.6, -30.1)	s = 1.05 (0.93, 1.16) b = -56.8 (-62.3, -51.2)
NH04	s = 0.44 (0.18, 0.70) b = -31.2 (-45.1, -17.2)	s = 0.54 (0.43, 0.66) b = -45.5 (-51.1, -39.9)	s = 0.37 (-0.09, 0.82) b = -44.3 (-66.3, -22.3)	s = 0.79 (0.51, 1.07) b = -54.0 (-67.5, -40.4)
NH05	s = 0.57 (0.49, 0.66) b = -35.8 (-40.1, -31.5)	s = 0.75 (0.57, 0.93) b = -44.4 (-53.4, -35.5)	s = 0.81 (0.60, 1.03) b = -51.6 (-62.0, -41.1)	s = 0.92 (0.54, 1.30) b = -61.5 (-80.0, -43.0)
NH06	s = 0.52 (-0.06, 1.10) b = -32.5 (-62.0, -2.97)	s = 0.86 (0.73, 0.99) b = -46.6 (-52.8, -40.4)	s = 0.60 (0.28, 0.92) b = -48.0 (-63.7, -32.3)	s = 0.49 (0.23, 0.75) b = -47.6 (-60.3, -35.0)
NH07	s = 0.33 (0.08, 0.58) b = -19.5 (-31.7, -7.3)	s = 0.98 (0.61, 1.36) b = -51.6 (-69.7, -33.4)	s = 0.42 (0.34, 0.50) b = -25.4 (-29.4, -21.4)	s = 0.83 (0.67, 0.98) b = -49.4 (-57.0, -41.8)
NH08	s = 0.43 (-0.09, 0.94) b = -25.3 (-51.2, 0.57)	s = 0.22 (-0.19, 0.63) b = -24.5 (-44.7, -4.3)	s = 0.40 (0.08, 0.72) b = -44.2 (-59.9, -28.5)	s = 0.12 (-0.41, 0.64) b = -31.8 (-57.5, -6.0)
NH09	s = 0.52 (0.34, 0.70) b = -32.8 (-42.1, -23.6)	s = 0.72 (0.57, 0.86) b = -42.0 (-49.3, -34.8)	s = 0.40 (0.17, 0.63) b = -38.6 (-49.8, -27.4)	s = 0.79 (0.56, 1.01) b = -56.7 (-67.6, -45.7)
NH10	s = 0.53 (0.24, 0.82) b = -33.3 (-48.6, -18.1)	s = 0.93 (0.51, 1.34) b = -54.1 (-74.5, -33.7)	s = 0.39 (0.10, 0.68) b = -40.6 (-54.8, -26.5)	s = 0.46 (0.08, 0.85) b = -43.7 (-62.3, -25.1)
NH11	s = 0.47 (0.19, 0.76) b = -31.3 (-45.6, -17.0)	s = 0.75 (0.62, 0.88) b = -52.1 (-58.6, -45.7)	s = 0.47 (0.22, 0.72) b = -43.4 (-55.5, -31.2)	s = 0.44 (-0.05, 0.94) b = -42.6 (-67.4, -17.9)
NH12	s = 0.67 (0.48, 0.87) b = -42.3 (-53.3, -31.4)	s = 0.62 (0.21, 1.03) b = -45.9 (-66.0, -25.9)	s = 0.53 (0.31, 0.75) b = -43.9 (-54.6, -33.2)	s = 0.29 (0.04, 0.53) b = -41.0 (-52.8, -29.3)
NH13	s = 0.63 (0.39, 0.86) b = -28.1 (-39.6, -16.6)	s = 0.97 (0.75, 1.19) b = -52.9 (-63.6, -42.2)	s = 0.40 (0.11, 0.68) b = -39.0 (-52.9, -25.1)	s = 0.65 (0.26, 1.04) b = -49.6 (-68.6, -30.7)

using other methods in experimental animals and humans. Compression rates estimated from the slope of velocity-intensity I/O functions at medium-to-high stimulation levels (40-90 dB SPL) recorded from the BM in chinchilla cochleae varied between 0.2 and 0.5 dB/dB (Ruggero et al., 1997). Psychoacoustical compression estimates between 0.15 to 0.35 dB/dB for NH listeners were reported in several studies (e.g. Fereczkowski et al., 2017a,b; Nelson et al., 2001; Oxenham and Plack, 1997; Plack et al., 2004). Compression estimates using group averaged DPOAE level-growth functions (without source separation) in NH listeners at moderate stimulus levels (50 to 70 dB SPL) were shown to be about 0.2 dB/dB (e.g. Dorn et al., 2001; Neely et al., 2003, 2009). Comparing these different results, the similarity of the compression estimates from the different methods may suggest that similar aspects of peripheral auditory processing are reflected by the respective measures.

However, despite the similarity, the shapes of the growth functions obtained

with EFRs and the growth functions obtained behaviorally differed. In the EFR level-growth functions of the NH listeners, a breakpoint was observed at 500, 1000 and 2000 Hz between the compressive growth at low stimulus levels and the levels at which the level-growth function saturated. This is consistent with results from Picton et al. (2007), who reported similar patterns when using multi-frequency stimuli even though saturation was only reported in the case of the medium carrier frequencies (1000 and 2000 Hz) but not at 500 Hz and 4000 Hz. Picton et al. (2007) argued that the saturation at high sound pressure levels may result from an interaction of the responses to the different frequency components, as no saturation was observed when measuring the EFR level-growth function using a single SAM tone at 2000 Hz. A breakpoint has also typically been observed in the psychoacoustic studies, where a two-slopes fit has been used to estimate BM I/O functions (e.g.; Nelson et al., 2001; Plack, 2013; Plack and Skeels, 2007; Plack et al., 2004). In these studies, the slope at the lower levels (up to about 30 dB SPL) has typically been set to one (assuming a linear growth of the respective BM I/O functions), and a compressive slope has been fitted to the data at higher stimulus levels. The level at which the fitted functions with the two slopes cross is then often referred to as the breakpoint. However, the breakpoint observed in the EFR studies reflects a saturation and occurs at much higher input sound pressure levels than the breakpoints reported in the behavioral studies, and therefore should not be compared. Furthermore, in contrast to the behavioral compression estimates, no sign of a linear growth at the lower stimulus intensities of the EFR level-growth functions was found in the present study in any of the listeners at any frequency.

For the HI listeners, the medians of the slopes at the non-impaired frequencies (500, 1000 and 2000 Hz) were not statistically different from the slopes obtained in the NH listeners at the same frequencies. The slopes at 4000 Hz (where the listeners showed a mild hearing loss) were higher, i.e. the function was less compressive, and statistically different from the corresponding slopes in the NH group, consistent with the reduced compression observed with other methods (Neely et al., 2009; Nelson et al., 2001; Oxenham and Plack, 1997; Ruggero and Rich, 1991).

At 4000 Hz, it was not possible to measure statistically significant EFR magnitudes at low stimulus levels. This was true also for those HI listeners where the complete EFR level-growth function could be obtained in non-impaired frequency regions (see panel D in Fig. 2.4). The lack of significant responses at

low levels at 4000 Hz is consistent with the high correlations between behavioral detection thresholds and EFR thresholds both in NH and HI individuals reported in various studies (e.g.; Dimitrijevic et al., 2002; Herdman and Stapells, 2001, 2003; Picton et al., 2005; Rance and Rickards, 2002; Van Maanen and Stapells, 2005).

2.4.2 Compression estimates based on DPOAE level-growth functions in NH listeners

The median compression estimates obtained from DPOAEs in the present study (Fig. 2.2C) varied between 0.5 and 0.74 dB/dB. These results are consistent with corresponding compression estimates reported in Neely et al. (2009) and similar to the less compressive growth reported in Popelka et al. (1993). They differ, however, from the estimates obtained in several other studies. Dorn et al. (2001) reported compression estimates of 0.15-0.35 dB/dB and Neely et al. (2003) found compression values gradually varying between 1 dB/dB at near-threshold levels to about 0.25 dB/dB at higher stimulus levels up to about 70 dB SPL. These values are more consistent with the compression rates obtained using psychoacoustics, direct recording in animals, and the median EFR slopes of the present study.

Despite the similar median value of compression in Neely et al. (2009) and the present study, the within-frequency variability of the estimated DPOAE slopes at 500, 1000 and 2000 Hz (but not at 4000 Hz) was substantially larger in Neely et al. (2009) than in the present study. Based on the large variability of the individual DPOAE level-growth functions, Neely et al. (2003) already concluded that DPOAE slopes do not provide robust estimates of compression in the individual listener. Consistent with this, Rodríguez et al. (2011) reported that cochlear compression estimates using DPOAEs were not correlated with behavioral compression estimates in the same individual listeners. One reason for the large variability in the estimated DPOAE slopes in Neely et al. (2003, 2009) could be that the two sources of DPOAE generation (Kalluri and Shera, 2001; Shera et al., 1999) were not separated. In fact, Mauermann and Kollmeier (2004) demonstrated that DPOAE source separation can reduce the variability of individual DPOAE level-growth functions, consistent with the results from the present study. This and other factors, such as the chosen primary level paradigm and the discrete-versus-sweeping presentation mode of the primaries,

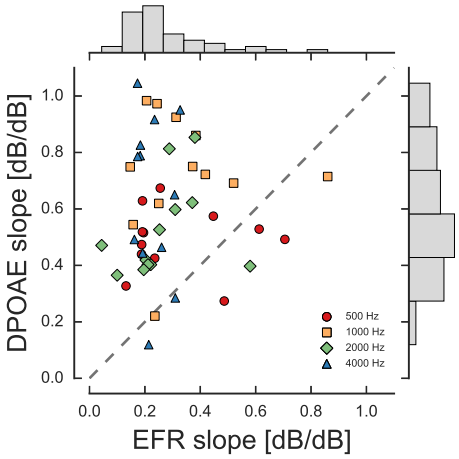


Figure 2.6: Comparison between EFR slopes and DPOAE slopes in the NH listeners for 500 Hz (red circles), 1000 Hz (orange squares), 2000 Hz (green diamonds) and 4000 Hz (blue triangles). The dashed dark-gray line show a 1:1 ratio. The marginal histograms show the common distribution of EFR (top) and DPOAE slopes (right).

can contribute to the variability in the DPOAE data (Long et al., 2008) and make a comparison across studies challenging.

2.4.3 Comparison of compression estimates based on EFR and DPOAE level-growth functions

Figure 2.6 shows a comparison of the slopes derived from the EFR level-growth functions and the DPOAE level-growth functions for the same group of NH listeners. The marginal histograms indicate that the DPOAE slopes across frequencies were more broadly distributed than the EFR slopes, which were centered around 0.2 dB/dB. The variability at a given frequency was similar for the DPOAE and the EFR slopes, except for the 4000-Hz condition where the DPOAE slopes varied more strongly than the EFR slopes (see Fig. 2.2A and C).

In general, the compression estimates obtained from the DPOAE slopes were higher (less compressive) than the compression estimates obtained from the EFR slopes. One reason for this difference could be the different stages along the auditory pathway at which DPOAEs and EFRs are generated. DPOAEs are generated at the level of the BM (Shera, 2004) whereas EFRs elicited by SAM tones modulated at rates of about 80-100 Hz are thought to be generated mainly at the level of the auditory brainstem (Herdman et al., 2002a). Assuming a compressive growth at the DPOAE generator site, the more compressively growing EFR level-growth functions may reflect a combination of cochlear- and retro-cochlear (brainstem) compression. Another possible reason could be the choice of the stimulus level paradigm which might affect the slope of the

DPOAE level-growth function. Previous studies (e.g. Mauermann and Kollmeier, 2004; Neely et al., 2003, 2009) considered L_1 to be equal to $0.4 \cdot L_2 + 39$ (in dB SPL) as suggested by Kummer et al. (1998). Boege and Janssen (2002) argued that for this level paradigm, the generation site of the DPOAE happens close to the place where the second primary (f_2) traveling wave peaks, and that the obtained DPOAE thus reflects the compressive growth of the BM at f_2 (i.e. "on-frequency"). In contrast, in the present study where equal-level primaries were used, the region of maximum overlap is at a location apical to the f_2 place (i.e. "off-frequency"), which could explain why the DPOAE level-growth functions obtained here as well as in Popelka et al. (1993) were less compressive than those found in the previous studies.

2.4.4 On- versus off-frequency contributions to compression estimates

The compressive growth of BM I/O functions measured locally in non-human animal models reflects "on-frequency" responses at a given site on the BM. At "off-frequency" places, BM I/O functions have been demonstrated to grow linearly (e.g.; Ruggero et al., 1997). Thus, in order to estimate on-frequency cochlear compression, the obtained response (with any method) needs to be dominated by on-frequency processing. At low intensities, a SAM tone excites a narrow region of the BM and the AN. Thus, the EFR response is likely to be dominated by the activity of a small population of neurons tuned to the carrier frequency of the SAM tone. Thus, the response may indeed be dominated by the on-frequency response. However, at medium and high stimulus intensities, the excitation pattern on the BM broadens and a larger population of AN (and brainstem) neurons tuned to frequencies remote from the carrier frequency contribute to the gross synchronized activity and might contribute to the measured EFR.

This is supported by direct recordings in animal models where AN fibers have been shown to follow the periodicity of high intensity tones with frequencies below the fibers' best frequency (e.g.; Anderson et al., 1951; Kiang and Moxon, 1973). Joris and Yin (1992) reported that at high stimulus levels, off-frequency AN fibers showed a higher synchrony to SAM tones than on-frequency fibers and thus represented a stronger response to modulations. Parthasarathy et al. (2016) studied the effect of a second, high-frequency SAM carrier onto the encoding of a low-frequency SAM carrier through EFRs in rats. They concluded that the presence of the second SAM tone basal to the place of the main SAM

tone stimulus led to a reduction of the EFR to the lower-frequency component due to a reduced recruitment of basal AN fibers. Off-frequency contributions to the EFR recorded at higher levels in a multi-frequency paradigm may thus underlie the saturation effect observed both in the present and previous studies for stimulus levels above 60 dB SPL (Picton et al., 2007) (see Fig. 2.1 A-C). In the case of the 4000-Hz carrier, no higher-frequency carrier that could interfere with the EFR was present, such that no saturation effect was observed. Thus, the slopes derived from EFR level-growth functions may not reflect purely place-specific peripheral compression since they might be generated due to a mixture of on- and off-frequency neural activity.

The same concern can be expressed when interpreting cochlear compression estimated using psychoacoustics or DPOAEs. Regarding DPOAEs, the distortion source of the emission is usually simplified as a single source located at the peak of the traveling wave envelope, although many distortion sources might be induced in the region where the two primaries traveling waves overlap (Shera, 2004). The extent of potential off-frequency contributions to the non-linear component of the DPOAE is unknown. The slopes derived from DPOAE level-growth functions, particularly in the higher-level regions, may be affected by off-frequency excitation. Regarding behaviorally obtained estimates of BM I/O functions, on-frequency and off-frequency maskers have been used in a forward-masking paradigm. Wojtczak and Oxenham (2009) showed that using the high-level off-frequency maskers in the temporal masking curve (TMC) and growth of masking (GOM) paradigm may lead to an overestimation of compression by as much as a factor of 2.

2.5 Summary and Conclusion

Median values of peripheral compression estimates based on the EFR level-growth functions in the NH listeners of the present study were consistent with previously reported cochlear compression rates estimated from invasively recorded BM velocity-intensity functions in non-human mammals, behaviorally derived BM I/O functions using forward masking and group average DPOAE level-growth functions. In the HI listeners, the EFR compression estimates at the frequency with reduced sensitivity (4000 Hz) showed reduced compression consistent with the hypothesis of reduced BM compression in the case of a sensorineural hearing loss. The measurements of the EFR magnitudes were fairly

repeatable which could allow for its potential use in clinical applications. The slopes of the EFR level-growth functions were smaller than those derived from DPOAE level-growth functions in the same NH listeners, suggesting that either several sources of compression at cochlear and retro-cochlear levels contribute to the EFR response, or that the level paradigm used in the DPOAE measurements might be assessing off-frequency growth functions, which could result in less compressive estimates. In any case, the mechanisms underlying the generation of the respective responses (EFR versus DPOAE) are largely different and caution must be taken regarding the interpretation of the results obtained across methods. Generally, it should be emphasized that the similarity of the median compression estimates obtained using the different methods might not imply that the same function of the system is assessed. In particular, the role of on-frequency vs off-frequency contributions to the responses in the different paradigms requires further investigation as well as the robustness of the respective methods when considering compression measures in the individual listener. In the case of the EFR measures, the compressive slopes most likely do not reflect place-specific cochlear compression but rather strong and level-dependent contributions from off-frequency neural populations. The use of computational models may provide insights into the representation of a given stimulus at cochlear and brainstem levels and the transformation between such representations and the respective (i.e. EFR versus DPOAE versus behavioral) outcome measures.

Acknowledgments

The authors wish to acknowledge the support of Glenis Long and Simon Henin from The City University of New York (CUNY) regarding the DPOAE recording and analysis. We thank James M. Harte from the Interacoustics Research Unit (IRU) for his valuable comments and discussions, as well as the help provided by Julia E. Rieck also from IRU regarding the EFR repeatability. This work was supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) at the Technical University of Denmark (DTU).

Supplementary material

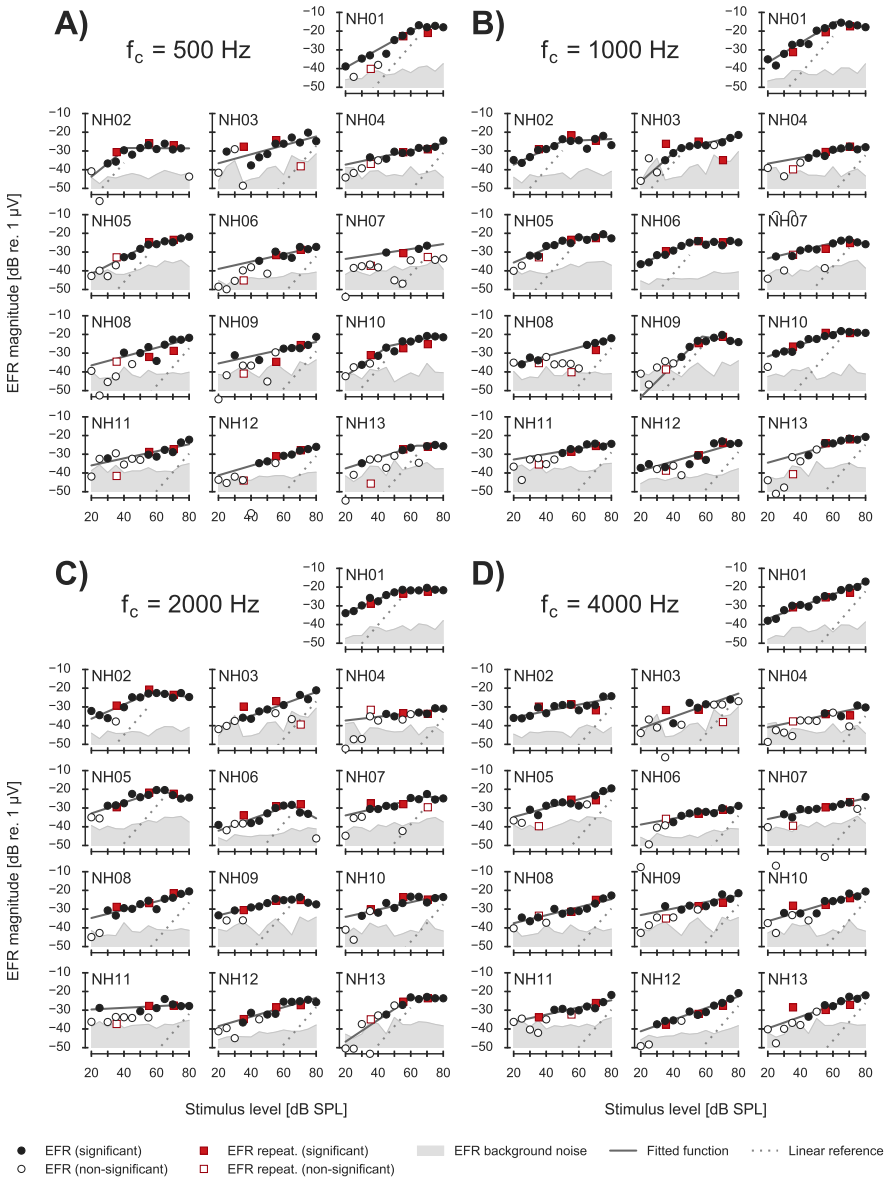


Figure 2.7: EFR level-growth function recorded in all NH listeners. EFR magnitudes are represented as filled symbols in case of a statistically significant response (positive F-test), and as open symbols in case of statistically non-significant (negative F-test) responses. Circles indicate the EFR magnitudes recorded in the first recording session and squares indicate EFR magnitudes recorded in the second recording session. EEG background noise estimates are shown as the gray shaded area. The best fitted curve is represented as a continuous dark-gray line. A linear reference with slope of 1 dB/dB is indicated by the dotted line.

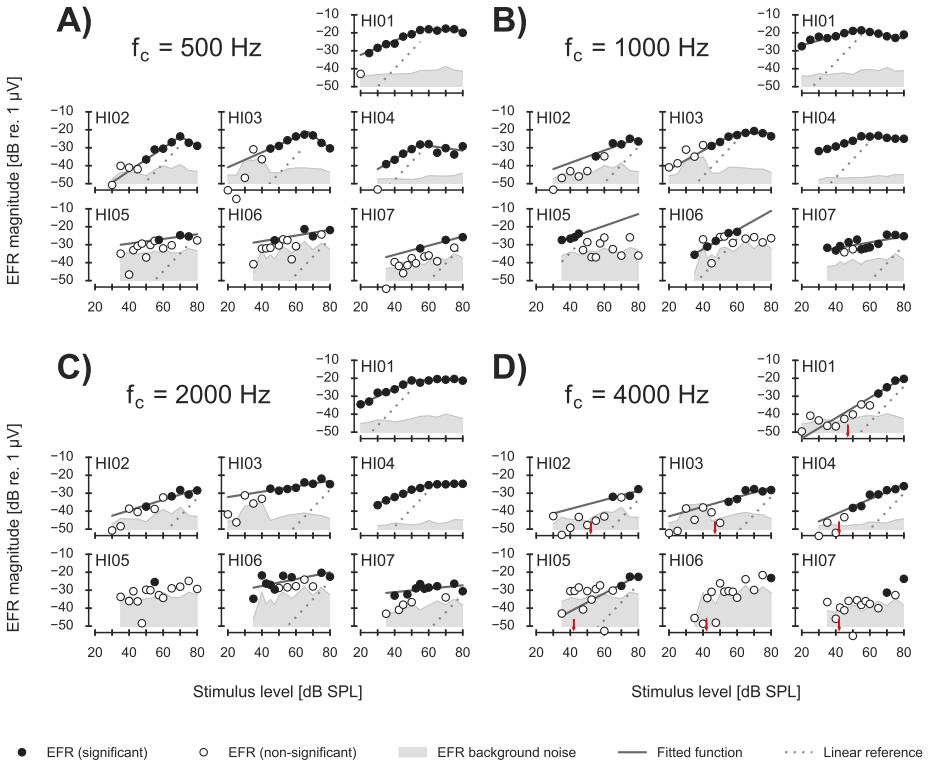


Figure 2.8: EFR level-growth function recorded in all HI listeners. EFR magnitudes are represented as circles, being solid when the response was statistically significant and open when was not statistically significant. EEG background noise estimates are shown as a light-gray shaded area. The best fitted curve is represented as a continuous dark-gray line. A linear reference with slope of 1 dB/dB is shown as a dotted line. The small red arrow in panel D indicates the behavioral hearing threshold of the listener at 4000 Hz in dB SPL.

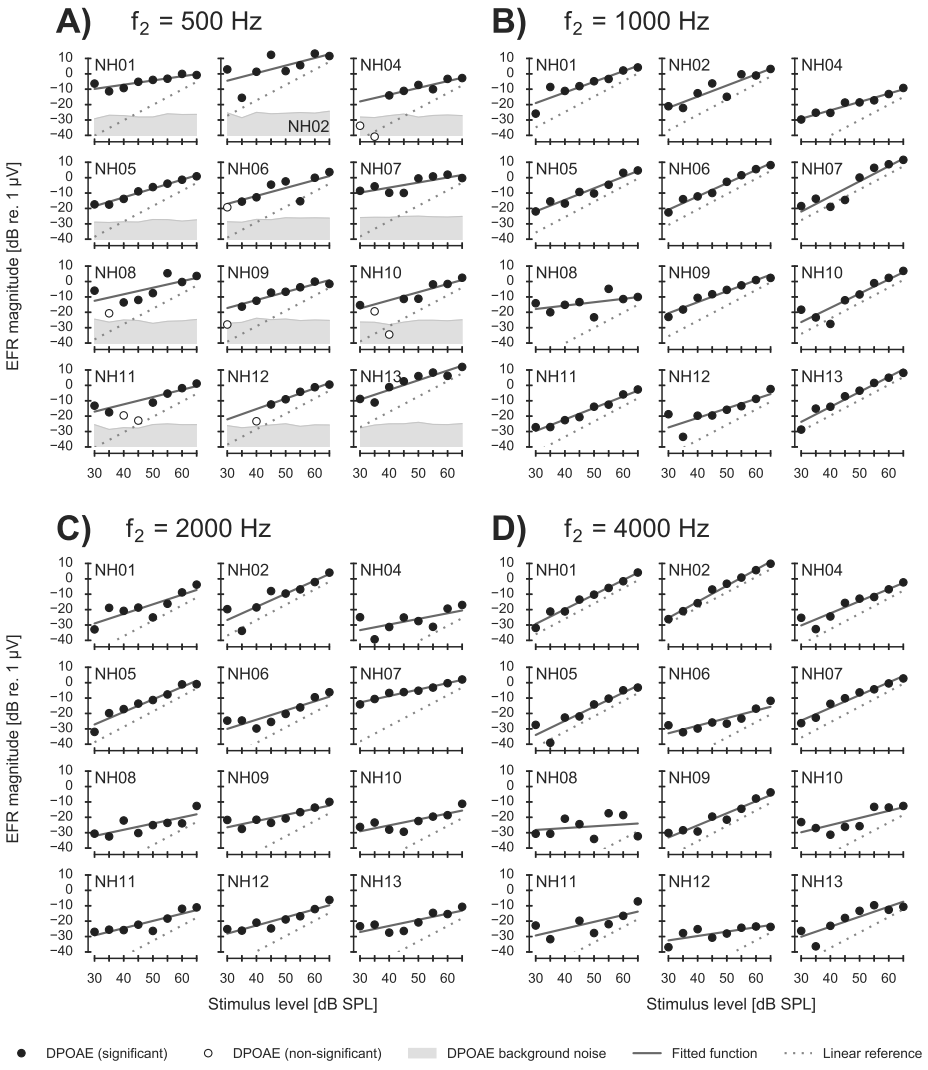


Figure 2.9: DPOAE level-growth function recorded in all NH listeners. EFR magnitudes are represented as circles, being solid when the response was statistically significant and open when was not statistically significant. EEG background noise estimates are shown as a light-gray shaded area. The best fitted curve is represented as a continuous dark-gray line. A linear reference with slope of 1 dB/dB is shown as a dotted line.

3

Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve[b]

Abstract

The healthy auditory system enables communication in challenging situations with high levels of background noise. Despite normal sensitivity to pure-tones, many listeners complain about having difficulties in such situations. Recent animal studies demonstrated that noise over-exposure that produces temporary threshold shifts can cause the loss of auditory nerve (AN) fiber synapses (i.e. cochlear synaptopathy), assumed to be selective for medium- and low-spontaneous rate (SR) fibers. In the present study, envelope following response (EFR) level-growth functions were recorded in normal-hearing (NH) threshold and mildly hearing-impaired (HI) listeners with threshold elevation at frequencies above 2 kHz. EFRs were elicited by sinusoidally amplitude modulated (SAM) tones with carrier frequency of 2 kHz, modulated at 93 Hz and modulation depths of 85% (strong) and 25% (shallow). Whereas the EFR level-growth functions for strongly modulated tones were similar for all listeners, EFR level-growth functions for shallowly modulated tones were found to be reduced at medium stimulation levels in some of the NH threshold listeners and saturated in the HI listeners for the whole level range. A phenomenological model of the AN was used to investigate the effects of off-frequency contributions (i.e. away from the characteristic place of the stimulus) and the differential loss of different AN fiber types on EFR level-growth functions. The model simulations suggested that: (1) EFR are domi-

[b] This chapter is based on Encina-Llamas et al. (under review[b]).

nated by the activity of high-SR fibers at all stimulus intensities, and (2) EFR at medium-to-high stimulus levels are dominated by the off-frequency contribution. Postulated synaptopathy led to simulations generally consistent with the recorded data, but a substantial amount of loss of all types of AN fibers had to be included within the model framework.

3.1 Introduction

It is well known that noise over-exposure can impair the auditory system by producing a sensorineural hearing loss in terms of a permanent elevation of the pure-tone thresholds. This has led to the interpretation that sound stimulation only producing a temporary threshold shift (TTS), but not a permanent threshold shift (PTS), does not permanently damage the auditory system. However, it has been reported that, despite normal sensitivity to pure tones, some listeners complain about having difficulties in challenging acoustical situations (Hind et al., 2011; Kumar et al., 2007; Saunders and Haggard, 1989; Tremblay et al., 2015).

Recent animal studies have shown that noise over-exposure producing TTS can in fact lead to the loss of auditory-nerve (AN) fiber synapses, without damaging the sensitive hair cells in the cochlea (Kujawa and Liberman, 2009). As this neuronal degeneration does not result in a PTS, it has been termed "hidden" hearing loss (Schaette and McAlpine, 2011). Kujawa and Liberman (2009) demonstrated in mice that "hidden" hearing loss, or more accurately "cochlear “synaptopathy” (for a review, see Liberman and Kujawa, 2017), resulting from carefully controlled noise exposure, did not alter hearing threshold. It was further shown that the level-growth function of distortion-product otoacoustic emissions (DPOAE) remained unaffected in the same mice. This indicated that the outer hair cells (OHC) were not damaged as a result of the noise exposure. However, the amplitude of the auditory brainstem response (ABR) wave-I was reduced at supra-threshold sound pressure levels. Wave-I is considered to reflect the action potentials of the auditory nerve, and should therefore be sensitive to a loss of AN fiber synapses. Furman et al. (2013) suggested that a selective loss of medium- and low-spontaneous rate (SR) fibers could account for the reduction of supra-threshold ABR wave-I magnitudes, while still preserving normal thresholds. However, Lobarinas et al. (2013) reported that, even in the case of a

substantial loss of inner hair cells (IHC) and AN fibers, behavioral pure-tone thresholds remained unchanged, suggesting that a major loss of high-SR fibers would not be reflected as a PTS either.

Noise-induced synaptopathy has to date been observed in several non-human mammalian species, such as mice (Furman et al., 2013; Kujawa and Liberman, 2009), guinea pigs (Lin et al., 2011; Liu et al., 2012), rats (Lobarinas et al., 2016) and rhesus macaques (Valero et al., 2017). Cochlear synaptopathy has also been reported as a natural phenomena in the normally aging (non-exposed) mouse ear (Sergeyenko et al., 2013). Noise exposure seems to accelerate such a natural degeneration of the AN (Fernandez et al., 2015). In humans, there is some evidence of such age-related synaptopathy (Makary et al., 2011; Viana et al., 2015). In addition, Viana et al. (2015) suggested that, similarly as in mice (Sergeyenko et al., 2013), the loss of peripheral axons in normal-aging humans is significantly greater than the loss of spiral ganglion cells (SGC). The argument is that SGC survive for months after the loss of their peripheral axons (Kujawa and Liberman, 2015). However, direct evidence of noise-induced synaptopathy in humans has not yet been reported, and the potential perceptual consequences have remained unknown (Plack et al., 2014), despite attempts in large studies to identify them (e.g.; Grose et al., 2017; Le Prell et al., 2017; Lopez-Poveda et al., 2017; Prendergast et al., 2016a).

Animal studies have shown evidence suggesting that synaptopathy is reflected in electroencephalographic (EEG) evoked response measurements, such as ABR (wave-I) (Furman et al., 2013; Kujawa and Liberman, 2009) or envelope following responses (EFR) (Shaheen et al., 2015). Some researchers have attempted to relate changes in evoked responses to self-reported estimates of noise exposure in humans (Prendergast et al., 2016a,b). To date, no correlation has been found. However, noise exposure scores derived from self-reported questionnaires of lifetime noise exposure rely on the subjective recall of noisy events and are based on numerous assumptions limiting their reliability (Coughlin, 1990). Other studies have found correlations between evoked responses and behavioral measures of temporal processing at supra-threshold levels in individual normal-hearing (NH) threshold listeners. Whereby the results from the poorly performing listeners were hypothesized to be linked to the loss of medium- and low-SR fibers (Bharadwaj et al., 2015; Mehraei et al., 2016). The inconclusive outcome of the human studies may reflect the challenge to directly assess the status of the AN synapses in humans, in contrast to other animals

in which invasive synaptic counts are considered ethically acceptable. Non-invasive evoked responses can be performed both in humans and non-human animals. Comparing these measures across different species could help to connect careful experimentally induced synaptopathy in non-human animals to its (potential) presence in humans. However, evoked responses measured using surface (scalp) electrodes represent the far-field sum of the activity of large populations of neurons, which might not be sensitive to specific local neuronal damage.

In the present study, EFRs were measured as a function of stimulus level (EFR level-growth functions)using strongly and shallowly modulated pure tones in listeners with NH threshold and in hearing-impaired (HI) listeners. It was hypothesized (via heuristic reasoning) that a preferential loss of medium- and low-SR fibers should reduce the EFR magnitudes at high supra-threshold stimulus levels, whereas the responses at lower levels should remain unaffected. Thus, depending on the presence or absence of medium- and low-SR fibers, the slope of the EFR level-growth functions should differ. It was further hypothesized that such a reduction or slope change should be more pronounced in the EFR responses elicited by the shallowly modulated tones than the strongly modulated tones. This was argued because high-intensity shallowly modulated stimuli are assumed to be preferentially encoded by the medium- and low-SR fibers (Bharadwaj et al., 2015; Bharadwaj et al., 2014). EFR level-growth functions at both modulation depths were also recorded in HI listeners, but the stimulus was presented only at a frequency where their audiogram was within the normal range to increase the likelihood of synaptopathy. It has been proposed that cochlear synaptopathy might be a precursor of posterior hair-cell damage (Kujawa and Liberman, 2015; Liberman and Kujawa, 2017; Sergeyenko et al., 2013). It was assumed that listeners who already show a threshold elevation (and therefore hair-cell dysfunction) at higher audiometric frequencies potentially suffer from synaptopathy at lower normal threshold audiometric frequencies.

As the history of noise exposure in both the NH threshold listeners and the HI listeners in this study is unknown, and estimates of lifetime noise exposure have failed in previous studies to predict cochlear synaptopathy in humans (e.g.; Prendergast et al., 2016a,b), the present study focused on EFR level-growth functions in individual differences and their potential relation to cochlear synaptopathy. In order to assist with the interpretation and the potential effect of synaptopathy on the obtained EFRs, a state-of-the-art computational model

of the AN was used to study the effects of a differential loss of the different AN fiber types on the EFR level-growth functions.

3.2 Methods

3.2.1 Listeners

A total of 13 adult listeners participated in this study, separated into 9 normal-hearing (6 males and 3 females, 26 ± 2.4 years) and 4 hearing-impaired (3 males and 1 female, 60.5 ± 6.7 years) listeners. All NH threshold listeners had thresholds below 15 dB hearing level (HL) at octave frequencies between 125 and 8000 Hz. All HI listeners were selected to have normal hearing (threshold ≤ 20 dB HL) below 3000 Hz and a mild hearing loss at 4000 Hz and above, with audiometric thresholds between 20 and 45 dB HL.

3.2.2 Apparatus

The EFR recordings were performed in a dark, double-walled soundproof and electrically shielded booth, where the listeners were laying on a comfortable clinical bed. The listeners watched a silent movie and were instructed to relax and avoid unnecessary movement. The recording and data analysis routines were implemented in MATLAB (The MathWorks, Inc., Natick, Massachusetts, USA). All acoustic stimuli were generated in MATLAB and presented via an RME Fireface UCX soundcard at a sampling rate of $f_{s|\text{stimulus}} = 48$ kHz using 24 bit encoding. The stimuli were presented through a pair of ER-3A insert earphones (Etymotic Research Inc.), with the contralateral ear blocked with a foam earplug.

EFRs were recorded using a Biosemi ActiveTwo system (sampling rate $f_{s|\text{EFR}} =$ 4096 Hz, 24 bits). Sixty-four active pin-type electrodes were used following the 10-20 system (American Clinical Neurophysiology Society, 2006). The results shown in this study represent the Cz-P10 potential in response to right-ear stimulation, and the Cz-P9 potential in response to left-ear stimulation (similar to vertex to ipsi- and to contra-mastoid montage respectively). Ground is achieved by a "Common Mode voltage" driven by two electrodes (DRL and CMS) placed at the center of the parieto-occipital coronal line (on either side of electrode POz). Conductive electrode gel was applied and the offset voltage was stabilized at < 20 mV for each electrode. The recorded EEG signals were downsampled by a factor of 2 which implied a hardware implemented 5^{th} order sinc response

low-pass filter with a −3 dB point at approximately 410 Hz. The EEG data were stored to hard disk for offline analysis.

3.2.3 EFR recordings and analysis

The EFR data were recorded in a single session, which lasted approximately two hours. The EFR level-growth functions were recorded in the NH threshold listeners using input levels in the range from 34 to 87 dB sound pressure level (SPL). In all NH threshold listeners, the right ear was stimulated. In the HI group, the ear which better fulfilled the selection criteria was chosen as recording ear.

A single sinusoidally amplitude modulated (SAM) tone was used as thestimulus, with $\mathrm{SAM}(t) = A \cdot \sin\left(2\pi f_c\, t\right) \cdot \left(\frac{1+m \cdot \sin\left(2\pi f_m\, t\right)}{2}\right)$, whereby A, f_c, f_m, $m \in [0,1]$ and t represents the amplitude, the carrier frequency, the modulation frequency, the modulation index and time respectively. The SAM tone had a carrier frequency (f_c) of 2005 Hz (referred as 2000 Hz throughout this work) and a modulation frequency (f_m) of 93 Hz. Two modulation depths (m) were used: "strong" ($m = 85\%$) and "shallow" ($m = 25\%$). The stimuli were calibrated using a B&K 4157 ear simulator to the desired root mean squared (RMS) level. The stimuli were digitally generated as 1-s long epochs and continuously presented to the listener in a loop, where a trigger signal marked the beginning of a new epoch for later averaging. The total stimulus duration depended on the stimulus intensity to achieve a statistically significant EFR signal-to-noise ratio (SNR). Table 3.1 shows the stimulus duration used for each input level in the EFR recordings. The recorded EEG data were analyzed as described in *chapter 2*. In short, the recorded EEG epochs were 1) band-pass filtered between 60 and 400 Hz, 2) rejected if the absolute value of their voltage amplitude exceeded ± 80µV, 3) weighted averaged to increase the EFR SNR as in John et al., 2001, and 4) concatenated in epochs of 16-s duration to achieve a higher frequency resolution in the EFR spectrum. The recorded EFR magnitude was detected from an estimate of the background EEG noise in the range of 3 Hz below and above the modulation frequency (96 bins) if $p \leq 0.01$ was achieved in an F-test statistical measure (Dobie and Wilson, 1996; Picton et al., 2003).

Table 3.1: Duration of EFR stimuli for each used input level.

Input level [dB SPL]	34	40	46	54	60	66	71	77	81	87
Duration [min]	10.0	8.5	8.5	7.0	7.0	6.5	5.5	5.5	5.5	5.5

The statistical analysis was performed in R 3.2.2 (R Core Team 2015) using a linear mixed-effects model. The model was fitted using the "lme4" R-package, v1.1.13 (Bates et al., 2015) and the p-values were calculated using the Satterthwaite approximation of the "lmerTest" R-package, v2.0.33 (Kuznetsova et al., 2015). The model analyses were conducted with three fixed effects variables: the level of stimulation as continuous independent variable, and the modulation depth and hearing status as categorical independent variables. The listeners were treated as random effects.

3.2.4 AN model

A humanized phenomenological AN model, implemented in MATLAB, was used to simulate the activity of the AN (Zilany et al., 2009, 2014). The model fibers were tuned to 200 characteristic frequencies (CF) ranging from 0.2 to 20 kHz; corresponding to equally-spaced positions in the basilar membrane (BM) according to the cochlear frequency map for humans (Greenwood, 1990). A non-uniform distribution of AN fibers per CF (or IHC) was implemented according to the distribution reported in Spoendlin and Schrott (1989), with a total number of AN fibers of about 32000 for the healthy auditory system. About 190 AN fibers synapses were independently computed at each CF. In the framework of the model, synaptopathy was simulated by computing a lower number of AN fiber synapses at each CF. Frequency-specific synaptic loss was implemented by fixing a given percentage of loss of fibers at single CFs, which were interpolated using a shape-preserving piecewise cubic Hermite interpolating polynomial evaluated over the complete range of modeled CFs. Hair-cell impairment was implemented by fitting the listener's audiogram using the *fitaudiogram* MATLAB function implemented by Zilany et al. (2009). As the distribution of the different AN fiber types at each CF is unknown in humans, the distribution reported from cats was used: 61% of high-SR fibers, 23% of medium-SR fibers and 16% of low-SR fibers (Liberman, 1978). Model simulations were performed using the same stimuli as in the human EFR recordings but with a duration of 1.2-s. Stimulus levels ranged from 10 to 100 dB SPL, in steps of 5 dB.

The model allows for control of the IHC and outer hair-cell (OHC) function independently, and provides the deterministic IHC voltage and synaptic output of each AN fiber type separately. The same IHC voltage at each CF was used to drive the stochastic synapse- and spike generator models (see Zilany et al., 2009), which was executed as many times as corresponding to the number of

AN fibers per IHC. The resulting synaptic outputs for each AN fiber type were summed to obtain the population response of this fiber type at each CF, which is comparable to the peri-stimulus time histogram (PSTH) used for experimental data. In order to analyze the steady-state encoding of the modulation, the 1-s long steady-state response, excluding on- and offsets were analyzed. A Fast Fourier Transform (FFT) was performed on the resulting synaptic output and the magnitude value at the modulation frequency bin was considered the simulated EFR.

The model's synaptic response was analyzed in populations corresponding to 1/3-octave frequency bands (CF bands) to investigate the contribution of each population to the total simulated AN EFR. The on-frequency (at or near the CF of the stimulus) simulated synaptic response was computed by summing the PSTH responses of all the CF within the frequency band centered at 2 kHz. Similarly, contributions from the off-frequency bands centered at 3 and 7 kHz were calculated. Figure 3.1 shows an example of the simulated synaptic output. Panel A shows the response of the simulated AN at three cycles of the modulation frequency representing the sum of the three AN fiber types using a strongly modulated SAM tone at 80 dB SPL stimulus level (see the animation in the *digital version* for levels from 5 to 100 dB SPL in steps of 5 dB). Panels B-E show the simulated synaptic output at the output of the 1/3-octave band centered at 2 kHz (on-frequency, D), at the output of the 3-kHz (C) and the 7-kHz band (B), as well as summed across the entire frequency range (E). The summed synaptic output (E) was used to compute the simulated EFR to be compared to the recorded EFR.

Previous studies have attempted to simulate steady-state responses, such as EFRs (Rønne, 2013) or frequency following responses (FFRs) (Dau, 2003), by convolving the simulated response of an AN model with a unitary response (e.g.; Melcher and Kiang, 1996) that reflected the contributions of different neural population along the auditory brainstem to the far-field evoked potential. In the present study, as synaptopathy occurs at the level of the AN, and for simplicity, only AN activity was considered (Zilany et al., 2009, 2014). It was then assumed that the codification of the envelope at the level of the AN would be similar to the recorded EFRs. It has been suggested though that EFRs to 80-100 Hz modulations are mainly generated at the level of the brainstem (Herdman et al., 2002a). However, Parthasarathy et al. (2016) showed good consistency between EFR recordings in rats and simulated EFRs using the cat version of the AN model

of Zilany et al. (2009, 2014).

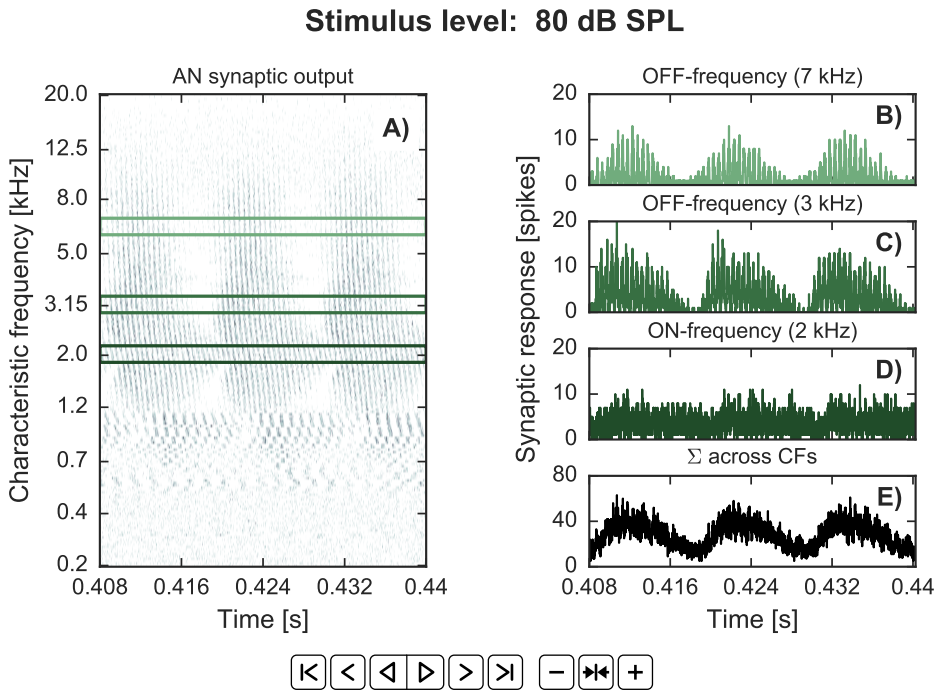


Figure 3.1: Simulated synaptic output of the AN model obtained using a SAM tone with $f_c = 2$ kHz, $f_m = 93$ Hz and $m = 0.85$. A) Simulated AN synaptic output at CFs from 0.2 to 20 kHz for three cycles of the f_m at the steady-state part of the response. The green rectangles illustrate the on-frequency (2 kHz) and the off-frequency bands (3 and 7 kHz). B-D) Synaptic outputs at three different $^1/_3$-octave bands. E) Synaptic output after summing across CFs. B-C) Off-frequency response at the bands centered at 7 and 3 kHz respectively. D) On-frequency response at the 2 kHz band. *Paper version*: Stimulus level presented at 80 dB SPL. *Digital version*: Stimulus levels from 5 to 100 dB SPL, in steps of 5 dB (*use the control buttons to navigate through the animation*).

3.3 Results

The data reported in this chapter is publicly available here (Encina-Llamas et al., 2017b).

3.3.1 EFR level-growth functions in human listeners

Figure 3.2 shows the complete set of EFR level-growth functions for the NH threshold (A) and the HI (B) listeners. The recorded EFR magnitudes are shown as circles, represented in dB relative to 1μV (blue for $m = 85\%$, red for $m = 25\%$). Filled symbols represent statistically significant EFRs magnitudes (positive F-test), open symbols represent non-significant (negative F-test) responses. The

estimated EEG background noises for each modulation depth are depicted as thin lines with consistent color labeling.

For both listener groups, EFR magnitudes obtained with the strongly modulated stimuli (blue) were larger than those obtained with the shallowly modulated tones (red). However, different trends were observed in the EFR level-growth functions across listeners, particularly for the shallowly modulated tones. In the case of the NH threshold listeners (Fig. 3.2, A), the results have been organized gradually from patterns showing monotonic and parallel EFR level-growth functions (i.e. listeners NH01 or NH02) to patterns showing non-monotonic level-growth functions (i.e. listeners from NH07 to NH09). In particular, for listener NH09, the EFR magnitudes for the strongly modulated tones grew monotonically with a single linear slope throughout the whole level range. This subject was considered as a potentially synaptopathic listener within the NH threshold group. In contrast, the responses to the shallowly modulated tones initially grew with a single slope up to 55 dB SPL, showed a decrease of the EFR magnitudes from 55 to 70 dB SPL and a recovery above 70 dB SPL, with comparable EFR magnitudes between 80 to 90 dB SPL as for the strongly modulated tones.

For the HI listeners (Fig. 3.2, B), the EFR level-growth functions for the strongly modulated tones grew monotonically with a similar slope as for the NH threshold listeners. The EFR level-growth function for the shallowly modulated tones showed, however, a strongly compressive or even saturating growth. Figure 3.3 shows boxplots of the fitted slopes for both modulation depths in the NH threshold and the HI listeners. A post-hoc statistical analysis using a mixed linear model revealed that the estimated slopes in the NH threshold listeners were not statistically different for the two modulation depths ($t_{148} = 0.723$, $p = 0.4704$). However, the EFR level-growth function for the shallowly modulated tones in the HI listeners tones were significantly lower than the slopes for the strongly modulated tones ($t_{48} = 3.646$, $p = 0.00042$).

3.3.2 Simulating EFR level-growth functions in human listeners with and without hair-cell dysfunction

Figure 3.4 shows the simulated EFR level-growth functions for A) a NH listener, B) a HI listener accounting for the threshold elevation with dysfunctional OHCs only, and C) a HI listener accounting for the impairment with a combination

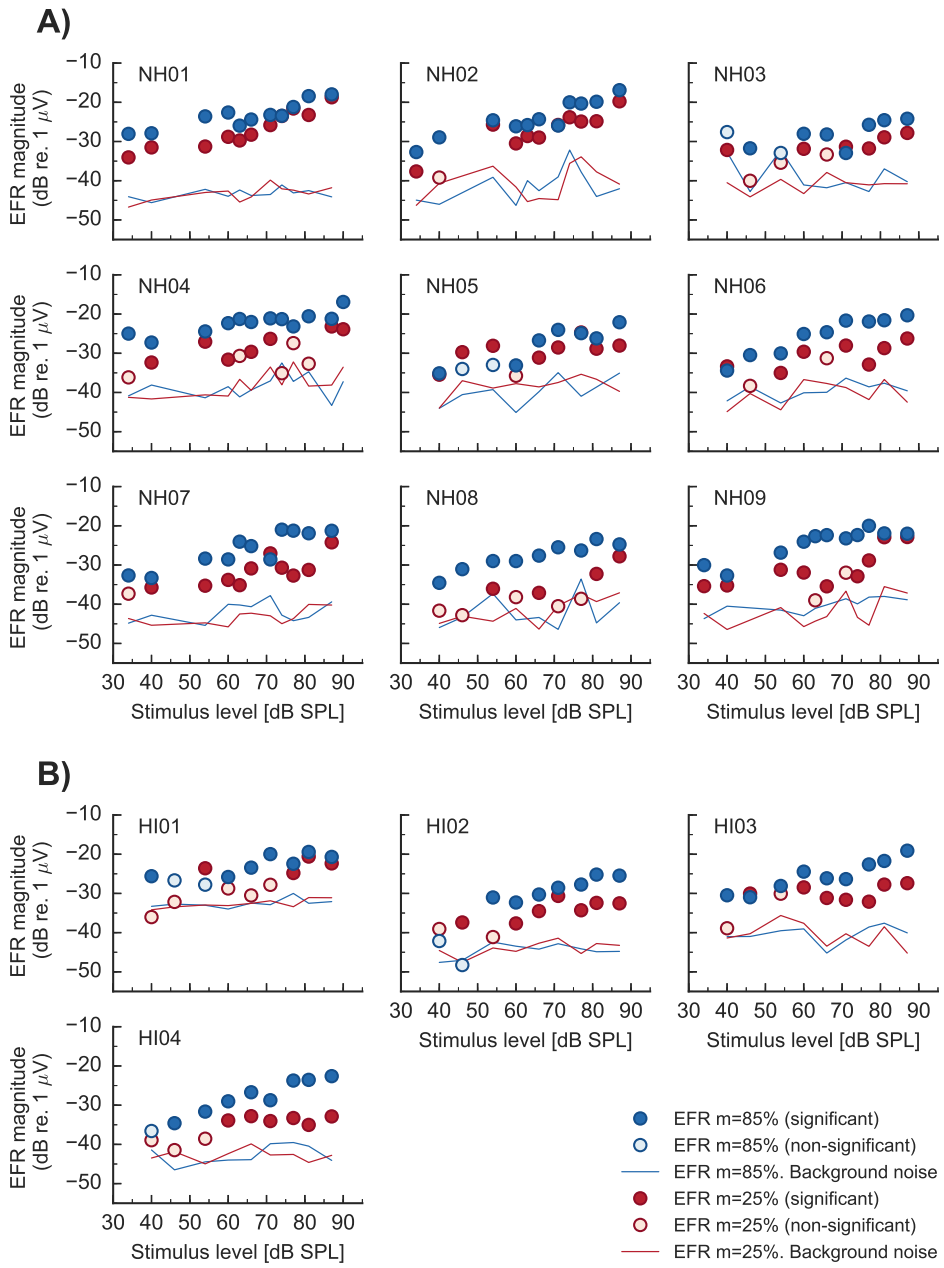


Figure 3.2: EFR level-growth function recorded in A) NH threshold and B) HI listeners using strongly (blue) and shallowly (red) modulated tones. EFR magnitudes in dB relative to 1μV are represented as filled circles in case of a statistically significant response (positive F-test), and as open circles in case of statistically non-significant (negative F-test) responses. EEG background noises estimates for each modulation depth are shown as thin lines with consistent color labeling.

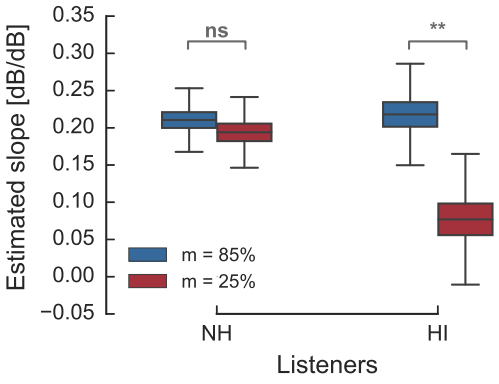


Figure 3.3: Estimated slopes of the EFR level-growth functions recorded in the NH threshold and HI listeners using strongly (blue) and shallowly (red) modulated tones. Tukey boxplots, where the bottom and top of the box are the first and third quartiles respectively, and the band inside the box is the second quartile (the median). Whiskers show 1.5 of the interquartile range (IQR) of the lower and upper quartile (*** corresponds to a $p\text{-value} \leq 0.001$, ns corresponds to a $p\text{-value} > 0.05$).

of one third of IHCs and two thirds of OHCs, as suggested in the literature (Lopez-Poveda and Johannesen, 2012; Spongr et al., 1997). The representation is similar to Figure 3.2, but with the simulated EFR magnitudes expressed in arbitrary units (a.u.) in dB. The circles represent the simulated EFRs for a NH listener (reprinted in all panels for comparison), while the squares represent the simulated EFR for a HI listeners. The blue and red symbols represent EFR magnitudes obtained with the strongly and the shallowly modulated tones, respectively. The audiogram from the listener HI04, who was selected as a HI representative listeners based on the gradual organization, was considered to adjust the parameters of the IHC and OHC processes in the model.

The simulated EFR level-growth functions for the NH listener (panel A) showed a parallel and monotonic growth over the input level range used in the EFR recordings (35-90 dB SPL, un-shaded area). The EFR magnitudes for the strongly modulated tones were larger than for the shallowly modulated tones. In general, the model simulations were able to capture the trend observed in the recorded EFR level-growth functions in some NH threshold listeners (i.e. NH01 and NH02, see Fig. 3.2).

In the simulated EFR level-growth functions for the HI listener when accounting for the impairment with a pure OHC dysfunction (panel B), no difference with the simulated results for the NH listener was found. If at least one third of the threshold elevation was assigned to IHC dysfunction and the remaining two thirds to OHC dysfunction (panel C), the only difference to the NH simulation was a small reduction at input level between 60 and 85 dB SPL, amounting to about 3 dB both for the strongly- and shallowly modulated tones. A similar result was found when assuming a hearing loss due to IHCs dysfunction only

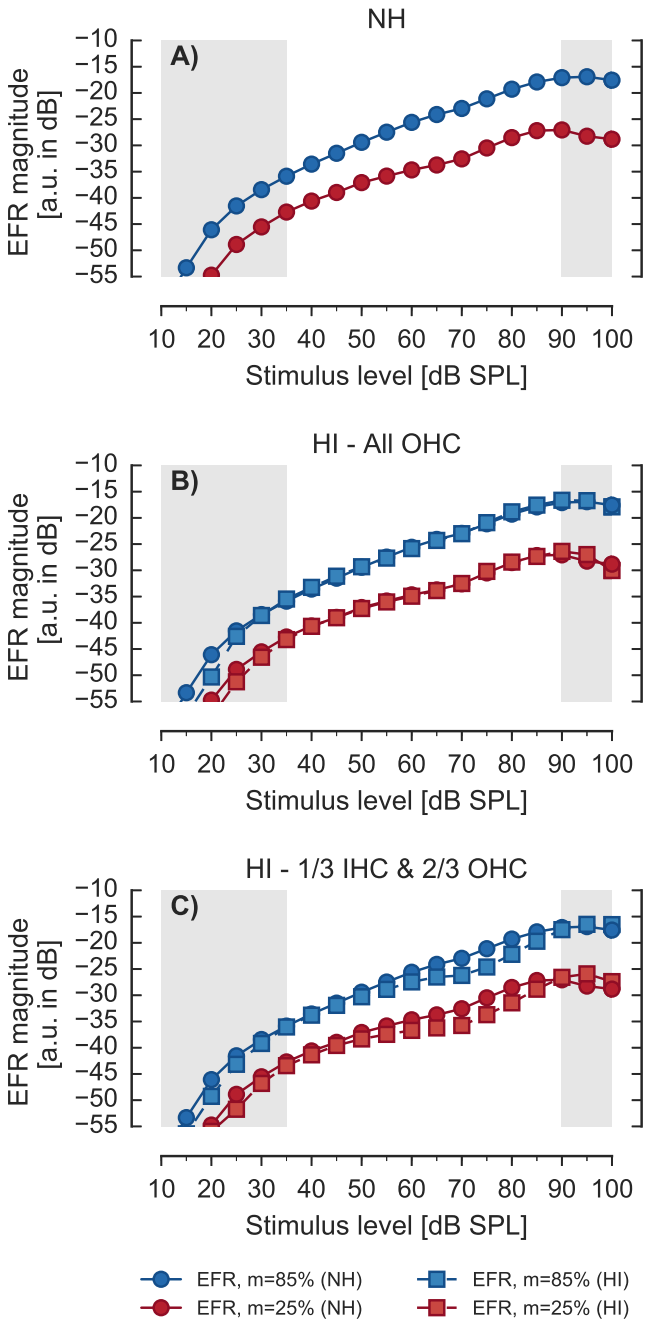


Figure 3.4: Simulated AN EFR level-growth functions in NH and HI listeners. Squares indicate the simulated EFRs with different combinations of hair-cell dysfunction, whereas the circles show the EFR level-growth function in a simulated NH listener as reference. A) Simulation for a NH listener. B) Simulation for a HI listener accounting for the threshold elevation with a pure OHC dysfunction. C) Simulation for a HI listener accounting for the threshold elevation with a combination of one third of IHC and two thirds of OHC dysfunction.

(not shown). This is at odds with the recorded EFR level-growth functions in HI listeners where both a loss of sensitivity and a saturation were found in the response to shallowly modulated tones (see Fig. 3.2).

3.3.3 Simulating EFRs in NH threshold listeners and HI listeners with postulated synaptopathy

Figure 3.5 shows the simulated EFR for A) a NH threshold listener after assuming a complete loss of medium- and low-SR fibers at all CFs (an AN with only high-SR fibers); for B) a NH threshold listener including an empirically chosen synaptic loss that best approximated the results obtained from the NH threshold listener NH09; and for C) a HI listener including an empirically chosen loss of synapses that best approximated the results obtained from the HI listener HI04. The representation is the same as in Figure 3.4, but the squares represent the simulated EFRs including a postulated synaptic loss.

The simulated EFR level-growth functions with a loss of 100% of medium- and low-SR fibers (panel A, squares) were nearly the same as the EFR level-growth functions in the reference simulation (circles), with a small decrement of less than 1.5 dB for both modulation depths. The simulation that best approximated the non-monotonic growth found for some NH threshold listeners (i.e. NH09) required a frequency-specific loss of all types of AN fibers (panel B) and, more specifically, a loss of up to 85% in the range from 3000 to 4000 Hz. To be able to simulate EFR level-growth functions that are similar to those of the listener HI04 (panel C), a substantial loss of AN fiber synapses was required to be included in addition to the hair-cell dysfunction. A total loss of all three types of AN fiber synapses above 6 kHz was considered as well as significant losses of synapses (of $> 60\%$) at frequencies above 600 Hz. While the simulated EFRs in the NH threshold and the HI cases (panels B and C respectively) fitted the trend in the recorded data (see Fig. 3.2), the impact was the same for the strongly and shallowly modulated tones. In contrast to the model simulations, the recorded EFRs showed a stronger reduction for the shallowly modulated tones compared to the strongly modulated tones.

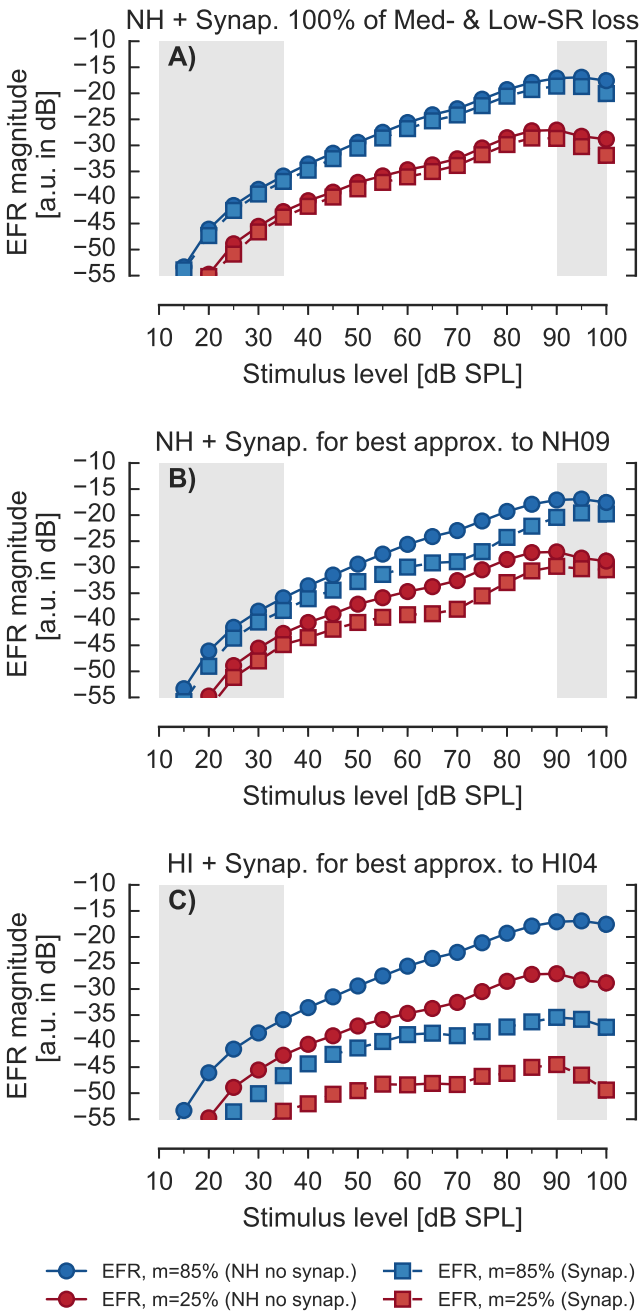


Figure 3.5: Simulated AN EFR level-growth functions in NH threshold and HI listeners with additional postulated synaptopathy. Squares indicate the simulated EFRs with synaptopathy, whereas the circles show the NH listener simulated response as reference (same as Fig. 3.4, A)). A) Simulation with a loss of 100% of medium- and low-SR fibers. B) Simulation with a frequency-specific loss of all types of fiber synapses to better approximate the response obtained in NH09 in Figure 3.2, A. C) Simulation with a frequency-specific loss of all types of fiber synapses to better approximate the response obtained in HI04 in Figure 3.2, B.

Table 3.2: Percentage of all three types of additional AN fiber loss at different CFs implemented in the model in the NH threshold and the HI listeners simulations.

Frequency [kHz]	0.1	0.3	0.6	1	2	3	4	6	10	13	16	19
Loss of AN fibers [%] in NH simulations	0	0	0	0	20	65	85	20	0	0	0	0
Loss of AN fibers [%] in HI simulations	0	0	60	70	70	80	95	100	100	100	100	100

3.4 Discussion

3.4.1 EFR level-growth functions from strongly and shallowly modulated SAM tones

It was hypothesized, via an heuristic argument, that synaptopathy produces differences in the EFR level-growth functions within a homogeneous group of young NH threshold listeners. It was further hypothesized that synaptopathy leads to non-monotonic or saturating EFR level-growth functions for the shallowly modulated SAM tones. This hypothesis was based on the assumption that high intensity sounds are encoded by the activity of medium- and low-SR fibers because high-SR are saturated at those high stimulation levels (Liberman, 1978; Yates, 1990), similarly proposed by Bharadwaj et al. (2015) and Bharadwaj et al. (2014). Indeed, the individual results of the NH threshold listeners demonstrated different EFR level-growth functions for shallowly modulated tones (Fig. 3.2, A); and more similar functions for strongly modulated tones. For instance, the EFR level-growth function for the shallowly modulated tones grew monotonically with a single slope for listener NH01, whereas in NH09 the EFRs grew non-monotonically, reducing its magnitude at medium stimulus levels but recovering at higher levels. The EFR magnitude reduction at 65-70 dB SPL was about 10 dB. This difference might reflect a truly physiological difference between those two subjects because it is probably larger than the intrinsic EFR variability. EFR test-retest variability results at 70 dB SPL from SAM tones at f_c = 2000 Hz, f_m = 93 Hz and m = 85% reported in *chapter 2* were no larger than 5 dB, much smaller than the difference seen across NH threshold listeners in Figure 3.2. It is assumed though that the difference in modulation depth in both studies would not affect the test-retest variability, although there is no available data to our knowledge showing test-retest variability in EFRs from SAM tones of different modulation depths. Whether the different patterns observed in the

EFR level-growth functions in listener NH01 versus NH09 are due to a loss of AN fiber synapses is unknown.

Individual differences in the EFR level-growth functions for shallowly modulated tones were also observed for the HI listeners (Fig. 3.2, B). Whereas the EFR functions for strongly modulated tones grew monotonically with a single slope (similarly to the strongly modulated EFRs in the NH threshold listeners), the EFR magnitudes for shallowly modulated tones did not vary much across stimulus level, leading to a saturated growth function. The change in slope was shown to be statistically significant (linear mixed-effects model, Fig. 3.3). The question whether the different patterns observed from the EFR level-growth functions in both NH threshold and HI listeners could be due to the loss of AN fiber synapses was investigated in the framework of a computational model.

3.4.2 A model of the auditory nerve to investigate individual differences in EFR level-growth functions

Under the assumption that the EFR is sufficiently described by the summation of the instantaneous firing rate of AN fibers across frequency and fiber type, the simulations might help to shed some light on the contributions of CF bands and fiber type to the overall response. Figure 3.6 shows the simulated EFR level-growth functions for a NH listener, separately for each fiber type (rows) and CF bands (columns). Column A (panels A1-A4) shows the synaptic output summed across all CF bands. Column B (panels B1-B4) shows the output of the band centered at 2 kHz (on-frequency band). Column C (panels C1-C4) shows the output of the band centered at 3 kHz, and column D (panels D1-D4) the output of the band centered at 7 kHz.

The analysis of the simulated AN activity in different CF bands (columns) showed that the EFRs at medium-to-high stimulus levels were not purely due to activity in the on-frequency band (column B), but had strong contributions of AN neural populations located more basally (i.e. at higher CF bands, columns C and D). Within each frequency band, the response showed a bell-shaped curve, horizontally shifted along the stimulus level axis for more distal CF bands, consistent with the synchrony-level functions recorded from single neurons in the AN of the cat reported in Joris and Yin (1992). Hence, at higher stimulation levels, the off-frequency contributions dominated the total EFR magnitudes in the model framework, leading to the overall monotonic growth observed when

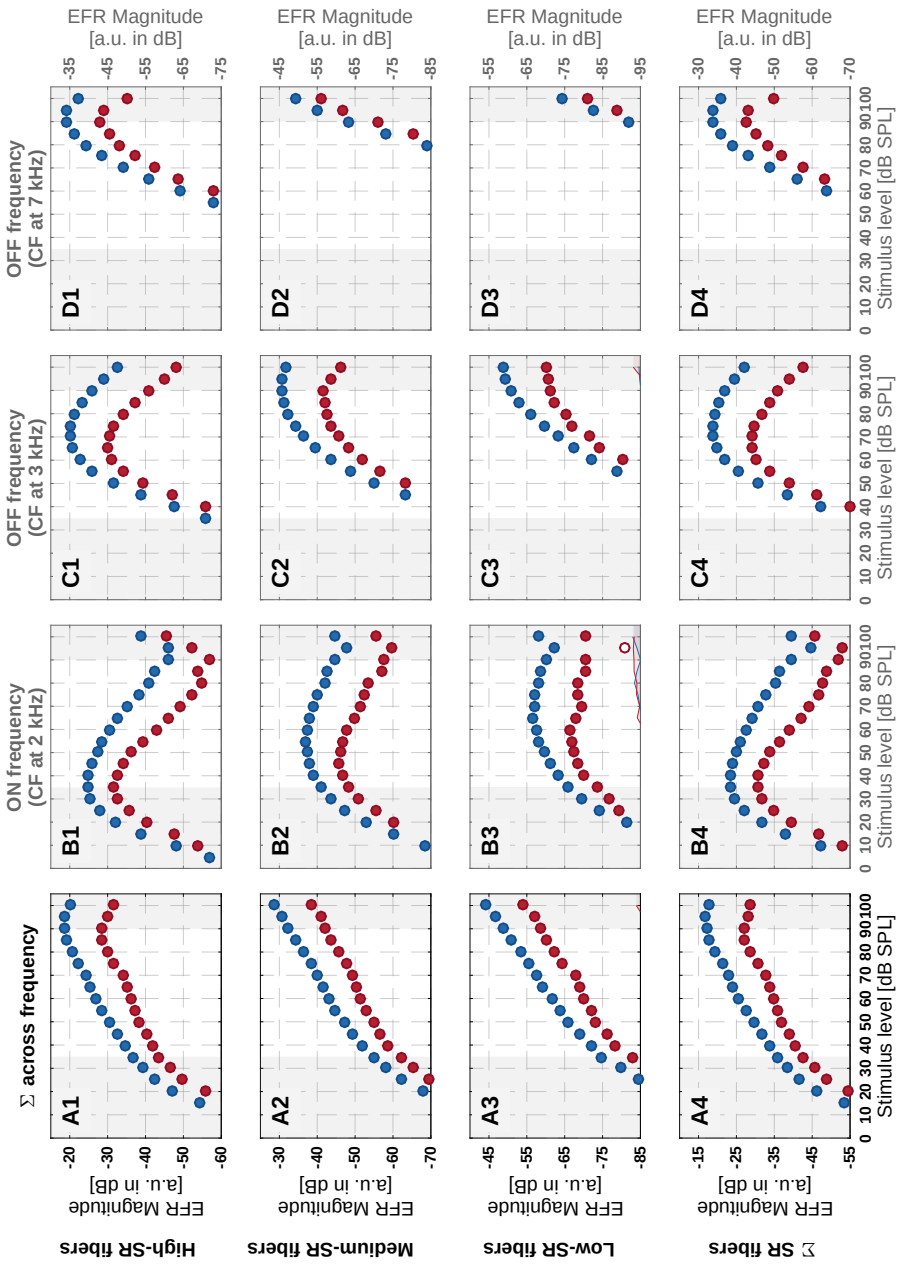


Figure 3.6: Simulated AN EFR level-growth functions in a NH listener separately for different CF bands and for each AN fiber type. Solid circles represent statistically significant EFR magnitudes in dB and open circles represent non-significant responses. Blue markers show responses for strongly modulated stimuli and red markers for shallowly modulated stimuli. The thin lines represent the background noise. The columns show the EFR level-growth functions centered at different CF bands and the rows show the simulated results at the different AN SR fiber types.

summing across CFs.

Similarly, when analyzing the contributions of the different types of AN fibers (rows), the simulated EFRs were dominated by the high-SR fibers in the whole stimulus level range. For simplicity, focusing on the on-frequency band (column B), even though the medium-SR (B2) and low-SR (B3) fibers level-growth functions did not reduce as much as the high-SR (B1) fibers with increasing stimulation level, their overall firing rates to the modulation frequency of the SAM tone were lower than those of the high-SR fibers (notice the different vertical axis scales). Thus, the medium- and low-SR fibers contribute very little to the summed response (B4). The underlying reason for this is mainly due to the implemented uneven distribution of fiber types with the high-SR fibers dominating in number (Liberman, 1978). In conclusion, the model simulations suggest that, for low stimulus intensities, the envelope was encoded by the high-SR fibers tuned to on-frequency CFs. As stimulus level increases, the modulations were predominantly encoded by the high-SR fibers at off-frequency bands (see also Fig. 3.1).

The above model observations explain why the simulated EFR level-growth functions do not change significantly when a pure OHC dysfunction was imposed (Fig. 3.4, B). A loss of gain (OHC dysfunction) might only affect the response when presenting a stimulus with spectral content at CFs with OHC damage (on-frequency processing). The AN fiber tuning curves associated to CFs with OHC dysfunction will each show a reduction of sensitivity at the tip of the tuning curve (broader frequency selectivity). However, the tails (off-frequency) will be largely unaffected (Liberman and Dodds, 1984). Thus, the off-frequency excitation of high CF AN fibers through stimuli at a lower CF with normal sensitivity will be the same regardless of OHC loss. In contrast, when at least one third of the hearing threshold elevation was assigned to dysfunction of the IHCs (Fig. 3.4, C), a small reduction in the simulated EFR level-growth function was found at stimulus levels corresponding to CFs bands at which the off-frequency contributions were maximal. Nevertheless, no combination of hair-cell impairment led to simulated EFR level-growth functions similar to the recorded ones in the HI listeners.

The model framework was used to explore how synaptic loss might be used to explain the non-monotonic patters observed in some NH threshold listeners (Fig. 3.2, A) and the strong saturation in the HI listeners (Fig. 3.2, B). It is not intended here to claim that a given listener suffers from synaptopathy. The

purpose was to investigate the potential effects of synaptopathy on the EFR level-growth functions through the modeling framework. The simulated EFR level-growth functions were insensitive to a selective loss of medium- and low-SR fibers (Fig. 3.5, A), as a reduction of less than 1.5 dB was obtained in the simulated EFRs after including a complete loss of the medium- and low-SR fibers at all CFs. However, it was possible to account for the non-monotonic growth found in some NH threshold listeners (i.e. NH09) by reducing the number of all types of AN fibers (including high-SR fibers loss), mainly centered in the CF band in which the response peaks at medium levels (Fig. 3.6, column D). To obtain simulated EFRs that better approximated the results from the HI listeners, a large loss of all three types of AN fiber synapses had to be included in a broad CF range. This was in addition to the hair-cell dysfunction (see Table 5.1). These observations are at odds with the interpretation of the data in Furman et al. (2013) that only medium- and low-SR fibers are affected by synaptopathy. However, the model simulations are consistent with previous findings (Paul et al., 2016), where a certain degree of high-SR fiber loss had to be included to similar model simulations to account for the differences observed in the EFR magnitudes recorded in NH threshold listeners with and without tinnitus, which has also been related to synaptopathy (Schaette and McAlpine, 2011). Furthermore, a reanalysis of the data from Furman et al. (2013) concluded that in fact there was a loss of high-SR fibers at a ratio of about 1 : 3 with loss of low- and medium-SR fibers (Marmel et al., 2015).

The model was, however, not able to capture all the details of the EFR level-growth functions observed in the recorded data. The implementation of either hair-cell dysfunction and synaptopathy within the model framework affected similarly the predicted EFR level-growth functions for both strongly and shallowly modulated tones (see Figs. 3.4 and 3.5). In contrast, the recorded EFR level-growth functions for strongly modulated tones were similar to all listeners, whereas the individual differences were observed only on the EFRs for shallowly modulated tones (see Fig. 3.2). However, consistent with the AN model simulations, Shaheen et al. (2015) reported a significant reduction in the EFR level-growth functions of synaptopathic mice when using strongly modulated SAM tones with modulation frequencies between 800 to 1000 Hz. The EFRs in mice showed group delays consistent with generators between the AN and the cochlear nucleus. In the present study, a modulation frequency of 93 Hz was used to elicit the EFRs in humans, that are assumed to be mainly dominated

by brainstem sources (Herdman et al., 2002a; Kuwada et al., 2002). Brainstem processing, such as central gain mechanisms (Chambers et al., 2016; Möhrle et al., 2016), affecting differently strongly and shallowly modulated stimuli, may explain the inconsistency of the human data versus the animal data and the model simulations.

Nevertheless, the model simulation suggested two main conclusions: (1) the EFR level-growth functions at medium-to-high stimulation levels are strongly dominated by the contributions from off-frequency neuronal activity, and (2) there must be a significant degree of loss of high-SR fibers for synaptopathy to be reflected in the EFR. First, the interpretation of the role of the medium- and low-SR fibers on encoding temporal fluctuations at high stimulus levels, based on the rate-level curves of the different AN fibers types (Liberman, 1978; Yates, 1990), has led to different hypotheses to study synaptopathy in humans (including the present study and others like i.e. Bharadwaj et al. (2015), Bharadwaj et al. (2014), and Mehraei et al. (2016)). The rate-level functions are derived from direct recordings in single AN neurons, and therefore provide information regarding AN neuronal activity at on-frequency stimulation. However, electrophysiological evoked responses rely on large neural population to encode sounds. As synaptopathy affects supra-threshold processing, high sound stimulation levels that produce a broad excitation of the AN are commonly used, and thus the contribution of AN neurons tuned to off-frequency CFs should be carefully considered in the design of future hypotheses. In fact, Paul et al. (2016) suggested the use of background noise to minimize the effect of off-frequency contributions. Second, the results from Furman et al. (2013) suggesting that synaptopathy was selective to medium- and low-SR fibers have been interpreted as if noise exposure did not damage high-SR fibers at all, although a reanalysis of this data by Marmel et al. (2015) suggested otherwise. Under the framework of the model, this implies that synaptopathy should produce no change in the EFRs. However, EFRs were shown to be reduced in synaptopathic mice relative to the unexposed animals (Shaheen et al., 2015). Therefore, either the effect of synaptopathy on the EFRs was not well captured in the model simulations, or the lack of vulnerability of the high-SR fibers to synaptopathy reported in mice may not be directly transferable to other species, such as humans. An alternative explanation could be that high- versus medium- and low-SR fibers were found to be more evenly distributed in mice (Taberner and Liberman, 2005) than in cats (Liberman, 1978), which may result in a stronger impact of the loss

of medium- and low-SR fiber in the EFR obtained in mice. In order to study the potential consequences of synaptopathy in humans, auditory evoked potential studies in parallel in humans and in non-human species where synaptopathy has been characterized, together with the use of species-specific computational models are desirable.

3.5 Conclusions

EFR level-growth functions recorded from a homogeneous group of young NH threshold listeners demonstrated individual differences for strongly and shallowly modulated tones, indicating differences in neural supra-threshold encoding of envelope modulations. Similar differences for mild HI listeners measured at an audiometrically normal center frequency supported the idea of coexisting hearing loss and supra-threshold deficits at frequencies of normal sensitivity.

A model of AN activity was shown to account for the trends observed in the EFR level-growth recorded in NH threshold listeners. To account for the non-monotonic trends obtained in the EFR level-growth functions for some NH threshold listeners, a loss of all types of AN fibers needed to be implemented in the model. An exclusive loss of medium- and low-SR fiber did not have an impact on the simulated EFR level-growth functions, in contrast to suggestions in the literature. The same was found for the postulated synaptopathy in HI listeners, where a large loss of all three AN fiber types had to be included in a very broad frequency range.

Overall, the simulations suggest that EFRs are dominated by high-SR fibers, and that off-frequency neurons increasingly contribute to the EFR with increasing stimulus level. The finding that the envelope is better encoded at off-frequency CFs (rather than on-frequency) when SAM tones are presented at high stimulus levels should be generally considered when using EFRs to investigate supra-threshold coding.

Acknowledgments

The authors want to thank Le Wang and Graham Voysey from Boston University (BU) for their valuable comments and discussions regarding the model implementation and analysis. I would like to express our sincere gratitude to Laurel

Carney for her wise advice regarding the analysis of the model outcome and the fascinating discussions during her stay in Denmark about the role of the different types of AN fibers, the interpretation of the rate-level functions and the processing of modulations in the AN, the CN and the IC. This work was supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) at the Technical University of Denmark (DTU). The authors want to acknowledge the support provided by the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience.

4

Simulating envelope following responses using the model of Verhulst *et al.* (2015)

Abstract

Auditory evoked potentials, such as envelope following responses (EFR), provide an objective tool to noninvasively investigate the function of the auditory system. Such physiological measures can in principle be compared across species, allowing the translation of insights gained from invasive studies in non-human animals to humans. In the current understanding, EFRs arise from the contribution of many different neuronal sources located at different stages along the auditory pathway. These in turn respond in a phase-locked manner to the envelope of the acoustic stimulus. Thus the interpretation of EFRs is not trivial. Computational models can assist in analyzing such results, as the contribution of, for example, the different stages along the auditory pathway, the different characteristic frequencies (CF), the different types of neurons (i.e. high- vs medium vs low-spontaneous rate (SR) auditory nerve (AN) fibers) and the effect of hair-cell or AN fiber damage (i.e. synaptopathy) can be systematically incorporated in the model. In the present study, the model proposed by Verhulst et al. (2015) was used to simulate EFRs for a large range of stimulus levels using strongly and shallowly modulated tones. The analysis of the model outcome was focused on the relation between the on- and off-frequency contributions (i.e. near or away from the characteristic frequency of the stimulus, respectively). Overall, the model results showed that the high-SR fibers dominated the total summed response across fiber types and that the off-frequency CF contributions dominated the total response at moderate-to-high stimulus levels. However, the model simulations demonstrated an unrealistic over-representation of the

steady-state AN activity at high CFs when medium-to-high stimulus levels were used. The causes of such enhanced activity at the high frequencies were investigated and a suggestion for improving the model based on a different fitting of one of the model parameters was made, which resulted in a more balanced and realistic response of the AN.

4.1 Introduction

In the previous *chapter 3* based on the study by Encina-Llamas et al. (under review[b]), a well-established model of the auditory nerve (AN) (Zilany et al., 2009, 2014) was used to study the growth-level functions of envelope following responses (EFR). The contributions of each type of spontaneous rate (SR) AN fiber as well as the on- versus off-frequency (i.e. near or away from the characteristic frequency of the stimulus respectively) contributions to the total EFR were investigated. In addition, the effect of a loss of AN fiber synapses (synaptopathy) onto the EFR level-growth functions was considered. The model by Zilany et al. (2009, 2014) does not provide information regarding the processing occurring at the level of the basilar membrane (BM) in the cochlea, but has been designed to successfully account for physiological results directly recorded from neurons in the AN. In this phenomenological model, the BM is represented by a time-varying level-dependent filterbank, coupled to an inner hair-cell (IHC) transduction and the AN synapse model. The combined model was designed to faithfully represent AN firing rated to a broad range of stimuli, but not to represent more peripheral physiology accurately, i.e. BM velocity. Such filter-based BM models cannot account for experimentally relevant effects like otoacoustic emissions (OAE) because the system is simulated without interaction between frequency channels (Epp et al., 2010). Nonlinear time-domain transmission-line cochlear models are an alternative to the filter-based models (e.g.; Epp et al., 2010; Verhulst et al., 2012). Verhulst et al. (2015) coupled a nonlinear transmission-line cochlear model with models of the IHC (Jepsen et al., 2008; Shamma et al., 1989; Sumner et al., 2002; Zhang et al., 2001), the AN-synapse (Meddis, 1986; Westerman and Smith, 1988; Zhang and Carney, 2005) and functional models of the ventral cochlear nucleus (VCN) and the inferior colliculus (IC) (Carney et al., 2015; Nelson and Carney, 2004), so that responses from the cochlea (i.e. OAE) to the brainstem (i.e. auditory brainstem responses, ABR)

could be simulated.

The model of Verhulst et al. (2015) was mainly focused on obtaining accurate brainstem responses to transient and broadband stimuli (i.e. ABRs), in particular level-dependent changes of ABR wave-V latency. Previous models have accounted for brainstem evoked responses by using the model by Zilany et al. (2009, 2014) as a preprocessor in connection with a unitary response (e.g.; Melcher and Kiang, 1996), intended to reflect the transfer function of different neuronal contributions along the auditory brainstem to the far-field evoked potential (Dau, 2003; Rønne, 2013; Rønne et al., 2012). Although this approach was successful for simulating ABR wave-V amplitudes (Rønne et al., 2012), it provided inaccurate results on the simulated ABR latencies. These were however correctly captured by the model of Verhulst et al. (2015).

The purpose of the present study was to investigate whether the model of Verhulst et al. (2015) could also be used for simulating EFRs, with the ultimate goal of studying the effect of synaptopathy onto the EFR level-growth functions, as similarly done in *chapter 3*, but with inclusion of a more plausible model of the basilar membrane. Given that the model by Zilany et al. (2009, 2014) and the model by Verhulst et al. (2015) provide comparable results at the level of the AN, the additional BM outputs provided by the latter model (i.e. BM velocity and OAEs) would be very useful for analyzing and suggesting novel stimulus paradigms that could simultaneously measure EFRs and OAEs (Purcell et al., 2003; Sanchez and Epp, 2015). EFRs elicited by sinusoidally modulated (SAM) tones at modulation rates of about 80-100 Hz are believed to be mainly generated by neuronal sources at the level of the brainstem, corresponding to the ABR wave-V (Picton et al., 2003). However, the ABR is an evoked response to a transient acoustic stimulus whereas the EFR represent a steady-state neuronal response, which was out of the initial scope of the model by Verhulst et al. (2015). As discussed previously (see *chapter 3*), the EFRs are a gross electroencephalographic (EEG) potential on which off-frequency contributions may be critical at high stimulation levels. As the model by Verhulst et al. (2015) provides outputs at different stages of the auditory pathway, a systematic analysis starting at the level of the BM and onward was performed focusing on simulated neuronal population responses and off-frequency contributions of neural responses.

4.2 Methods

4.2.1 Implementation of the model by Verhulst *et al.* (2015)

The model described in Verhulst et al. (2015) was used through the *Corti* modeling environment implemented in Python 3.3 by Voysey and Encina-Llamas (2016). One thousand characteristic frequencies (CF) or frequency channels ranging from 20 Hz to 20 kHz were used, equally-spaced along the cochlea according to the frequency map for human (Greenwood, 1990). Each segment of the nonlinear transmission-line model (Verhulst et al., 2012) representing the cochlear partition corresponded to one CF. The output of the BM model (BM velocity) at each CF was passed to the IHC model from which the IHC receptor potential was computed assuming one IHC per cochlear segment. The resulting simulated IHC receptor potential ($V_{\mathrm{IHC,n}}$) [1] at each CF was passed to the AN synapse model for each AN fiber type (i.e. high-SR, medium-SR and low-SR) which provided the mean instantaneous spiking rate for each fiber type (this model does not account for spiking probability). At this stage, a matrix of the form CF x Time samples was obtained for each fiber type containing the fiber AN response. To account for the different AN fiber distribution, it was assumed that 19 AN fiber synapses innervate each IHC, distributed into 13 high-SR fibers (68,4%), 3 medium-SR fibers (15,8%) and 3 low-SR fibers (15,8%) (Liberman, 1978). A uniform number of synapses per IHC across CF was assumed. The resulting AN fiber response matrices for each different fiber type were scaled by their corresponding fiber proportion (e.g. the resulting matrix from the high-SR fiber was multiplied by $^{13}/_{19}$, the ones from the medium- and low-SR by $^{3}/_{19}$). The three resulting AN matrices (for each fiber type) were summed together to provide the instantaneous firing rate AN population response ($r_{\mathrm{AN},n}(t)$), which was passed on to the VCN model and then to the IC model (Carney et al., 2015; Nelson and Carney, 2004).

Model simulations were performed using the same stimuli as in the human EFR recordings reported in *chapter 3* (f_c = 2000 Hz, f_m = 93 Hz) but with a duration of 0.3-s. Stimulus levels ranged from 5 to 110 dB sound pressure level (SPL), in steps of 5 dB. In order to analyze the steady-state response at the modulation frequency, the onset and offset parts of the simulated responses at the different peripheral stages (i.e. BM, IHC and AN) were removed so that the

[1] The same notation as in Verhulst et al. (2015) was adopted throughout the present manuscript.

steady-state response contained an integer number of cycles of the modulation frequency (lasting about 0.226 s). For the analysis of the on- versus off-frequency responses, the AN synaptic output at the on-frequency channel (2 kHz) and at one off-frequency channel (7 kHz) were considered.

4.3 Results

4.3.1 Simulated basilar-membrane (BM) and auditory nerve (AN) response

Figure 4.1 shows the simulated BM velocity (panel A) and IHC receptor potential (panel B) obtained using a SAM tone with carrier frequency $f_c = 2$ kHz, modulation frequency $f_m = 93$ Hz and modulation depth $m = 85\%$. The response is shown for a stimulus level of 90 dB SPL (for a stimulus level ranging from 5 to 110 dB SPL in steps of 5 dB in the form of an animation in the *digital version*) and for four cycles of the f_m at the steady-state part of the response at CFs from 20 Hz to 20 kHz.

At 90 dB SPL, the BM (panel A) showed a response which was spread towards frequencies higher than f_c (i.e. towards basal locations), with the broader excitation at the peaks of the modulation cycles. At this stimulation level, the maximal BM velocity amplitudes were were located at CFs slightly higher but still near on-frequency (f_c). At the troughs of the modulation frequency, the BM velocity became minimal at all CF frequencies, including the on-frequency channels. The red and blue striping pattern showed the fine-structure (f_c) whereas the four big areas of high activity represented the four modulation cycles (f_m). Along a range of levels, excitation patterns were very narrow and frequency specific at low stimulus levels and broadened with increasing level (see animation in the *digital version*). In particular at stimulus levels above 60 dB SPL, the broadening rate accelerated.

At 90 dB SPL, the IHC receptor potential showed a similar response than the BM velocity but with an excitation located particularly towards higher CFs, and slightly towards lower CFs. The maximal IHC voltage at such high stimulation levels was located slightly above the CF of f_c (on-frequency). However, in contrast to the BM response, the IHC receptor potential at or near the on-frequency CF fluctuated less over time, and the troughs of the modulation frequency were not so clearly represented. In contrast, at higher off-frequency CFs, the mod-

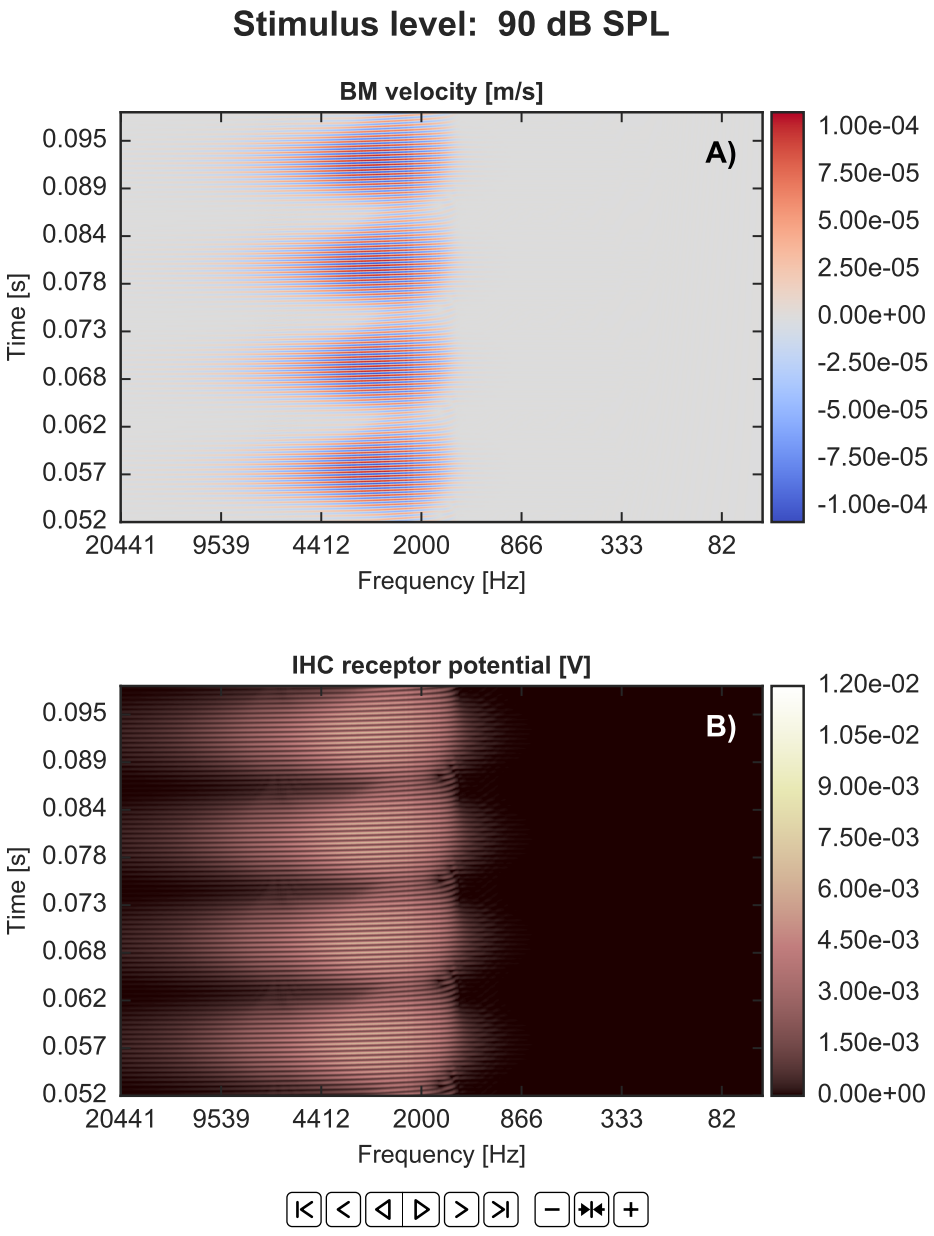


Figure 4.1: Simulated (A) BM velocity and (B) IHC receptor potential obtained using a SAM tone with $f_c = 2$ kHz, $f_m = 93$ Hz and $m = 85\%$. Four cycles of the modulation frequency (f_m) at the steady-state part of the response at CFs from 20 Hz to 20 kHz are shown (high frequencies to the left and low frequencies to the right). *Paper version*: Stimulus level presented at 90 dB SPL. *Digital version*: Stimulus levels from 5 to 110 dB SPL, in steps of 5 dB (*use the control buttons to navigate through the animation*). The color scale in (A) varies with stimulus level for visualization purposes, whereas in (B) is kept constant so that the responses across stimulus levels can be compared.

ulated pattern could clearly be observed. The fine-structure was visible as a striping pattern. For increasing stimulus level, a spread of excitation could be observed (see also animation in the *digital version*). Whereas the excitation patterns of the BM velocity (panel A) and the IHC potential (panel B) were very similar at stimulus levels up to about 35-40 dB SPL, the IHC patterns showed higher amplitudes at higher CFs than the BM at higher stimulus levels, particularly at levels above 55-60 dB SPL.

Figure 4.2 shows the simulated AN response using a SAM tone with carrier frequency $f_c = 2$ kHz, modulation frequency $f_m = 93$ Hz and modulation depth $m = 85\%$. Panels A-D show the response of the high-SR fibers, medium-SR fibers, low-SR fibers and the sum of all three fiber types for a stimulus level of 90 dB SPL, respectively, for four cycles of f_m at the steady-state part of the response at CFs from 20 Hz to 20 kHz (see the animation in the *digital version* for levels from 5 to 110 dB SPL in steps of 5 dB).

The AN response at 90 dB SPL showed very broad activity spreading towards CFs much higher than f_c for all three fiber types, especially for the high-SR fibers (panel A) and the sum across fiber types (panel D). The broad excitation patterns for the high-SR fibers (A) and the sum across fiber type (D) showed significantly larger responses at very high off-frequency CFs (above about 10 kHz) than at CFs near the on-frequency. This was also found for the medium-SR fibers (B) but with smaller amplitudes. The response after summing across all three fiber types (D) was very similar to the response for the high-SR fiber (A), meaning that the high-SR fibers dominated the total AN response (compare the color scales of all panels). Similar to the IHC response, and in contrast to the BM response, the AN response at or near the on-frequency CFs showed little fluctuation at f_m over time, whereas the modulation was well represented at the higher off-frequency CFs. At very low stimulation levels the AN activity was only present at the high-SR fibers panel (A) and the excitation was very narrow and frequency specific around the 2 kHz CF (see also the animation in the *digital version*). As the stimulus intensity increased, first the medium-SR (B), and ultimately the low-SR fibers (C) started to respond. With increasing stimulus level, the AN activity, especially for the high-SR fibers (A), broadened rapidly recruiting AN neurons tuned to high CFs. Such spread of the high-SR fibers (A), and therefore also the sum across fiber types (D), showed activity above 10 kHz already at stimulation levels of 60 dB SPL. Hence, at those medium-to-high stimulus levels, the relative AN activity at the high off-frequency CFs versus the

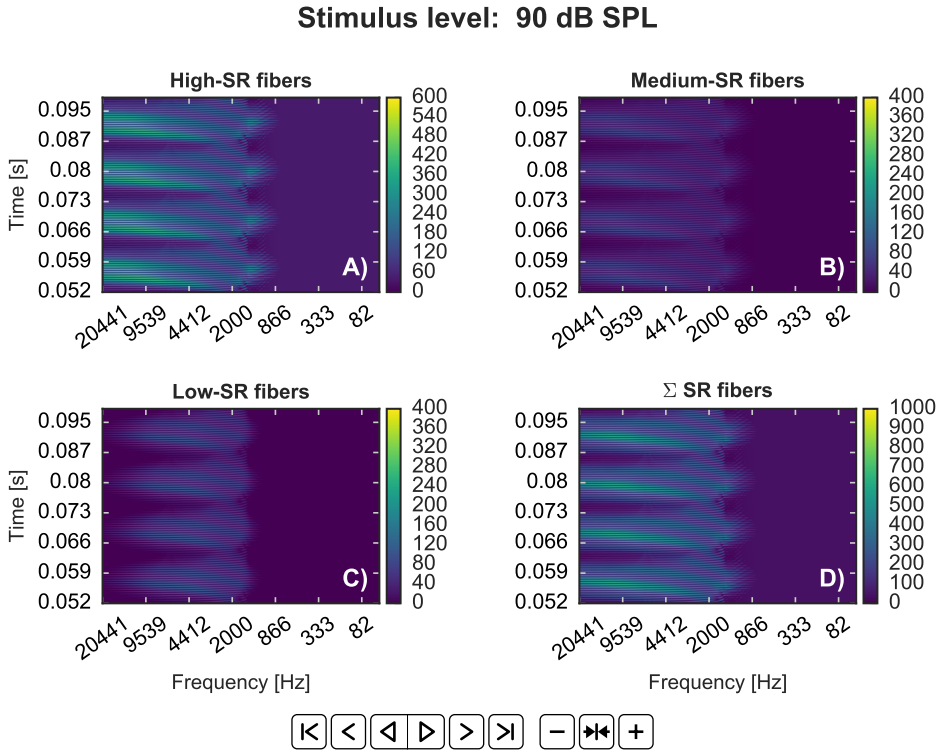


Figure 4.2: Simulated AN synaptic output (mean instantaneous firing rate, spikes/s) obtained using a SAM tone with $f_c = 2$ kHz, $f_m = 93$ Hz and $m = 85\%$. The AN response is shown for (A) the high-SR fibers, (B) the medium-SR fibers, (C) the low-SR fibers and (D) the sum of all three fiber types. Four cycles of the modulation frequency (f_m) at the steady-state part of the response at CFs from 20 Hz to 20 kHz are shown (high frequencies to the left and low frequencies to the right). *Paper version*: Stimulus level presented at 90 dB SPL. *Digital version*: Stimulus levels from 5 to 110 dB SPL, in steps of 5 dB (*use the control buttons to navigate through the animation*).

on-frequency ones were larger than the corresponding relative IHC response and the corresponding relative BM response.

4.3.2 On- versus off-frequency contributions to simulated auditory nerve responses

Figure 4.3 shows the simulated instantaneous firing rate of the AN obtained at on-frequency (2 kHz) and off-frequency (7 kHz) using the same stimuli as in Figure 4.2 at a stimulus level of 90 dB SPL (animated from 5 to 110 dB SPL in the *digital version*). Panels A-D show the response of the high-SR fibers, medium-SR fibers, low-SR fibers, and the sum of all three fiber types, respectively. The panels with the number 1 correspond to the on-frequency response for each fiber type, and panels with number 2 correspond to the off-frequency response.

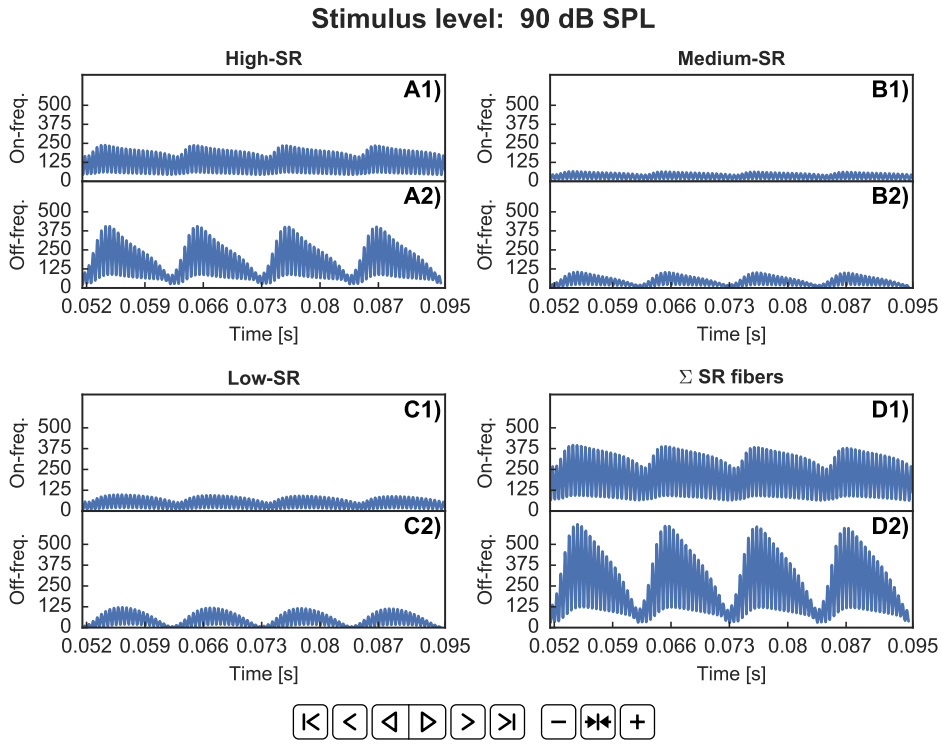


Figure 4.3: Simulated AN synaptic output (mean instantaneous firing rate, spikes/s) on-frequency (2 kHz) and off-frequency (7 kHz), obtained using a SAM tone with $f_c = 2$ kHz, $f_m = 93$ Hz and $m = 85\%$. The AN response is shown for (A1-A2) the high-SR fibers, (B1-B2) the medium-SR fibers, (C1-C2) the low-SR fibers and (D1-D2) the sum of all three fiber types. The panels with the number 1 correspond to the on-frequency ($f_c = 2$ kHz) response for each fiber type whereas number 2 correspond to the off-frequency (7 kHz) response. Four cycles of the modulation frequency (f_m) at the steady-state part of the response are shown. *Paper version*: Stimulus level presented at 90 dB SPL. *Digital version*: Stimulus levels from 5 to 110 dB SPL, in steps of 5 dB (*use the control buttons to navigate through the animation*).

The same time windows as in Figures 4.1 and 4.2 are shown with four cycles of f_m at the steady-state part of the response.

The AN synaptic response at 90 dB SPL showed different patterns at the on-frequency CF (D1) than at the off-frequency CF (D2). As visible in Figure 4.2, such high stimulus levels neurons tuned to frequencies much higher than the CF of the SAM tone (off-frequencies) were excited. The response at the on-frequency CF for all three types of AN fibers and the sum across fiber type (A1-D1) was more saturated than the corresponding response at the off-frequency CF (A2-D2). Therefore, at high stimulation levels, the modulation was better encoded at the off-frequency CF than at the on-frequency CF. The simulated AN response for the high-SR fiber (A) was larger than the response for the medium-

SR fiber (B) and the low-SR fiber (C), so that the sum across fiber type (D) was dominated by the high-SR response.

Similar to Figure 4.2, the AN activity for low stimulus levels was only present at the on-frequency CF for the high-SR fibers panel (A1). As the stimulus intensity increase (see animation in the *digital version*), first the on-frequency CF for the medium-SR fibers and then the low-SR fibers (C1) start to increase their firing rate following both, the SAM tone envelope and fine-structure. Whereas the AN synaptic output (e.g. for the high-SR fiber, A1) up to about 30 dB SPL followed the sinusoidal envelope, the simulated AN response at higher stimulus levels was shown to be skewed towards the onset of each modulation cycle. At about 50 dB SPL, first the off-frequency CF for the high-SR fibers (A2), and then for the medium- and low-SR fibers start responding. Above this level the simulated response at the on-frequency CF becomes more saturated with increasing level and the envelope better represented at the off-frequency CF in all three fiber types responses (A2-C2) and in the sum across fiber types (D2). At very high stimulus levels (above 85 dB SPL), the skewed response was also found at the off-frequency CF.

4.4 Discussion

4.4.1 Over-represented simulated AN response at high frequencies

The model simulations revealed three key observations: 1) the activity of the high-SR fibers dominated the total AN response; 2) the modulation of the SAM tones (i.e. the envelope) was better represented at off-frequency CFs than on-frequency CF for high stimulation levels; and 3) the spread of excitation seemed to be enhanced at the level of the AN versus the IHC, and at the level of the IHC versus the BM.

The first two observations are consistent with results obtained using the model of Zilany et al. (2009, 2014) in *chapter 3*. It could be clearly observed in Figures 4.2 and 4.3 that the simulated AN response summed across fiber types (D) was very similar to the response for the high-SR fiber (A), indicating its dominance in the total AN response. As discussed in chapter 3, this was due to the uneven distribution of the three SR fiber types based on data from the cat AN (Liberman, 1978), on which the number of high-SR fibers is larger than the medium- and low-SR fibers. At high stimulation levels, the AN simulation

results suggested that the modulation was better encoded at the off-frequency CF than at the on-frequency CF (Figure 4.3). This is also consistent with the AN simulations using the model of Zilany et al. (2009, 2014) reported in the chapters 3 and 5 in the present thesis, and it is also consistent with invasive AN recordings in cats reported by Joris and Yin (1992). The IHC transduction characteristic, which shows a steep growth at low stimulation levels and a strong saturation at high levels (Russell et al., 1986), flattens out the envelope of the AN response at on-frequency CFs and allows large fluctuations at off-frequency CFs.

The model results using the model by Verhulst et al. (2015) showed that the spread of excitation towards high CFs with increasing stimulus levels seemed enhanced at the level of the AN over the responses from the IHC. More accurately, what the model simulation showed was not that the excitation patterns became broader from one stage to another along the pathway, but rather that the relative response at high stimulation levels between the on- versus off-frequencies shifted from being dominated by the near on-frequency CFs in the BM response (Fig. 4.1, A) to a stronger weight of the off-frequency response in the IHC response (Fig. 4.1, B) and finally being dominated by the very high off-frequency CFs in the AN response (Fig. 4.2, D).

The IHC transduction could explain the enhancement of the off-frequency CFs in the IHC receptor potential (Fig 4.1, B) at high stimulation levels (i.e 90 dB SPL) over the simulated BM excitation (Fig 4.1, A). In this comparison, although the off-frequency response was more clearly represented in the IHC response than in the BM response, the larger activity across the frequency axis was still centered near the on-frequency CF (but slightly towards the higher frequencies). However, when comparing the response of the IHC (Fig 4.1, B) with the response of the AN firing rate (Fig. 4.2, D) at the same high stimulus levels, the maximum AN activity was centered at the higher CF. Such AN responses are inconsistent with the simulations reported in chapter 3 using the model of Zilany et al. (2009, 2014), on which the maximal firing rate was centered at near the on-frequency CFs (but slightly towards the higher frequencies). It is important that the concepts of envelope synchrony and firing rate are not confused. Both the models described in Verhulst et al. (2015) and Zilany et al. (2009, 2014) are consistent with the point that, at high stimulus levels, the envelope is better represented at high frequency off-frequency CFs (i.e. better envelope synchrony, and therefore EFR) compared to the on-frequency CF.

However, the model of Zilany et al. (2009, 2014) showed that the firing rate (i.e. the absolute number of spikes within a second) was similar at the medium CFs where the SAM tone was presented (near on-frequency) and much lower at very high and low CFs. The model of Verhulst et al. (2015), on the other hand showed that the simulated AN firing rate at high stimulus levels was much larger at very high (basally located) CFs than everywhere else. This simulations using the model described in Verhulst et al. (2015) are not supported by any physiological data to the author's knowledge, suggesting that it might require some additional adjustments before being applicable for the simulation of EFRs.

4.4.2 A suggestion for improving the AN response at high frequencies

The simulated instantaneous firing rate of the AN synapses in the model of Verhulst et al. (2015) ($r_{\mathrm{AN},n}(t)$) depends on the product of two variables: the concentration of synaptic neurotransmitters in the immediate store ($\mathrm{CI}_n(t)$) and the neurotransmitter vesicle release permeability to the immediate store ($\mathrm{PI}_n(t)$). Several time and frequency dependent equations govern the response of $\mathrm{CI}_n(t)$ and $\mathrm{PI}_n(t)$ and therefore the AN response (see the equations (6-10) and (A1-A16) in Verhulst et al. (2015)). However, of all the parameters listed in the AN synapse part of Table I in Verhulst et al. (2015), only the parameter $A_{\mathrm{ss},n}$ is CF-dependent. This parameter stands for the AN steady-state firing rate at saturation, that is the maximum instantaneous firing rate of a given AN fiber in the steady-state part of the response. This parameter was derived by fitting a 1$^{\mathrm{st}}$-order polynomial to the data reported in Figure 17 of Liberman (1978). The parameter $A_{\mathrm{ss},n}$ is directly or indirectly related in the equations to $\mathrm{CI}_n(t)$ and $\mathrm{PI}_n(t)$, and might be very influential on the AN synaptic output when simulation steady-state responses at high stimulation levels where many AN neurons are driven in their saturating firing rate regions.

The fitting reported in Verhulst et al. (2015) was

$$A_{\mathrm{ss},n;\ \mathrm{Original}} = 150 + (\mathrm{CF}_n/100) \tag{4.1}$$

in which the frequency vector CF_n was linear. However, the raw data shown in Liberman (1978) was plotted on a logarithmic axis in the abscissa. Therefore, the fitting from Verhulst et al. (2015) represented in the correct logarithmic axis became an exponential fit, where the low and medium CFs were under-

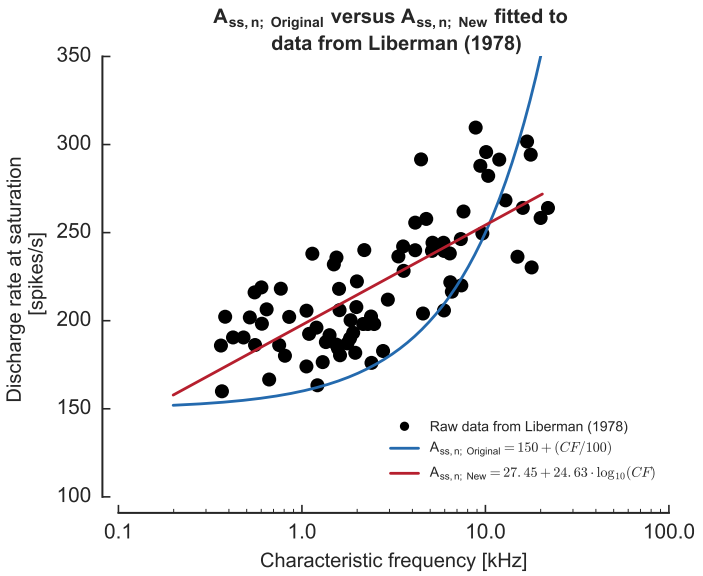


Figure 4.4: Fits to describe the parameter $A_{ss,n}$ (discharge rate at saturation, spikes/s) implemented in Verhulst et al. (2015) (blue line) and as suggested in the present study (red line). The black circles show (reprinted with permission) the raw data directly recorded from AN neurons from the cat reported in Liberman (1978).

estimated and the very high frequencies were over-estimated. Figure 4.4 shows the raw data reprinted from Liberman (1978) as black circles and the blue continuous line shows the fit used in Verhulst et al. (2015). The red continuous line represents a new model using the logarithmic CF axis:

$$A_{ss,n;\ New} = 27.45 + 24.63 \cdot \log_{10}(CF_n) \tag{4.2}$$

This model follows more accurately the frequency-dependent trend observed in the raw data.

The effects on the AN synaptic output are shown in Figure 4.5. Panels A-D show the output as a functions of CF of the main variables that determine the AN synaptic response in the original model of Verhulst et al. (2015) at a particular time sample (t = 0.068 s) in the steady-state part of the response at the maximum of a modulation cycle for a stimulus level of 95 dB SPL. Panels E-H show the same responses when using the parameter $A_{ss,n}$ with the implementation suggested in the present study. Panels A and E, show the response of the IHC receptor potential (V_{IHC}) as a blue solid line and the IHC receptor potential threshold ($V_{TH} = 50\mu V$) as a red dashed line below which the AN output remains at SR for the high-SR fibers (see Verhulst et al. (2015) for more details). Panels B and F show the neurotransmitter vesicle release permeability to the immediate store ($PI_n(t)$), which is the only stimulus-dependent parameter in the AN model (through the IHC response V_{IHC}). The parameter $PI_n(t)$ depended on $A_{ss,n}$ and

the effect of the suggested fit is visible when comparing panels B and F. Although the trends in panels B and F are similar, and that they are very similar to the IHC responses, the value of $\mathrm{PI}_n(t)$ in the suggested implementation (F) was slightly lower than the original implementation at the high frequency extreme (left). In addition, the first lower peak and mainly the second higher peak in the response of panel F were larger than the ones in panel B.

Panels C and G show the concentration of synaptic neurotransmitters in the immediate store ($\mathrm{CI}_n(t)$). The parameter $\mathrm{CI}_n(t)$ also depended on $A_{\mathrm{ss},n}$ but the effect of the suggested implementation mainly affected its response at the lower frequencies, which did not modify the AN output at those frequencies after being multiplied by $\mathrm{PI}_n(t)$. In the high frequency part of $\mathrm{CI}_n(t)$ (left), the response for both implementations were very similar, but the response at the very high frequency extreme (left) was slightly lower in the newly suggested implementation(G) than in the original implementation (C).

Finally, the product of the responses of $\mathrm{CI}_n(t)$ and $\mathrm{PI}_n(t)$ resulted in the AN synaptic output $r_{\mathrm{AN},n}(t)$ (for Verhulst et al. (2015) implementation, panels B$\cdot$C = D; for the suggested implementation, panels F$\cdot$G = H). In the original implementation from Verhulst et al. (2015) (panel D) the maximal AN response was at the high frequency extent on which the firing rate was about 150 spikes/s larger than the response near the on-frequency CFs (second smaller peak). Such over-representation of the AN response at the higher frequencies arose from the exponential fit of the parameter $A_{\mathrm{ss},n}$ (see Fig. 4.4) which was mitigated when using the suggested fit of $A_{\mathrm{ss},n}$ (panel H). In the suggested implementation, the maximal AN activity at high stimulation levels was similar at all frequencies. The AN synaptic response at the higher CFs (panel H) were reduced with respect to the AN response shown in panel D. At the same time, the response near the on-frequency CFs in panel H was larger relative to the corresponding response in panel D. This difference between the original implementation and the suggested one is clearly due to the modification in the fitting of the parameter $A_{\mathrm{ss},n}$ because it goes in the same direction, on which the low-to-medium frequencies in the new fitting were accentuated while the high frequencies were attenuated.

A complementary view of the effect of the modification in the fitting of the parameter $A_{\mathrm{ss},n}$ is shown in Figure 4.6. This shows the maximum firing rate in the steady-state part of the AN response at different CFs for both implementations of $A_{\mathrm{ss},n}$ when using a SAM tone with $f_c = 2000$ Hz and $f_m = 93$ Hz at 95 dB SPL. The blue symbols and line show the maximum AN firing rate when using the

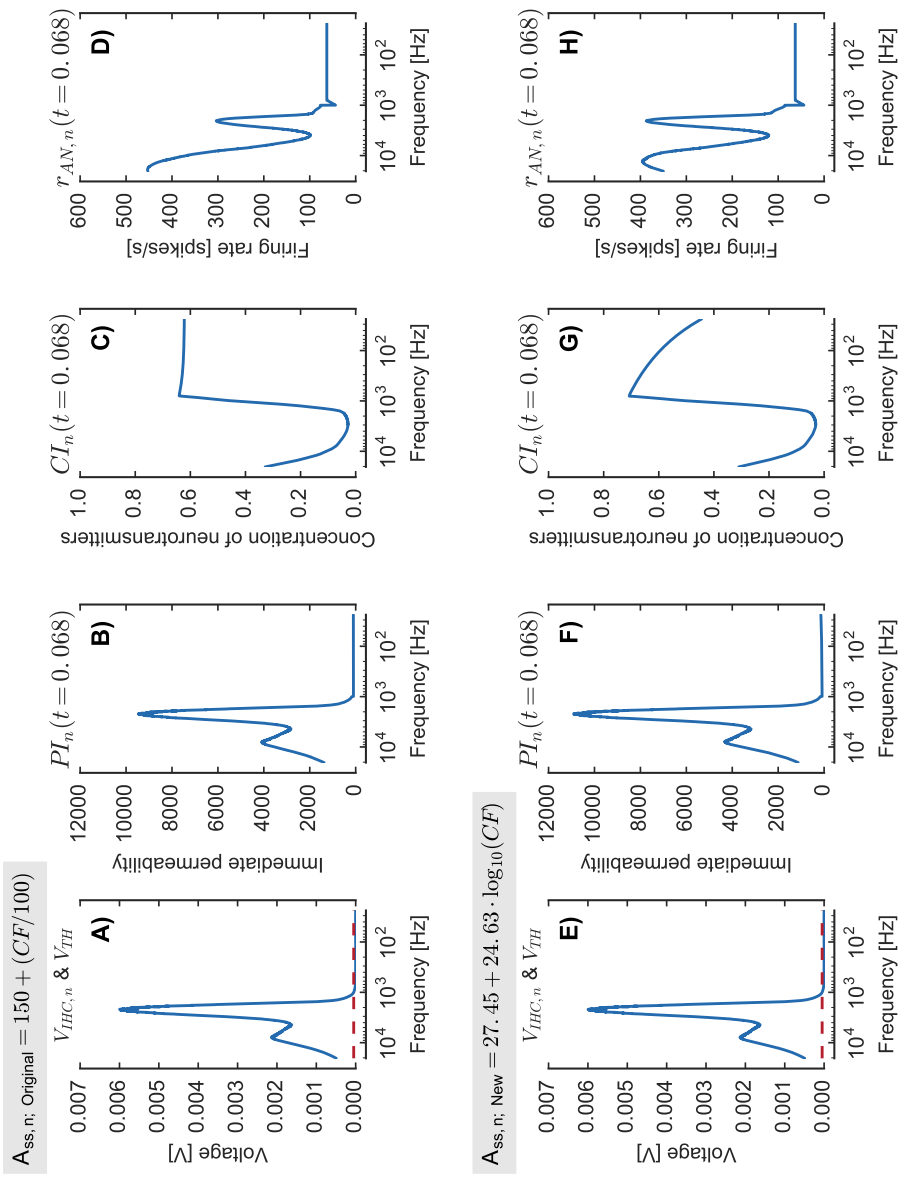


Figure 4.5: Comparison of the effect of the (A-D) original implementation of the parameter $A_{\mathrm{ss},n}$ ($A_{\mathrm{ss},n;\,\mathrm{Original}} = 150 + (\mathrm{CF}_n/100)$) and the (E-H) newly suggested implementation ($A_{\mathrm{ss},n;\,\mathrm{New}} = 27.45 + 24.63 \cdot \log_{10}(\mathrm{CF}_n)$) on the variables $\mathrm{PI}_n(t)$, $\mathrm{CI}_n(t)$ and $r_{\mathrm{AN},n}(t)$ in the AN synapse module. Simulation using a SAM tone with $f_c = 2000$ Hz, $f_m = 93$ and $m = 85\%$ presented at a stimulus level of 95 dB SPL at the time $t = 0.068$ s corresponding to a peak of a modulation cycle in the steady-state part of the response.

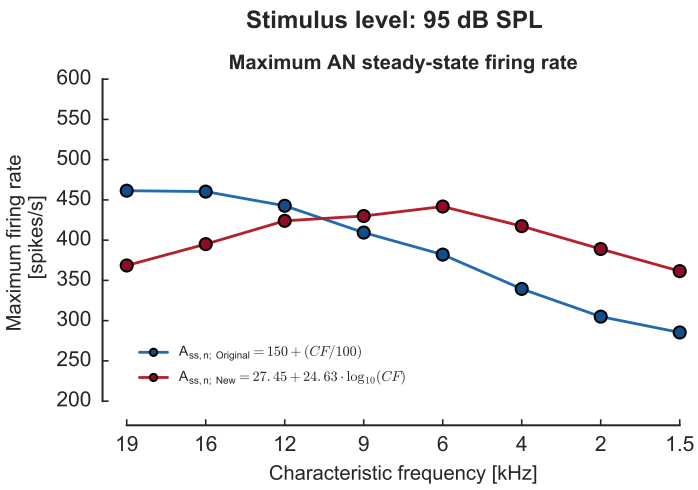


Figure 4.6: Maximum firing rate at the steady-state part of the AN response at different CFs for the original implementations of $A_{ss,n}$ in Verhulst et al. (2015) and the newly suggested implementation using a SAM tone with f_c = 2000 Hz, f_m = 93 Hz and m = 85% presented at a stimulation levels of 95 dB SPL.

original implementation of Verhulst et al. (2015). The red symbols and line show the same simulated results for the new implementation of $A_{ss,n}$. In the original implementation, the maximal AN firing rate was skewed towards the higher frequencies (left) for high stimulus levels. This led to an increase in the firing rate with increased stimulus level. In the results from the suggested implementation of $A_{ss,n}$, the maximum AN firing rate is shown to be more uniform across CFs peaking at 6 kHz. This is much closer to the on-frequency CF of 2 kHz, and therefore more plausible.

The suggested modification to the model of Verhulst et al. (2015) improved the simulated steady-state responses at the level of the AN when using medium-to-high stimulation levels. However, it is unknown whether this change leads to a reduced performance by the model at the upcoming stages (e.g. the VCN) or when using other types of stimuli (e.g. transient response or responses to stimuli at low intensities). The complexity of such nonlinear models does not allow a general conclusion based on only a limited evaluation as in the present study. A detailed analysis and evaluation of the effect of the suggested change for improvement of the model of Verhulst et al. (2015) in response to high stimulus level steady-state responses would be required to quantify the effects. As there was no immediate benefit of using the Verhulst et al. (2015) model over the model of Zilany et al. (2009, 2014), and considering that the original implementation of Verhulst et al. (2015) model did not provide accurate enough results at the level of the AN, it was decided not to use this model to study the EFRs in its current form.

4.5 Conclusion

The simulations using the model of Verhulst et al. (2015) for a SAM tone with carrier frequency of 2000 Hz and a modulation frequency of 93 Hz at stimulus levels ranging from 5 to 110 dB SPL led to similar general conclusions as the ones derived using the model of Zilany et al. (2009, 2014) reported in chapters 3 and 5. Overall, the simulations suggested that: 1) the high-SR fibers dominate the steady-state AN response, and 2) the encoding of the envelope is better represented at high CF off-frequency channels than near the on-frequency CFs when using high stimulation levels. However, the simulated AN activity at the high CFs in the model of Verhulst et al. (2015) to medium-to-high intensity stimuli was shown to be unrealistically over-represented in comparison to the simulated results from the model of Zilany et al. (2009, 2014) and not supported by direct physiological observations of the AN activity. It was suggested that the parameter $A_{\mathrm{ss},n}$ in the model of Verhulst et al. (2015) representing the AN steady-state firing rate at saturation may have been inaccurately fitted, and a new fit was suggested. After implementing the new fit, it was shown that the AN response at the higher CFs channels was mitigated and the maximum firing rate became more balanced and even across CFs. It remains however to be investigated whether such a modification of the model alters the model response at other stages (i.e. the response in the VCN or the IC) or under other stimulus conditions (i.e. the response to transient stimuli or to low intensity sounds).

Acknowledgments

This work was done at the Auditory Neuroscience Laboratory in the Center for Computational Neuroscience and Neural Technology (CompNet) at Boston University (BU), MA (USA) under the supervision of Prof. Barbara Shinn-Cunningham. We thank Le Wang and Graham Voysey from BU for their valuable comments and discussions regarding the model implementation and analysis. The authors want to acknowledge the support provided by the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience. This work was supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) at the Technical University of Denmark (DTU).

5

Envelope following response in mice with and without cochlear synaptopathy investigated through a computational model of the auditory nerve[c]

Abstract

Cochlear synaptopathy, defined as the loss of auditory nerve (AN) fiber synapses without hair-cell damage, has been proposed as a potential source of supra-threshold deficits reported in listeners with normal audiometric thresholds. Despite the fact that synaptopathy has been demonstrated in several non-human mammalian species, its presence in humans is still controversial. Many authors have failed to find similar patterns in electrophysiological responses in synaptopathic non-human animals and in humans. In the present study, EFR level-growth functions for strongly and shallowly modulated tones were recorded in noise-exposed and non-exposed mice, and compared to simulated EFRs using a computational model of the AN. Noise-induced synaptopathy in mice led to a reduction of EFR magnitudes at medium stimulus levels, equally affecting the responses for strongly and shallowly modulated tones. The observed patterns were similar to EFR level-growth functions reported previously in normal-hearing human listeners for shallowly modulated tones. The model simulations indicate that the EFR magnitudes were dominated by the activity of the high-spontaneous rate fibers and that the off-frequency neurons (i.e. away from the characteristic place of the stimulus) increasingly contributed to the EFR with increasing stimulus level. Although the general trend of

[c] This chapter is based on Parthasarathy et al. (2017a).

the recorded EFRs and the overall effect of synaptopathy in mice could be accounted for by the model simulations, the shape of the EFR level-growth functions did not fully agree with the recordings, probably caused by the differences in AN tuning between the model (cat) and the data (mouse). A mouse version of the computational model is needed to account for the data and to contribute to the understanding of synaptopathy across species.

5.1 Introduction

It has been demonstrated in non-human mammals that sound overexposure results in a permanent loss of synapses between the inner hair cells (IHC) and the auditory nerve (AN) fibers (e.g.: Furman et al., 2013; Kujawa and Liberman, 2009; Lin et al., 2011; Liu et al., 2012; Lobarinas et al., 2016; Valero et al., 2017). In animal models, such a loss of AN fiber synapses, often referred to as cochlear synaptopathy (Kujawa and Liberman, 2009) or "hidden" hearing loss (Schaette and McAlpine, 2011), is not manifested as a permanent reduction of sensitivity as measured by distortion product otoacoustic emissions (DPOAE) or auditory brainstem responses (ABR) thresholds. In mice, Shaheen et al. (2015) reported that the envelope following responses (EFR), a gross electroencephalographic (EEG) potential elicited by populations of neurons that fire synchronously (phase-locked) to the envelope of an acoustic stimulus, were reduced in animals with noise-induced synaptopathy. In humans, several studies have attempted to investigate synaptopathy (Bharadwaj et al., 2015; Grose et al., 2017; Le Prell et al., 2017; Lopez-Poveda et al., 2017; Mehraei et al., 2016; Prendergast et al., 2016a,b), leading to inconsistent results and contradictory conclusions. Since it is not possible to directly quantify synaptopathy by a count of AN synapses in humans, human studies have to rely on indirect measurements and assumptions about the method used (see *chapter 3* for a more detailed discussion).

Nevertheless, Bharadwaj et al. (2015) reported a weak but statistically significant correlation between EFRs recorded as a function of modulation depth, using sinusoidally amplitude modulated (SAM) tones presented at high suprathreshold levels, and behavioral amplitude modulation detection thresholds. They argued that a loss of medium- and low-spontaneous rate (SR) AN fibers (Furman et al., 2013; Liberman, 1978) may explain the poorer performance in

the amplitude modulation detection task and the reduced EFR amplitudes in individual listeners. Chapter 3 reported individual differences in EFR level-growth functions recorded for shallowly modulated SAM tones in a homogeneous group of young and normal-hearing (NH) threshold (as defined by the audiogram) listeners. The reduction of the EFR magnitudes observed in a subset of the NH threshold listeners could be simulated using a computational model of the AN (Zilany et al., 2009, 2014) when it included synaptopathy. Controversially, this required a significant loss of high-SR fibers in addition to a loss of medium- and low-SR fibers. However, the question whether the individual differences within the NH threshold listeners group reported in *chapter 3* were due to a loss of AN fiber synapses still needs to be answered. On one hand, the computational model provided limited and simplified information regarding the function of the AN based on data from the cat. On the other hand, the recorded EFRs, which were elicited using modulation frequencies around 100 Hz in human listeners where the presence of synaptopathy cannot be controlled, reflected gross synchronized neuronal activity mainly generated at the level of the brainstem. In order to evaluate whether the model simulations and the reduction of the EFR magnitudes reported in *chapter 3* were related to a loss of AN fiber synapses, a similar study showing EFRs in an animal model were the presence of synaptopathy can be controlled and both the model simulations and the recordings represent corresponding neuronal generators was performed.

In the present study, EFR level-growth functions were measured in mice, following the experimental paradigm of *chapter 3*. The measurements were performed by Aravindakshan Parthasarathy, PhD at the Massachusetts Eye and Ear Infirmary - Harvard Medical School (MEEI - HMS), in Boston (MA), USA. Inspired by the study in humans, EFR level-growth function for strongly and shallowly modulated SAM tones were measured in noise-exposed (synaptopathic) and control young mice. As in *chapter 3*, an animal version (cat) of the computational AN model was used to simulate the effect of synaptopathy on the EFR level-growth functions, with focus on the dominance of high-SR fibers and the off-frequency contributions (i.e. away from the characteristic place of the stimulus) at supra-threshold responses. The combination of EFR recordings in animals and computational modeling can help to find common trends in electrophysiological responses in synaptopathic mice and humans to bridge the gap between the different species.

5.2 Methods

5.2.1 Experimental animals

Male 32 week old CBA/CaJ mice (7 control, 9 exposed) were tested in this study. Mice were exposed awake and unrestrained to octave-band noise (8-16 kHz) at 100 dB SPL for 2 hours to produce noise-induced synaptopathy without a permanent elevation of hearing threshold (see also Shaheen et al., 2015)). Control mice were not exposed to synaptopathic noise, and were of the same age, gender and strain, born and raised in the same colony in relatively quiet conditions (Sergeyenko et al., 2013).

5.2.2 Physiological tests

All physiological experiments were conducted in an acoustically and electrically shielded and heated chamber. Specific details of the used instrumentation can be found in Sergeyenko et al. (2013) and at the Engineering Resources website from MEEI - HMS. Mice were anesthetized with ketamine and xylazine and their ears were evaluated under microscopic observation before physiological testing (Sergeyenko et al., 2013). The electrophysiological recordings (ABRs and EFRs) were performed using subdermal needle electrodes at the vertex and ventrolateral to the pinna of the stimulated ear with a ground reference electrode at the base of the tail (sampling rate $f_s = 25$ kHz, 24 bits resolution).

ABRs were recorded using 5 ms tone pips with 0.5 ms $\cos^2$ rise-fall presented at a rate of 30 per second in alternating polarity at log-spaced frequencies from 5.6 to 45.2 kHz using subdermal needle electrodes. Stimuli were presented from 25 to 90 dB sound pressure level (SPL) in steps of 5 dB to obtain the ABR level-growth functions. The recorded ABRs were amplified (10000X), band-pass filtered (0.3-3 kHz) and ensemble-averaged (1024 samples/level) to increase the signal-to-noise ratio (SNR). ABR waveforms were analyzed offline and identified and quantified (peak to following through, p-p) by visual inspection. ABR wave-I threshold was defined as the lowest stimulus level (in dB SPL) at which a wave-I could be identified.

DPOAEs at the distortion product $2f_1 - f_2$ were recorded in response to two primary tones (f_1 and f_2) presented at a ratio of $f_2/f_1 = 1.2$. The frequency of f_2 was the same as the test frequencies used in the ABR measurements, and its stimulus level was 10 dB below the f_1 level. DPOAE level-growth functions were

obtained at f_2 stimulus levels from 20 to 80 dB SPL, in steps of 5 dB. The DPOAE level was defined as the amplitude at the frequency bin corresponding to the frequency $2f_1 - f_2$ in the recorded averaged spectrum. The DPOAE threshold was defined as the f_2 primary level (in dB SPL) required to elicit a criterion DPOAE of -5 dB SPL at each test frequency.

EFRs were obtained using 200 ms long SAM[a] tones with 5 ms $\cos^2$ ramps presented 200 times at a repetition rate of 3.1/s. Two carrier frequencies (presented serially, not simultaneously) were used: 12.14 kHz (non-synaptopathic frequency) or 30.5 kHz (synaptopathic frequency). The modulation frequency was $f_m = 1024$ Hz. Two modulation depths were used: $m = 85\%$ (referred here to as strongly modulated tones) and $m = 25\%$ (referred to as shallowly modulated tones). Stimuli were presented at levels ranging from 20 to 80 dB SPL in steps of 5 dB to obtain the EFR level-growth functions. The recorded EFRs were analyzed offline (see Encina-Llamas et al., under review[b], for details). In short, the 200 repetitions were ensemble-averaged and band-pass filtered between 800 and 2000 Hz (30^{th}-order Butterworth). The averaged time-domain waveforms were transformed into the frequency domain (FFT) and the EFR magnitude (in dB re 1μV) was obtained from the bin of the spectrum corresponding to f_m. To ensure a minimum SNR, the EFR magnitude was compared to an estimate of the background electroencephalographic (EEG) noise in the range of 120 Hz below and above f_m (46 bins). The amplitude was defined significant if $p \leq 0.01$ was achieved in an F-test statistical measure with 2, 92 degrees of freedom (Dobie and Wilson, 1996; Picton et al., 2003).

5.2.3 AN model

A phenomenological model of the cat AN, implemented in MATLAB (The MathWorks, Inc., Natick, Massachusetts, USA), was used to simulate the activity of the different AN fiber types (Zilany et al., 2009, 2014). A similar implementation of the model as reported in *chapter 3* was used, but with some modifications to adapt it to the cat (e.g. distribution of AN fibers per IHC. See the Methods sections in *chapter 3*). Since the model was applied to simulate data from mice, the simulation results need to be discussed based on the assumptions underlying the specific choice of parameters.

[a] $A \cdot \sin\left(2\pi f_c\, t\right) \cdot \left(\frac{1+m \cdot \sin\left(2\pi f_m\, t\right)}{2}\right)$.

The model fibers were tuned to 200 characteristic frequencies (CF) ranging from 0.6 to 40 kHz, corresponding to equally-spaced positions on the basilar membrane (BM) according to the cochlear frequency map for cat (Greenwood, 1990; Liberman, 1982). A non-uniform distribution of AN fibers per CF (or IHC) was implemented according to the distribution reported in Liberman et al. (1990). Since only 200 CFs were simulated, the number of nerve fibers was multiplied by a factor of 10 to obtain a total number of about 50000 for the healthy auditory system (Liberman et al., 1990). At medium frequencies where the density was highest, about 280 AN fibers synapses were independently computed at each CF. Synaptopathy was simulated by computing a lower number of AN fibers at each CF. Frequency-specific synaptic loss was implemented by defining a given percentage of loss of fibers at single CFs, and interpolating the remaining number of synapses using a shape-preserving piecewise cubic Hermite interpolating polynomial evaluated over the complete distribution of AN fibers along the CF vector. The ensemble of AN fibers were subdivided into three types based on their spontaneous rate (SR) firing in quiet as reported in Liberman (1978), resulting in 61% high-SR fibers, 23% medium-SR fibers 16% low-SR fibers.

To link the EFRs measured in mice to the simulated EFRs in the model of cat AN, the stimulus frequencies were adapted for the simulations by matching the place along the cochlear across the two species. The SAM carrier frequencies used in the EFR recordings in mice were at 12.14 kHz (non-synaptopathic frequency) and 30.5 kHz (synaptopathic frequency), which corresponds to a distance of 1.85 and 3.55 mm from the base of the mouse cochlea respectively (with a total cochlear length of 5.13 mm) (Müller et al., 2005). Transformed to the same relative positions along the cat cochlea (9.00 and 17.34 mm from the base, for a total length of 25 mm), the carrier frequencies were set to 2.28 and 12.93 kHz, respectively (Greenwood, 1990; Liberman, 1982). The modulation frequency were also transform to each relative cochlear positions, resulting in $f_m = 392$ Hz at the non-synaptopathic frequency and $f_m = 782$ Hz at the synaptopathic frequency. The same two modulation depths of $m = 85\%$ and $m = 25\%$ were used. The stimuli had levels between 10 and 100 dB SPL in steps of 5 dB and a total duration of 1.2 s.

The simulated synaptic responses were processed as described in *chapter 3*, from which the EFR were extracted, and a similar analysis in 1/3-octave frequency bands was performed to investigate the on- versus off-frequency contributions

and the relative contribution of the different types of AN to the total simulated EFR. The on-frequency band for the synaptopathic frequency stimulation was centered at 2.28 kHz, and the off-frequency bands were centered at 3.6 and 10 kHz. For the non-synaptopathic frequency, the on-frequency band was centered at 12.93 kHz, and the off-frequency bands were centered at 15 and 21 kHz.

5.3 Results

5.3.1 Validating noise-induced synaptopathy in mice

Figure 5.1 shows the physiological measures recorded on the noise exposed and non-exposed mice to validate the presence of synaptopathy in the noise exposed animals. The black symbols and lines show the results obtained from the control animals, and the gray symbols and lines show the results from the noise exposed animals. Panels A and B show the minimum stimulus level in dB SPL needed to obtain a detectable DPOAE or ABR wave-I, respectively, at different frequencies (the DPOAE or ABR thresholds). Panels C-F show DPOAE or ABR wave-I amplitude recorded as a functions of stimulus level. Panels C and D show the DPOAE and the ABR wave-I level-growth functions at the non-synaptopathic frequency (12.1 kHz), respectively. Panels E and F show the DPOAE and ABR wave-I level-growth functions at the synaptopathic frequency (30.5 kHz), respectively.

All measures were similar for the exposed and the non-exposed mice (panels A-E), except for the ABR level-growth functions at the synaptopathic frequency (panel F). All animals show no loss of sensitivity in either of the measures, indicating no permanent elevation of hearing threshold. The ABR amplitudes of the exposed mice at 30.5 kHz showed however a reduction at medium-to-high stimulation levels. These results are consistent with the results reported in the noise-induced synaptopathy literature in mice (e.g.; Fernandez et al., 2015; Furman et al., 2013; Kujawa and Liberman, 2009; Kujawa and Liberman, 2015).

5.3.2 EFR level-growth functions in exposed versus non-exposed mice

Figure 5.2 shows the group-averaged EFR level-growth functions recorded at the non-synaptopathic (panel A, $f_c = 12.1$ kHz) and at the synaptopathic frequency (panel B, $f_c = 30.5$ kHz) for the noise-exposed (circles, solid lines) and control (squares, dashed lines) mice. The EFR magnitudes are shown in dB relative to

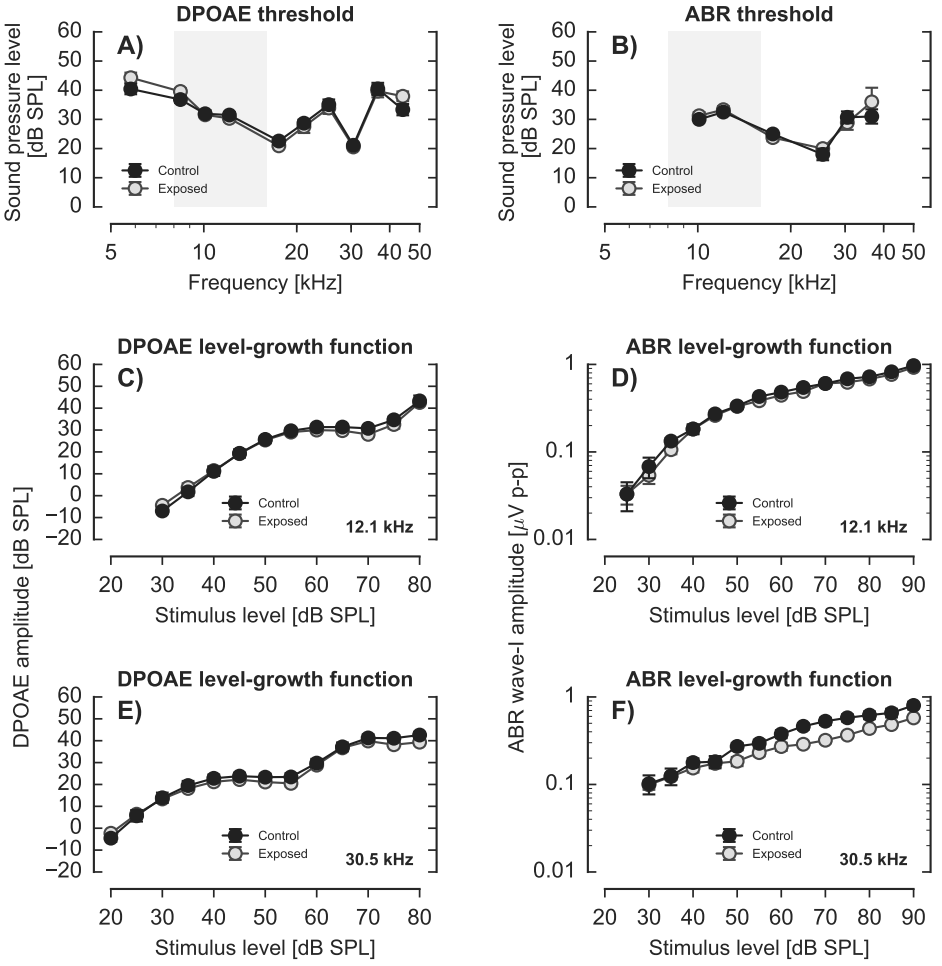


Figure 5.1: DPOAE and ABR wave-I thresholds and DPOAE and ABR wave-I level-growth functions recorded in exposed (black symbols and lines) and non-exposed (gray symbols and lines) mice. Panels A) and B) show the DPOAE threshold and the ABR wave-I thresholds at different frequencies. The gray shaded area indicate the bandwidth of the noise used to induce the synaptic loss. The panel C shows the DPOAE level-growth functions at 12.1 kHz (non-synaptopathic frequency) and the panel D shows the ABR wave-I level-growth functions at the same frequency expressed in μV peak to following through, p-p. The panel E and D show the DPOAE and the ABR level-growth functions at 30.5 kHz (synaptopathic frequency) respectively. Data are means ± standard error of the mean (SEM).

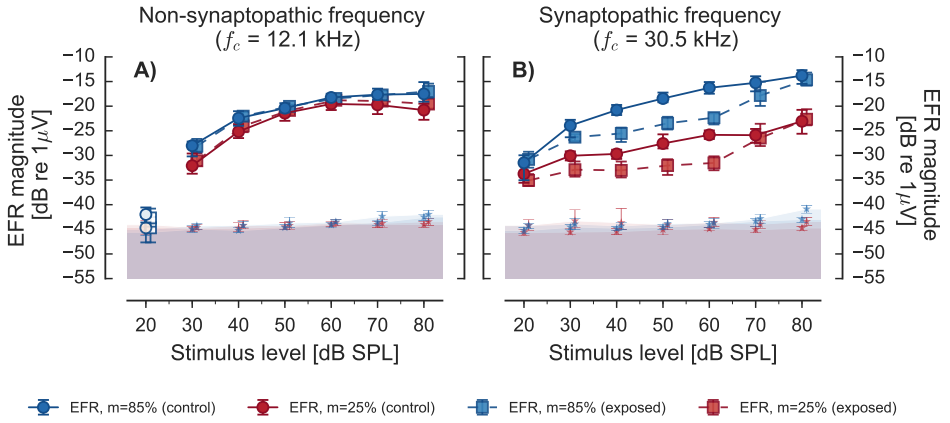


Figure 5.2: EFR level-growth function recorded at the non-synaptopathic (Panel A), $f_c = 12.1$ kHz) and synaptopathic (Panel B), $f_c = 30.5$ kHz) frequencies for exposed (circles, solid lines) and non-exposed (squares, dashed lines) mice using strongly (blue) and shallowly (red) modulated tones. EFR magnitudes in dB relative to $1\mu V$ are solid in case of a statistically significant response (positive F-test) or open in case of statistically non-significant (negative F-test) responses. EEG background noises estimates for each modulation depth are shown as as asterisks and shaded areas with consistent color labeling.

1µV for the strongly modulated (blue) and for the shallowly modulated stimuli (red). Filled symbols represent statistically significant (positive F-test), and open symbols represent non-significant (negative F-test) EFRs magnitudes. The error bars on the markers show 95% confidence intervals. The estimated EEG background noises for each modulation depth are depicted as blue asterisks (strongly modulated) or red asterisks (shallowly modulated) and shaded areas of corresponding colors.

Group-averaged EFR level-growth functions at the non-synaptopathic frequency (panel A) grew monotonically with increasing stimulation level for the strongly modulated tones. For the shallowly modulated tones, the EFR level-growth functions grew monotonically up to 60 dB SPL, and saturated and slightly decreased at higher stimulus levels. The EFR magnitudes for strongly and shallowly modulated tones were very similar, with slightly higher EFR amplitudes for the strongly modulated tones. For each modulation depth, the EFR level-growth functions for the exposed mice (squares, dashed lines) were virtually identical to that for the non-exposed (circles, solid lines) mice.

At the synaptopathic frequency (panel B), the EFR level-growth functions for the exposed mice (squares, dashed lines) showed a different trend than the EFR level-growth functions for the non-exposed (circles, solid lines) mice. For the

non-exposed animals, the EFR level-growth functions grew monotonically with stimulus intensity for both modulation depths, similar to the non-synaptopathic frequency. The EFR magnitudes for the strongly modulated tones (blue circles, solid line) were higher than for the shallowly modulated tones (red circles, solid line). In the exposed mice, the EFR magnitudes were reduced with respect to the non-exposed EFR level-growth functions at medium stimulation levels (approximately from 40 to 70 dB SPL) with a very shallow slope. At the lowest and highest stimulus levels, the EFR magnitudes for the exposed mice were very similar to the those for the non-exposed animals. The reduction of the EFR magnitudes at medium stimulus levels affected the EFRs for both modulation depths. These results are consistent with the EFR level-growth functions for 100% modulated SAM tones reported in Shaheen et al. (2015).

5.3.3 Simulated EFR level-growth functions in non-exposed animals

Figures 5.3 and 5.4 show simulated EFR level-growth functions at a frequency corresponding to the "non-synaptopathic" ($f_c = 2.28$ kHz, corresponding to 12.1 kHz in mice) and the "synaptopathic" ($f_c = 12.93$ kHz, corresponding to 30.5 kHz in mice) frequencies, but without including a loss of synapses. The EFR magnitudes are expressed in arbitrary units (a.u.) in dB. The columns show the EFR level-growth functions at different 1/3-octave bands. Panels A1-A4 show the synaptic output summed across all the CFs, panels B1-B4 show the output at the on-frequency band, and panels C1-C4 and D1-D4 show the output of two off-frequency bands. The vertical axis on the right of the figure (gray) correspond to columns B-D (which are smaller than the EFR magnitudes in column A since they only show a band-limited response). The rows show the EFR level-growth functions for the different types of SR fibers (A1-D1 for high-SR, B2-D2 for medium-SR, and A3-D3 for low-SR fibers), and the sum of all three types of SR fibers (A4-D4). The un-shaded areas mark the stimulus level range recorded in the mice.

The simulated EFR level-growth functions showed similar patterns as the model simulations reported in *chapter 3* for human listeners. The simulated EFR level-growth functions showed different trends when summed across CF bands (columns A) compared to when analyzed in single CF bands (columns B-D). The EFR level-growth functions summed across the CF bands grew monotonically with increasing stimulation level, whereas the functions in one CF band showed non-monotonic growth with level. In all SR fiber types and frequency bands of

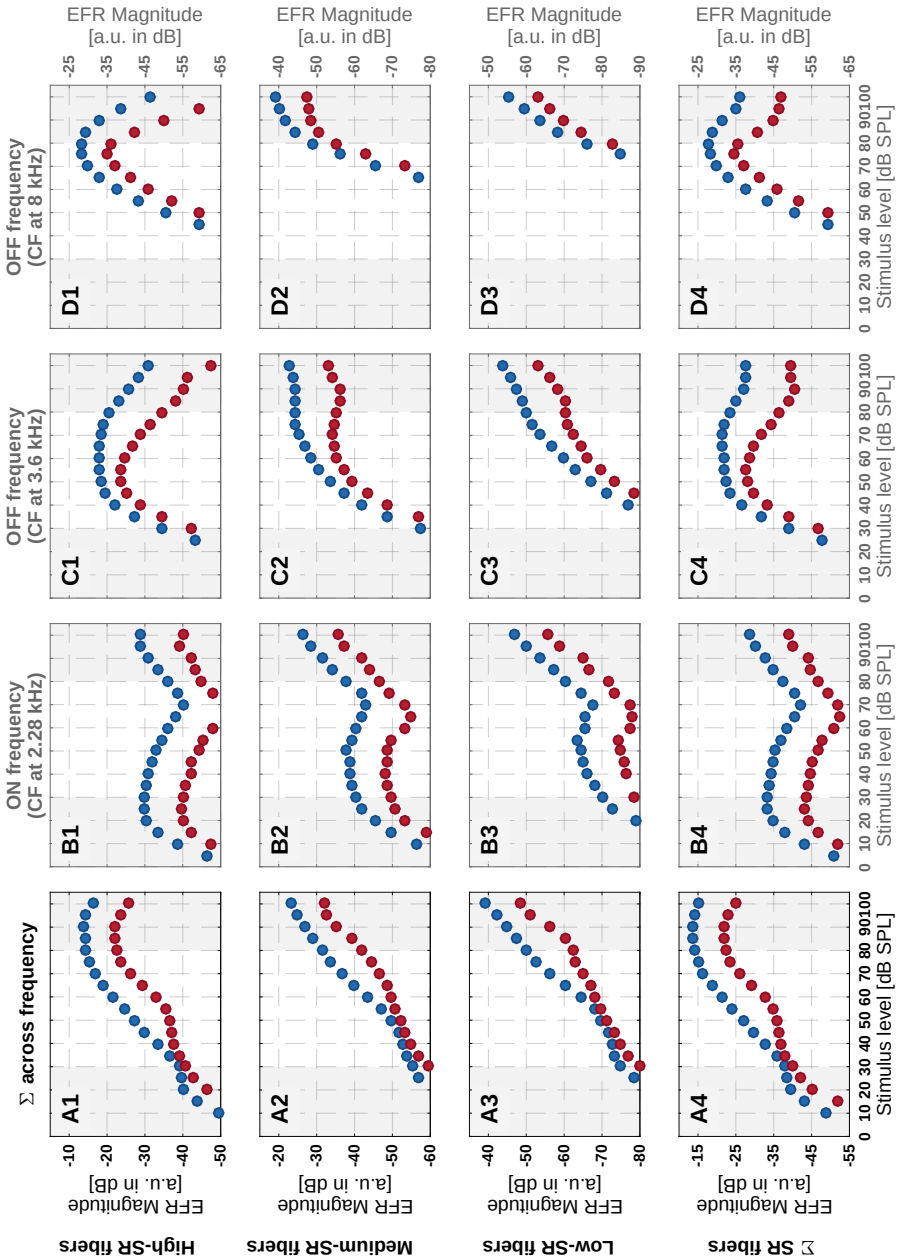


Figure 5.3: Simulated EFR level-growth functions at the non-synaptopathic frequency using the AN model of the cat without including a loss of synapses. Solid circles represent EFR magnitudes in arbitrary units in dB. Blue markers show responses for the strongly modulated stimuli and red markers for the shallowly modulated stimuli. The columns show the EFR level-growth functions at different CF bands (on- and off-frequency bands) and the rows show the simulated results at the different types of fiber of the AN (see the text for a detailed description). The un-shaded areas indicate the level range used in the EFR recordings in mice shown in Figure 5.2

analysis, the simulated EFR magnitudes for the strongly modulated tones were always higher than the EFRs for the shallowly modulated tones.

At the non-synaptopathic frequency (Fig. 5.3), the EFR level-growth function of the on-frequency CF band (panel B4) initially grew steeply at low stimulus levels with a peak around 25 dB SPL, then decreased up to about 60-70 dB SPL, and then grew again for higher stimulus levels. Only looking at the high-SR fibers, the simulated level-growth functions (panel B1) were very similar to the one summed across fiber type (panel B4). The responses for the medium- and low-SR fibers showed a similar trend but smoother and with much smaller EFR magnitudes (note differently scaled axes). The similarity of the simulated EFR level-growth functions for the high-SR fibers (panel B1) and summed across CF bands (panel B4) indicate that the medium- and low-SR fibers only contribute little to the summed response. At off-frequency CF bands (Fig. 5.3, columns C and D), the response was similar to the on-frequency CF band, but shifted horizontally on the stimulus level axis. The magnitudes in the off-frequency CF bands were higher, due to the higher density of AN fibers at those bands (higher number of synapses per IHC) (Liberman et al., 1990).

At the synaptopathic frequency (Fig. 5.4), the EFR level-growth functions showed a similar behavior as at the non-synaptopathic frequency, but with some qualitative differences. When summed across CF bands, the simulated EFR at the synaptopathic frequency showed a more strictly monotonic growth, compared to the non-synaptopathic frequency (panels A4). In the on-frequency CF band (B4), the summed response was dominated by the high-SR fibers (as in the non-synaptopathic frequency), but only up to levels of about 70 dB after which the response of the high-SR fibers decreases while the response of the medium- and high SR fibers increased.

5.3.4 Simulated EFR level-growth functions in exposed animals

Figure 5.5 shows the simulated EFR level-growth functions at the non-synaptopathic (Panel A) and the synaptopathic (Panel B) frequencies, including an implemen-

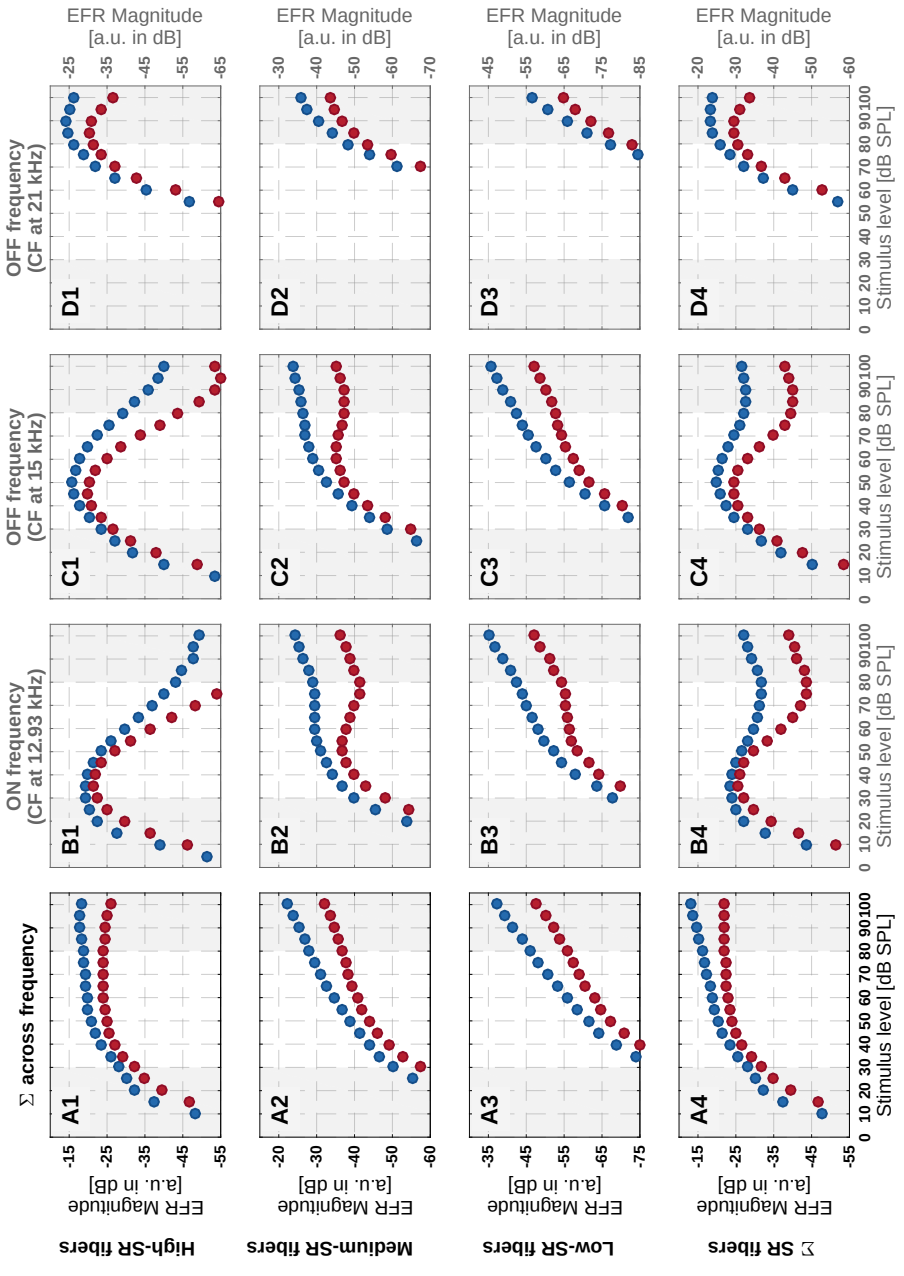


Figure 5.4: Same as Figure 5.3, but at the synaptopathic frequency using the AN model of the cat.

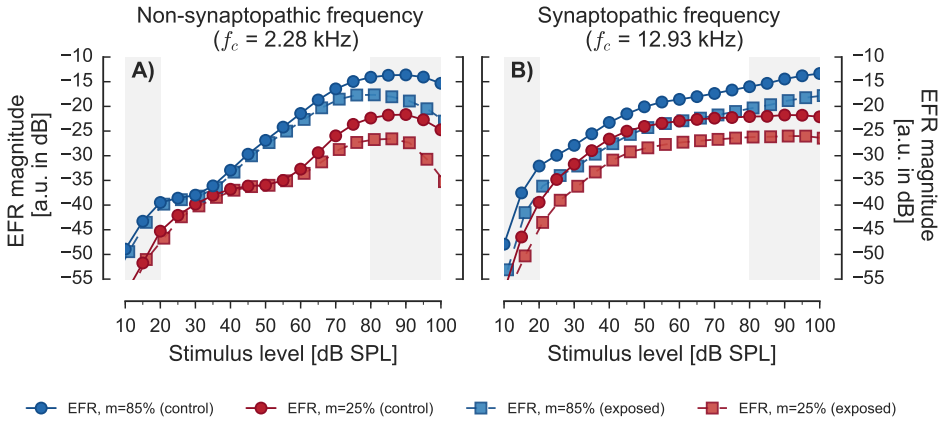


Figure 5.5: Simulated AN EFR level-growth function at the synaptopathic (Panel A), ($f_c = 2.28$ kHz) and non-synaptopathic (Panel B), $f_c = 12.93$ kHz) frequencies in exposed (circles, solid lines) and non-exposed mice (squares, dashed lines) using strongly (blue) and shallowly (red) modulated tones. The EFR magnitudes are expressed in dB in arbitrary units (a.u.). The un-shaded areas indicate the level range used in the EFR recordings in mice shown in Figure 5.2.

tation of synaptopathy [b]. The panels show the model response summed across CFs and SR fiber type (same as in panel A4 in Figs. 5.3 and 5.4). The simulated EFR level-growth functions of the "non-exposed", healthy model (without synaptopathy) are shown as circles and solid lines, while the EFR level-growth functions of the "exposed", synaptopathic model are shown as squares and dashed lines. The un-shaded areas indicate the level range used in the EFR recordings in mice shown in Figure 5.2.

Consistent with the model simulations reported in *chapter 3*, a selective loss of medium- and low-SR fibers did not produce significant differences between the non-synaptopathic (non-exposed) and synaptopathic (exposed) simulated EFR level-growth functions (results not shown). Therefore, the loss of synapses was equally included in the model to all three types of fibers. The frequencies used in the computational model were adapted from the mouse to the cat cochlea (Greenwood, 1990; Liberman, 1982).

[b] At the time of writing this manuscript, the histological analysis and the synapse counting was not finished. However, the type of mice (CBA/CaJ) and age (32 weeks old), and the type of noise (8-16 kHz), the level (100 dB SPL) and the duration (2 hours) at which they were exposed, were identical to what used in Fernandez et al. (2015). Therefore, the percentage of loss of synapses reported in Fernandez et al. (2015) was used in the model simulations in the present study assuming that the mice used here might show a similar loss of AN fiber synapses. Table 5.1 shows the fiber loss in percentage per frequency implemented in the model to simulate synaptopathy.

Table 5.1: Implemented percentage of all three types of AN fiber loss at different CFs in the model simulation of synaptopathy.

Frequency in the mouse [kHz]	5.6	8	11.3	16	22	32	45	64
Frequency in the cat [kHz]	0.31	0.91	1.97	3.93	7.14	14.1	25.9	48.3
Loss of synapses [%]	0	5	4	4	31	39	47	38

The simulated EFR level-growth functions at the non-synaptopathic frequency ($f_c = 2.28$ kHz, Fig. 5.5, A) after including a loss of synapses (squares, dashed lines) were very similar to the EFRs from the healthy simulation (circles, solid lines). At the higher stimulus levels, the simulated EFR magnitudes of the synaptopathic model were reduced relative to the response of the healthy model. At the synaptopathic frequency ($f_c = 12.93$ kHz, Fig. 5.5, B), the EFR level-growth functions after including the loss of synapses in the model (squares, dashed lines) were reduced in the whole stimulus level range from respect to the healthy model simulation (circles, solid lines).

5.4 Discussion

5.4.1 EFR level-growth functions in exposed and non-exposed animals

The EFR level-growth functions recorded in mice were lower in magnitude for shallowly modulated tones than for strongly modulated tones. This was the cased at both the non-synaptopathic and the synaptopathic frequency, for both groups of mice. This reduction in EFR magnitude with modulation depth is consistent with previous literature on EFR in humans (Lins et al., 1995; Picton et al., 2003). The effect is larger for the higher (synaptopathic) frequency than for the lower (non-synaptopathic) frequency. This, presumably, is due to differences in frequency selectivity at these frequencies because while the hearing thresholds of the mouse are similar at both frequencies, the AN tuning evaluated as $Q_{10\ \mathrm{dB}}$ value is larger at the higher frequency (Taberner and Liberman, 2005). When comparing the non-exposed to the exposed animals, no difference was found at the non-synaptopathic frequency, while the exposed mice showed reduced EFR magnitudes at medium-to-high stimulus levels. Since the exposed animals did not show any sign of outer hair-cell loss (see Fig. 5.1), this effect is not due to changes in sensitivity (Kujawa and Liberman, 2009). The reduction

in EFR magnitude in the synaptopathic frequency is consistent with previous studies measuring EFR in mice using SAM tones with a modulation depth of 100% (Shaheen et al., 2015). In the light of AN rate-level functions, and assuming that shallow modulations at high intensities are encoded by medium- and low-SR fibers, the consistent reduction of EFR amplitude at the synaptopathic frequency for both modulation depths does not match the interpretation that only medium- and low-SR fibers located around that CF of the carrier frequency of the SAM tone (on-frequency) encode the intensity fluctuations (but not the high-SR fibers because its firing rate is saturated). This would result in a different reduction in EFR magnitude for the strongly versus shallowly modulated tones.

Different trends in individual EFR level-growth functions measured in a group of young NH threshold human listeners were reported in *chapter 3*. Some listeners showed, consistent with the results in mice, a reduction in EFR magnitudes at medium stimulus intensities. The reduction in the human results was however only for the shallowly modulated tones and not for the strongly modulated tones, which is inconsistent with the mice data. Since all listeners were normal hearing according to the audiogram and the state of the AN synapses is not directly accessible in humans, the interpretation of these results depend on the assumption that synaptopathy exists in human listeners and that it can be measured using the methods applied (see *chapters 3* and *6* for a detailed discussion). In addition, the EFR recorded in mice were elicited by SAM tones with high modulation frequencies ($f_m = 1024$ Hz), believed to be generated between the level of the AN and the cochlear nucleus (Shaheen et al., 2015). In contrast, the human data was obtained with lower modulation frequencies ($f_m = 93$ Hz), believed to arise mainly from brainstem generators (Herdman et al., 2002a; Kuwada et al., 2002). Such difference in the neuronal sources generating the EFR in relation with additional processing in the brainstem (i.e. central gain mechanisms, Chambers et al., 2016; Möhrle et al., 2016), could explain the observed differences between the mice and the human results.

5.4.2 Simulated EFR level-growth functions

The model results of the present study are in qualitative agreement (at least in terms of the differences seen) with the EFRs recorded from non-exposed and exposed mice at the synaptopathic and non-synaptopathic frequencies. The overall shape of the level-growth functions were different, but the similarity

at the non-synaptopathic frequency and the differences at the synaptopathic frequency between exposed and non-exposed animals could well be accounted for. Since the model used was based on parameters for the AN of the cat, the transposition of the stimulus frequencies and the similarity in tuning and AN fiber density and distribution are important factors that require consideration when interpreting the simulated results.

The analysis of the model simulations in *chapter 3* let to two key observations: 1) the off-frequency neural activity dominated the total EFR magnitude at medium-to-high stimulation levels; and 2) the activity of high-SR fibers always dominated the total EFR magnitude at all stimulus levels. This is consistent with the modeling analysis from the present study (see Fig. 5.3 and 5.4). In the present study, the simulated EFR level-growth functions summed across CFs grew monotonically whereas the EFR level-growth functions at the on- and off-frequency CF bands showed non-monotonic growth, and the simulated EFR magnitudes for the strongly modulated SAM tones were always higher than the EFR magnitudes for the shallowly modulated tones.

Although the general trends of the simulated EFR level-growth functions at the non-synaptopathic frequency in mice (Fig 5.3), at the synaptopathic frequency in mice (Fig. 5.4) and at the tested frequency in NH threshold human listeners (Fig. 3.6 in *chapter 3*) are consistent, slightly different patterns could be observed. For example, comparing the simulated EFR magnitudes resulting from the sum across CFs and fiber type (panel A4), the EFR level-growth functions grew monotonically with a single linear slope in the human simulations (Fig. 3.6, A4 in *chapter 3*), whereas the EFR magnitudes grew monotonically but in a more "curved" manner at the non-synaptopathic frequency in the mice simulations (Fig 5.3, A4). The EFR functions followed approximately a logarithmic growth at the synaptopathic frequency in the mice simulations (Fig 5.4, A4). Different patterns could also be observed in the simulated EFR level-growth functions at the on- and off-frequency bands.

The core of the computational model was essentially the same in the present study (cat model) and in *chapter 3* (humanized model). As the total EFR magnitude, in particular at high stimulation levels, depended on the recruitment of neurons tuned to higher frequencies (off-frequency contributions), the precise tuning properties of those neurons determine and strongly influence the pattern of the simulated EFR level-growth function. Similarly, the local tuning properties at a particular on- and off-frequency bands might determine the

simulated EFR level-growth functions at such frequency band.

Taberner and Liberman (2005) reported AN tuning properties expressed as mean $Q_{10\ \mathrm{dB}}$ values in 5 mammalian species, including mouse and cat, showing a very consistent tuning properties across species at a particular CF. Therefore, assuming that the AN tuning properties define the pattern of the EFR level-growth functions, the simulated EFR magnitudes using SAM tones with a given carrier frequency in the cat model would be closer to the recorded EFRs in mice at the same frequency. However, according to the model analysis, the off-frequency contributions dominated the response at high stimulation levels. Knowing that the maximum frequency in the cat model was 40 kHz, the simulated EFR level-growth functions obtained executing the cat model at the synaptopathic frequency used in the mice recordings ($f_c = 30.5$ kHz) produced inconsistent results (not shown) due to insufficient off-frequency contributions. However, interestingly, the simulated EFR level-growth functions at the synaptopathic frequency (Fig. 5.4, A4) were in good agreement with the recorded EFR level-growth functions at the non-synaptopathic frequency (Fig. 5.2, A), both for the strongly and shallowly modulated responses. The carrier frequency at the non-synaptopathic frequency in the recordings in mice ($f_c = 12.1$ kHz) was coincidentally very close to the carrier frequency at the synaptopathic frequency in the simulations from the cat model ($f_c = 12.93$ kHz), which reinforce the idea that the correct AN tuning properties might lead to more accurate model simulations.

The challenge of using a computational model based on the cat AN to simulate mice data exposed above led to a disagreement already between the healthy model simulations and the recorded EFRs in non-exposed animals. Consequently, the model simulations after imposing a loss of AN fiber synapses did not agree either (Fig. 5.5). However, although the patterns in the EFR level-growth functions were in disagreement, the general effect of synaptopathy at the non-synaptopathic versus the synaptopathic frequencies was consistent with what observed in the recorded EFRs (Fig. 5.2). At the non-synaptopathic frequencies, the simulated EFR level-growth functions after imposing the loss of synapses shown in Table 5.1 were almost identical to the EFR magnitudes obtained from the healthy model, except for the small reduction at the higher stimulus levels. At the synaptopathic frequency, the model simulation showed a reduction of the EFR magnitudes for both modulation depths when a loss of AN fiber synapses is included, consistent with the EFRs recorded in mice (Fig. 5.2,

B). However, such reduction of the EFR magnitudes in the simulations was also obtained at the higher stimulation levels, in contrast to the recovery at seen at those high stimulus levels in the recorded EFRs. It is believed that adapting the AN model to the mouse, implementing the correct AN tuning and CF range, would lead to a more accurate model simulations.

5.4.3 On the need of a species specific computational model to study synaptopathy

A major challenge when studying synaptopathy, or other peripheral cellular damage in the auditory system (i.e. IHC vs OHC loss), consists on the impossibility of directly assess neuronal or hair-cell damage, in contrast to the studies performed in non-human animals. In order to face this challenge, similar investigations performed in parallel in human listeners and non-human mammals would be desirable. Although behavioral studies are feasible in some non-human animals for particular tasks, they are generally very time consuming and costly. On the other hand, objective physiological methods allow for a more efficient inter-species comparison of peripheral auditory processing. Electrophysiological evoked responses are generated by the synchronized activity of multiple neuronal sources, difficulting its interpretation. Moreover, the surrounding anatomy to the neural generators is morphologically different from species to species, leading to significant variations of the dominating sources on a given surface electrode configuration. Computational modeling might be a good tool to analyze and interpret such physiological responses and transfer knowledge across species. However, accurate computational models for each particular species are needed to satisfactory compare the results. For example, in the work documented here, although the general trends obtained from the EFR level-growth functions recorded in mice could be explained by a computational model of the cat AN, the simulations were not accurate enough.

It is discussed above that both the AN tuning properties and the available frequency range of off-frequency contributions might play an important role when simulating EFRs at the level of the AN. Therefore, as the mouse is to date the best characterized animal model to study synaptopathy (e.g., Chambers et al., 2016; Fernandez et al., 2015; Furman et al., 2013; Jensen et al., 2015; Kujawa and Liberman, 2009; Kujawa and Liberman, 2015; Sergeyenko et al., 2013; Shaheen et al., 2015; Valero et al., 2017), we are in need of an accurate mouse AN

computational model. One could assume that the phenomenological properties of the cochlear filtering, the IHC transduction and the AN adaptation might be comparable in cats and mice, and thus a simple adaptation of the middle ear filtering (Dong et al., 2013; Saunders and Summers, 1982), AN tuning properties (Taberner and Liberman, 2005) and the correct frequency range (Müller et al., 2005) to the mouse would lead to a more accurate outcome. Besides this, the computational model simulations from both the present chapter and in *chapter 3* led to the conclusion that the high-SR fibers dominate the encoding of the stimulus envelope (i.e. the EFR) at the level of the AN. This is due to the distribution of the three different AN fiber types, which were implemented in the model as 61% of high-SR fibers versus a 39% of medium- and low-SR fibers based on the cat data from Liberman (1978). However, Taberner and Liberman (2005) reported that the distribution of AN fiber types in CBA/CaJ mice is not bi-modal as in cats and that the high-SR fibers represented a 49% of the total AN population, lower than the cat. Therefore, after incorporating those changes to a mouse version of the AN computer model, it would be expected that a preferential loss of medium- and low-SR fibers would have an effect onto the simulated EFRs due to the more equal distribution of AN fiber types in mice. This could potentially reconcile the contradiction between the mice recordings presented in this study (Fig. 5.2) and the preferential loss of medium- and low-SR fibers suggested in Furman et al. (2013).

5.5 Conclusion

The EFR level-growth functions recorded in noise exposed and non-exposed mice were larger for strongly modulated tones than for shallowly modulated tones. There was no difference in the EFR level-growth functions recorded in noise exposed and non-exposed mice at the frequency where there was not loss of synapses (non-synaptopathic frequency). However, at the synaptopathic frequency, the EFR magnitudes in the exposed animals were reduced from respect to the responses of the control animal group at medium stimulation levels, affecting equally to the responses for both strongly and shallowly modulated tones. At the lower and higher stimulation levels, the EFR magnitudes in both mice groups at the synaptopathic frequency were similar.

A computational model based on AN data from the cat (Zilany et al., 2009, 2014) was used to simulate the EFR level-growth functions obtained from mice.

Overall, the simulations suggested that: 1) the EFR magnitudes are dominated by the high-SR fibers, and 2) the off-frequency neurons increasingly contribute to the EFR with increasing stimulus level. This is consistent with the findings from *chapter 3* where a humanized version of the model was used. However, although the model could account for the general trends obtained from the mice EFR recordings (larger EFR magnitudes for strongly modulated tones) and the overall effect of synaptopathy in the EFRs (reduction of EFR magnitudes affecting equally to both modulation depths), the simulated patterns of the EFR level-growth functions were not in fully agreement with the recorded ones. This might be due to the inaccurate AN tuning characteristics in the model (cat) with respect to the recordings (mice), although the stimuli were adapted from one species to the other. A mouse version of the model on which the correct AN tuning and frequency range would be considered may provide with more accurate results. As the mouse is the best characterized species in the study of synaptopathy, we are in need of an accurate mouse version of the AN model in order to transfer knowledge across species.

Acknowledgments

The mice recordings in this study were done at the Massachusetts Eye and Ear Infirmary, Harvard Medical School in Boston, MA, USA. We want to sincerely thank to Dr. Sharon G. Kujawa for allowing the collaboration between our groups and in particular to Dr. Aravindakshan Parthasarathy for collecting the mice data and for very fruitful discussions regarding synaptopathy, EFRs and AN coding in rodents. This work was supported by the Oticon Centre of Excellence for Hearing and Speech Sciences (CHeSS) at the Technical University of Denmark (DTU). The authors want to acknowledge the support provided by the Erasmus Mundus Student Exchange Network in Auditory Cognitive Neuroscience.

6

General discussion

6.1 Summary

The present thesis describes the potential of envelope following responses (EFR) to evaluate hearing function at supra-threshold stimulus levels. Firstly, EFR level-growth functions were proposed as a tool to estimate compression in the peripheral auditory system. Secondly, they could be used as a potential biomarker for detecting the loss of auditory nerve (AN) fiber synapses (i.e. cochlear synaptopathy). Peripheral compression estimates from EFR level-growth functions were obtained for normal-hearing (NH) and hearing-impaired (HI) listeners. These estimates were compared, in the same individual listeners, to estimates of compression from distortion-product otoacoustic emission (DPOAE) level-growth functions. EFR level-growth functions for strong- and shallow modulation depths were used to investigate the integrity of the AN in individual NH threshold and HI listeners. The HI listeners were tested at an audiometric frequency showing normal sensitivity. Additionally, a computational model of the AN was used to investigate the individual patterns obtained in the EFR level-growth functions. The model analysis focused on the contributions of on- versus off-frequencies (i.e. near or away from the characteristic frequency of the stimulus) to the total EFR. The contributions of the different types of AN fibers (i.e. high-, medium- and low-spontaneous rate (SR) fibers) were also investigated. In order to circumvent the ambiguities when investigating cochlear synaptopathy in humans, EFR level-growth functions were also recorded in noise exposed and non-exposed mice. In this case, synaptopathy can be directly assessed. An animal version of the same computational model was used to interpret the results in mice and relate these to the human data. Finally, the use of an alternative model of the auditory periphery was also considered. The representation of the basilar membrane (BM) motion could potentially be more accurate in this second model, as the BM is modeled as a transmission line versus a filterbank. Unfortunately, the representation of the AN activity

to steady-state stimuli was shown to be physiologically implausible and, even though a suggestion for improvement was made, the use of the model was rejected.

6.2 Estimating peripheral compression using envelope following responses: revisited

Direct invasive measurements of BM velocity-intensity functions in non-human mammals have demonstrated that the compressive growth of the BM is an on-frequency phenomena. It thus does not occur at off-frequency BM locations (i.e. basal and apical places relative to the characteristic place of the stimulus), where the BM growth is linear (Robles and Ruggero, 2001; Ruggero et al., 1997). Any non-invasive method aiming to accurately estimate the same cochlear compression as that obtained through invasive techniques must ensure to assesses such on-frequency processing only. In the case that significant off-frequency contributions are present in the total response of such given method, a combination of on-frequency (compressive) and off-frequency (linear) sources would be represented in the metric. Thus, the response would not represent the same on-frequency BM processing, compromising the accuracy of the metric due to the spurious response. One of the main observations from the model analysis described in *chapter 3* led to the conclusion that EFRs recorded using medium-to-high stimulus levels might not only be contaminated but completely dominated by off-frequency contributions. Therefore, according to the model simulations, EFR level-growth functions may not assess on-frequency processing exclusively. Consequently, the compressive slope derived from the fit on the EFR level-growth functions might not be related to the compressive growth of the BM found in invasive recordings in non-human mammals.

It was discussed in *chapter 2* that the concerns regarding the on- versus off-frequency contributions to the compression estimates could be expanded to other indirect methods, such as the psychoacoustical estimations of BM input/output (I/O) functions and the DPOAE level-growth functions. In short, the generation of DPOAEs is commonly simplified as a single source at the peak of the traveling wave envelope of the higher frequency primary (f_2), whereas more realistically it is produced by the combination of many distortion sources in the region where the traveling waves of the two primaries overlap (Shera, 2004). With regard to behavioral methods, on- and off-frequency maskers presented at

high stimulation levels are used in forward-masking paradigms to estimate BM I/O functions. These methods rely on strong assumptions. First, it is generally assumed that the off-frequency activity can be ignored by presenting noise maskers at moderate intensities at frequency bands away from the region of interest. Second, it is assumed that it is possible to independently assess purely on- and off-frequency BM processing to derive, for instance, on- and off- temporal masking curves. This assumes that the conceptual idea of a "perceptual filter" corresponds to a specific place where the BM is excited. This is clearly not true, as the perceptual representation of, for instance, a pure tone presented in the presence of low-pass noise used to psychoacoustically derive an "auditory filter" will change due to "off-frequency listening" relative to the Δf between the noise band skirt and the tone (Moore, 2001; Patterson, 1974). However, the excitation pattern that the tone produces on the BM is the same (or perhaps slightly suppressed) regardless of the position of the noise band. Therefore, the conceptual idea of an "auditory filter" does not correspond to a fixed place on the BM, because the "auditory filter" can shift but the BM excitation pattern is fixed in place. Third, it is assumed that the BM I/O function can be derived by a simple linear subtraction of those curves. All in all, the computational model analysis suggested that the interpretation of the processing of the peripheral auditory system excited by high intensity stimuli must be considered with caution.

Nevertheless, even though the model simulations suggested a dominance of the off-frequency contributions in the EFR at high stimulus levels, the median values of peripheral compression estimated using EFRs were consistent with estimates of compression obtained from other methods. For instance, the estimates coincide with the slope obtained from invasive recordings of basilar membrane (BM) velocity-intensity functions in non-human mammals (Ruggero et al., 1997), from BM I/O functions derived behaviorally in human listeners (Nelson et al., 2001; Oxenham and Plack, 1997) and from group-averaged DPOAE level-growth functions (Dorn et al., 2001; Neely et al., 2003, 2009). However, the similarity of median values obtained using different methods (i.e. correlation) does not imply that the same underlying mechanism is being assessed (i.e. causality), and therefore caution must be taken when interpreting the results obtained with the different methods.

6.3 On the challenge of investigating the potential presence of cochlear synaptopathy in human listeners

Chapters 3 and *5* investigated the potential use of EFRs to assess synaptopathy in human listeners. Noise-induced synaptopathy has been demonstrated in several non-human mammal species, such as mice (Furman et al., 2013; Kujawa and Liberman, 2009), guinea pigs (Lin et al., 2011; Liu et al., 2012), rats (Lobarinas et al., 2016) and rhesus macaques (Valero et al., 2017). This has been possible through a carefully calibrated and controlled noise exposure, and subsequent post-mortem histological examination of the status of the AN synapses using confocal microscopy techniques. In humans, it is clearly not ethically possible to recreate such noise exposure experiments, carefully controlled to ensure only temporal threshold shifts (TTS) and leading to synaptopathy, tested by subsequent histological examination. Moreover, the window at which noise overexposure produces synaptopathy with a TTS but not a permanent threshold shift (PTS) is very narrow (Liu et al., 2012) and varies across species (Kujawa and Liberman, 2009; Lin et al., 2011; Lobarinas et al., 2016; Valero et al., 2017). Other than a histological examination of post-mortem donated human temporal bones (Makary et al., 2011; Viana et al., 2015), it is also obviously not possible to directly assess the status of AN synapses. Furthermore, with unclear noise exposure histories and a widely heterogeneous population [1], it remains a challenge to find noise-induced synaptopathy in humans. Whereas the boundaries of the window at which noise overexposure induces synaptopathy are very strict in the animal models due to their genetic identity, the greater variability in human listeners might lead to more diffused transitions. Therefore, the noise dose that might induce synaptopathy in humans is unknown and more difficult to control. It is currently unclear: (1) if noise-induced synaptopathy exists in humans; (2) that if it exists, if there are any functional consequences of it; and (3) how a practically reliable diagnostic test could be constructed to separate between synaptopathic and non-synaptopathic listeners.

In order to address the challenge of indirectly assessing the status of AN synapses in humans, different approaches have been adopted. One strategy has been to explore correlations between multiple physiological and/or behavioral tests, alongside an estimate of lifetime noise exposure based on a self-reported

[1] Recall that most animal studies were carried out on a genetically restricted homogeneous population (e.g.; Kujawa and Liberman, 2009)

questionnaire in a large sample of listeners (Grose et al., 2017; Le Prell et al., 2017; Lopez-Poveda et al., 2017; Prendergast et al., 2016a,b). Another approach correlated electrophysiological responses with responses from behavioral measurements in smaller samples of listeners, based on the heuristic interpretation of the role of the different types of AN fibers on encoding supra-threshold sounds and the effect of a preferential loss of medium- and low-SR fibers (Bharadwaj et al., 2015; Mehraei et al., 2016). In addition, Mehraei et al. (2016) also included similar electrophysiological responses in mice, where synaptopathy could be directly quantified. Unfortunately, both approaches have so far led to strongly different conclusions. Bharadwaj et al. (2015) and Mehraei et al. (2016) found individual differences that could be explained by the hypothesized effect of synaptopathy. In contrast, Grose et al. (2017), Le Prell et al. (2017), Lopez-Poveda et al. (2017), and Prendergast et al. (2016a,b) did not find any correlation in any of the tested conditions that could lead to the conclusion that noise-induced synaptopathy was present. As indicated above, it should be considered that susceptibility to noise-induced damage may vary significantly among individuals. This might arise from the differential activation of the olivocochlear efferent system mediated by the expression of specific proteins, which has led to vernacular expressions such as "steel" or "tough" versus "glass" or "tender" ears (Luebke and Foster, 2002). In fact, the middle ear muscle reflex (MEMR) was shown to be attenuated in synaptopathic mice, and has been proposed as a potential metric in the early detection of cochlear synaptopathy. In order to prevail over the challenges discussed here, novel approaches should be considered.

Unique to the studies in this thesis, computational modeling of the AN fiber firing patterns has been used to interpret the individual differences obtained in EFR level-growth functions, both from a homogeneous group of young NH listeners and from mild HI listeners. Our approach allows for a more objective analysis of the potential effect of synaptopathy in the electrophysiological measurements. The interpretation of such effects is not based on the assumed role of different types of AN fiber, but rather on a well-established computational model that is able to account for a large collection of physiological data directly recorded in the AN (Zilany et al., 2009, 2014). Nevertheless, it should be emphasized again that such models are limited by definition, as they are built to represent *certain* aspects of experimental data in the hope to generalize towards other data. Thus, caution must be taken regarding the interpretation of the simulated results. In general, the simulations obtained with computa-

tional models aiming to recreate physiological responses should be evaluated using direct physiological data (from literature). Extraordinary care should be taken when using a "humanized" version of a model that was designed and parameterized using data from another species (in our case the cat), as such evaluation with corresponding physiological data is not possible. For instance, *chapter 4* described the attempt of using the model of Verhulst et al. (2015) to simulate EFRs as a function of stimulus level, the use of which was disregarded since physiologically implausible simulated results were obtained. Additionally, similar EFR level-growth functions were recorded in exposed and unexposed mice [2]. An animal version of the same computational model was used to perform a similar investigation as was done for humans but based on the EFRs recorded in mice. Similar electrophysiological responses recorded in humans and non-human animals, where synaptopathy can be directly assessed, were compared. The computational model was intended to be used as the needed bridge to connect the "clear" findings from non-human animals to the more "arguable" results from humans listeners Nevertheless, the combination of human and non-human animal electrophysiology with computational modeling may help to overcome the challenges we are facing.

Furman et al. (2013) suggested that noise-induced cochlear synaptopathy is selective to medium- and low-SR fibers. The observation from the modeling work described in *chapters 3* and *5* (and similarly in Paul et al., 2016) indicated that a significant loss of high-SR fibers had to be included to affect the simulated EFRs. A simplified and strict interpretation of the results reported by Furman et al. (2013) led to hypotheses regarding the differential role of the different types of AN fibers at supra-threshold stimulus levels (Bharadwaj et al., 2015; Mehraei et al., 2016). Such hypotheses were based on an heuristic interpretation of the AN fibers rate-level curves (e.g.; Joris and Yin, 1992; Liberman, 1978; Sachs et al., 1989; Sokolowski et al., 1989; Yates et al., 1990). High-SR fibers show rate-level functions that rapidly increase in a range of about 20-30 dB above threshold and saturate at higher stimulus levels. Such saturation has been interpreted as if the high-SR fiber's response is the same at medium-to-high stimulus levels, and thus does not encode fluctuations (e.g. envelope encoding). Following this interpretation, assuming that synaptopathy only affects medium- and low-SR fibers and that the high-SR fibers do not encode

[2] The mice data was recorded by Dr. Aravindakshan Parthasarathy at the Massachusetts Eye and Ear Infirmary, Harvard Medical School in Boston, MA (USA)

the envelope at high stimulus levels, the excitation of the auditory system with shallowly modulated tones should selectively assess the activity of medium- and low-SR fibers. The heuristic hypothesis proposed in *chapters 3* of the present thesis to investigate synaptopathy, inspired by previous work, also assumed this interpretation. However, in the light of the model simulations presented in *chapters 3*, *4* and *5*, and supported by direct physiological recordings in the AN of the cat (Joris and Yin, 1992), it seems that the concept of rate has been misinterpreted. Firing rate is defined as the absolute number of spikes per time frame (e.g. one second). However, it does not provide information about how the spikes are distributed within the time frame. For example, consider an AN fiber that generates 100 spikes within one second in response to a 20 Hz tone (sine wave with phase 0 rad)[3]. This fiber has a rate of 100 spikes/s. However, the 100 spikes can either be uniformly distributed within the time frame (one spike every 10 ms) or can have 5 spikes at $t = 50 \cdot n + 12.5$ ms, $n = 0, 1, 2, ...$ (at the positive cycles of the sinusoid). Both distributions result in the same overall rate, but while the former response does not encode the waveform fluctuations, the latter does. A more appropriate metric to understand envelope codification of AN fibers is synchrony. The computer model simulations and physiological responses in AN fibers (Joris and Yin, 1992) are consistent in revealing that synchrony-intensity functions to amplitude modulated tones decrease with increasing stimulus level at supra-threshold levels. Such decrement in level is less pronounced for medium- and low-SR fibers than for high-SR fibers. The physiological responses are obtained from direct recordings in single AN neurons at the characteristic frequency (CF) of the tone, and thus they represent purely on-frequency processing. However, at medium-to-high stimulus levels, the synchrony responses of off-frequency neurons (including the high-SR) are larger than that of the on-frequency response. Therefore, as the modulation is better encoded off-frequency, it is feasible to assume that the system might make use of the information wherever it is better represented. It is interesting to observe that even though this phenomena was already reported in the literature (Joris and Yin, 1992), none of the previously described hypotheses considered it. Joris and Yin (1992) explicitly wrote:

«To understand the response of the population of AN fibers to an AM signal, it is instructive to consider responses to an AM complex with f_c off-CF. [...]

[3] This example pretends to illustrate the point and not to be physiologically accurate. The values are therefore unrealistic.

Except for a horizontal shift along the SPL abscissa, the phase locking to f_m resembles that obtained with f_c at CF. Similar results were obtained in seven other fibers as well as with f_c placed below CF (not shown) in seven fibers.»

Nevertheless, the computer model simulations revealed this observation very clearly. The use of computational models as a refining tool for the generation of novel hypotheses is further discussed in section 6.3.2.

Chapter 3 and *5* reported EFR level-growth functions for strongly (85%) and shallowly (25%) modulated SAM tones recorded from human listeners and mice. Most of the NH threshold listeners showed EFR level-growth functions growing monotonically and in parallel with increasing stimulus level, with larger EFR magnitudes for the strongly than for the shallowly modulated tones. As discussed previously, cochlear synaptopathy cannot be assessed to date in humans, but it is however possible to clearly differentiate between noise-exposed (synaptopathic) and unexposed (non-synaptopathic) mice. Nevertheless, it was observed that some listeners within the same NH threshold group showed a reduction of the EFR magnitudes at medium stimulation levels for shallowly modulated tones. A similar reduction of EFR magnitudes at medium stimulation levels were observed at the synaptopathic frequency in noise-exposed mice. The EFR level-growth functions recorded in the HI listeners also showed different patterns for strongly than for shallowly modulated tones. In the HI listeners, the stimulus was presented at a frequency where their audiogram was within the "normal" range. The EFR level-growth functions for the strongly modulated tones were similar to the corresponding ones recorded in NH threshold listeners. A strong saturation was observed, on the other hand, for shallowly modulated tones. In contrast to humans, the EFR magnitude reduction due to synaptopathy in mice was equally present in the responses for both modulation depths. Consistent with the data from mice, the simulated EFR level-growth functions using the model by Zilany et al. (2009, 2014) also showed a reduction due to synaptopathy, equal for both strongly and shallowly modulated tones. In line with this difference between potentially synaptopathic humans and certain synaptopathic mice, EFR recorded at a fixed supra-threshold level as a function of modulation depth (EFR depth-growth functions) also show opposing patterns. Bharadwaj et al. (2015) reported that the slope of EFR depth-growth functions was steeper in NH threshold listeners which also showed more difficulties in the behavioral detection of amplitude modulations. Similarly, in *chapter 3* both the NH threshold listeners considered as potentially synaptopathic and

the HI listeners showed a larger difference between the EFR magnitude for strongly and shallowly modulated (steeper slope) tones. This was relative to those NH threshold listeners considered as non-synaptopathic (parallel monotonic growth). However, Parthasarathy et al. (2017b) reported the opposite trend in age-induced synaptopathic mice. The slope of EFR depth-growth functions was steeper in young (non-synaptopathic) mice than in old (synaptopathic) mice, because the EFR for strongly modulated tones showed larger magnitudes in non-synaptopathic ears.

Aiming to explain these differences between humans and mice (and the model simulations), one needs to consider the stimuli used. In the human data, a modulation frequency of 93 Hz was used, whereas in mice this was 1024 Hz. Mice are sensitive to much higher frequencies than humans (Müller et al., 2005), which allowed the use of higher carrier frequencies and in consequence higher modulation frequencies. too. Shaheen et al. (2015) reported that EFRs recorded in mice using modulation frequencies around 1000 Hz showed a group delay of 2 ms, which corresponds to the latency of auditory brainstem responses (ABR) wave-II. Bushy cells of the cochlear nucleus are though to be the generator of ABR wave-II according to lesion studies in the cat (Melcher and Kiang, 1996). In contrast, EFRs elicited by modulation frequencies around 100 Hz in humans are believed to be generated more centrally, with main contributing neuronal sources located at the level of the brainstem (Herdman et al., 2002a; Kuwada et al., 2002). In addition, it has been reported that synaptopathy leads to a reduction of the ABR wave-I amplitude (associated to the AN activity) at high stimulus levels but not of the ABR wave-IV or wave-V (associated with activity in the brainstem) (Möhrle et al., 2016; Shaheen et al., 2015). This effect was similarly reported to occur in NH listeners suffering from tinnitus (Schaette and McAlpine, 2011). Recently, Chambers et al. (2016) reported that central gain mechanisms can partially restore firing rate at the level of the midbrain (and the auditory cortex) after AN fiber deafferentation. It could be speculated that such gain in the brainstem neural circuitry is more effective with restoring stronger temporal fluctuations, such as transient stimuli to elicit ABRs or strongly modulated tones to elicit EFRs, but partially fails when shallowly modulated stimuli are presented. The generation of EFRs and ABRs have been interrelated in the literature (Bohórquez and Ozdamar, 2008; Galambos et al., 1981; Roß et al., 2002), suggesting that EFR might arise from the linear superposition of periodically evoked transient activity, although this is still under

discussion among the scientific community (Bidelman, 2015; Galbraith and Brown, 1990). However, very little is know regarding how strongly and shallowly modulated stimuli are encoded at the different neuronal centers along the auditory pathway, and whether cochlear synaptopathy or other type of neuronal deafferentation and restoration mechanisms such as central gain may interfere. Further research combining animal physiology, human electrophysiology and computational modeling is needed to clarify the encoding of strongly and shallowly modulations in synaptopathic ears.

6.3.1 EFRs at supra-threshold stimulus levels: A general observation

EFR are currently used in clinical applications, typically referred to as auditory steady-state responses (ASSR), mainly to obtain an objective evaluation of hearing thresholds (e.g. Dimitrijevic et al., 2002; Michel and Jørgensen, 2016; Ozdek et al., 2010; Picton et al., 2005). EFRs have also been proposed as a potential tool to diagnose cochlear dead regions (Wilding et al., 2011), defined as regions in the cochlea with absent or highly dysfunctional IHCs. The assessment of cochlear dead regions have been traditionally based on the measurement of psychoacoustical tuning curves (PTC) (Moore, 2001). A shift of the tuning curve tip towards higher frequencies (i.e.detuning) has been associated with the presence of a dead region. An objective version of the PTC, named electrophysiological tuning curves (ETC), was proposed by Markessis et al. (2009). In order to be applied in the clinic, where minimizing the recording time is crucial, Wilding et al. (2011) proposed an alternative to the threshold-based ETC on which the probe was a SAM tone presented at a high stimulation level (50 dB sensation level, SL). Even though NH listeners without any sign of cochlear dead regions were tested, Wilding et al. (2011) reported tip shifts (of about 250 Hz) towards higher frequencies, which are associated to cochlear dead regions in the PTCs literature (Moore, 2001). However, considering the model simulations in the healthy system shown in *chapters 3*, *chapter 4* and *5* of the present thesis, such detuning towards higher frequencies in the ETCs would not be due to hair-cell dysfunction but because the modulations of the SAM are better encoded at high frequency off-frequency CFs in the AN when high stimulation levels are used. According to the simulations from the AN models, at high stimulation levels the stimulus envelope is poorly encoded at on-frequency CFs due to the saturating effect of the IHC transduction (input-output characteristics) at high input intensities (Russell et al., 1986), but better encoded at neurons tuned to

off-frequency CFs. Thus, in general, considering that evoked potentials reflect the response of population of neurons, it should be empathized that frequency-specific stimuli presented at high levels do not produce place- (BM) and CF- (AN) specific responses but broad stimulation of the hearing system.

6.3.2 Allying heuristic innovation with computational power: a suggestion for the generation of novel hypotheses

The process of generating novel hypotheses in the field of hearing research is commonly based on an heuristic interpretation of the function of the auditory system. In order to do so, the researcher must have a pre-established idea of the function of the system, that is, an internal model of the auditory system. Computational models that intend to describe a small part of the system (e.g. the periphery) are yet extraordinarily complex and highly nonlinear. In order to strive for the highest scientific standards when attempting to construct new hypotheses, to rely only on an heuristic interpretation of the function of the auditory system seems risky because it is difficult to carefully control for all the potential interactions and forecast an accurate output. More often than not, the generation of new hypotheses implies the extrapolation of knowledge obtained during the study of one part of the system towards unknown parts. For example, in the context of this thesis, the interpretation of sharp frequency selectivity in NH listeners derived using low-intensity stimuli is commonly maintained when exciting the auditory system with high-intensity stimuli. Similarly, HI listeners are assumed to have "broadened filters", even when stimulated with tones at frequencies of normal sensitivity. Such extrapolation totally disregards the off-frequency excitation AN fibers tuned to high CFs. These AN fibers tuned to CF higher than that of the tone are excited through their low-frequency tails, and therefore the disruption of their on-frequency active mechanism (i.e. cochlear amplifier) due to OHC dysfunction does not alter their excitation, because such excitation is not an on-frequency processing. Thus, the excitation of the AN will be the same in NH and in HI listeners, and at high stimulation levels the "auditory filters" are not sharp because the excitation is broad in all cases. In our hypothesis regarding the study of synaptopathy, the heuristic interpretation of the rate-level curves of the different types of AN fibers was misinterpreted, and thus, its effect on the EFR was not entirely accurate. In addition, the off-frequency activity on the generation of EFRs was not considered either in the

original hypothesis.

The generation of novel ideas (i.e. innovation) generally occurs when *a priori* unrelated knowledge becomes connected in a new way and reconsidered. A computational model cannot do such astonishing cognitive process, but can serve as a refining tool and to explore potential failures in the proposed hypothesis. The use of computational tools was not easily accessible in the past, but luckily this is not the case anymore. Therefore, the combination of innovative heuristic reasoning with the power of computational models might help on providing less risky and more accurate hypotheses. Nevertheless, the limitations of computer models must be present at all time. In case of doubtful or unexpected simulation outcomes, the model response should be contrasted to empirically-obtained data (e.g. direct physiological recordings in non-human animals). The computer model should be used as a non-intelligent tool that can suggest involuntarily omitted considerations.

6.4 Perspectives

The off-frequency contributions seem to dominate the magnitude of the EFR at medium-to-high stimulation levels. This observation can be directly derived from the simulations of both computational models reported in the present thesis in *chapter 3*, *4* and *5*, and is also consistent with the saturation observed in the EFR level-growth functions obtained from the multi-frequency stimulus recorded in NH listeners shown in *chapter 2*. EFRs recorded in rats and model simulations reported in Parthasarathy et al. (2016) are also consistent with this interpretation. The verification of the role of off-frequency contributions in the EFR is of great interest because, on one hand, might define the limitations of using EFRs at medium-to-high stimulus levels on assessing frequency-specific processing; and on the other hand allows the investigation of remote places of the cochlear processing by means of high intensity modulated tones presented at lower frequencies. The simulations from the humanized version of the model by Zilany et al. (2009, 2014) were shown to provide accurate results to account for EFR level-growth functions recorded in NH threshold listeners. The use of off-frequency maskers or a second SAM tone presented at a higher carrier frequency than the target could be used to further study the effect of the off-frequency contributions on the total EFR. Parameters such as the phase of the modulation frequency of the second higher frequency SAM tones have been

shown to interfere in the total recorded EFR (Guérit et al., 2017). A combination of experimental electrophysiological work in both human listeners and non-human mammals, with additional computational model simulations might help to deepen the understanding of EFR generation.

It was investigated in the present thesis whether cochlear synaptopathy could be reflected in the EFR level-growth functions by combining electrophysiological recordings in humans and mice. A humanized computational model of the AN was used to simulate the different individual patterns observed within the homogeneous group of young NH threshold listeners and the mild HI listeners, on which a postulated loss of AN fiber synapses was assumed. A cat version of the AN model was used to simulate the EFR level-growth functions recorded in noise exposed and unexposed mice. It was shown that the species difference between cat and mouse led to unsatisfactory simulated results. However, the analysis of the model simulations suggested that the implementation of the correct BM tuning for each species in the model framework might lead to satisfactory results. As the mouse is the most accurately characterized animal model regarding the study of synaptopathy (both noise- and age-induced), a mouse version of the AN model is desirable. Even though many model parameters could be modified to achieve a very accurate mouse AN model, it could be assumed that the general phenomenological processing of the current version of the animal model can be extrapolated from species to species. Thus, the adaptation of the tuning parameters in the BM section of the model, the adjustment of the correct frequency range at which mice are sensitive to, and the re-implementation of the middle ear filter based on physiological data from the mouse could be enough to provide a first approximation to the mouse AN activity.

Cochlear synaptopathy has not been demonstrated to date in humans. Many studies have failed due to the impossibility of finding statistically different responses (either behavioral or physiological) because the human listeners cannot be correctly grouped as being synaptopathic and non-synaptopathic. This does not imply that cochlear synaptopathy is not present in the human ear. In the case that the method used for grouping between synaptopathic and non-synaptopathic human listeners is not sensitive to AN synapse damage, the groups might not be clearly defined. As a consequence, the outcome of any used metric might be inconclusive because the potential effect of synaptopathy will be vanished producing a non-significant effect due to such inaccurate dif-

ferentiation of listeners. More formally, it should be considered that a random reshuffling of two statistically different probability distributions leads to a statistically non-significant difference, in particular if the effect size is not large. Moreover, it should be reckoned with that the heterogeneity across species and within the same species when wild-type (genetically diverse) individuals are examined, as in humans, adds more complexity to the investigation. Alternatively, the reversed approach on which the listeners are grouped based on their manifested deficits could be considered. In addition, alternative effects derived from the loss of AN fiber synapses should be taken into consideration. For instance, specially low-SR fibers may play a role in eliciting the medial olivocochlear reflex (MOCR: Liberman, 1988, 1991; Ye et al., 2000) and the middle-ear muscle reflex (MEMR: Kobler et al., 1992; Liberman and Kiang, 1984; Rouiller et al., 1986), which have been shown in animal models to be affected by synaptopathy (Valero et al., 2015). Notwithstanding, an interdisciplinary approach as proposed in the present thesis, might be the most efficient strategy to study cochlear synaptopathy in humans: The combination of concurrent non-invasive physiological (e.g. electrophysiology, magnetoencephalography, magnetic resonance imaging) studies in human listeners and in non-human animals where synaptopathy can be directly assessed with invasive techniques, with the additional support of species-specific computational model simulations, might be a promising route to follow.

Bibliography

American Clinical Neurophysiology Society (2006). "Guideline 5: Guidelines for Standard Electrode Position Nomenclature". In: *Am. J. Electroneurodiagnostic Technol.* 46.3.

Anderson, D., J. Rose, and J. Hind (1951). *Temporal position of discharges in single auditory nerve fibers within the cycle of a line wave stimulus: frequency and intensity effects*. DOI: 10.1121/1.1912474.

Ashmore, J. (2008). "Cochlear outer hair cell motility". In: *Physiol. Rev.* Pp. 173–210. DOI: 10.1152/physrev.00044.2006..

Bates, D., M. Mächler, B. M. Bolker, and S. C. Walker (2015). "Fitting linear mixed-effects models using lme4". In: *J. Stat. Softw.* 67.1, pp. 1–48. DOI: 10.18637/jss.v067.i01.

Bharadwaj, H. M., S. Masud, G. Mehraei, S. Verhulst, and B. G. Shinn-Cunningham (2015). "Individual Differences Reveal Correlates of Hidden Hearing Deficits". In: *J. Neurosci.* 35.5, pp. 2161–2172. DOI: 10.1523/JNEUROSCI.3915-14.2015.

Bharadwaj, H. M., S. Verhulst, L. Shaheen, M. C. Liberman, and B. G. Shinn-Cunningham (2014). "Cochlear neuropathy and the coding of supra-threshold sound." In: *Front. Syst. Neurosci.* 8.February, p. 26. DOI: 10.3389/fnsys.2014.00026.

Bidelman, G. M. (2015). "Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR". In: *Hear. Res.* 323, pp. 68–80. DOI: 10.1016/j.heares.2015.01.011.

Bland, J. M. and D. G. Altman (1986). "Statistical methods for assessing agreement between two methods of clinical measurement". In: *Lancet* 327.8476, pp. 307–310. DOI: 10.1016/S0140-6736(86)90837-8.

Boege, P. and T. Janssen (2002). "Pure-tone threshold estimation from extrapolated distortion product otoacoustic emission I/O-functions in normal

and cochlear hearing loss ears". In: *J. Acoust. Soc. Am.* 111.4, p. 1810. DOI: 10.1121/1.1460923.

Bohórquez, J. and O. Ozdamar (2008). "Generation of the 40-Hz auditory steady-state response (ASSR) explained using convolution." In: *Clin. Neurophysiol.* 119.11, pp. 2598–607. DOI: 10.1016/j.clinph.2008.08.002.

Carney, L. H., T. Li, and J. M. McDonough (2015). "Speech Coding in the Brain: Representation of Vowel Formants by Midbrain Neurons Tuned to Sound Fluctuations". In: *eNeuro* 2.4, pp. 1–12. DOI: 10.1523/ENEURO.0004-15.2015.

Chambers, A. R. et al. (2016). "Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation". English. In: *Neuron* 89.4, pp. 867–879. DOI: 10.1016/j.neuron.2015.12.041.

Coughlin, S. S. (1990). "Recall bias in epidemiologic studies". In: *J. Clin. Epidemiol.* 43.1, pp. 87–91. DOI: 10.1016/0895-4356(90)90060-3.

D'Haenens, W. et al. (2008). "Auditory steady-state responses in normal hearing adults: a test-retest reliability study." In: *Int. J. Audiol.* 47.8, pp. 489–98. DOI: 10.1080/14992020802116136.

Dallos, P. (2008). "Cochlear amplification, outer hair cells and prestin." In: *Curr. Opin. Neurobiol.* 18.4, pp. 370–6. DOI: 10.1016/j.conb.2008.08.016.

Dandy, W. (1934a). "Effects on hearing after subtotal section of the cochlear branch of the auditory nerve". In: *Bull. Johns Hopkins Hosp.* 55.

Dandy, W. (1934b). "The effect of hemisection of the cochlear branch of the human auditory nerve. Preliminary report". In: *Bull. Johns Hopkins Hosp.* 54.

Dau, T. (2003). "The importance of cochlear processing for the formation of auditory brainstem and frequency following responses." In: *J. Acoust. Soc. Am.* 113.2, pp. 936–950.

Dimitrijevic, A. et al. (2002). "Estimating the audiogram using multiple auditory steady-state responses." In: *J. Am. Acad. Audiol.* 13.4, pp. 205–24.

Dobie, R. A. and M. J. Wilson (1996). "A comparison of t test, F test, and coherence methods of detecting steady-state auditory-evoked potentials, distortion-product otoacoustic emissions, or other sinusoids". In: *J. Acoust. Soc. Am.* 100.4 PART 1, pp. 2236–2246.

Dong, W. and E. S. Olson (2013). "Detection of cochlear amplification and its activation". In: *Biophys. J.* 105.4, pp. 1067–1078. DOI: 10.1016/j.bpj.2013.06.049.

Dong, W., P. Varavva, and E. S. Olson (2013). “Sound transmission along the ossicular chain in common wild-type laboratory mice”. In: *Hear. Res.* 301.Sp. Iss. SI, pp. 27–34. DOI: 10.1016/j.heares.2012.11.015.

Dorn, P. A., D Konrad-Martin, S. T. Neely, D. H. Keefe, E Cyr, and M. P. Gorga (2001). “Distortion product otoacoustic emission input/output functions in normal-hearing and hearing-impaired human ears”. In: *J. Acoust. Soc. Am.* 110.6, pp. 3119–3131.

Encina-Llamas, G., T. Dau, and B. Epp (2017a). *Estimates of peripheral compression using envelope following responses [Data set]. Zenodo.* DOI: 10.5281/ZENODO.844834.

Encina-Llamas, G., J. M. Harte, T. Dau, B. G. Shinn-Cunningham, and B. Epp (2017b). *Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve [Data set]. Zenodo.* DOI: 10.5281/zenodo.844850.

Encina-Llamas, G., T. Dau, and B. Epp (under review[a]). “Estimates of peripheral compression using envelope following responses”. In: *JARO - J. Assoc. Res. Otolaryngol.*

Encina-Llamas, G., J. M. Harte, T. Dau, B. G. Shinn-Cunningham, and B. Epp (under review[b]). “Investigating the effect of synaptopathy on envelope following responses using a model of the auditory nerve”. In: *JARO - J. Assoc. Res. Otolaryngol.*

Epp, B., J. L. Verhey, and M. Mauermann (2010). “Modeling cochlear dynamics: Interrelation between cochlea mechanics and psychoacousticsa)”. In: *J. Acoust. Soc. Am.* 128.4, pp. 1870–1883. DOI: 10.1121/1.3479755.

Ernst, M. D. (2004). “Permutation Methods: A Basis for Exact Inference”. In: *Stat. Sci.* 19.4, pp. 676–685. DOI: 10.1214/088342304000000396.

Fereczkowski, M., M. L. Jepsen, T. Dau, and E. N. MacDonald (2017a). “Investigating time-efficiency of forward masking paradigms for estimating basilar membrane input-output characteristics”. In: *PLoS One* 12.3. Ed. by M. S. Malmierca, e0174776. DOI: 10.1371/journal.pone.0174776.

Fereczkowski, M., M. L. Jepsen, T. Dau, and E. N. MacDonald (2017b). *PLoS_One_Gap_method.* Dataset on Zenodo. DOI: 10.5281/ZENODO.344536.

Fernandez, K. A., P. W. C. Jeffers, K. Lall, M. C. Liberman, and S. G. Kujawa (2015). “Aging after noise exposure: acceleration of cochlear synaptopathy in "recovered" ears.” In: *J. Neurosci.* 35.19, pp. 7509–20. DOI: 10.1523/JNEUROSCI.5138-14.2015.

Fisher, R. A. (1935). *The design of experiments*. Edinburgh: Oliver and Boyd.

Furman, A. C., S. G. Kujawa, and M. C. Liberman (2013). "Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates." In: *J. Neurophysiol.* 110.3, pp. 577–86. DOI: 10.1152/jn.00164.2013.

Galambos, R, S Makeig, and P. J. Talmachoff (1981). "A 40-Hz auditory potential recorded from the human scalp". In: *Proc. Natl. Acad. Sci. U. S. A.* 78.4, pp. 2643–2647.

Galbraith, G. C. and W. S. Brown (1990). "Cross-correlation and latency compensation analysis of click-evoked and frequency-following brain-stem responses in man". In: *Electroencephalogr. Clin. Neurophysiol. Evoked Potentials* 77.4, pp. 295–308. DOI: 10.1016/0168-5597(90)90068-O.

Greenwood, D. D. (1990). "A cochlear frequency-position function for several species—29 years later". In: *J. Acoust. Soc. Am.* 87.6, pp. 2592–2605. DOI: 10.1121/1.399052.

Grose, J. H., E. Buss, and J. W. Hall (2017). "Loud Music Exposure and Cochlear Synaptopathy in Young Adults: Isolated Auditory Brainstem Response Effects but No Perceptual Consequences". In: *Trends Hear.* 21, p. 233121651773741. DOI: 10.1177/2331216517737417.

Guérit, F., J. Marozeau, and B. Epp (2017). "Linear combination of auditory steady-state responses evoked by co-modulated tones". In: *Poster Present. Assoc. Res. Otolarngology*. Vol. 142. 4, EL395–EL400. DOI: 10.1121/1.5007757.

Guinan, J. (2013). "New insights into cochlear amplification." In: *Biophys. J.* 105.4, pp. 839–40. DOI: 10.1016/j.bpj.2013.07.016.

Gummer, A. W., J. Meyer, G. Frank, M. P. Scherer, and S. Preyer (2002). "Mechanical transduction in outer hair cells." In: *Audiol. Neurootol.* 7.1, pp. 13–6. DOI: 46856.

Herdman, A. T. and D. R. Stapells (2001). "Thresholds determined using the monotic and dichotic multiple auditory steady-state response technique in normal-hearing subjects". In: *Scand. Audiol.* 30.1, pp. 41–49. DOI: 10.1080/010503901750069563.

Herdman, A. T. and D. R. Stapells (2003). "Auditory steady-state response thresholds of adults with sensorineural hearing impairments". In: *Int. J. Audiol.* 42.5, pp. 237 –248.

Herdman, A. T., O. Lins, P. Van Roon, D. R. Stapells, M. Scherg, and T. W. Picton (2002a). "Intracerebral Sources of Human Auditory Steady-State Responses". In: *Brain Topogr.* 15.2, pp. 69–86. DOI: 10.1023/A:1021470822922.

Herdman, A. T., T. W. Picton, and D. R. Stapells (2002b). “Place specificity of multiple auditory steady-state responses”. In: *J. Acoust. Soc. Am.* 112.4, pp. 1569–1582.

Hind, S. E., R. Haines-Bazrafshan, C. L. Benton, W. Brassington, B. Towle, and D. R. Moore (2011). “Prevalence of clinical referrals having hearing thresholds within normal limits.” EN. In: *Int. J. Audiol.* 50.10, pp. 708–16. DOI: 10.3109/14992027.2011.582049.

Hofstetter, P., D. Ding, N. Powers, and R. J. Salvi (1997). “Quantitative relationship of carboplatin dose to magnitude of inner and outer hair cell loss and the reduction in distortion product otoacoustic emission amplitude in chinchillas”. In: *Hear. Res.* 112.1-2, pp. 199–215. DOI: 10.1016/S0378-5955(97)00123-8.

Jensen, J. B., A. C. Lysaght, M. C. Liberman, K. Qvortrup, and K. M. Stankovic (2015). “Immediate and delayed cochlear neuropathy after noise exposure in pubescent mice”. In: *PLoS One* 10.5. DOI: 10.1371/journal.pone.0125160.

Jepsen, M. L., S. D. Ewert, and T. Dau (2008). “A computational model of human auditory signal processing and perception”. In: *J. Acoust. Soc. Am.* 124.1, pp. 422–438. DOI: 10.1121/1.2924135.

Jerger, J, T. A. Oliver, and F Pirozzolo (1990). “Impact of central auditory processing disorder and cognitive deficit on the self-assessment of hearing handicap in the elderly.” In: *J. Am. Acad. Audiol.* 1.2, pp. 75–80.

John, M. S. and T. W. Picton (2000). “MASTER: a Windows program for recording multiple auditory steady-state responses”. In: *Comput. Methods Programs Biomed.* 61.2, pp. 125–150.

John, M. S., A Dimitrijevic, and T. W. Picton (2001). “Weighted averaging of steady-state responses”. In: *Clin. Neurophysiol.* 112.3, pp. 555–562.

John, M. S., O. G. Lins, B. L. Boucher, and T. W. Picton (1998). “Multiple auditory steady-state responses (MASTER): stimulus and recording parameters”. In: *Int. J. Audiol.* 37.2, pp. 59–82.

Joris, P. X. and T. C. T. Yin (1992). “Responses to amplitude-modulated tones in the auditory nerve of the cat”. In: *J. Acoust. Soc. Am.* 91.1, p. 215. DOI: 10.1121/1.402757.

Kalluri, R. and C. A. Shera (2001). “Distortion-product source unmixing: A test of the two-mechanism model for DPOAE generation”. In: *J. Acoust. Soc. Am.* 109.2, pp. 622–637.

Kiang, N. Y. S. and E. C. Moxon (1973). “Tails of Tuning Curves of Auditory-Nerve Fibers”. In: *J. Acoust. Soc. Am.* 54.1, p. 274. DOI: 10.1121/1.1977980.

King, K. and D. Stephens (1992). “Auditory and psychological factors in ’auditory disability with normal hearing’.” In: *Scand. Audiol.* 21.2, pp. 109–114. DOI: 10.3109/01050399209045990.

Knight, R. D. and D. T. Kemp (2001). “Wave and place fixed DPOAE maps of the human ear”. In: *J. Acoust. Soc. Am.* 109.4, p. 1513. DOI: 10.1121/1.1354197.

Kobler, J. B., J. J. Guinan, S. R. Vacher, B. E. Norris, and J. Guinan Jr (1992). “Acoustic reflex frequency selectivity in single stapedius motoneurons of the cat”. In: *J. Neurophysiol.* 68.3, pp. 807–817.

Kujawa, S. G. and M. C. Liberman (2006). “Acceleration of age-related hearing loss by early noise exposure: evidence of a misspent youth.” In: *J. Neurosci.* 26.7, pp. 2115–23. DOI: 10.1523/JNEUROSCI.4985-05.2006.

Kujawa, S. G. and M. C. Liberman (2009). “Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss.” In: *J. Neurosci.* 29.45, pp. 14077–85. DOI: 10.1523/JNEUROSCI.2845-09.2009.

Kujawa, S. G. and M. C. Liberman (2015). “Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss”. In: *Hear. Res.* 330.Part B, pp. 191–199. DOI: 10.1016/j.heares.2015.02.009.

Kumar, G., F. Amen, and D. Roy (2007). “Normal hearing tests: is a further appointment really necessary?” In: *J. R. Soc. Med.* 100.2, p. 66. DOI: 10.1258/jrsm.100.2.66-a.

Kummer, P., T. Janssen, and W. Arnold (1998). “The level and growth behavior of the 2 f1f2 distortion product otoacoustic emission and its relationship to auditory sensitivity in normal hearing and cochlear hearing loss”. In: *J. Acoust. Soc. Am.* 103.6, pp. 3431–3444. DOI: 10.1121/1.423054.

Kuwada, S, R Batra, and V. L. Maher (1986). “Scalp potentials of normal and hearing-impaired subjects in response to sinusoidally amplitude-modulated tones”. In: *Hear. Res.* 21.2, pp. 179–192.

Kuwada, S., J. S. Anderson, R. Batra, D. C. Fitzpatrick, N. Teissier, and W. R. D’Angelo (2002). “Sources of the scalp-recorded amplitude-modulation following response.” In: *J. Am. Acad. Audiol.* 13.4, pp. 188–204.

Kuznetsova, A., P. B. Brockhoff, and R. H. B. Christensen (2015). *lmerTest: tests in linear mixed effects models. R package version 2.0-20.*

Le Prell, C. G., S. K. Grinn, J. Baker, and K. Wiseman (2017). “Effects of recreational noise on evoked potential amplitude and other auditory test metrics”. In: *J. Hear. Sci. - XXV IERASG Bienn. Symp.* P. 55.

Liberman, M. C. and L. W. Dodds (1984). “Single neuron labeling and chronic cochlea pathology. III. Stereocilia damage and alterations in threshold tuning curves”. In: *Hear. Res.* 16.1, pp. 54–74. DOI: 10.1016/0378-5955(84)90025-X.

Liberman, M. C. (1978). “Auditory-nerve response from cats raised in a low-noise chamber”. In: *J. Acoust. Soc. Am.* 63.2, p. 442. DOI: 10.1121/1.381736.

Liberman, M. C. (1982). “The cochlear frequency map for the cat: Labeling auditory-nerve fibers of known characteristic frequency”. In: *J. Acoust. Soc. Am.* 72.5, pp. 1441–1449. DOI: 10.1121/1.388677.

Liberman, M. C. (1988). “Physiology of cochlear efferent and afferent neurons: Direct comparisons in the same animal”. In: *Hear. Res.* 34.2, pp. 179–191. DOI: 10.1016/0378-5955(88)90105-0.

Liberman, M. C. (1991). “Central projections of auditory-nerve fibers of differing spontaneous rate. I. Anteroventral cochlear nucleus”. In: *J. Comp. Neurol.* 313.2, pp. 240–258. DOI: 10.1002/cne.903130205.

Liberman, M. C. and S. G. Kujawa (2017). “Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms”. In: *Hear. Res.* DOI: 10.1016/j.heares.2017.01.003.

Liberman, M. C., L. W. Dodds, and S. Pierce (1990). “Afferent and efferent innervation of the cat cochlea: quantitative analysis with light and electron microscopy.” In: *J. Comp. Neurol.* 301.3, pp. 443–60. DOI: 10.1002/cne.903010309.

Liberman, M. and N. Y. S. Kiang (1984). “Single-neuron labeling and chronic cochlear pathology. IV. Stereocilia damage and alterations in rate- and phase-level functions”. In: *Hear. Res.* 16.1, pp. 75–90. DOI: 10.1016/0378-5955(84)90026-1.

Lin, H. W., A. C. Furman, S. G. Kujawa, and M. C. Liberman (2011). “Primary neural degeneration in the Guinea pig cochlea after reversible noise-induced threshold shift.” In: *J. Assoc. Res. Otolaryngol.* 12.5, pp. 605–16. DOI: 10.1007/s10162-011-0277-0.

Lins, O. G. and T. W. Picton (1995). “Auditory steady-state responses to multiple simultaneous stimuli”. In: *Electroencephalogr. Clin. Neurophysiol.* 96.5, pp. 420–432.

Lins, O. G., P. E. Picton, T. W. Picton, S. C. Champagne, and A. Durieux-Smith (1995). "Auditory steady-state responses to tones amplitude-modulated at 80–110 Hz". In: *J. Acoust. Soc. Am.* 97.5, pp. 3051–3063. DOI: 10.1121/1.411869.

Liu, L. et al. (2012). "Silent damage of noise on cochlear afferent innervation in guinea pigs and the impact on temporal processing." In: *PLoS One* 7.11, e49550. DOI: 10.1371/journal.pone.0049550.

Lobarinas, E., R. Salvi, and D. Ding (2013). "Insensitivity of the audiogram to carboplatin induced inner hair cell loss in chinchillas." In: *Hear. Res.* 302, pp. 113–20. DOI: 10.1016/j.heares.2013.03.012.

Lobarinas, E., C. Spankovich, and C. G. Le Prell (2016). "Evidence of "hidden hearing loss" following noise exposures that produce robust TTS and ABR wave-I amplitude reductions". In: *Hear. Res.* DOI: 10.1016/j.heares.2016.12.009.

Long, G. R., C. L. Talmadge, and J. Lee (2008). "Measuring distortion product otoacoustic emissions using continuously sweeping primaries". In: *J. Acoust. Soc. Am.* 124.3, pp. 1613–1626. DOI: 10.1121/1.2949505.

Lopez-Poveda, E. A. and A. Alves-Pinto (2008). "A variant temporal-masking-curve method for inferring peripheral auditory compression". In: *J. Acoust. Soc. Am.* 123.3, pp. 1544–1554. DOI: 10.1121/1.2835418.

Lopez-Poveda, E. A. and P. T. Johannesen (2012). "Behavioral estimates of the contribution of inner and outer hair cell dysfunction to individualized audiometric loss". In: *JARO - J. Assoc. Res. Otolaryngol.* 13.4, pp. 485–504. DOI: 10.1007/s10162-012-0327-2.

Lopez-Poveda, E. A., P. T. Johannesen, B. C. Buzo, F. M. Rønne, N. H. Pontoppidan, and J. M. Harte (2017). "On the value of brief sound audiometry as a diagnostic tool for cochlear synaptopathy". In: *Poster Present. Assoc. Res. Otolarngology*.

Luebke, A. E. and P. K. Foster (2002). "Variation in inter-animal susceptibility to noise damage is associated with alpha 9 acetylcholine receptor subunit expression level." In: *J. Neurosci.* 22.10, pp. 4241–7. DOI: 20026409.

Makary, C. A., J. Shin, S. G. Kujawa, M. C. Liberman, and S. N. Merchant (2011). "Age-related primary cochlear neuronal degeneration in human temporal bones." In: *J. Assoc. Res. Otolaryngol.* 12.6, pp. 711–7. DOI: 10.1007/s10162-011-0283-2.

Markessis, E. et al. (2009). “Frequency tuning curves derived from auditory steady state evoked potentials: a proof-of-concept study.” In: *Ear Hear.* 30.1, pp. 43–53. DOI: 10.1097/AUD.0b013e31818fbb7a.

Marmel, F., M. A. Rodríguez-Mendoza, and E. A. Lopez-Poveda (2015). “Stochastic undersampling steepens auditory threshold/duration functions: implications for understanding auditory deafferentation and aging.” In: *Front. Aging Neurosci.* 7, p. 63. DOI: 10.3389/fnagi.2015.00063.

Mauermann, M. and B. Kollmeier (2004). “Distortion product otoacoustic emission (DPOAE) input/output functions and the influence of the second DPOAE source”. In: *J. Acoust. Soc. Am.* 116.4, Part 1, pp. 2199–2212.

Meddis, R (1986). “Simulation of mechanical to neural transduction in the auditory receptor”. In: *J. Acoust. Soc. Am.* 79.3, pp. 702–711. DOI: 10.1121/1.393460.

Mehraei, G. et al. (2016). “Auditory Brainstem Response Latency in Noise as a Marker of Cochlear Synaptopathy”. In: *J. Neurosci.* 36.13, pp. 3755–3764. DOI: 10.1523/JNEUROSCI.4460-15.2016.

Mehraei, G. (2016). “Auditory brainstem response latency in noise as a marker of cochlear synaptopathy”. PhD thesis. Harvard - MIT.

Melcher, J. R. and N. Y. S. Kiang (1996). “Generators of the brainstem auditory evoked potential in cat III: Identified cell populations”. In: *Hear. Res.* 93.1-2, pp. 52–71. DOI: 10.1016/0378-5955(95)00200-6.

Michel, F. and K. F. Jørgensen (2016). “Comparison of threshold estimation in infants with hearing loss or normal hearing using auditory steady-state response evoked by narrow band CE-chirps and auditory brainstem response evoked by tone pips”. In: *Int. J. Audiol.* 56.2, pp. 1–7. DOI: 10.1080/14992027.2016.1234719.

Möhrle, D. et al. (2016). “Loss of auditory sensitivity from inner hair cell synaptopathy can be centrally compensated in the young but not old brain”. In: *Neurobiol. Aging* 44, pp. 173–184. DOI: 10.1016/j.neurobiolaging.2016.05.001.

Moore, B. C. J. (2001). “Dead Regions in the Cochlea: Diagnosis, Perceptual Consequences, and Implications for the Fitting of Hearing Aids”. In: *Trends Amplif.* 5.1, pp. 1–34. DOI: 10.1177/108471380100500102.

Müller, M. and D. Robertson (1991). “Relationship between tone burst discharge pattern and spontaneous firing rate of auditory nerve fibres in the guinea

pig". In: *Hear. Res.* 57.1, pp. 63–70. DOI: 10.1016/0378-5955(91)90075-K.

Müller, M., K. Von Hünerbein, S. Hoidis, and J. W. T. Smolders (2005). "A physiological place-frequency map of the cochlea in the CBA/J mouse". In: *Hear. Res.* 202.1-2, pp. 63–73. DOI: 10.1016/j.heares.2004.08.011.

Neely, S. T., M. P. Gorga, and P. A. Dorn (2003). "Cochlear compression estimates from measurements of distortion-product otoacoustic emissions". In: *J. Acoust. Soc. Am.* 114.3, pp. 1499–1507.

Neely, S. T., T. A. Johnson, J. Kopun, D. M. Dierking, and M. P. Gorga (2009). "Distortion-product otoacoustic emission input/output characteristics in normal-hearing and hearing-impaired human ears." In: *J. Acoust. Soc. Am.* 126.2, pp. 728–38. DOI: 10.1121/1.3158859.

Neff, W. D. (1947). "The effects of partial section of the auditory nerve." In: *J. Comp. Physiol. Psychol.* 40.4, pp. 203–15. DOI: 10.1037/h0056931.

Nelson, D. A., A. C. Schroder, and M. Wojtczak (2001). "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners". In: *J. Acoust. Soc. Am.* 110.4, pp. 2045–2064.

Nelson, P. C. and L. H. Carney (2004). "A phenomenological model of peripheral and central neural responses to amplitude-modulated tones". In: *J. Acoust. Soc. Am.* 116.4, p. 2173. DOI: 10.1121/1.1784442.

Nuttall, A. L. and D. Dolan (1996). "Steady-state sinusoidal velocity responses of the basilar membrane in guinea pig". In: *J. Acoust. Soc. Am.* 99.3, p. 1556. DOI: 10.1121/1.414732.

Oxenham, A. J. and C. J. Plack (1997). "A behavioral measure of basilar-membrane nonlinearity in listeners with normal and impaired hearing". In: *J. Acoust. Soc. Am.* 101.6, pp. 3666–3675.

Oxenham, A. J. (2016). "Predicting the Perceptual Consequences of Hidden Hearing Loss". In: *Trends Hear.* 20, pp. 1–6. DOI: 10.1177/2331216516686768.

Ozdek, A., M. Karacay, G. Saylam, E. Tatar, N. Aygener, and M. H. Korkmaz (2010). "Comparison of pure tone audiometry and auditory steady-state responses in subjects with normal hearing and hearing loss." In: *Eur. Arch. oto-rhino-laryngology* 267.1, pp. 43–9. DOI: 10.1007/s00405-009-1014-8.

Parthasarathy, A., J. Lai, and E. L. Bartlett (2016). "Age-Related Changes in Processing Simultaneous Amplitude Modulated Sounds Assessed Using Envelope Following Responses". In: *JARO - J. Assoc. Res. Otolaryngol.* 17.2, pp. 119–132. DOI: 10.1007/s10162-016-0554-z.

Parthasarathy, A., G. Encina-Llamas, B. G. Shinn-Cunningham, and S. G. Kujawa (2017a). “Temporal processing deficits due to noise-induced synaptopathy studied using envelope following responses”. In: *Poster Present. Assoc. Res. Otolarngology*. Baltimore, MA, USA.

Parthasarathy, A., E. Y. Smith, and S. G. Kujawa (2017b). “Temporal processing deficits in age-related cochlear synaptopathy.” In: *Poster Present. Assoc. Res. Otolarngology2*.

Patterson, R. D. (1974). “Auditory filter shape.” In: *J. Acoust. Soc. Am.* 55.4, pp. 802–9. DOI: 10.1121/1.1914603.

Paul, B. T., I. C. Bruce, and L. E. Roberts (2016). “Evidence that hidden hearing loss underlies amplitude modulation encoding deficits in individuals with and without tinnitus”. In: *Hear. Res.* DOI: 10.1016/j.heares.2016.11.010.

Picton, T. W. (2010). *Human Auditory Evoked Potentials*. Plural Pub.

Picton, T. W., C. R. Skinner, S. C. Champagne, A. J. C. Kellett, and A. C. Maiste (1987). “Potentials evoked by the sinusoidal modulation of the amplitude or frequency of a tone”. In: *J. Acoust. Soc. Am.* 82.1, pp. 165–178.

Picton, T. W., M. S. John, A. Dimitrijevic, and D. Purcell (2003). “Human auditory steady-state responses”. In: *Int. J. Audiol.* 42.4, pp. 177–219.

Picton, T. W., A. Dimitrijevic, M.-C. Perez-Abalo, and P. Van Roon (2005). “Estimating audiometric thresholds using auditory steady-state responses.” In: *J. Am. Acad. Audiol.* 16.3, pp. 140–56. DOI: 10.3766/jaaa.16.3.3.

Picton, T. W., P. van Roon, and M. S. John (2007). “Human auditory steady-state responses during sweeps of intensity.” In: *Ear Hear.* 28.4, pp. 542–57. DOI: 10.1097/AUD.0b013e31806dc2a7.

Plack, C. J. (2013). “Cochlear compression: recent insights from behavioural experiments”. In: *Adv. Exp. Med. Biol.* 787, pp. 31–38. DOI: 10.1007/978-1-4614-1590-9_4.

Plack, C. J. and V. Skeels (2007). “Temporal integration and compression near absolute threshold in normal and impaired ears”. In: *J. Acoust. Soc. Am.* 122.4, pp. 2236–44. DOI: 10.1121/1.2769829.

Plack, C. J., V. Drga, and E. A. Lopez-Poveda (2004). “Inferred basilar-membrane response functions for listeners with mild to moderate sensorineural hearing loss”. In: *J. Acoust. Soc. Am.* 115.4, pp. 1684–1695.

Plack, C. J., D. Barker, and G. Prendergast (2014). “Perceptual consequences of "hidden" hearing loss.” In: *Trends Hear.* 18. DOI: 10.1177/2331216514550621.

Popelka, G. R., P. A. Osterhammel, L. H. Nielsen, and A. N. Rasmussen (1993). "Growth of distortion product otoacoustic emissions with primary-tone level in humans". In: *Hear. Res.* 71.1-2, pp. 12 –22.

Prendergast, G. et al. (2016a). "An Investigation of Hidden Hearing Loss in Young Adults with Normal Hearing". In: *Poster Present. Assoc. Res. Otolarngology*.

Prendergast, G. et al. (2016b). "Effects of noise exposure on young adults with normal audiograms I: Electrophysiology". In: *Hear. Res.* 344, pp. 1–14. DOI: 10.1016/j.heares.2016.10.028.

Purcell, D. W., M. S. John, and T. W. Picton (2003). "Concurrent measurement of distortion product otoacoustic emissions and auditory steady state evoked potentials". In: *Hear. Res.* 176.1-2, pp. 128–141. DOI: 10.1016/S0378-5955(02)00770-0.

Rance, G. and F. Rickards (2002). "Prediction of hearing threshold in infants using auditory steady-state evoked potentials." In: *J. Am. Acad. Audiol.* 13.5, pp. 236–45.

Rance, G. and A. Starr (2015). *Pathophysiological mechanisms and functional hearing consequences of auditory neuropathy*. DOI: 10.1093/brain/aww270.

Recio, A. and N. Rich (1998). "Basilar-membrane responses to clicks at the base of the chinchilla cochlea". In: *J. Acoust. Soc. Am.* 103.4, pp. 1972–1989. DOI: 10.1121/1.421377.

Rhode, W. and L. Robles (1974). "Evidence from Mössbauer experiments for nonlinear vibration in the cochlea". In: *J. Acoust. Soc. Am.* 55.3, pp. 588–596.

Rhode, W. S. and A. Recio (2000). "Study of mechanical motions in the basal region of the chinchilla cochlea". In: *J. Acoust. Soc. Am.* 107.6, p. 3317. DOI: 10.1121/1.429404.

Robinette, M. S. and T. J. Glattke (2002). *Otoacoustic Emissions: Clinical Applications*. Thieme Publishers Series. Thieme.

Robles, L and M. A. Ruggero (2001). "Mechanics of the mammalian cochlea". In: *Physiol. Rev.* 81.3, pp. 1305–1352.

Rodríguez, J., S. T. Neely, W. Jesteadt, H. Tan, and M. P. Gorga (2011). "Comparison of distortion-product otoacoustic emission growth rates and slopes of forward-masked psychometric functions." In: *J. Acoust. Soc. Am.* 129.2, pp. 864–75. DOI: 10.1121/1.3523340.

Rønne, F. M. (2013). "Modeling auditory evoked potentials to complex stimuli". PhD thesis. Technical University of Denmark, p. 141.

Rønne, F. M., T. Dau, J. Harte, and C. Elberling (2012). "Modeling auditory evoked brainstem responses to transient stimuli." In: *J. Acoust. Soc. Am.* 131.5, pp. 3903–13. DOI: 10.1121/1.3699171.

Rose, J. E., J. E. Hind, D. J. Anderson, and J. F. Brugge (1971). "Some effects of stimulus intensity on response of auditory nerve fibers in the squirrel monkey". In: *J-Neurophysiol* 34.4, pp. 685–699.

Roß, B., T. W. Picton, and C. Pantev (2002). "Temporal integration in the human auditory cortex as represented by the development of the steady-state magnetic field". In: *Hear. Res.* 165.1-2, pp. 68–84. DOI: 10.1016/S0378-5955(02)00285-X.

Rouiller, E. M., R. Cronin-Schreiber, D. M. Fekete, and D. K. Ryugo (1986). "The central prodjections of intracellularly labeled auditory nerve fibers in cats: An analysis of terminal morphology". In: *J. Comp. Neurol.* 249.2, pp. 261–278. DOI: 10.1002/cne.902490210.

Ruggero, M. a. and N. C. Rich (1991). "Furosemide alters organ of corti mechanics: evidence for feedback of outer hair cells upon the basilar membrane." In: *J. Neurosci.* 11.4, pp. 1057–1067. DOI: 10.1016/j.biotechadv.2011.08.021.Secreted.

Ruggero, M. a., N. C. Rich, A. Recio, S. S. Narayan, and L. Robles (1997). "Basilar-membrane responses to tones at the base of the chinchilla cochlea". In: *J. Acoust. Soc. Am.* 101.4, pp. 2151–2163.

Russell, I. J., A. R. Cody, and G. P. Richardson (1986). "The responses of inner and outer hair cells in the basal turn of the guinea-pig cochlea and in the mouse cochlea grown in vitro". In: *Hear. Res.* 22.1-3, pp. 199–216. DOI: 10.1016/0378-5955(86)90096-1.

Ryan, a and P Dallos (1975). "Effect of absence of cochlear outer hair cells on behavioural auditory threshold." In: *Nature* 253.5486, pp. 44–46. DOI: 10.1038/253044a0.

Sachs, M. B., R. L. Winslow, and B. H. Sokolowski (1989). "A computational model for rate-level functions from cat auditory-nerve fibers". In: *Hear. Res.* 41.1, pp. 61–69. DOI: 10.1016/0378-5955(89)90179-2.

Sanchez, R. H. and B. Epp (2015). "Simultaneous measurement of auditory-steady-state responses and otoacoustic emissions to estimate peripheral compression". In: *Proc. ISAAR 2015 Individ. Hear. Loss – Charact. Model. Compens. Strateg. 5th Symp. Audit. Audiol. Res.* Vol. 5. August, pp. 421–428.

Saunders, G. H. and M. P. Haggard (1989). “The clinical assessment of obscure auditory dysfunction–1. Auditory and psychological factors.” In: *Ear Hear.* 10.3, pp. 200–8.

Saunders, J. C. and R. M. Summers (1982). “Auditory structure and function in the mouse middle ear: An evaluation by SEM and capacitive probe”. In: *J. Comp. Physiol.* 146.4, pp. 517–525. DOI: 10.1007/BF00609448.

Schaette, R. and D. McAlpine (2011). “Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model.” In: *J. Neurosci.* 31.38, pp. 13452–7. DOI: 10.1523/JNEUROSCI.2156-11.2011.

Schuknecht, H. F. and R. C. Woellner (1955). “An Experimental and Clinical Study of Deafness from Lesions of the Cochlear Nerve”. In: *J. Laryngol. Otol.* 69.02, pp. 75 –97. DOI: 10.1017/S0022215100050465.

Sergeyenko, Y., K. Lall, M. C. Liberman, and S. G. Kujawa (2013). “Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline.” In: *J. Neurosci.* 33.34, pp. 13686–94. DOI: 10.1523/JNEUROSCI.1783-13.2013.

Shaheen, L. A., M. D. Valero, and M. C. Liberman (2015). “Towards a Diagnosis of Cochlear Neuropathy with Envelope Following Responses”. In: *J. Assoc. Res. Otolaryngol.* DOI: 10.1007/s10162-015-0539-3.

Shamma, S. A., R. S. Chadwick, W. J. Wilbur, K. A. Morrish, and J. Rinzel (1989). “A biophysical model of cochlear processing: Intensity dependence of pure tone responses”. In: *J. Acoust. Soc. Am.* 80.1, pp. 133–145. DOI: 10.1121/1.394173.

Shera, C. A., J. Guinan, and J. J. Guinan Jr (1999). “Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs”. In: *J. Acoust. Soc. Am.* 105.2 Pt 1, pp. 782–798.

Shera, C. A. (2004). “Mechanisms of mammalian otoacoustic emission and their implications for the clinical utility of otoacoustic emissions.” In: *Ear Hear.* 25.2, pp. 86–97. DOI: 10.1097/01.AUD.0000121200.90211.83.

Sokolowski, B. H., M. B. Sachs, and J. L. Goldstein (1989). “Auditory nerve rate-level functions for two-tone stimuli: Possible relation to basilar membrane nonlinearity”. In: *Hear. Res.* 41.2-3, pp. 115–123. DOI: 10.1016/0378-5955(89)90005-1.

Spoendlin, H and A Schrott (1989). “Analysis of the human auditory nerve”. In: *Hear. Res.* 43.1, pp. 25–38. DOI: 10.1016/0378-5955(89)90056-7.

Spongr, V. P., D. G. Flood, R. D. Frisina, and R. J. Salvi (1997). "Quantitative measures of hair cell loss in CBA and C57BL/6 mice throughout their life spans." In: *J. Acoust. Soc. Am.* 101.6, pp. 3546–53. DOI: 10.1121/1.418315.

Starr, A., T. W. Picton, Y. Sininger, L. J. Hood, and C. I. Berlin (1996). "Auditory neuropathy". In: *Brain* 119.3, pp. 741–753. DOI: 10.1093/brain/119.3.741.

Stebbins, W. C., J. E. Hawkins Jr, L.-G. Johnsson, and D. B. Moody (1979). "Hearing Thresholds with Outer and Inner Hair Cell Loss". In: *Am. J. Otolaryngol.* 1.1, pp. 15–27. DOI: 10.1107/S0567739468000896.

Sumner, C. J., E. A. Lopez-Poveda, L. P. O'Mard, and R. Meddis (2002). "A revised model of the inner-hair cell and auditory-nerve complex". In: *J. Acoust. Soc. Am.* 111.5, p. 2178. DOI: 10.1121/1.1453451.

Taberner, A. M. and M. C. Liberman (2005). "Response properties of single auditory nerve fibers in the mouse." In: *J. Neurophysiol.* 93.1, pp. 557–69. DOI: 10.1152/jn.00574.2004.

Trautwein, P., P. Hofstetter, J. Wang, R. Salvi, and A. Nostrant (1996). "Selective inner hair cell loss does not alter distortion product otoacoustic emissions". In: *Hear. Res.* 96.1-2, pp. 71–82. DOI: 10.1016/0378-5955(96)00040-8.

Tremblay, K. L. et al. (2015). "Self-reported hearing difficulties among adults with normal audiograms: the beaver dam offspring study." In: *Ear Hear.* 36.6, e290–9. DOI: 10.1097/AUD.0000000000000195.

Valero, M., J. Burton, S. Hauser, T. Hackett, R. Ramachandran, and M. Liberman (2017). "Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta)". In: *Hear. Res.* DOI: 10.1016/j.heares.2017.07.003.

Valero, M. D., K. E. Hancock, and M. C. Liberman (2015). "The Middle Ear Muscle Reflex in the Diagnosis of Cochlear Neuropathy." In: *Hear. Res.* DOI: 10.1016/j.heares.2015.11.005.

Van Maanen, A. and D. R. Stapells (2005). "Comparison of multiple auditory steady-state responses (80 versus 40 Hz) and slow cortical potentials for threshold estimation in hearing-impaired adults". In: *Int. J. Audiol.* 44.11, pp. 613 –624. DOI: 10.1080/14992020500258628.

Verhulst, S., T. Dau, and C. A. Shera (2012). "Nonlinear time-domain cochlear model for transient stimulation and human otoacoustic emission". In: *J. Acoust. Soc. Am.* 132.6, pp. 3842–3848. DOI: 10.1121/1.4763989.

Verhulst, S., H. M. Bharadwaj, G. Mehraei, C. A. Shera, and B. G. Shinn-Cunningham (2015). "Functional modeling of the human auditory brainstem response to

broadband stimulation". In: *J. Acoust. Soc. Am.* 138.3, pp. 1637–1659. DOI: 10.1121/1.4928305.

Viana, L. M. et al. (2015). "Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue". In: *Hear. Res.* 327, pp. 78–88. DOI: 10.1016/j.heares.2015.04.014.

Voysey, G. and G. Encina-Llamas (2016). "Corti: Version 0.9". PhD thesis. DOI: 10.5281/ZENODO.57111.

Wang, J., N. L. Powers, P. Hofstetter, P. Trautwein, D. Ding, and R. Salvi (1997). "Effects of selective inner hair cell loss on auditory nerve fiber threshold, tuning and spontaneous and driven discharge rate". In: *Hear. Res.* 107.1-2, pp. 67–82. DOI: 10.1016/S0378-5955(97)00020-8.

Westerman, L. A. and R. L. Smith (1988). "A diffusion model of the transient response of the cochlear inner hair cell synapse". In: *J. Acoust. Soc. Am.* 83.6, p. 2266. DOI: 10.1121/1.396357.

Wever, E. G. and W. D. Neff (1947). "A further study of the effects of partial section of the auditory nerve." In: *J. Comp. Physiol. Psychol.* 40.4, pp. 217–26. DOI: 10.1037/h0062713.

Wilding, T. S., C. M. McKay, R. J. Baker, and K. Kluk (2012). "Auditory steady state responses in normal-hearing and hearing-impaired adults: an analysis of between-session amplitude and latency repeatability, test time, and F ratio detection paradigms." In: *Ear Hear.* 33.2, pp. 267–78. DOI: 10.1097/AUD.0b013e318230bba0.

Wilding, T., C. McKay, R. Baker, T. Picton, and K. Kluk (2011). "Using the auditory steady state response to record response amplitude curves. A possible fast objective method for diagnosing dead regions." In: *Ear Hear.* 32.4, pp. 485–97. DOI: 10.1097/AUD.0b013e31820a77e2.

Wojtczak, M. and A. J. Oxenham (2009). "Pitfalls in behavioral estimates of basilar-membrane compression in humans." In: *J. Acoust. Soc. Am.* 125.1, pp. 270–281. DOI: 10.1121/1.3023063.

Yates, G. K. (1990). "Basilar membrane nonlinearity and its influence on auditory nerve rate-intensity functions". In: *Hear. Res.* 50.1-2, pp. 145 –162. DOI: 10.1016/0378-5955(90)90041-M.

Yates, G. K., I. M. Winter, and D. Robertson (1990). "Basilar membrane nonlinearity determines auditory nerve rate-intensity functions and cochlear dynamic range". In: *Hear. Res.* 45.3, pp. 203 –219. DOI: 10.1016/0378-5955(90)90121-5.

Ye, Y., D. G. Machado, and D. O. Kim (2000). "Projection of the marginal shell of the anteroventral cochlear nucleus to olivocochlear neurons in the cat". In: *J. Comp. Neurol.* 420.1, pp. 127–138. DOI: 10.1002/(SICI)1096-9861(20000424)420:1<127::AID-CNE9>3.0.CO;2-7.

Zhang, X. and L. H. Carney (2005). "Analysis of models for the synapse between the inner hair cell and the auditory nerve". In: *J. Acoust. Soc. Am.* 118.3, pp. 1540–1553. DOI: 10.1121/1.1993148.

Zhang, X., M. G. Heinz, I. C. Bruce, and L. H. Carney (2001). "A phenomenological model for the responses of auditory-nerve fibers: I. Nonlinear tuning with compression and suppression". In: *J. Acoust. Soc. Am.* 109.2, pp. 648–670. DOI: 10.1121/1.1336503.

Zhao, F. and D. Stephens (2007). "A critical review of King-Kopetzky syndrome: Hearing difficulties, but normal hearing?" In: *Audiol. Med.* 5.2, pp. 119–124. DOI: 10.1080/16513860701296421.

Zilany, M. S. A., I. C. Bruce, P. C. Nelson, and L. H. Carney (2009). "A phenomenological model of the synapse between the inner hair cell and auditory nerve: long-term adaptation with power-law dynamics." In: *J. Acoust. Soc. Am.* 126.5, pp. 2390–412. DOI: 10.1121/1.3238250.

Zilany, M. S. A., I. C. Bruce, and L. H. Carney (2014). "Updated parameters and expanded simulation options for a model of the auditory periphery." In: *J. Acoust. Soc. Am.* 135.1, pp. 283–6. DOI: 10.1121/1.4837815.

Contributions to Hearing Research

Vol. 1: *Gilles Pigasse*, Deriving cochlear delays in humans using otoacoustic emissions and auditory evoked potentials, 2008.

Vol. 2: *Olaf Strelcyk*, Peripheral auditory processing and speech reception in impaired hearing, 2009.

Vol. 3: *Eric R. Thompson*, Characterizing binaural processing of amplitude-modulated sounds, 2009.

Vol. 4: *Tobias Piechowiak*, Spectro-temporal analysis of complex sounds in the human auditory system, 2009.

Vol. 5: *Jens Bo Nielsen*, Assessment of speech intelligibility in background noise and reverberation, 2009.

Vol. 6: *Helen Connor*, Hearing aid amplification at soft input levels, 2010.

Vol. 7: *Morten Løve Jepsen*, Modeling auditory processing and speech perception in hearing-impaired listeners, 2010.

Vol. 8: *Sarah Verhulst*, Characterizing and modeling dynamic processes in the cochlea using otoacoustic emissions, 2010.

Vol. 9: *Sylvain Favrot*, A loudspeaker-based room auralization system for auditory research, 2010.

Vol. 10: *Sébastien Santurette*, Neural coding and perception of pitch in the normal and impaired human auditory system, 2011.

Vol. 11: *Iris Arweiler*, Processing of spatial sounds in the impaired auditory system, 2011.

Vol. 12: *Filip Munch Rønne*, Modeling auditory evoked potentials to complex stimuli, 2012.

Vol. 13: *Claus Forup Corlin Jespersgaard*, Listening in adverse conditions: Masking release and effects of hearing loss, 2012.

Vol. 14: *Rémi Decorsière*, Spectrogram inversion and potential applications for hearing research, 2013.

Vol. 15: *Søren Jørgensen*, Modeling speech intelligibility based on the signal-to-noise envelope power ration, 2014.

Vol. 16: *Kasper Eskelund*, Electrophysiological assessment of audiovisual integration in speech perception, 2014.

Vol. 17: *Simon Krogholt Christiansen*, The role of temporal coherence in auditory stream segregation, 2014.

Vol. 18: *Márton Marschall*, Capturing and reproducing realistic acoustic scenes for hearing research, 2014.

Vol. 19: *Jasmina Catic*, Human sound externalization in reverberant environments, 2014.

Vol. 20: *Michał Fereczkowski*, Design and evaluation of individualized hearing-aid signal processing and fitting, 2015.

Vol. 21: *Alexandre Chabot-Leclerc*, Computational modeling of speech intelligibility in adverse conditions, 2015.

Vol. 22: *Federica Bianchi*, Pitch representations in the impaired auditory system and implications for music perception, 2016.

Vol. 23: *Johannes Zaar*, Measures and computational models of microscopic speech perception, 2016.

Vol. 24: *Johannes Käsbach*, Characterizing apparent source width perception, 2016.

Vol. 25: *Gusztáv Löcsei*, Lateralized speech perception with normal and impaired hearing, 2016.

Vol. 26: *Suyash Narendra Joshi*, Modelling auditory nerve responses to electrical stimulation, 2017.

Vol. 27: *Henrik Gerd Hassager*, Characterizing perceptual externalization in listeners with normal, impaired and aided-impaired hearing, 2017.

Vol. 28: *Richard Ian McWalter*, Analysis of the auditory system via synthesis of natural sounds, speech and music, 2017.

Vol. 29: *Jens Cubick*, Characterizing the auditory cues for the processing and perception of spatial sounds, 2017.

Vol. 30: *Gerard Encina-Llamas*, Characterizing cochlear hearing impairment using advanced electrophysiological methods, 2017.

The end.

To be continued...

The healthy auditory system can enable communication in complex acoustical scenarios. The cells responsible for transforming sound into electrical spikes that the brain can interpret degenerate endogenously or due to insult, leading to sensorineural hearing impairment. It is known that pure-tone audiometry does not entirely reflect damage of inner hair cells and/or the auditory nerve. This may lead to inaccurate diagnostics. Therefore, novel diagnostic methods capable of detecting all the tiny dysfunctions that today remain hidden are needed. This PhD project investigated the use of advanced electrophysiological methods to evaluate supra-threshold processing in the peripheral auditory system.

DTU Electrical Engineering
Department of Electrical Engineering

Ørsteds Plads
Building 348
DK-2800 Kgs. Lyngby
Denmark
Tel: (+45) 45 25 38 00
Fax: (+45) 45 93 16 34
www.elektro.dtu.dk